
Studying protein-DNA interactions *in vitro* and *in vivo* using single-molecule photoswitching

Stephan Uphoff

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



The Queen's College
University of Oxford
Michaelmas Term 2012

Disclaimer

I hereby declare that the work in this thesis is that of the candidate alone, except where indicated in the text, and as described below.

- Chapter 2 is an extended version of my publications (Uphoff et al., 2010) and (Uphoff et al., 2011) including their text and figures. Achillefs Kapanidis edited the text of these manuscripts.
- The single-molecule FRET data processing software (including image registration, PSF detection, and fitting) was developed by Seamus Holden and the data visualisation software by Kristofer Gryte.
- The dual-colour TIRF microscope setup for single-molecule FRET was designed and built by Ludovic Le Reste and Johannes Hohlbein.
- The confocal microscope for single-molecule FRET was designed and built by Geraint Evans and Johannes Hohlbein.
- Catabolite activator protein for single-molecule FRET experiments was produced and labelled by Javier Periz.
- Klenow fragment and DNA polymerase I for single-molecule FRET experiments were produced and labelled by Catherine M. Joyce and Timothy Craggs.
- The multi-colour TIRF microscope for super-resolution microscopy and three-colour single-molecule FRET was built in cooperation with Justin Pinkney.
- Fixed Cos-7 cells were provided by Argyris Papantonis.
- Fixed mouse embryonic fibroblast cells were provided by Chen Sok Lam.
- *E.coli* cell strains were generated by Rodrigo Reyes-Lamothe and Anjana Badrinarayanan.
- Chapter 5 is an extended version of my publication (Uphoff et al., 2013), including text and figures. Achillefs Kapanidis edited the text of this manuscript.
- Chapter 6 includes data and figures that I originally created for the publication by Badrinarayanan et al. (2012).

Stephan Uphoff
October 2012

STUDYING PROTEIN-DNA INTERACTIONS *in vitro* AND *in vivo* USING SINGLE-MOLECULE PHOTOSWITCHING

Stephan Uphoff, The Queen's College
(Department of Condensed Matter Physics)

Thesis submitted for the degree of Doctor of Philosophy
University of Oxford, Michaelmas Term 2012.

ABSTRACT

Protein-DNA interactions govern the fundamental cellular processes of DNA replication, transcription, repair, and chromosome organisation. Despite their importance, the detailed molecular mechanisms of protein-DNA interactions and their organisation in the cell remain elusive. The complexity of molecular biology demands new experimental concepts that resolve the structural and functional diversity of biomolecules.

In this thesis, I describe fluorescence methods that give a direct view on protein-DNA interactions at the single-molecule level. These methods employ photoswitching to control the number of active fluorophores in the sample. Förster Resonance Energy Transfer (FRET) measures the distance between a donor and an acceptor fluorophore to report on biomolecular structure and dynamics *in vitro*. Because a single distance gives only limited structural information, I developed “switchable FRET” that employs photoswitching to sequentially probe multiple FRET pairs per molecule. Switchable FRET resolved two distances within static and dynamic DNA constructs and protein-DNA complexes. Towards application of switchable FRET, I investigated aspects of the nucleotide selection mechanism of DNA polymerase.

I further explored application of single-molecule imaging in the complex environment of the living cell. Photoswitching was used to resolve the precise localisations of individual fluorophores. I constructed a super-resolution fluorescence microscope to image fixed cellular structures and track the movement of individual fluorescent fusion proteins in live bacteria. I applied the method to directly visualise DNA repair processes by DNA polymerase I and ligase, generating a quantitative account of their repair rates, search times, copy numbers, and spatial distribution in the cell. I validated the approach by tracking diffusion of replisome components and their association with the replication fork. Finally, super-resolution microscopy showed dense clusters of SMC (Structural Maintenance of Chromosomes) protein complexes *in vivo* that have previously been hidden by the limited resolution of conventional microscopy.

Acknowledgements

I am very grateful to my supervisor Achilles Kapanidis for his enthusiastic support and mentoring, visionary project ideas, and our exciting conference trips.

I would like to thank everyone in the Kapanidis group for the great atmosphere and companionship. Our office activities certainly provided a healthy balance to the solitary dark hours at the microscope. I would like to thank Seamus Holden for all the help when I first started in the lab and his invaluable contributions to our joint research projects. I much acknowledge David Sherratt's guidance on my journey from physics towards biology and the opportunities he gave me to present my research. I am very grateful to Rodrigo Reyes-Lamothe; I benefited much from his expertise in molecular biology. Ludovic Le Reste, Johannes Hohlbein, Geraint Evans, and Justin Pinkney are thanked for building magnificent microscopes. I acknowledge Sarah Matthews, Federico Garza de Leon, Anjana Badrinarayanan, Timothy Craggs, Catherine Joyce, Nigel Grindley, Javier Periz, Kristofer Gryte, Robert Crawford, Alexandra Tomescu, David Yadin, Chen Sok Lam, Argyris Papantonis, Sebastian van de Linde, and Mike Heilemann for providing samples, software, and help with experiments.

I owe my most sincere gratitude to The MathWorks for funding and particularly Jos Martin for supervision during my internships. The Queen's College is thanked for providing accommodation and funding to attend conferences. I am very grateful for my German National Academic Foundation doctorate fellowship and their support throughout my studies.

To Ana Ineyda Namburete: Thank you for your help and care.

An Rainer und Anette und Sonja: Ich bin so dankbar für eure Hilfe während meines ganzen Bildungsweges.

Stephan Uphoff,
Oxford, October 2012

Publications

- Uphoff, S., Reyes-Lamothe, R., Garza de Leon, F., Sherratt, D.J., Kapanidis, A.N. Single-molecule DNA repair in live bacteria. Proceedings of the National Academy of Sciences USA (2013 in press).
- Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M.C., Sherratt, D.J. Architecture and action of Structural Maintenance of Chromosome complexes in vivo. *Science* 338, 528-531 (2012).
- Holden, S.J., Uphoff, S., Kapanidis, A.N. DAOSTORM: A general method for super-resolution imaging of very dense samples. *Nature Methods* 8, 279-280 (2011).
- Uphoff, S., Gryte, K., Evans, G., Kapanidis, A.N. Improved temporal resolution and linked hidden Markov modeling to probe multiple switchable single-molecule FRET acceptors. *ChemPhysChem* 12, 571-579 (2011). With design of the journal cover.
- Holden, S.J., Uphoff, S. et al. Defining the limits of single-molecule FRET resolution in TIRF microscopy. *Biophysical Journal* 99, 3102-3111 (2010).
- Uphoff, S., Holden, S.J., Le Reste, L., Periz, J., van de Linde, S., Heilemann, M., Kapanidis, A.N. Monitoring multiple distances within a single molecule using switchable FRET. *Nature Methods* 7, 831-836 (2010). Highlighted in *Nature Chem. Biol.* and in *Chemical & Engineering News*.

Talks

- Super-resolution imaging: Bacteria and beyond. Microscopy and Analysis Conference, Bath, 2013.
- Measuring enzyme copy numbers and single-molecule reaction rates in living cells. American Society for Cell Biology Annual Meeting, San Francisco, 2012.
- Super-resolution microscopy in living cells. Postech International Symposium on Bioimaging, Pohang, 2012.
- Super-resolution imaging of replication and repair. Society for General Microbiology Autumn Conference, Warwick, 2012.
- Single-molecule DNA repair and replication in live bacteria. 73rd Harden Conference Machines on Genes II, Oxford, 2012.
- Single-molecule DNA repair in live bacteria. Biophysical Society Meeting, San Diego, 2012.

- Single-molecule DNA replication and repair in live bacteria, Department of Biochemistry Seminar, Oxford, 2012.
- Single-molecule DNA repair in live bacteria. Biological Physics Seminar, Oxford, 2012.
- Monitoring multiple distances within a single molecule using switchable FRET. Single-molecule Biophysics Meeting, Aspen, 2011.
- Physics meets biology, one molecule at a time. Queens College Symposium, Oxford, 2010.
- Measuring multiple distances within a single molecule using switchable FRET. Biological Physics Seminar, Oxford, 2010.
- Measuring single-molecule structure and dynamics using advanced FRET methods. Single-molecule biophysics Summer School, Leuven, 2010.
- Measuring multiple distances within a single molecule or protein-DNA complex. UK RNA Polymerase Workshop, Nottingham, 2010.
- Measuring multiple distances within a single molecule using switchable FRET. Biophysical Society Meeting, San Francisco, 2010.
- Measuring multiple distances within a single molecule using switchable FRET. PicoQuant Single-molecule Spectroscopy Workshop, Berlin, 2009.

Contents

1	Introduction	1
1.1	Protein-DNA interactions	2
1.2	Single-molecule fluorescence microscopy	8
1.3	Summary of the Project	17
2	Switchable FRET	19
2.1	Introduction	19
2.2	Proof-of-principle	22
2.3	Further development of switchable FRET	30
2.4	Applications of switchable FRET	40
2.5	Conclusion	46
2.6	Materials and methods	49
3	Conformational states of DNA polymerase I	57
3.1	Introduction	57
3.2	Binding of Klenow fragment to immobilised DNA	61
3.3	Conformational dynamics of Klenow fragment	63
3.4	Comparing Klenow fragment and full-length DNA polymerase I	63
3.5	Conclusion	67
3.6	Materials and methods	68

4	Super-resolution fluorescence microscopy	71
4.1	Introduction	71
4.2	Microscope design and construction	73
4.3	Stochastic Optical Reconstruction Microscopy (STORM)	79
4.4	Photoactivated Localization Microscopy (PALM) and single-particle tracking	89
4.5	Counting protein copy numbers	97
4.6	Conclusion	100
5	Single-molecule DNA repair in live bacteria	101
5.1	Introduction	101
5.2	DNA polymerase I and ligase in undamaged cells	107
5.3	Direct observation of DNA repair	113
5.4	DNA repair rates	117
5.5	Conclusion	129
6	Imaging DNA replication and chromosome organising proteins in live bacteria	131
6.1	Introduction	131
6.2	Organisation of the E.coli replisome	133
6.3	Imaging replisome components	136
6.4	Structural Maintenance of Chromosome proteins	141
6.5	Imaging MukBEF proteins	143
7	Conclusion and Outlook	149

List of abbreviations and important variables

ALEX: Alternating-Laser Excitation; AOTF: acousto optical tunable filter; AP site: apurinic or apyrimidinic DNA site; bp: basepair; BER: base excision repair; BME: β -mercaptoethanol; BSA: bovine serum albumin; CAP: catabolite activator protein; CDF: cumulative distribution function; δ : clamp loader subunit; D : accurate diffusion coefficient; D^* : apparent diffusion coefficient; dNTP: deoxyribonucleoside triphosphate; dsDNA: double-stranded DNA; ϵ : PolIII subunit; E : accurate FRET efficiency; E^* : apparent FRET efficiency; *E.coli*: *Escherichia coli*; EMCCD: Electron Multiplying Charged Coupled Device camera; F_{DD} : donor-excited donor emission; F_{DA} : donor-excited acceptor emission; F_{AA} : acceptor-excited acceptor emission; Fis: factor for inversion stimulation protein; FRET: Förster Resonance Energy Transfer; GFP: green fluorescent protein; KF: Klenow fragment; λ : wavelength; Lig: DNA ligase; LB: Luria Broth; M9: glycerol minimal medium with amino acids and vitamins; MEA: mercaptoethylamine; MMS: methyl methanesulfonate; MSD: mean-squared displacement; NA: numerical aperture; NER: nucleotide excision repair; nt: nucleotide; *oriC*: origin of replication; PAmCherry: photoactivatable mCherry; PALM: Photoactivated Localization Microscopy; PBS: phosphate-buffered saline; PEG: polyethyleneglycol; Pol: DNA polymerase I; Pol III: DNA polymerase III; PSF: point-spread-function; R_0 : Förster radius; S : stoichiometry; sem: standard error of the mean; SMC: Structural Maintenance of Chromosomes; ssDNA: single-stranded DNA; std: standard deviation; STORM: Stochastic Optical Reconstruction Microscopy; *ter*: replication terminus region; TIRF: Total Internal Reflection Microscopy.

Chapter 1

Introduction

Current progress in the life sciences is driven to a large extent by method development. Novel experimental approaches promise to uncover the fundamental mechanisms of life that have been hidden within biological complexity. Heterogeneity is at the heart of biological complexity. Even seemingly homogeneous biological ensembles show variation between their individual components, ranging from phenotypic differences between clonal cells to structural heterogeneity of chemically identical molecules. Traditional ensemble experiments measure time- and population-averaged properties that can not directly report on stochastic and transient processes. These limitations are now being overcome by single-cell and single-molecule methods that resolve the individual players in the ensemble. In this thesis, I discuss the development of single-molecule fluorescence microscopy methods to study the structure and function of proteins and their complexes with DNA. A central feature of my work is the use of photoswitching, the activation and deactivation of fluorophores by light, to separate detection of single molecules over time. Chapter 2 is about switchable FRET, a novel method to investigate molecular structure and dynamics by monitoring multiple distances within a single molecule. Switchable FRET relies on photoswitching

to measure different fluorophores in a sequential manner. The method is demonstrated on several synthetic DNA constructs and a protein-DNA complex. The following chapter 3 investigates aspects of the structure and conformational dynamics of DNA polymerase I in complex with DNA. The results provide a foundation for future application of switchable FRET to this system. The concepts and instrumentation behind switchable FRET, namely single-molecule imaging and photoswitching, are also central to super-resolution fluorescence microscopy methods. Photoswitching controls the number of fluorescent molecules within a cell, enabling the localisation of isolated individual molecules with nanometre resolution, thus overcoming the optical resolution limit. Chapter 4 shows my implementation and further development of this technique. The following chapter 5 describes my application of super-resolution microscopy to directly observe DNA repair activities of DNA polymerase I and ligase, using the bacterium *Escherichia coli* as a model system. This approach provides a quantitative description of a reaction pathway *in vivo*. Finally, I focus on bacterial DNA replication and chromosome organising proteins in chapter 6, confirming the performance and versatility of the super-resolution microscopy method.

The following introduction provides the basic biological background and methodological concepts relevant to the whole thesis, while more detailed information is given at the beginning of each chapter.

1.1 Protein-DNA interactions

DNA

The genetic information of all cellular organisms is encoded in deoxyribonucleic acid (DNA). DNA consists of two linear polymer strands, each containing a sequence of nucleotide (nt) units. The DNA backbone links adjacent units via

phosphodiester bonds between the C-3 and C-5 carbons of their deoxyribose sugar rings. The asymmetry of the 5' C-5-phosphate terminus and 3' C-3-OH terminus gives DNA strands a chemical direction, with antiparallel arrangement of the twin strands. A glycosidic bond from the sugar ring attaches one of the four nucleobases, the purines adenine (A) and guanine (G), or the pyrimidines thymine (T) and cytosine (C). X-ray crystallography revealed the famous double-helical structure of the twin DNA strands that hold together by hydrogen bonding between the nucleobases, such that the paired bases are on the inside of the polymer and the backbone chains face outwards (Watson and Crick, 1953). Adenine links to thymine via two hydrogen bonds, and guanine links to cytosine via three hydrogen bonds. The duplex has a diameter of 2 nm, twists at a periodicity of 3.4 nm with a helical rise of 0.34 nm per nucleotide. The two strands are closer on one side of the helix than on the other, forming a major and a minor helical groove that expose the base pairs (bp).

The double-stranded DNA structure provides chemical stability for long-term information storage in the linear sequence of A,G,T,C letters. However, spontaneous decay of glycosidic bonds, oxidation or alkylation of bases, as well as enzymatic digestion give rise to several thousand DNA lesions per human cell per day (Lindahl, 1993). The genomic stability is safeguarded over evolutionary time scales by DNA repair mechanisms; the subject of chapter 5. The DNA structure also reveals how genetic information can be readout and duplicated. Watson and Crick (1953) wrote “It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material”. In DNA replication, the complementary sequences of the two DNA strands each serve as a template for two identical DNA copies. Separation of the double helix exposes the bases and allows a DNA polymerase enzyme to synthesise the new strand according to the template sequence as dic-

tated by base pairing. Chapters 3 and 5 explore functional aspects of a model DNA polymerase *in vitro* and *in vivo*; while chapter 6 focuses on the DNA replication machinery.

Cellular DNA is organised into highly compacted chromosomes. Physical effects including DNA super-coiling and phase-separation between DNA and other molecules in the cytoplasm drive global chromosome condensation (Woldringh and Nanninga, 2006), while particular DNA-binding proteins wrap, bend, and clamp DNA segments for specific compaction and regulation of DNA transactions (Dillon and Dorman, 2010; Nasmyth and Haering, 2005). The second part of chapter 6 focuses on such Structural Maintenance of Chromosomes complexes.

Proteins

Genetic information is stored in units of genes that instruct protein synthesis. Proteins are amino acid polymers (polypeptides) with a primary role in cellular processes, serving enzymatic, structural, and regulatory functions. Transcription of genes by RNA polymerase generates messenger ribonucleic acid (mRNA) as a template for protein synthesis during translation. In most organisms, the genetic code instructs for 20 different amino acids, whose arrangement in a sequence is specific to each protein type. The function of a protein is governed by the structure of the polypeptide chain fold and dynamic conformational changes. To fulfil their diverse function, proteins interact and form complexes with other proteins, nucleic acids, and small ligands, often in a modular fashion. Protein-DNA interactions are paramount to the main cellular processes of replication, transcription, and chromosome maintenance. Proteins like transcription factors, recombinases, and restriction endonucleases recognise specific DNA sequences. They likely employ a combination of three-dimensional diffusion and one-dimensional sliding on DNA to find a particular sequence within the vast excess of chromosomal

DNA (Halford and Marko, 2004; Elf et al., 2007; Tafvizi et al., 2011; Hammar et al., 2012). DNA binding proteins can bend, loop, and distort the DNA structure to perform their function or flag recruitment of other proteins (Kapanidis et al., 2001; Dillon and Dorman, 2010). DNA nicks, gaps, breaks, mismatches, and chemical lesions are recognised by DNA repair proteins such as DNA polymerases, ligases, nucleases, and glycosylases (Friedberg, 2006). Despite extensive research, folding and function of proteins can not be fully predicted from the gene sequence. Chapter 2 introduces a novel experimental method to probe the structure and dynamics of purified nucleic acids, proteins, and their complexes in solution (Uphoff et al., 2010, 2011).

However, proteins do not function in isolation but within the crowded cellular environment. The spatial distribution and activity of proteins is dictated by interactions with other molecules and changes according to the cell cycle, and cellular responses to the environment. Chapters 5 and 6 explore protein-DNA interactions that are key to DNA repair, replication, and chromosome organisation in the context of the living cell.

Diffusion

Molecular collision by diffusion constitutes the first step of general bi-molecular reactions, including protein-DNA interactions. The fundamental mode of diffusion is Brownian motion; the random movement of particles due to stochastic collisions with other particles. Here, I briefly discuss the basic laws of Brownian motion that are central to this thesis, in particular the mean-squared displacement formula.

The average force of viscous resistance with drag coefficient γ on a moving particle with velocity v is given by Stoke's law $F = -\gamma v$. The simplest model for stochastic forces by random molecular collisions adds to this average force a noise

term $f(t)$ whose time average is zero ($\langle f(t) \rangle = 0$) and temporal correlation is short-lived compared to the observation time scale: $\langle f(t)f(t+\tau) \rangle = \lambda\delta(\tau)$. The coefficient λ corresponds to the amplitude of the fluctuations. The Langevin equation describes the motion of a particle with mass m under the stochastic force $-\gamma v + f(t)$, equivalent to Newton's law:

$$m \frac{dv}{dt} = -\gamma v + f(t) . \quad (1.1)$$

Solving this equation in the one-dimensional case shows that the average velocity decays exponentially: $\langle v(t) \rangle = v_0 \cdot \exp(-\gamma \cdot t/m)$. The mean-squared velocity, however, approaches

$$\langle v^2(t) \rangle = \frac{\lambda}{2m\gamma} . \quad (1.2)$$

On the other hand, statistical mechanics states that the average kinetic energy is related to the thermal energy $k_B T$ in equilibrium:

$$\frac{1}{2} m \langle v^2(t) \rangle = \frac{1}{2} k_B T . \quad (1.3)$$

Comparing these two equations gives the fluctuation-dissipation relation between the amplitude λ of the fluctuating forces and the viscous drag γ :

$$\lambda = 2\gamma k_B T . \quad (1.4)$$

A more experimentally accessible observable than the mean-squared velocity is the mean-squared position, obtained by integrating $\langle v^2(t) \rangle$:

$$\langle x^2(t) \rangle = 2 \frac{k_B T}{\gamma} t = 2Dt . \quad (1.5)$$

The mean-squared displacement of a particle from the original position increases

linearly in time with diffusion coefficient $D = k_B T / \gamma$. The n -dimensional case has an extra factor n . Using the drag coefficient, $\gamma = 6\pi\eta R$, of a medium with viscosity η on a sphere with radius R gives the Einstein-Stokes relation:

$$D = \frac{k_B T}{6\pi\eta R} . \quad (1.6)$$

The probability distribution of a particle to move from $x = 0$ to position x in time t is given by:

$$p(x, t) = \frac{1}{4\pi Dt} \cdot \exp\left(\frac{-x^2}{4Dt}\right) . \quad (1.7)$$

Assuming a cellular cytoplasm with viscosity $\eta = 0.01 \text{ Pa} \cdot \text{s}$ at 20°C , a protein with $R \sim 10 \text{ nm}$ has a diffusion coefficient of $D \sim 2 \mu\text{m}^2\text{s}^{-1}$. Diffusion is hence an efficient mode of transport in *E.coli* cells with a length of 2-4 μm .

Direct association of a protein with a particular interaction partner is governed by the diffusion coefficients of both partners. For protein-DNA interactions, however, diffusion of DNA loci with $D \sim 10^{-5} \mu\text{m}^2\text{s}^{-1}$ inside bacteria (Elmore et al., 2005) is negligible compared to the diffusion of most proteins. The diffusion-limited search time of a single protein for N stationary DNA sites with association radius a in a volume V is given by (Halford and Marko, 2004):

$$t_{\text{search}} = \frac{V}{4\pi DaN} . \quad (1.8)$$

Using $V_{E.coli} = 10^{-18} \text{ m}^3$; $D = 2 \mu\text{m}^2\text{s}^{-1}$ and $a = 0.5 \text{ nm}$, gives a typical search time for a single DNA site of about 80 s.

These theoretical considerations provide a framework for the analysis of protein motion and interaction rates. However, because of the various specific and non-specific interactions and irregular shapes of molecules in the crowded and heterogenous cellular environment, protein diffusion coefficients and hence inter-

action rates can not be theoretically predicted. Chapters 4, 5, and 6 describe single-particle tracking experiments to measure diffusion coefficients and target search times of proteins in live *E.coli*.

1.2 Single-molecule fluorescence microscopy

Single-molecule approaches

The diverse chemical composition and modular assembly of biological polymers are at the heart of biological complexity. Even proteins or nucleic acids with identical amino acid or base sequences can adopt very different structures that interconvert stochastically. Conventional ensemble experiments neglect biological heterogeneity by measuring population-averaged properties. Uncovering the whole distribution of an observable within a population by observing individual molecules isolated in space or over time is the major achievement of single-molecule techniques (reviewed in Xie and Lu (1999); Selvin and Ha (2008); Vale (2008); Kapanidis and Strick (2009); van Oijen and Loparo (2010)). The probabilistic nature of chemical processes dictates that even artificially synchronised systems will diverge quickly when propagating through a series of reaction steps. Single-molecule techniques can extract time-trajectories of individual molecules to directly monitor complex kinetics and dynamics and to capture transient or rare intermediates. Single-molecule method development was the driving force behind pioneering discoveries, such as patch clamp recordings on single ion channels (Neher and Sakmann, 1976), direct observation of the rotation of F1-ATPase motor proteins (Noji et al., 1997), laser trap experiments revealing step sizes of single kinesin motor proteins (Svoboda et al., 1993), to name just a few. In particular, fluorescence microscopy has proven a powerful and versatile tool for measuring the motion of single fluorescently-labelled molecules (Vale et al., 1996;

Yildiz et al., 2003). Detection of single fluorescent nucleotide analogues is also central to a new generation of DNA sequencing approaches (Eid et al., 2009). The distance-dependence of Förster Resonance Energy Transfer (FRET) can be exploited to measure the separation between a donor and an acceptor fluorophore at the single-molecule level (Ha et al., 1996), directly reporting on biomolecular structure and dynamics (reviewed in Clegg (1992); Selvin (2000); Kapanidis and Strick (2009)). Chapter 2 shows an extension of single-molecule FRET to monitor multiple distances within a single molecule (Uphoff et al., 2010, 2011). Molecular heterogeneity is particularly pronounced in the complex and fluctuating environment of the cell. Many cellular processes are not dictated by average properties of large molecular ensembles, but involve few proteins or reaction sites (Xie et al., 2008). For instance, gene activation by an individual transcription factor can drastically change the cellular phenotype (Choi et al., 2008). Characterising the properties of single biomolecules *in vivo* is hence important for a molecular understanding of cell function. However, rendering single-molecule techniques compatible with live-cell application is challenging due to the lack of control over physiological conditions and difficulty of delivering probes into the cell. Initially, single-molecule imaging *in vivo* was limited to fluorescent proteins fused to slowly diffusing membrane proteins (Deich et al., 2004; Yu et al., 2006). Improved fluorescence detection sensitivity and imaging modalities now allow tracking the movement of single fluorescent proteins in the cytoplasm (Elf et al., 2007; English et al., 2011). The second part of this thesis exploits photoswitching of fluorophores for live-cell super-resolution fluorescence microscopy.

Resolution of optical microscopy

Optical microscopy is one of the most widely used tools in biological research because of the ability to observe specimen with minimal perturbation. Yet,

the wave-nature of light causes interference which limits the resolution of optical microscopy. The finite aperture of the microscope objective collects only light up to a maximum angle θ , restricting the extent of the diffraction pattern in the back focal plane of the objective. High spatial frequencies of the diffraction pattern are omitted, thus prohibiting a sharp reconstruction of the sample image, equivalent to the principle of a lowpass filter in optics or electronics. Consequently, a point source of light with wavelength λ in the sample appears as a blurred point-spread-function (PSF) in the microscope image. The resolution limit of a microscope is defined by the minimum distance of two point sources in the sample, such that their PSFs are separately resolved in the microscope image

$$\Delta x = \frac{1.22\lambda}{2 \cdot \text{NA}} . \quad (1.9)$$

The numerical aperture $\text{NA} = n \cdot \sin \theta < 1.5$ depends on the emission angle and the refractive index n between the sample and the objective. The resolution is therefore fundamentally limited to about half the wavelength of the observed light, i.e. > 250 nm for fluorescence applications.

Single-molecule fluorescence microscopy

Fluorescence microscopy employs fluorophores to specifically image molecules of interest with high contrast. Single-molecule sensitivity is reached by (i) maximising photon collection using high numerical aperture objectives with immersion oil, (ii) minimising loss of signal with high-quality optics and detectors, and (iii) reducing background noise with a small excitation or detection volume and laser excitation tailored to the fluorophore spectrum. Total Internal Reflection Fluorescence (TIRF) microscopy (Axelrod, 2001) limits excitation to fluorophores at the surface of the microscope slide. An evanescent optical field is generated at

the glass-water interface where the refractive index changes from $n_{\text{glass}} = 1.50$ to $n_{\text{water}} = 1.33$ if the excitation angle exceeds

$$\alpha > \sin^{-1} \left(\frac{n_{\text{water}}}{n_{\text{glass}}} \right) = 62.5^\circ . \quad (1.10)$$

TIRF microscopes typically employ a CCD camera with an electron-multiplying gain register for reduced readout noise. Confocal microscopy, on the other hand, uses a focused laser beam for excitation combined with a pinhole and an avalanche photodiode (APD) detector in the emission path to omit background noise from outside the focal volume.

Fluorophore types

The main characteristics that determine the choice of fluorophores for single-molecule fluorescence microscopy are the excitation and emission spectra, brightness, photostability, special photophysical properties, molecule size, and labelling procedures. Synthetic fluorophores such as cyanine and oxazine dyes are very small, bright, and range from the blue to the far red spectrum (e.g. Alexa488, Cy3, Cy5, Atto655, Cy7). Maleimide or NHS-ester modifications allow site-specific conjugation to proteins at cysteine residues or amino-modified DNA nucleotides, which is particularly important for FRET applications. However, these labelling procedures can not be directly performed *in vivo*, although live-cell conjugation techniques are under development (van de Linde et al., 2012). Fluorescent antibodies can be used to stain specific proteins in fixed and permeabilised cells. Discovery of the green fluorescent protein (GFP) revolutionised cell biology research with the possibility to genetically replace any protein by a fluorescent fusion with little change to the functionality and abundance in live cells (Tsien, 1998). Compared to synthetic fluorophores, however, fluorescent

proteins are less bright and photostable, far larger and can usually only label the C- or N-terminus of a protein.

Photophysics

Photophysical processes of fluorophores involve stochastic transitions between different quantum states. The mean lifetime of a state is the inverse of the sum of all rates that depopulate this state. Excitation of a fluorophore is usually achieved by absorption of a photon that matches the energy difference between the ground state and first excited singlet state of a fluorophore. The main discrete quantum states are superimposed with higher vibrational and rotational states that blur the energy gap, giving rise to a smooth excitation spectrum with a tail towards shorter wavelengths. The extinction coefficient describes the absorption of a fluorophore at a particular wavelength. Typical fluorophores spend several nanoseconds in the excited state before relaxation occurs through one of several competing pathways including vibrational relaxation, quenching by other molecules, Förster resonance energy transfer (FRET), or fluorescence emission. Bright fluorophores are characterised by a high fluorescence quantum yield, the fraction of emitted and absorbed photons. Relaxation into higher vibrational states of the ground state broadens the fluorescence spectrum towards longer wavelengths. The Stokes shift between the excitation and emission spectra allows filtering the weak fluorescence signal from the intense excitation light for microscopy and spectroscopy applications.

Instead of direct transition to the ground state, fluorophores can enter a triplet state of unpaired electron spins with a long lifetime of several microseconds. Persistent population of triplet states decreases the fluorescence rate and increases the likelihood of photobleaching by an irreversible chemical reaction from higher excited states (Eggeling et al., 2006). Triplet quenching chemicals like Trolox (6-

hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid), β -mercaptoethanol (BME), mercaptoethylamine (MEA), or combinations of reducing and oxidising agents (e.g. ascorbic acid and methylviologen) reduce triplet state lifetimes of synthetic fluorophores (Rasnik et al., 2006; Vogelsang et al., 2008; Cordes et al., 2009; Altman et al., 2011). Molecular oxygen is also an efficient triplet quencher but this reaction creates radical singlet oxygen which readily photobleaches the nearby fluorophore. Single-molecule fluorescence imaging *in vitro* hence relies on oxygen removal and addition of an alternative triplet quencher. The most common oxygen scavenging system employs a combination of glucose oxidase, glucose, and catalase (Selvin and Ha, 2008; Aitken et al., 2008). The reducing environment of the cell also lowers photobleaching rates (van de Linde et al., 2012).

Photoswitching and photoactivation

This thesis evolves around applications of photoswitching, the controlled activation and deactivation of fluorophores by light. Characterisation of such photophysical processes has been driven mainly by the realm of super-resolution fluorescence microscopy. Red cyanine dyes convert from their excited state into a long-lived dark state through a thiol addition to the polymethine bridge by agents like MEA or BME. Excitation in the UV to green spectral region recovers the fluorescent state, permitting reversible repetition of the photoswitching cycle (Bates et al., 2005; Heilemann et al., 2005; Dempsey et al., 2009). The photoswitching rates depend linearly on the excitation power, suggesting that the conversion reactions are driven by absorption of a single photon. Reactivation was strongly enhanced by excitation of a resonant donor fluorophore in close proximity (Bates et al., 2005).

A general redox reaction system for photoswitching of most synthetic fluorophore across the spectral range has been developed (Vogelsang et al., 2009). A reducing

agent like ascorbic acid reacts with the triplet state to create a non-fluorescent radical anion state. Subsequent oxidation by methylviologen or molecular oxygen recovers the singlet ground state. Oxidation and reduction potentials of the redox system are chosen to selectively react with the triplet and radical states, such that the photoswitching rates can be controlled by the excitation intensity and the redox concentrations. High reducer and oxidiser concentrations can be used to suppress blinking and photobleaching by reducing triplet state lifetimes (Vogelsang et al., 2008).

Fluorescent proteins with photoswitching capabilities have been developed by substituting protein residues of GFP and its variants. Photoactivatable fluorescent proteins such as PA-GFP (Patterson and Lippincott-Schwartz, 2002) or PAmCherry (Subach et al., 2009b) irreversibly convert from an initial non-fluorescent state to a fluorescent state upon irradiation with 405 nm light. Crystal structures of PAmCherry in the dark and bright states show that the photoactivation mechanism involves decarboxylation of a glutamic acid residue which results in the double-bond formation that completes the delocalised electron system of the fluorophore (Subach et al., 2009a). Other photoconvertible fluorescent proteins change emission spectra or switch reversibly between dark and bright states (reviewed in Lippincott-Schwartz and Patterson (2009)).

Photoswitching provides a general mechanism to reduce the concentration of active fluorophores without changing the concentration of labelled molecules. This concept is central to super-resolution microscopy techniques (Betzig et al., 2006; Huang et al., 2009), photochromic FRET imaging (Mao et al., 2008), and *in vitro* protein-binding assays (Loveland et al., 2012). Most applications demand a high intensity contrast between the bright and dark states and control over their lifetimes.

Förster Resonance Energy Transfer

Förster Resonance Energy Transfer (FRET) is a non-radiative interaction between a donor and an acceptor fluorophore (Förster, 1948). The electronic states of the excited donor and relaxed acceptor couple via a dipole-dipole interaction that scales with the inverse sixth power of the donor-acceptor separation R . The additional de-excitation pathway reduces the lifetime τ_D of the unquenched donor. The energy transfer rate is given by

$$k_T = \frac{1}{\tau_D} \left(\frac{R_0}{R} \right)^6 . \quad (1.11)$$

The Förster Radius R_0 is characteristic to a particular donor-acceptor pair and environment, depending on the overlap integral of the donor emission spectrum $f_D(\lambda)$ and the acceptor absorption spectrum $\epsilon_A(\lambda)$:

$$R_0 = \frac{9000 \cdot \ln(10) \cdot \phi_D \cdot \kappa^2}{128\pi^5 \cdot N_A \cdot n^4} \cdot \int_0^\infty f_D(\lambda) \epsilon_A(\lambda) d\lambda . \quad (1.12)$$

Here, N_A is the Avogadro number, n is the refractive index of the medium, and ϕ_D is the donor quantum yield. κ^2 describes the relative orientation of the donor-acceptor dipole moments, varying between 0 for crossed dipoles to 4 for parallel dipoles; isotropic rotation of both fluorophores gives $\kappa^2 = 2/3$. Following energy transfer, the excited acceptor dissipates excess vibrational energy and returns to the ground state by photon emission depending on the quantum yield. The FRET efficiency E is the probability of energy transfer following donor excitation, such that $E = 0.5$ for $R = R_0$

$$E = \frac{k_T}{k_T + (1/\tau_D)} = \frac{1}{1 + (R/R_0)^6} . \quad (1.13)$$

Application of FRET

The steep distance-dependence of the FRET efficiency for donor-acceptor separations close to R_0 can be exploited to measure nanometre distances and sub-nanometre distance changes within the molecular frame of reference at the ensemble level (Lakowicz, 2006) and the single-molecule level (Ha et al., 1996; Deniz et al., 1999; Roy et al., 2008; Holden et al., 2010). The common FRET pair Cy3B-Cy5 has a Förster radius of 6.7 nm (Uphoff et al., 2010), ideal for the size of many proteins. Single-molecule FRET is commonly measured by the donor and acceptor emission intensities F_{DD} and F_{DA} upon donor excitation. This gives the apparent FRET efficiency

$$E^* = \frac{F_{DA}}{F_{DD} + F_{DA}} . \quad (1.14)$$

Measurement of E^* is sufficient for most single-molecule FRET applications to monitor relative distance changes. Conversion of E^* into distances, however, requires measurement of R_0 including the rotational freedom of fluorophores and corrections for cross-talk between signals and detection biases (by measuring the detection-correction factor γ) to obtain the accurate FRET efficiency E (Lee et al., 2005). Structural interpretation of FRET distances further involves modeling of the fluorophore linkers (Woźniak et al., 2008; Muschielok et al., 2008). Incomplete labelling and photobleaching typically result in donor-only molecules with low apparent FRET efficiency. Conventional FRET measurements can not distinguish between these species and intact FRET pairs with a large donor-acceptor separation. Furthermore, FRET changes need not necessarily originate from distance changes but can also result from changes in photophysical characteristics of the fluorophores. Additional detection of acceptor fluorescence F_{AA} upon direct acceptor excitation identifies molecules with an intact FRET pair.

Alternating-laser excitation (ALEX) allows measurement of donor-acceptor stoichiometry and FRET efficiency almost simultaneously by switching rapidly between donor and acceptor excitation (Kapanidis et al., 2004). The stoichiometry ratio S is defined as

$$S = \frac{F_{\text{DD}} + F_{\text{DA}}}{F_{\text{DD}} + F_{\text{DA}} + F_{\text{AA}}} . \quad (1.15)$$

Donor-only molecules are characterised by $S \sim 1$ and acceptor-only molecules by $S \sim 0$. Depending on the relative excitation intensities, donor-acceptor molecules show intermediate stoichiometry values.

1.3 Summary of the Project

The aim of this thesis is the description of single-molecule fluorescence methods for the investigation of protein-DNA interactions *in vitro* and *in vivo*. These methods rely on photoswitching to control the number of active fluorophores. The project is situated at the interface of physics, chemistry, and biology. Physics relates to the development of instrumentation and data analysis methods that are required to detect, quantify, and manipulate the faint fluorescence signal of single molecules. I used chemical techniques for sample preparation, labelling reactions, and particular buffers to control fluorophore properties. Finally, the methods demonstrated in this thesis were developed to address central problems in molecular biology. The multi-disciplinary nature of my project relied on co-operations, most notably with Seamus J. Holden in Achillefs Kapanidis' group and Rodrigo Reyes-Lamothe in David Sherratt's group, and other contributions that are mentioned in the text.

Chapter 2

Switchable FRET

2.1 Introduction

The function of biomolecules is closely linked to their structure. Performing structural studies on single molecules allows the observation of heterogeneity between molecules, such as structural differences between static or slowly interconverting molecular subpopulations, dynamic heterogeneity associated with structural changes, and transient molecular interactions. The fundamental role of heterogeneity in biological systems (Tawfik, 2010) demands methods that maximise information from individual molecules. Fluorescence techniques have unique properties in this context, owing to labelling-specificity and instrumentation that allows measurement of single fluorophores in aqueous solution at room temperature. In particular, single-molecule FRET is widely used to probe biomolecular structure and dynamics by exploiting the distance-dependence of the dipole-dipole interaction between an excited donor and an acceptor fluorophore. Formation of detailed structural models or complex reaction pathways requires multiple FRET measurements with donors and acceptors attached at different positions (Mekler et al., 2002; Andrecka et al., 2008; Muschielok et al.,

2008; Woźniak et al., 2008; Balci et al., 2011). Yet, combining information from different measurements on heterogeneous systems is not straightforward; to address this challenge, multi-colour FRET methods have been developed that monitor multiple FRET pairs per molecule (Hohng et al., 2004; Clamme and Deniz, 2005). Moreover, improved understanding of photophysics and their control allows uninterrupted measurements of single immobilised organic fluorophores over minutes (Rasnik et al., 2006), permitting direct, real-time observation of dynamic heterogeneity, e.g., reaction cycles and non-equilibrium processes. Yet, multi-colour FRET techniques are technically and conceptually limited. Spectral overlap mixes the fluorescence signals and leads to complex energy transfer cascades which prevents direct interpretation of FRET efficiencies in terms of distances (Lee et al., 2007, 2010; Stein et al., 2011). Moreover, site-specific labelling with particular fluorophores is challenging, especially for multimeric proteins.

I have developed “switchable FRET”, the combination of reversible photoswitching of fluorophores (Bates et al., 2005; Heilemann et al., 2005) and single-molecule FRET to sequentially probe multiple donor-acceptor pairs within an individual molecule (Uphoff et al., 2010). Photoswitching separates the FRET interactions of different acceptors with a shared donor fluorophore over time. Switchable FRET complements the toolbox for probing structure and dynamics of individual molecules from more than one perspective, especially in heterogeneous systems (Fig. 2.1). I demonstrated the principle of switchable FRET using labelled DNA fragments and tested whether the concept is generally compatible with TIRF and confocal microscopy modalities, as well as various types of fluorophores (cyanine Cy5 and oxazine ATTO655) that undergo switching via different mechanisms (Heilemann et al., 2009b; Vogelsang et al., 2009). I optimised photoswitching conditions of cyanine fluorophore Cy5 for application in switchable FRET to increase the temporal resolution of the method by more than

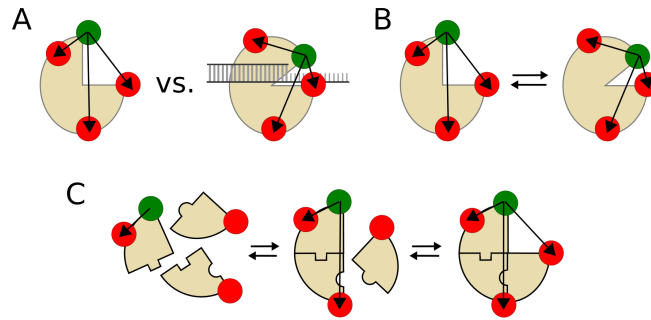


Figure 2.1: Multiple FRET pairs provide insight into heterogeneous biological systems from multiple perspectives. **(a)** Static structural heterogeneity; for example, distinct protein conformations in a ligand-bound and free form. **(b)** Dynamic structural heterogeneity; for example, conformational fluctuations associated with enzymatic function. **(c)** Dynamic and static structural heterogeneity; for example, transient formation and dissociation of multimolecular complexes.

an order of magnitude (Uphoff et al., 2011). This progress allows the method to be tuned to a wide range of time scales and facilitates application to various biological processes that occur on the sub-second time scale. To demonstrate the broad biological compatibility of switchable FRET, I probed two distances within a protein-DNA complex. Furthermore, I used switchable FRET to directly monitor conformational changes within single dynamic DNA Holliday junctions from two perspectives. This chapter presents a continuation of my MSc research project. For completeness, I included results (indicated in the figure caption) that I previously presented in my MSc thesis.

2.2 Proof-of-principle

Periodic switching

In switchable FRET, immobilised molecules labelled with a single donor and multiple identical acceptors are probed over time by reversibly switching acceptors between a dark state with an absorption spectrum that does not overlap with the donor emission spectrum (and does not participate in FRET) and a fluorescent state that participates normally in FRET. This concept promises to resolve multiple FRET pairs at different positions on a single molecule in a temporal fashion. Reversible switching can interrogate each pair many times for the study of structural changes within a molecule. In principle, switchable FRET can be realised either in a periodic or a stochastic manner. I first describe periodic photoswitching of Cy5, a popular cyanine fluorophore for super-resolution microscopy and single-molecule FRET spectroscopy; Cy5 can be photoswitched by cycling red (635 nm) and green (532 nm) excitation (Fig. 2.2) (Bates et al., 2005; Heilemann et al., 2005). In each cycle of periodic switchable FRET, red excitation switches all acceptors off, after which green excitation activates a single acceptor and probes its particular FRET efficiency. Different acceptors can thus be measured by sequential cycling.

Using TIRF microscopy¹, I studied immobilised 55 bp double-stranded DNA (ds-DNA) molecules, specifically labelled with a single Cy5 acceptor separated by 10 bp from a Cy3B donor (sequences and modifications of the DNA oligonucleotides are shown in Fig. 2.16). Fluorescence from the donor and the acceptor emission spectrum were split into two detection channels on an emCCD camera (Margeat et al., 2006) to extract the single-molecule fluorescence signals F_{AA} (Cy5 emission

¹The TIRF microscope for switchable FRET experiments was built by Ludovic Le Reste and Johannes Hohlbein; Materials and Methods section 2.6.

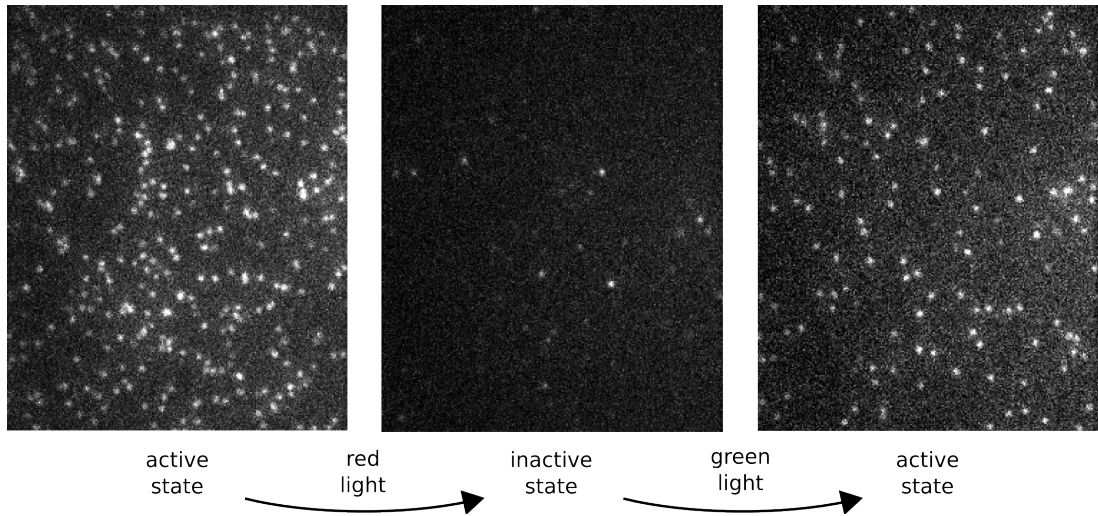


Figure 2.2: Photoswitching of Cy5 (previously presented in my MSc thesis). Single immobilised dsDNA molecules labelled with Cy5 appear as bright spots in the TIRF microscopy image. At the beginning of the experiment (left), the number of active Cy5 molecules is high. During the following 100 seconds of 635 nm excitation in the presence of MEA, almost all fluorophores get switched off (middle). Many Cy5 molecules can be restored to the active state after 5 seconds of additional 532 nm excitation.

upon 635 nm excitation), F_{DD} (Cy3B emission upon 532 nm excitation), and F_{DA} (Cy5 emission upon 532 nm excitation). Cycles of 9 s 635 nm excitation and 1 s 532 nm excitation reversibly switched the majority of Cy5 molecules off and on (Fig. 2.3A). Deactivation was very slow compared to activation, the latter occurring almost instantaneously for all acceptors during green excitation intervals (green highlighted segments, Fig. 2.3A). Irreversible photobleaching is apparent from the gradual decay of peak numbers of molecules in each cycle (~ 250 molecules at 5 s; ~ 130 molecules at 85 s). The fluorescence time-trace of a single Cy5 from the same dataset shows step-like appearance (~ 600 counts/100 ms) and disappearance of fluorescence emission which marks reversible photoswitching at the single-molecule level (Fig. 2.3B). Most single Cy5 activation events coincide with the beginning of a green excitation pulse, during which I observed Cy5 acceptor fluorescence owing to FRET by the nearby Cy3B donor (~ 250 counts/100 ms, green background). Thermal activation and residual green in-

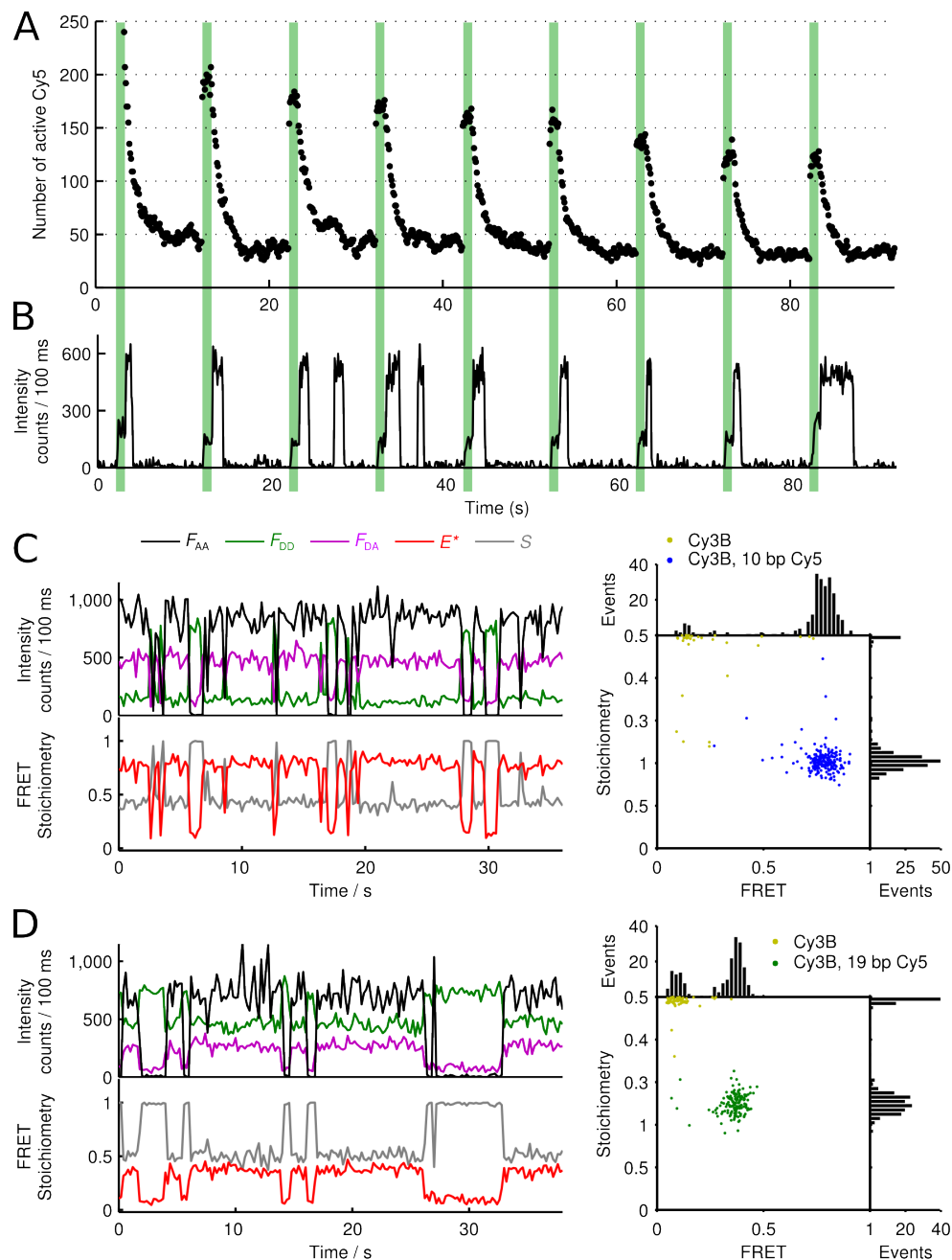


Figure 2.3: Reversible photoswitching of Cy5. **(A)** Periodic switching. Immobilised dsDNA molecules labelled with Cy5 and Cy3B at 10 bp separation were measured by cycles of 9 s red excitation (white background) and 1 s green excitation (green background) using TIRF microscopy. A localisation algorithm determined the number of active fluorescent Cy5 molecules in each 100 ms frame within the field of view. **(B)** Fluorescence intensity time-trace of an individual Cy5 molecule from the sample in (A). **(C, D)** Stochastic photoswitching. TIRF microscopy with 200 ms ALEX period induced stochastic and reversible photoswitching of Cy5 at (C) 10 bp and (D) 19 bp separation from Cy3B on immobilised dsDNA. Time-traces of fluorescence intensities F_{DD} (green line), F_{DA} (purple line), and F_{AA} (black line) are shown in the top graphs, corresponding FRET efficiency E^* (red line) and stoichiometry S (gray line) are plotted below. Data from individual time-traces (left) were plotted on E^* - S histograms for a single molecule (right). The two clusters in each histogram represent the active and the inactive acceptor states.

tensity during red excitation intervals induced a few stochastic activation events; these account for the offset of 40-50 active molecules in Fig. 2.3A.

Probing different acceptors on the same molecule sequentially requires activation rates to be tuned such that, on average, only a single acceptor is active during green excitation intervals. Here, green excitation fulfils a dual role, providing the FRET signal but also leading to undesirable high activation rates. Considering the additional complication of the dependence of the Cy5 activation rate on its distance from the donor (Bates et al., 2005), I conclude that periodic photoswitching with Cy5 is not practical for switchable FRET.

Stochastic switching

Instead, I tested stochastic photoswitching to separate active state intervals of multiple identical fluorophores on the same molecule such that they can be interrogated individually over time. I applied alternating-laser excitation (ALEX) (Kapanidis et al., 2004) at 532 nm and 635 nm with a period of 200 ms, such that the alternation rate exceeded the activation and deactivation rates of Cy5. Under these conditions, individual Cy5 molecules showed reversible stochastic photoswitching, apparent from spontaneous transitions between a bright state and a dark state in the F_{AA} acceptor intensity time-trace (Fig. 2.3C, black line). ALEX also allowed continuous observation of the apparent FRET efficiency, E^* , and the relative donor-acceptor stoichiometry, S , calculated from F_{DD} , F_{DA} , and F_{AA} according to equations 1.14 and 1.15. Because the donor is not switchable in our experiments, the stoichiometry ratio serves as a “counter” of the number of active acceptors at any time: donor-only states display high S values; states with a donor and a single active acceptor display intermediate S values; and states with one donor and multiple active acceptors display low S values. I

summarised the FRET and stoichiometry information of an individual molecule on a two-dimensional E^* - S histogram.

Active and inactive states of Cy5 in the presence of Cy3B display intermediate and high S values, respectively, yielding distinct clusters in the E^* - S histogram of a single molecule (Fig. 2.3C, right). In agreement with the Förster radius of 6.7 nm for Cy3B and Cy5 in this sample (Uphoff et al., 2010), their 10 bp (3.4 nm) separation led to a high FRET efficiency $E^* \sim 0.8$. Residual non-zero FRET efficiency in the inactive state represents leakage of donor fluorescence into the acceptor detection channel (Lee et al., 2005). The Cy3B fluorescence unquenches when Cy5 switches to the inactive state.

A different dsDNA construct with 19 bp separation between Cy3B and Cy5 also showed reversible stochastic photoswitching with ALEX, but a lower FRET value, $E^* \sim 0.35$, corresponding to the larger donor-acceptor separation (Fig. 2.3D). I observed that the inactive state of the distal 19 bp Cy5 was longer lived than the inactive state of the proximal 10 bp Cy5 using ALEX. In fact, photoactivation of Cy5 is not only induced by direct 532 nm excitation, but also by energy transfer of a nearby donor fluorophore (Bates et al., 2005). Photoactivation rates of a Cy5 acceptor in a switchable FRET experiment thereby depend on the distance to the donor fluorophore.

Demonstration of switchable FRET

To prove the ability of switchable FRET to resolve multiple FRET pairs at different positions on a single molecule in a temporal fashion, I prepared dsDNA labelled with two Cy5 acceptors at 10 bp and 19 bp separation from one Cy3B donor. One molecule could hence adopt four distinct photophysical states; namely (i) donor-only (denoted as 0-D0, where “0” stands for an acceptor in its inactive state and “-” denotes the long, 19 bp separation) when both acceptors

were inactive, (ii) state “A-D0” (low E^* , intermediate S) when only the distal acceptor was active; (iii) state “0-DA” (high E^* , intermediate S) when only the proximal acceptor was active; and (iv) state “A-DA” (high E^* , low S) when both acceptors were active.

A typical time-trace of fluorescence intensity, E^* , and S (Fig. 2.4A) from a single DNA molecule with one Cy3B and two Cy5 fluorophores shows that alternating red and green excitation caused stochastic photoswitching of the two acceptors, apparent from steps between three F_{AA} acceptor intensity levels. The intensity is maximal when both acceptors are active (~ 8.000 photon counts per 100 ms, state A-DA), intermediate during segments when only one acceptor is active (~ 4.000 photon counts per 100 ms, either state A-D0 or 0-DA), and low during donor-only segments (< 100 photon counts per 100 ms, 0-D0). Because the donor transfers energy only to active acceptors, F_{DD} is maximal for 0-D0 but becomes substantially quenched when an acceptor switches on; each of the four states is adopted several times. The donor is not affected by the switching conditions and does not blink during the course of the experiment.

The E^* - S histogram from the time-trace in Fig. 2.4A shows the expected four clusters representing the states 0-D0, A-D0, 0-DA, and A-DA (Fig. 2.4B, C). To identify states and extract statistical information, I separated the clusters using the standard clustering algorithm k-means. Data points were assigned to clusters by minimising the sum of distances to the corresponding centroid of their clusters in the E^* - S histogram. The identified distributions are narrow (standard deviation of $E^* \sim 0.03$), compared to typical diffusion-based FRET histograms (Lee et al., 2005). The two clusters of the single-pair FRET efficiencies ($E^* = 0.256$ for A-D0 and $E^* = 0.793$ for 0-DA) are clearly separated from states 0-D0 and A-DA and identified as individual states.

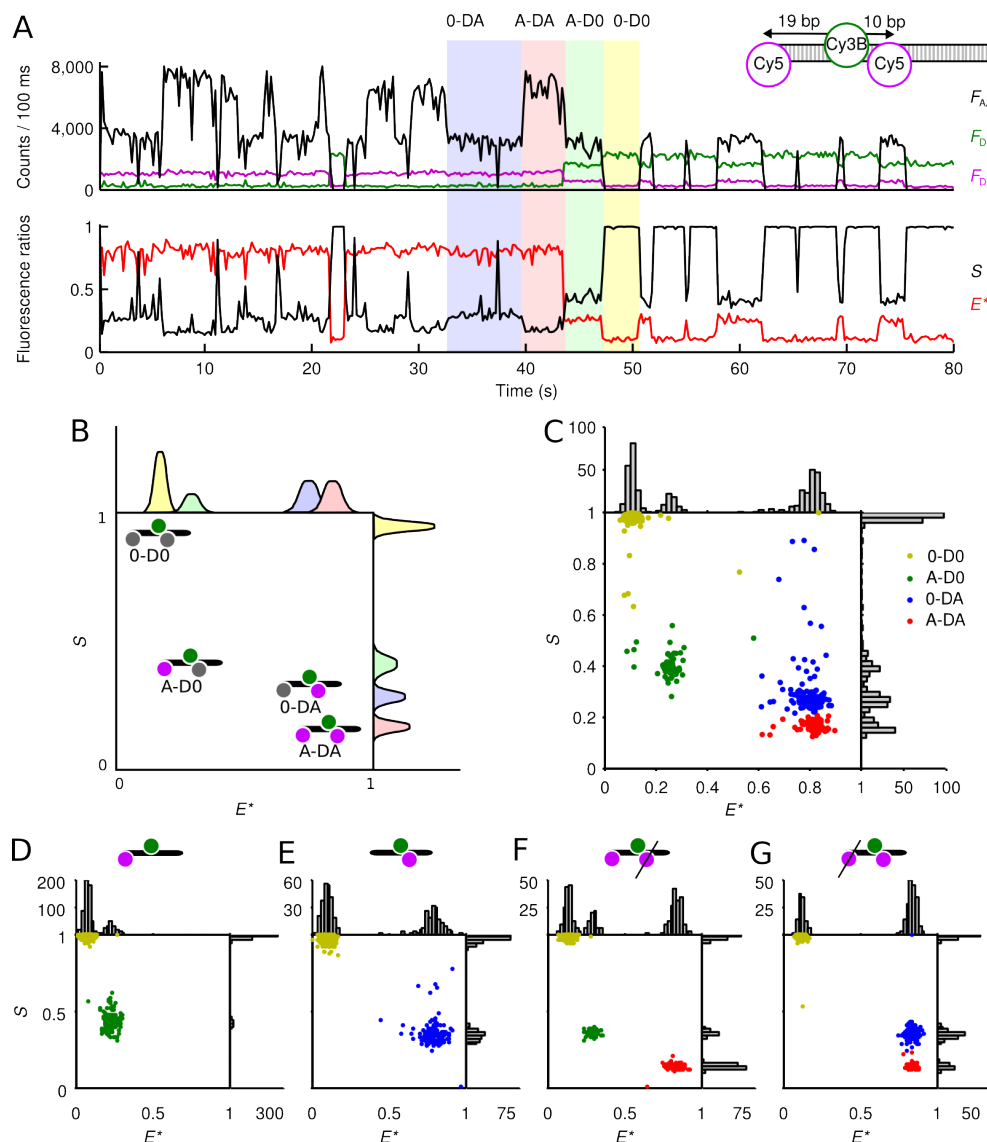


Figure 2.4: Proof-of-principle of switchable FRET (previously presented in my MSc thesis). (A) ALEX time-trace for an individual immobilised dsDNA molecule labelled with a single Cy3B donor separated from two Cy5 acceptors by 10 bp (high FRET) and 19 bp (low FRET). Top: ALEX causes stochastic photoswitching of Cy5 and yields fluorescence intensities F_{DD} (green line), F_{DA} (purple line), and F_{AA} (black line). Bottom: Time-traces of apparent FRET, E^* (red line), and stoichiometry, S (black line). The coloured backgrounds highlight a switching sequence of all four states. (B) Expected locations of the four clusters on the E^* - S histogram: 0-D0 at high S and low E^* (in yellow), A-D0 at intermediate S and low E^* (in green), 0-DA at intermediate S and high E^* (in blue), and A-DA at low S and high E^* (in red). (C) The E^* - S histogram of the data from (A) shows the four expected clusters, identified with a clustering algorithm. (D, E) Controls using DNA labelled with one donor and either the distal or proximal acceptor (experimental conditions as in (A)). (F, G) Controls with both acceptors under non-switching conditions. For the molecule in F, the proximal acceptor bleaches first, whereas in G, the distal acceptor bleaches first.

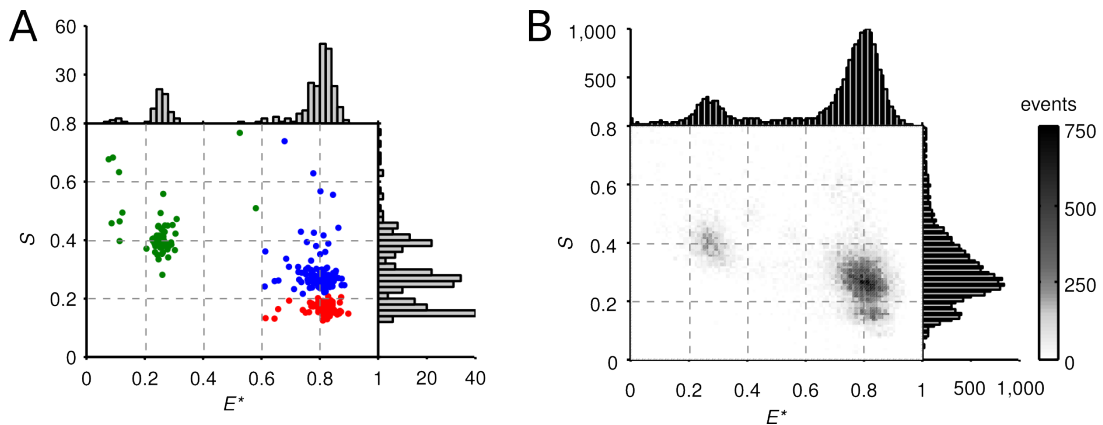


Figure 2.5: Single-molecule and combined E^* - S histogram for immobilised dsDNA labelled with Cy3B and two Cy5 at 10 and 19 bp separation (previously presented in my MSc thesis). (A) Single-molecule data from Fig. 2.4. Combined data of 147 molecules under identical conditions to (A). The higher abundance of the 0-DA state compared to the A-D0 state is due to the higher activation rate of the proximal acceptor; the deactivation rates are similar.

Controls

The controls with a single acceptor either at 10 bp or 19 bp separation from the donor showed clusters identical to those from the double-acceptor DNA, excluding the possibility that the single-pair FRET efficiency is altered by light absorption of the second acceptor in the inactive state (Fig. 2.4D, E). To test the requirement for switching to probe multiple distances, I measured the sample with both acceptors under non-switching conditions (Rasnik et al., 2006); in this case, irreversible sequential acceptor bleaching reported on only one of the two FRET pairs per molecule. If the proximal acceptor bleaches first, the 0-DA state is missing from the E^* - S histogram (Fig. 2.4F); when the distal acceptor bleaches first, the A-D0 state is missing (Fig. 2.4G).

To confirm that the information from a single molecule reflects a significant fraction of all molecules, I combined the E^* - S histograms of 147 molecules (Fig. 2.5). Due to irreversible photobleaching and the stochastic nature of photoswitching, 57% of fully labelled molecules adopted all four states. The mean E^* and S values differ slightly between molecules due to inter-molecule heterogeneity and

optical aberrations which broaden the distributions in the combined histogram (Holden et al., 2010). The match between the combined and single-molecule data demonstrates the robustness of switchable FRET.

2.3 Further development of switchable FRET

Buffer conditions for switchable FRET

Switchable FRET requires buffer conditions that induce photoswitching of acceptor fluorophores without perturbing stable donor fluorescence. To probe multiple acceptors over time, minimising donor photobleaching is crucial. A standard enzymatic oxygen scavenging system (Selvin and Ha, 2008) plus beta-mercaptoethanol (BME) or mercaptoethylamine (MEA), thiol-containing reducing agents, met these needs. In addition, BME and MEA act as triplet quenchers to minimise photobleaching of organic fluorophores.

I found, however, that Cy3B bleached within a few seconds of 532 nm excitation in freshly prepared switching buffer containing 14.3 mM BME while reversible photoswitching of Cy5 was very efficient. In agreement with the mechanism by which Trolox serves as an anti-bleaching reagent (Cordes et al., 2009), I found that incubation of BME in solution at 37°C for 24 h increased its performance as anti-bleaching reagent while preserving its function as an inducer of photoswitching; increased temperature accelerated this process. I also observed a gradual decrease in the deactivation rates of Cy5 upon storage of BME in solution at 4°C for several months. I hypothesise that, during incubation, an oxidised form of BME is generated which quenches the triplet state by a combination of reduction and oxidation according to the ROXS principle (Vogelsang et al., 2008), thereby inhibiting photobleaching. The consumption of the reduced form of BME during incubation and storage explains the decreased deactivation rates of Cy5. I

noticed similar behavior for MEA.

Fast switchable FRET using TIRF microscopy

Activation and deactivation of Cy5 are photo-induced processes, requiring population of excited states, and hence depend on the excitation intensity. I used this property to tune measurement time and temporal resolution of switchable FRET. Increased excitation intensities induced faster photoswitching while the higher photon count rate permitted faster laser alternation together with shorter camera exposure per frame. With little loss in data quantity and quality, measurement times could be adjusted from more than a minute down to ~ 1.5 seconds with corresponding temporal resolution of 200 ms down to 10 ms (Fig. 2.6). The camera readout rate became a limiting factor for data in Fig. 2.6c, which was acquired using 2x2 pixel binning and a reduced field of view.

Control of photoswitching rates is important for applying switchable FRET to various molecular processes operating on different time scales. Examination of transient short-lived structural states from multiple perspectives needs sufficiently fast switching rates to record different acceptors during the dwell time of a state. On the other hand, observation of rare events, slow processes, or repeated transitions requires long time-traces of single molecules without photobleaching. My results show that switchable FRET should be applicable to the whole range of processes to which conventional single-molecule FRET spectroscopy can be applied. Importantly, measuring more than two distances using switchable FRET will be facilitated by switching rates that generate a small on/off time ratio, so as to ensure that each acceptor is probed individually before irreversible photobleaching.

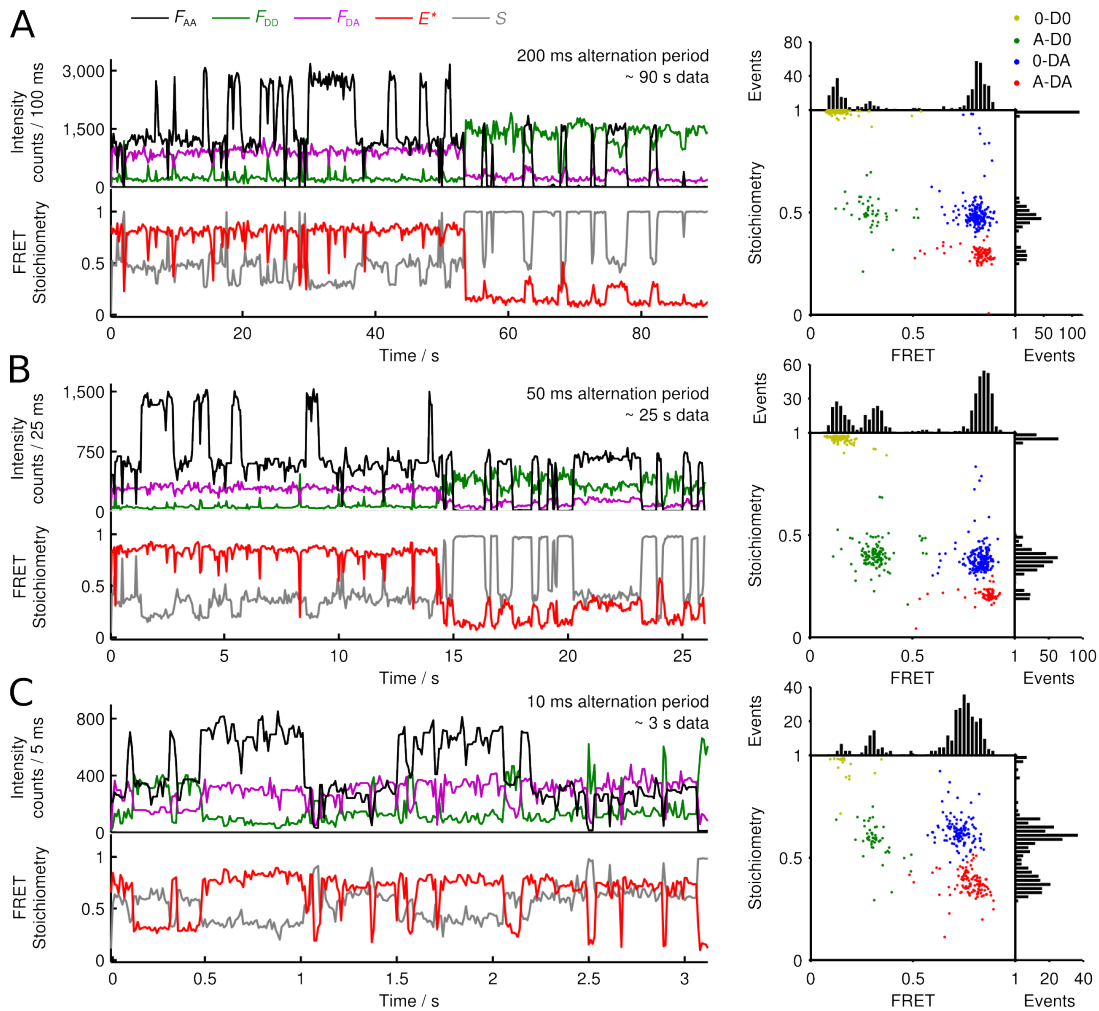


Figure 2.6: Switchable FRET with tunable temporal resolution. TIRF microscopy at different ALEX intensities and periods probed immobilised dsDNA labelled with Cy3B and two Cy5 at 10 and 19 bp separation. The laser alternation period equals the temporal resolution. Switching rates and temporal resolution increase from top to bottom. **(A)** 200 ms alternation period, 1 mW green and 1 mW red excitation. **(B)** 50 ms period, 2 mW green and 2 mW red excitation. **(C)** 10 ms period, 4 mW green and 4 mW red excitation. E^* - S histograms (right) represent data of the single-molecule time-traces (left).

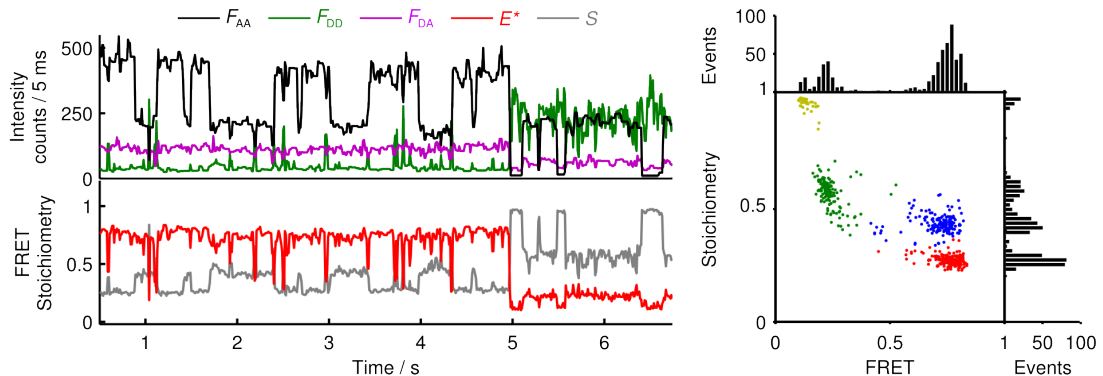


Figure 2.7: Switchable FRET by confocal microscopy. ALEX at 100 μ s alternation period probed a single immobilised dsDNA molecule labelled with two Cy5 at 10 and 19 bp separation from one Cy3B. Displayed data were integrated at 10 ms resolution. Colour representation as above.

Fast switchable FRET using confocal microscopy

To further extend the dynamic range of switchable FRET and potentially enable statistical photon trajectory analysis (Chung et al., 2009), I applied confocal scanning microscopy². Single immobilised molecules were automatically located after scanning a field of view using a piezo stage, positioned at the centre of the fixed laser excitation focus, and measured one at a time. I applied red (639 nm) and green (532 nm) ALEX at 100 μ s period and recorded photon arrival times continuously in the spectrally separated donor and acceptor detection channels by avalanche photodiodes. For display, fluorescence counts were integrated at 10 ms temporal resolution (Fig. 2.7). Switchable FRET experiments were performed using the same sample with two Cy5 fluorophores and buffer conditions as above.

Data quality using confocal microscopy was improved compared to TIRF microscopy at the highest temporal resolution while photobleaching and switching rates were similar. Focusing excitation onto a diffraction limited spot centred on a single molecule gives high local excitation intensity and optimal fluo-

²The confocal scanning microscope was built by Johannes Hohlbein and Geraint Evans; Materials and Methods section 2.6.

cence to background ratio, but at the sacrifice of parallel acquisition of molecules provided by TIRF-based imaging (although microscopy formats with multiple confocal spots could also be used). While the frame rate for TIRF microscopy is currently limited to > 1 ms (due to camera readout rates), confocal scanning microscopy can in principle achieve even higher temporal resolution before the fluorescence emission rate saturates at high excitation intensities. Compared to TIRF microscopy, no further image processing is needed in confocal microscopy. Combination of differences in quantum yield and detection efficiency of donor and acceptor caused the difference in S between the low and high FRET states (Lee et al., 2005); the extent of this difference varies between microscopes with distinct filter sets and detectors.

State identification

Switchable FRET reports on multiple FRET states originating from different acceptors per molecule. Data analysis is required to identify these states and extract information such as the mean E^* and S values. K-means clustering analysis relies on clear separation between clusters. Standard signal processing algorithms (such as moving median or frequency domain filters) fail to enhance the discontinuous stepping behavior which is characteristic of single-molecule time-traces. For data in Figs. 2.6 and 2.7, I employed a recently developed step-filter method to increase resolution of states (Little and Jones, 2010). The filter robustly approximates the data globally (as opposed to sliding-window algorithms) by a piece-wise constant curve. The algorithm minimises two terms: both the deviation of the solution from the input, and the difference between adjacent data points of the solution, which favors a step-like solution with piece-wise constant intervals. The single fitting parameter, here $\lambda = 2 \cdot 10^{-4}$, describes the trade-off between the two minimisations. For $\lambda = 0$, the data is repro-

duced exactly; whereas, for increasing λ , the fit tends towards a solution with pronounced piece-wise constant intervals.

However, state finding by data clustering is not optimal as it does not exploit the temporal information available from switchable FRET. Time-traces contain the temporal sequence of discrete instantaneous switching events which mark transitions between states. To this end, Kristofer Gryte extended hidden Markov Modeling (HMM) approaches (McKinney et al., 2006; Bronson et al., 2009) to link both FRET and stoichiometry observables. Linked HMM successfully identified transient states in the two-dimensional FRET-stoichiometry space and reconstructed their connectivity network that shows the conversions between states (Uphoff et al., 2011).

To characterise the resolution of FRET, I contributed to a detailed analysis of the sources of noise that broaden E^* distributions beyond the statistical noise expected from photon statistics (Holden et al., 2010). While individual molecules achieve the theoretically expected FRET resolution that is limited by photon counts only, heterogeneity between molecules limit the ability to distinguish distance changes to approximately the length of a single bp along the axis of DNA (0.34 nm).

Alternative acceptor

The on/off time ratio in switchable FRET experiments significantly differed for the proximal and distal Cy5. Due to its lower ratio, the distal Cy5 was probed mostly after the proximal Cy5 photobleached irreversibly. For this reason, distance-dependent activation rates limit the applicability of Cy5 as an acceptor in switchable FRET experiments with more than two acceptors.

To show the generality of switchable FRET, I tested an alternative photoswitchable acceptor ATTO655. This oxazine fluorophore switches through a combina-

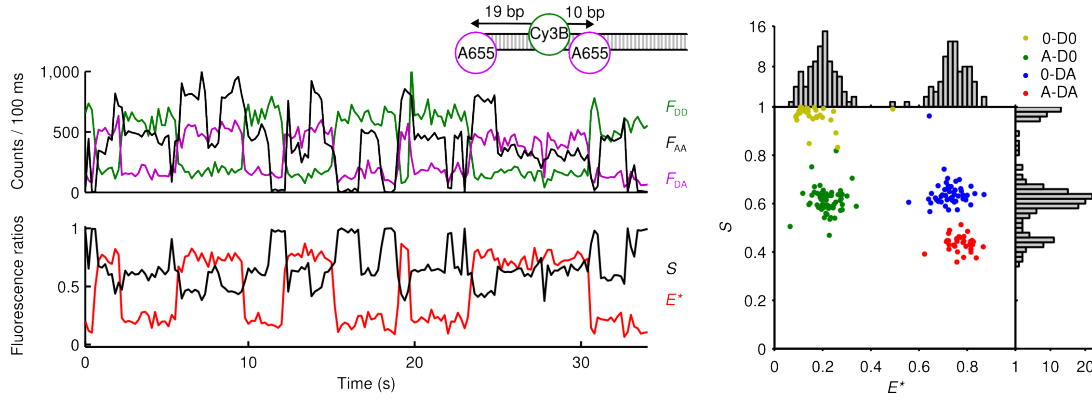


Figure 2.8: Switchable FRET with an alternative acceptor. ALEX time-trace for an individual immobilised DNA molecule labelled with a single Cy3B donor separated from two ATTO655 acceptors by 10 bp and 19 bp. The E^* - S histogram shows the four expected clusters: 0-D0 (in yellow), A-D0 (in green), 0-DA (in blue), and A-DA (in red).

tion of reduction and oxidation reactions (Heilemann et al., 2009b; Vogelsang et al., 2009). An example ALEX time-trace with two ATTO655 at 10 and 19 bp separation from one Cy3B shows the fluctuations in the F_{AA} signal due to stochastic acceptor photoswitching (Fig. 2.8). Under our experimental conditions, ATTO655 switched more rapidly than Cy5, capturing both FRET states in a shorter time and did not exhibit distance-dependent activation rates; both acceptors had the same on/off time ratio and were probed stochastically over time. Fortunately, a wide range of organic fluorophores that can be rendered photo-switchable is available (Dempsey et al., 2011) from which to choose the optimal fluorophore pair for switchable FRET experiments on a particular system.

Monte Carlo simulation of switchable FRET

To explore the limits of switchable FRET, I performed Monte Carlo simulations for molecules containing a single donor and an arbitrary number of acceptors. This work was previously presented in detail in my MSc thesis (Uphoff, 2010). Briefly, I modeled stochastic excitation, energy transfer, and photoswitching to create ALEX time-traces of single molecules with one donor and four acceptors

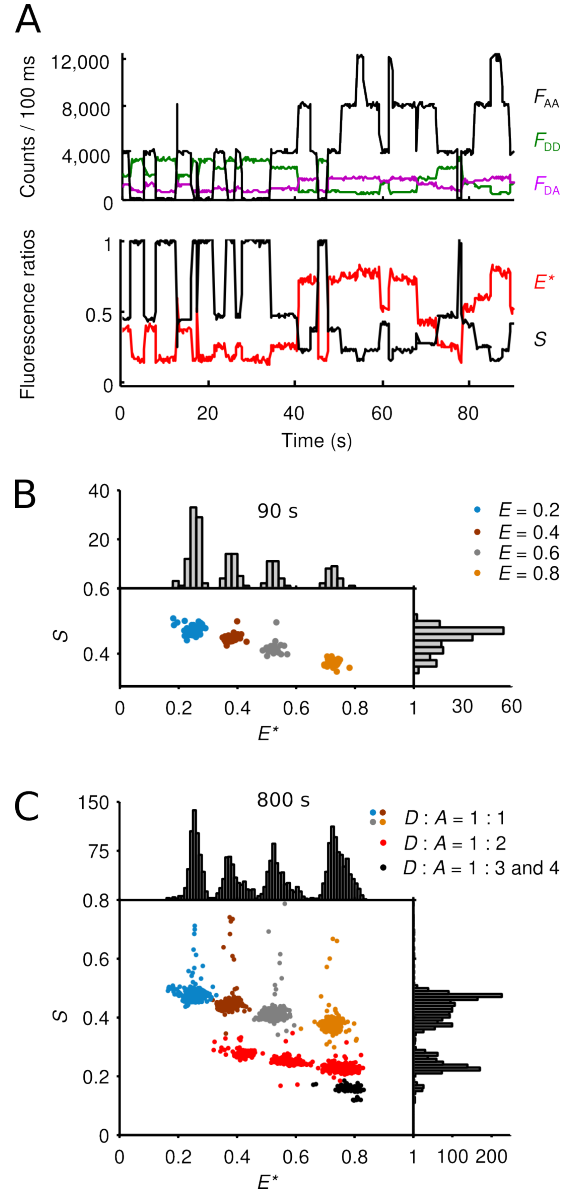


Figure 2.9: Monte Carlo simulation of switchable FRET with four acceptors (previously presented in my MSc thesis). **(A)** Simulated ALEX time-trace of a molecule with one donor and four acceptors with individual FRET efficiencies of $E = 0.2, 0.4, 0.6$, and 0.8 , respectively. The simulation parameters reflected the photophysical properties of Cy3B and Cy5 (Uphoff et al., 2010). **(B)** E^* - S histogram of the four single-pair FRET states of the data in A. The E^* values differ from the input E values because of the simulated donor leakage and the effect of the detection efficiency factor γ , which also causes the negative slope between the clusters. **(C)** The E^* - S histogram of a simulated time-trace of 800 s displays the four individual FRET states at intermediate S ($D:A = 1:1$). At lower S , there are six partly overlapping clusters with two active acceptors ($D:A = 1:2$), four overlapping clusters with three active acceptors ($D:A = 1:3$), and a few data points with four active acceptors ($D:A = 1:4$, at lowest S). The donor-only state was omitted for clarity.

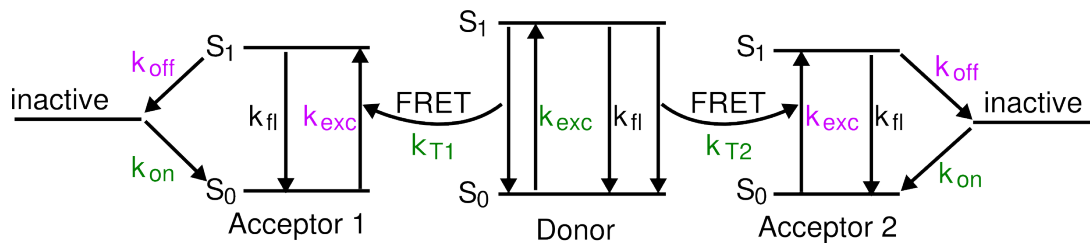


Figure 2.10: Energy diagram and model for switchable FRET with one donor and two acceptors (previously presented in my MSc thesis). Each fluorophore cycles between its ground state S_0 and excited state S_1 by excitation (k_{exc}) and fluorescence emission (k_{fl}). The donor transfers energy to either acceptor via FRET (k_T). The acceptors switch off by a transition to an inactive state (k_{off}) and switch on by a transition to the ground state (k_{on}).

with $E = 0.2, 0.4, 0.6$, and 0.8 (Fig. 2.9A). The resulting E^* - S histogram showed the expected four single-pair FRET clusters at the S segment from 0.3 to 0.6 ($D:A = 1:1$) (Fig. 2.9B). These results can be interpreted more easily than data representing four distances probed with multi-colour FRET using four different emission channels (Lee et al., 2010; Stein et al., 2011). In fact, the number of distances accessible to switchable FRET is fundamentally limited only by the FRET resolution, that is, the width of single FRET distributions. For illustration purposes, the E^* - S histogram of a simulated 800 s time-trace shows also rare states with multiple active acceptors (Fig. 2.9C).

My simulations also showed that a simple model (Fig. 2.10) accurately described the experimental switchable FRET data (Fig. 2.11). Widths of experimental and simulated FRET distributions were in good agreement, demonstrating that experimental noise can be explained by modeled photon shot noise, background noise, and camera amplification noise (Holden et al., 2010). The experimental stoichiometry distributions were twofold larger than the model predicted mostly owing to additional blinking of Cy5 and possibly fluctuations in laser intensities.

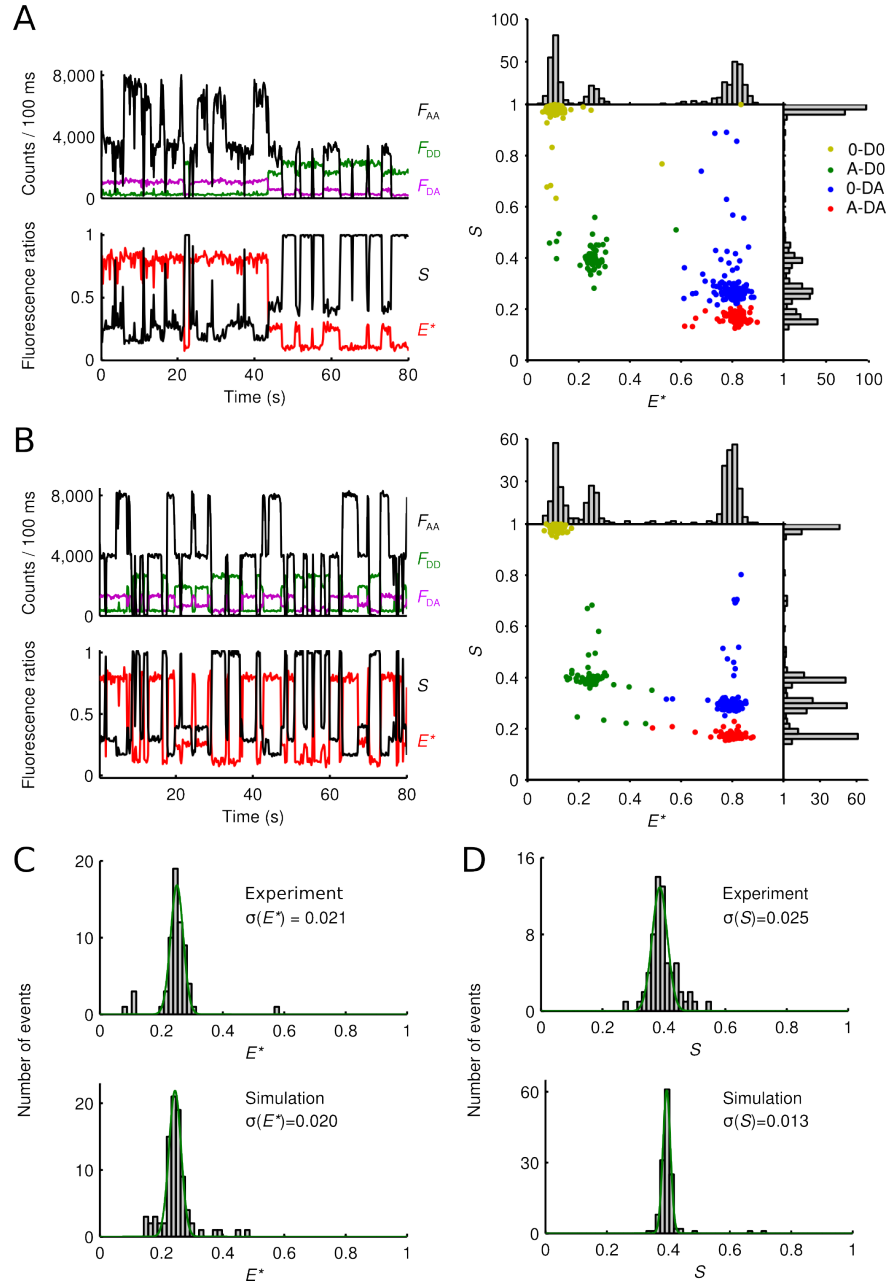


Figure 2.11: Direct comparison between simulation and switchable FRET experiment on dsDNA with Cy3B and two Cy5 acceptors (previously presented in my MSc thesis). **(A)** Experimental ALEX time-trace and E^* - S histogram. **(B)** Simulated ALEX time-trace and E^* - S histogram according to the model in Fig. 2.10. **(C)** Experimental (top) and simulated (bottom) FRET distributions of the cluster A-D0 from the data in A, B. The standard deviation was obtained from a Gaussian fit. **(D)** Stoichiometry distributions for cluster A-D0.

2.4 Applications of switchable FRET

Protein-DNA complex

To explore the potential of switchable FRET to probe structure and interactions of proteins and their complexes, I studied the complex of catabolite activator protein (CAP)³ with DNA. CAP is a global transcription activator that controls more than 100 genes in bacteria; it exerts its function through sequence-specific binding upstream of gene promotor regions in the presence of cyclic AMP (cAMP), followed by RNA polymerase recruitment to start transcription (Lawson et al., 2004). Protein crystallography (Schultz et al., 1991; Napoli et al., 2006) and ensemble FRET experiments (Kapanidis et al., 2001) showed that CAP bends DNA by 80-90° upon binding.

In my experimental design, I took advantage of the fact that CAP is a homodimeric protein, enabling introduction of two copies of switchable acceptor Alexa647 on a derivative of CAP that carries a single surface-exposed Cys residue per subunit (Fig. 2.12A). Alexa647 is a hydrophilic carbocyanine fluorophore with similar spectral and photophysical properties to Cy5 (Heilemann et al., 2008). CAP was mutated, produced, and labelled by Javier Periz. To distinguish between the two fluorophores on CAP in switchable FRET measurements, a Cy3B donor was incorporated on the 5' end of the top strand of a 39 bp dsDNA fragment containing the two-fold symmetric CAP recognition sequence. In the CAP-DNA complex with all three fluorophores present, two distinct FRET efficiencies are expected, reporting on the distances of ~ 40 Å and ~ 70 Å between the end of the DNA fragment and Cys17 on the proximal and distal subunit, respectively (Fig. 2.12A).

For single-molecule FRET measurements, doubly labelled CAP molecules were

³CAP is also known as cyclic AMP receptor protein (CRP)

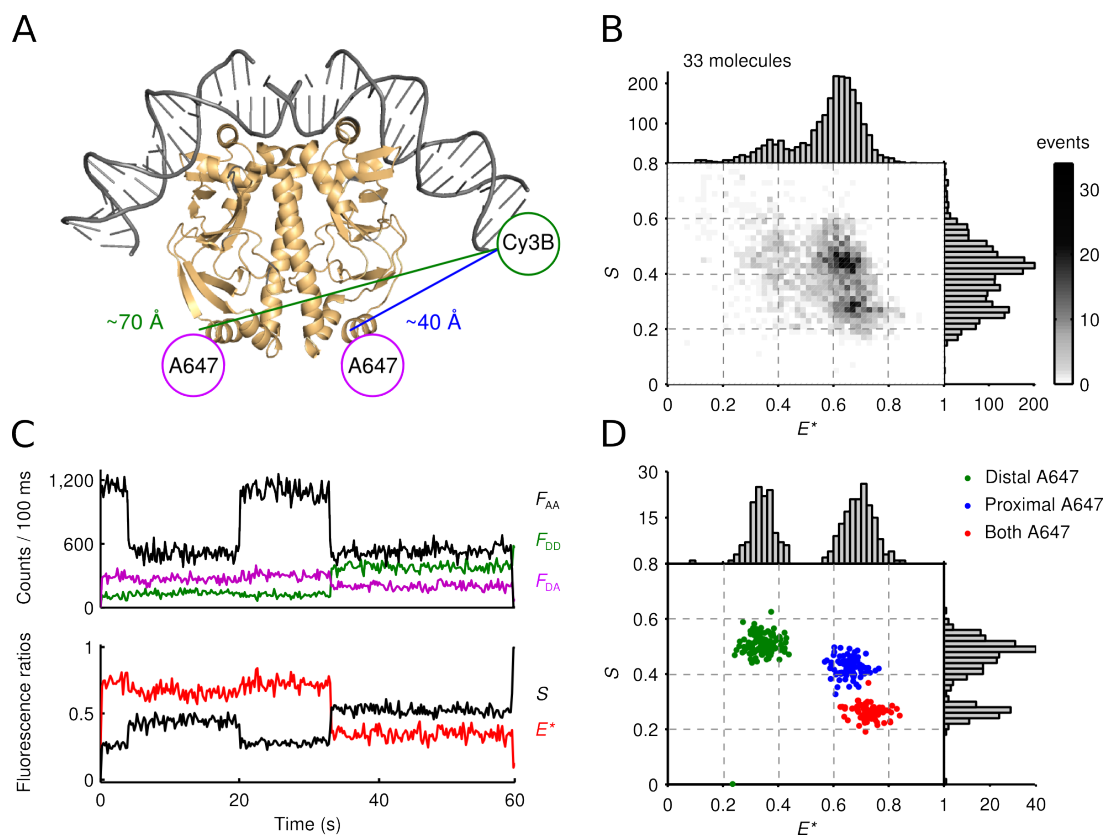


Figure 2.12: Switchable FRET monitors two distances within the CAP-DNA complex. **(A)** Crystal structure of the CAP-DNA complex (Napoli et al., 2006). CAP is in yellow, DNA is in grey, Alexa647 and Cy3B fluorophores are schematically represented as circles. Distances were measured between the DNA axis and the C- α of Cys17. **(B)** Combined single-molecule E^* - S histogram. **(C, D)** ALEX time-trace and corresponding E^* - S histogram.

incubated in the presence of $200\mu\text{M}$ cAMP with the immobilised DNA to form complexes on the microscope slide. TIRF imaging of stable CAP-DNA complexes resulted in colocalised donor and acceptor emission spots, and enabled switchable FRET measurements within single CAP-DNA complexes. ALEX time-traces exhibited stochastic switching of the two Alexa647 acceptors and their corresponding FRET states (Fig. 2.12B); the E^* - S histogram of this individual molecule shows a high and a low FRET state at intermediate stoichiometry, representing energy transfer to the acceptor on the proximal and distal subunit, respectively (Fig. 2.12C). The single-molecule data are in agreement with a combined single-molecule data under identical buffer conditions and laser intensities (Fig. 2.12D). These results demonstrate that switchable FRET can probe multiple distances within single protein-DNA complexes and they pave the way for similar work on other multimeric proteins (such as dimeric and multimeric transcription factors, helicases, and membrane proteins) which are difficult to label site-specifically with single fluorophores, but can be labelled with multiple identical fluorophores using standard techniques. Notably, many proteins have only few surface-exposed reactive Cys residues and as such, they are immediately amenable for use in switchable FRET experiments without any mutagenesis.

Conformational dynamics of DNA Holliday junctions

To explore the ability of switchable FRET to monitor molecular dynamics, I measured conformational changes within individual DNA Holliday junctions from two perspectives, facilitating the comparison to a multi-colour FRET study of a similar system (Hohng et al., 2004). In the presence of magnesium ions the Holliday junction adopts two alternative stacked X-like structures (McKinney et al., 2003; Karymov et al., 2005), the interconversion of which is central to DNA recombination (Lilley, 1997). I designed four DNA strands that hybridise such

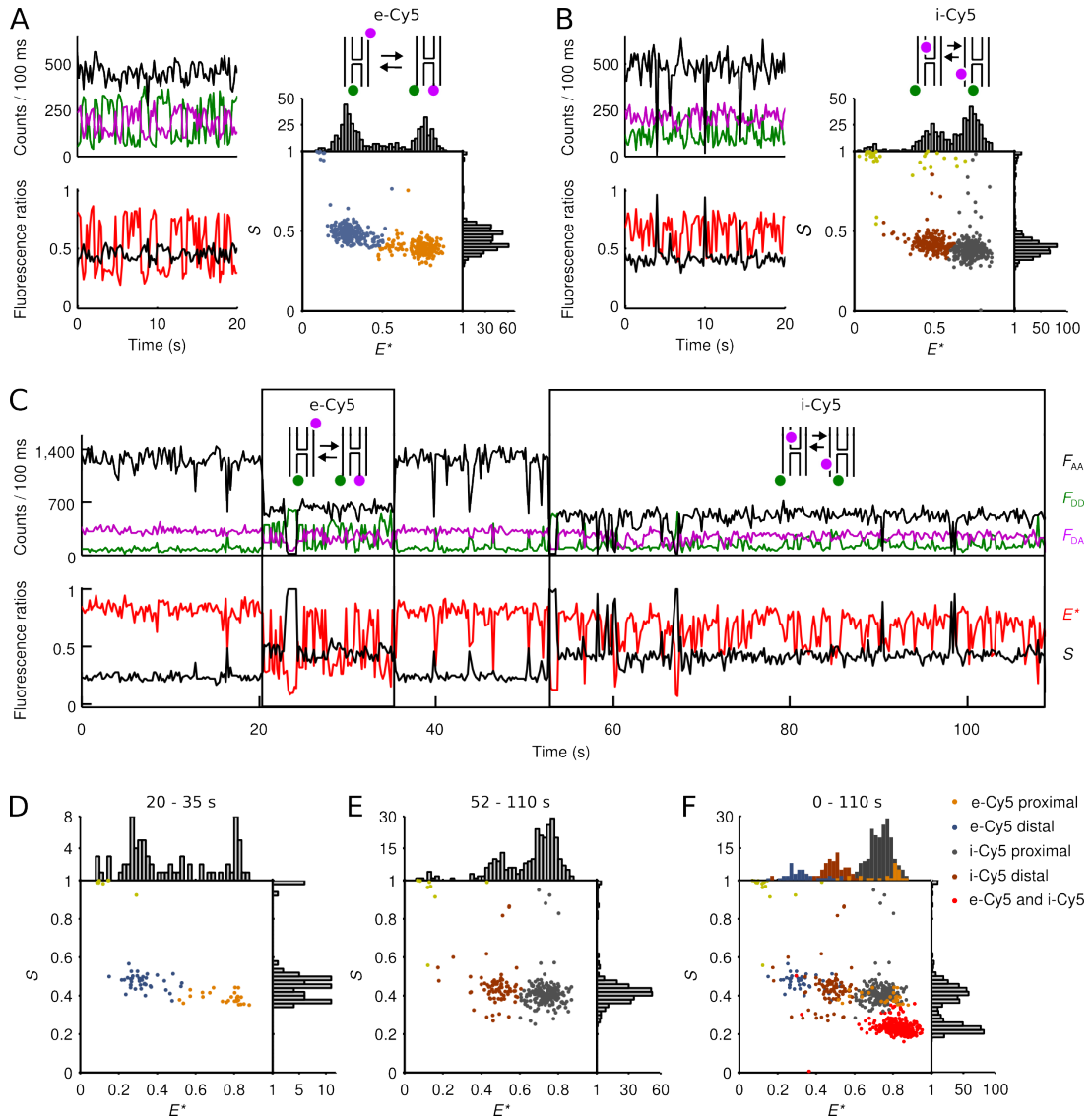


Figure 2.13: Holliday junction dynamics probed from two perspectives. **(A)** Only Cy3B and e-Cy5 present. **(B)** Only Cy3B and i-Cy5 present. Because frame rates and dynamic rates are similar, I observed a distribution of FRET values between the two states. **(C)** Switchable FRET measures structural changes within a single Holliday junction from two perspectives: e-Cy5 is active in the first box, i-Cy5 is active in the second box. **(D, E)** E^* - S histograms of the boxed regions in C. **(F)** E^* - S histogram of the full time-trace. K-means cluster analysis was performed on separate time-sections of the dataset to recover the states in D and E. Only single-pair FRET states are shown in the E^* histogram for clarity.

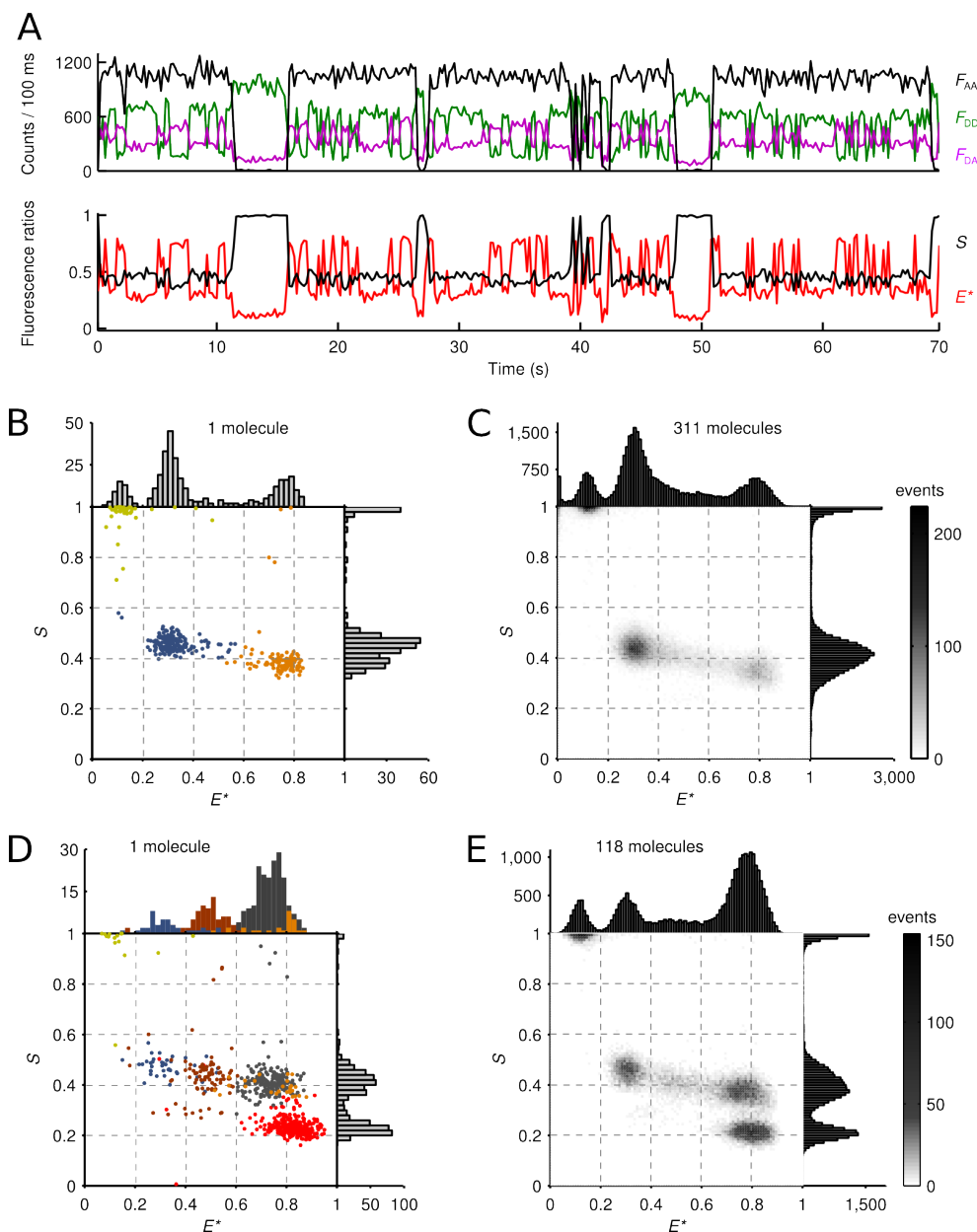


Figure 2.14: Holliday junction controls. **(A)** ALEX time-trace showing dynamics of a single Holliday junction from a control sample labelled with e-Cy5 and Cy3B only. **(B)** E^* - S histogram of the data in A. **(C)** Combined E^* - S histogram of 311 molecules of the control sample measured under identical conditions. **(D, E)** Single-molecule data from Fig. 5F compared to the combined data of 118 molecules that showed conformational dynamics.

that a Cy3B donor at the end of one branch exhibits FRET to an end-labelled Cy5 (“e-Cy5”) and an internally-labelled Cy5 (“i-Cy5”), each on a different branch (Fig. 2.13 schematics). Due to incomplete labelling, some molecules were only modified with a single FRET pair (Fig. 2.13A, B). For the pair of Cy3B and e-Cy5, fluctuations between a low and a high FRET state show conformational changes within one part of the Holliday junction ($E^* \approx 0.3$ and $E^* \approx 0.8$, dark blue and orange dots, Fig. 2.13A). Molecules carrying only Cy3B and i-Cy5 also show FRET fluctuations, but between two intermediate FRET states, consistent with the smaller distance range of the internally labelled acceptor ($E^* \approx 0.5$ and $E^* \approx 0.7$, brown and grey dots, Fig. 2.13B). In agreement with previous findings, a small conformer bias is indicated by the e-Cy5 pair preferring slightly the low FRET state while the i-Cy5 pair prefers slightly the high FRET state (Hohng et al., 2004).

Switchable FRET probes conformational changes of a single molecule from two perspectives provided by two switchable acceptors (Fig. 2.13C). When e-Cy5 and i-Cy5 are active at the same time (0-21 s), FRET is constant and high, showing that at least one of the acceptors is close to the donor in either conformation. Subsequently, photoswitching of i-Cy5 to the inactive state allows e-Cy5 to monitor dynamics of the Holliday junction from one perspective (21-35 s). After a time segment with both acceptors, only i-Cy5 remains active and probes conformational changes from a different perspective (55-110 s). E^* - S histograms of the boxed regions display clusters in agreement with the single-acceptor data (Figs. 2.13D-E). Although the high FRET states of e-Cy5 and i-Cy5 overlap significantly, temporal confinement by clear photoswitching events allows the separate analysis of different sections of the full time-trace to identify four distinct single-pair FRET states within one molecule (dark blue, brown, grey and orange clusters, Fig. 2.13F). Control experiments with a strictly doubly-labelled sample

(Cy3B and e-Cy5), as well as combined single-molecule data further support our results (Fig. 2.14).

Additional Monte Carlo simulations of switchable FRET confirmed that assignment of transient FRET states to multiple photoswitching acceptors is not limited to the regime where photoswitching is slow compared to the rates of structural dynamics (Fig. 2.15). If photoswitching and dynamics occur on the same timescale, transitions in FRET due to structural changes can be distinguished from photoswitching events that probe a different acceptor with a different FRET value, because the latter always involves an obligatory intermediate donor-only state or states with multiple active acceptors. Ultimately, in the regime where photoswitching is faster than the dynamic rates, each transient structure is examined from multiple perspectives, providing insight into the multi-dimensional trajectory of conformational changes.

2.5 Conclusion

Development of advanced experimental techniques contributes significantly to the rapid progress in the field of single-molecule biophysics (Hohlbein et al., 2010). Compared to multi-colour approaches, switchable FRET simplifies labelling, data acquisition and data analysis. The ability to probe several FRET pairs within single biomolecules will enable studies of their structure and dynamics in solution, complementing available crystallographic information. In the absence of prior information about expected distances, FRET states can still be assigned to the corresponding acceptor based on control experiments with single acceptors, while the fully labelled sample uncovers heterogeneity and dynamic interplay between states. In my MSc thesis, I have previously demonstrated application of switchable FRET to obtain accurate FRET efficiencies E for a

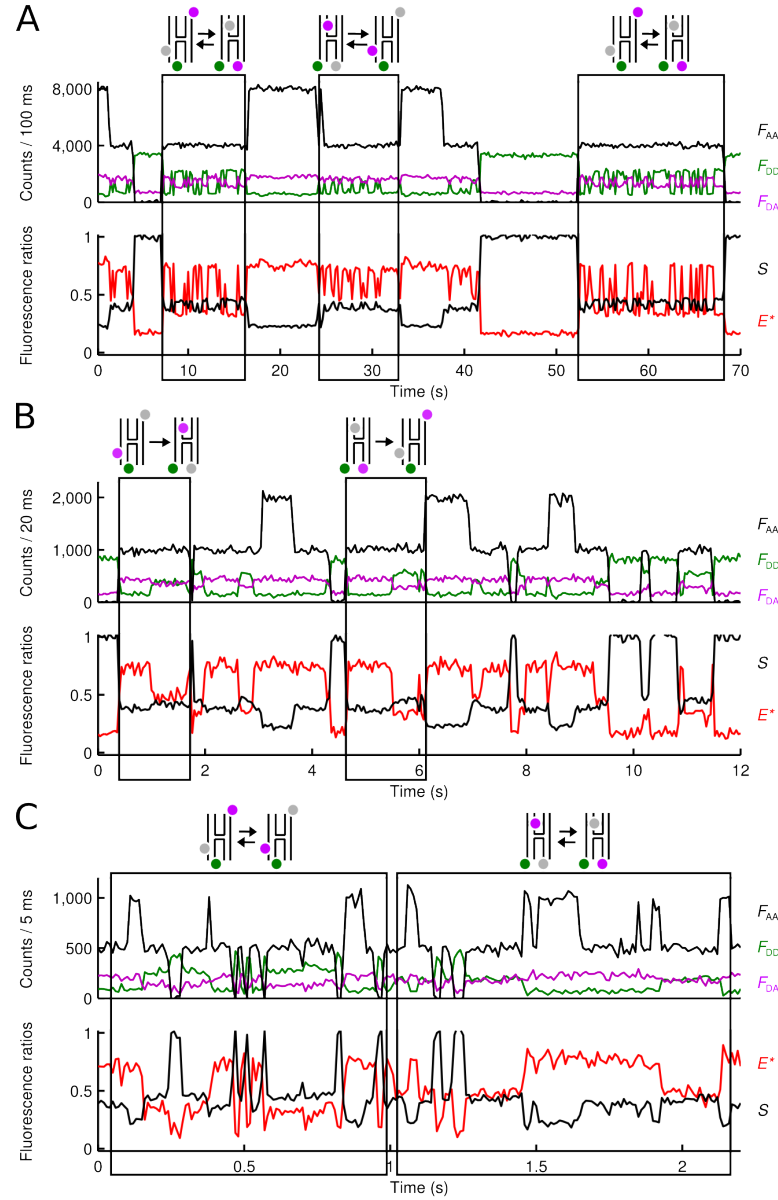


Figure 2.15: Simulation of Holliday junction dynamics with switchable FRET. **(A)** Slow photoswitching compared to dynamic rates. Boxed sections display a single acceptor probing multiple conformational changes (at 200 ms temporal resolution). **(B)** Similar photoswitching and dynamic rates. Boxed sections display a single acceptor probing individual conformational changes (at 40 ms temporal resolution). **(C)** Fast photoswitching compared to dynamic rates. Boxed sections display a single conformational state probed multiple times by two acceptors (at 10 ms temporal resolution).

single molecule (Uphoff, 2010), a feature important for experiments that generate constraints for structural analysis of biomolecular complexes. The ability to calculate accurate FRET from an individual molecule is useful when correction factors differ between molecules (for example, owing to changes in the local fluorophore environment). *De novo* protein structure prediction would benefit from multiple long-range FRET distance constraints instead of relying on sub-nanometre distance constraints from Nuclear Magnetic Resonance measurements (Bowers et al., 2000).

Enhanced control of photoswitching and compatibility with confocal scanning microscopy allows application of switchable FRET at a wide range of time scales. In contrast to TIRF microscopy, confocal scanning microscopy, as presented here, or with pulsed excitation, records photon arrival times; a feature that could be used to combine switchable FRET with statistical correlation and fluctuation analysis (Torella et al., 2011), as well as photon-by-photon analysis (Chung et al., 2009). Similar to super-resolution techniques (chapter 4), performance of switchable FRET depends primarily on controlled photoswitching. The number of FRET pairs measurable using switchable FRET is limited by the width of FRET distributions (Holden et al., 2010), provided that the on/off time ratio of each acceptor can be adjusted to be sufficiently small. Considering the rapid development in super-resolution techniques, which are also limited by small on/off time ratios, suitable fluorophores and imaging conditions for switchable FRET should allow measurement of more than two FRET pairs.

2.6 Materials and methods

Sample preparation

DNA oligonucleotides, labelled with Cy5, ATTO655, or amino-C6-dT-modified were purchased from IBA (Germany). DNA sequences and modifications are shown in Fig. 2.16. I labelled amino-modified strands with Cy3B-NHS, Cy5-NHS, or ATTO655-NHS. Briefly, 5 nmol of DNA oligonucleotide was labelled in 100 mM sodium tetraborate buffer at pH 9.0 with 50 nmol fluorophore-NHS ester. The labelling reaction was incubated at 50°C for 1 h before adding extra 50 nmol of fluorophore-NHS in DMSO to incubate at 37°C over night. Excess fluorophore was removed by ethanol precipitation and the labelled DNA was purified by 20% denaturing polyacrylamide gel electrophoresis. The labelled band was cut, eluted, and the urea salt was removed by centrifugation with Nanosep 3K spin columns (Pall), washed and resuspended in TE buffer (10 mM Tris-HCl and 1mM EDTA).

Complementary strands of DNA were annealed to form dsDNA by mixing equimolar amounts in annealing buffer (Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA), and heating to 95°C followed by slow cooling to 4°C. Fluorophores showed significant local rotational freedom as indicated by the steady-state anisotropies of 0.24 for Cy3B, 0.22 for the internally labelled Cy5, and 0.19 for the end-labelled Cy5.

Biotinylated dsDNA was immobilised on bovine serum albumin (BSA) passivated and neutravidin-coated coverslides, as described by Selvin and Ha (2008). Briefly, silicone gaskets (Grace Bio-Labs) were used as imaging chambers on No. 1.5 thickness coverslides (0.13 – 0.16 mm) and 30 μ l of 1 mg/ml biotinylated BSA was incubated for 5 minutes and washed with 100 μ l of T50 buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl). Subsequently, 30 μ l of 0.2 mg/ml neutravidin

solution was incubated for 2 minutes and washed with 100 μ l T50. I used 30 μ l of 10 pM biotinylated dsDNA and washed with 100 μ l T50 when the surface density reached 20 - 200 molecules per field of view ($25 \times 25 \mu\text{m}^2$). Samples were prepared immediately before each experiment.

To prevent non-specific adsorption of labelled proteins in CAP and DNA polymerase I experiments (chapter 3), I passivated microscope slides with polyethyleneglycol (PEG), as described (Selvin and Ha, 2008) and further optimised by Louise Aigrain. First, microscope slides were burned at 500°C for 1 h to remove fluorescent background particles. Burned slides were treated with acetone for 5 min; then 2% vectabond solution in acetone (Vector Laboratories) for 5 min and washed with dH₂O. Silicone gaskets were placed onto the nitrogen-dried slides. Each chamber was incubated for 3 h at room temperature with a mixture of 4 mg of NHS-PEG (for passivation) and 50 μ g biotin-PEG-NHS (for DNA immobilisation) in 20 μ l MOPS buffer, 50mM pH 7.5. Chambers were washed with PBS buffer and stored at 4°C for up to one week. Neutravidin coating and DNA binding was performed as for BSA-passivated slides.

Imaging was performed in either of two switching buffers: (i) phosphate-buffered saline (PBS) pH 7.4, 50 mM β -mercaptoethylamine (MEA); or (ii) 10 mM Tris-HCl pH 7.5, 10 mM NaCl, 14.3 mM beta-mercaptoethanol (BME). Imaging buffer was incubated at 37°C for 24 h and stored at -20°C. Oxygen was removed from the imaging buffer by adding 10% (w/v) glucose, 1 mg/ml glucose oxidase and 40 μ g/ml catalase directly before each experiment. For control experiments under non-switching conditions, the imaging buffer contained 0.5 mg/ml Trolox (6-hydroxy-2,5,7,8- tetramethylchroman-2-carboxylic acid) instead of MEA or BME. I sealed the imaging chamber from oxygen with a second coverslide.

Data acquisition

Imaging was performed on a custom-built total internal reflection fluorescence (TIRF) microscope designed and built by Ludovic Le Reste and Johannes Hohlbein. It features ALEX (Kapanidis et al., 2004; Margeat et al., 2006) with a red laser (635 nm, directly modulated diode laser, Cube model, Coherent) and a green laser (532 nm, continuous-wave solid state laser, Samba model, Cobolt; modulated using an acousto-optical modulator from AA Optics). The alternation period was 200 ms in all experiments unless noted differently. The green and red laser beams were coupled into the same single-mode optical fibre. At the fibre output, the laser beams were collimated, directed into an inverted microscope (IX71, Olympus) and focused (100x oil immersion objective, NA 1.4, Olympus) onto the sample under an angle allowing for total internal reflection. The fluorescence emission was collected by the same objective, separated from excitation light by a dichroic mirror (545 nm/650 nm, Semrock) and additional filters (545 nm longpass filter, Chroma; and 633/25 nm notch filter, Semrock). To allow for dual-colour detection, the emission light was split into a red and green channel (630 nm longpass dichroic mirror, Omega) and projected onto two separate regions of an EMCCD camera (iXon 897, Andor).

To extract single-molecule intensity time-traces, TIRF movies were analysed with custom written software in MATLAB⁴ (MathWorks) (Holden et al., 2010). First, the red and green emission channels were aligned based on calibration movies of immobilised fluorescence beads. The emission of single immobilised labelled DNA molecules in the red and green channel was automatically detected and linked using a peak-finding algorithm (Crocker and Grier, 1996). Peaks that did not meet ellipticity or nearest-neighbour criteria were excluded from further

⁴The software for single-molecule FRET analysis from TIRF data was written by Seamus Holden.

analysis (eccentricity < 0.25 , nearest neighbour > 8 pixels), thereby minimising artifacts due to molecules with overlapping point-spread functions (PSFs). The emission intensity of each molecule was estimated by fitting a two-dimensional Gaussian function to its PSF. This yields F_{DD} , F_{DA} , and F_{AA} to calculate E^* and S . Intensities were divided by the emCCD gain factor to obtain actual photon counts.

Note that experiments with different settings for the relative excitation intensities at 635 nm and 532 nm show a shift in S . Combined E^* - S histograms display the data of multiple movies under identical conditions. Populations in the E^* - S histogram were identified using the standard clustering algorithm k-means. The algorithm partitions the points on the two-dimensional histogram into clusters by minimising the sum of distances of points to the centroid of their cluster. Prior to k-means clustering, data in Figs. 2.6 and 2.7 were processed using the piece-wise constant step-filter by Little and Jones (2010) with $\lambda = 2 \cdot 10^{-4}$.

Confocal microscopy

Data in Fig. 2.7 were acquired using an epifluorescence confocal scanning microscope designed and built by Johannes Hohlbein and Geraint Evans. Green (532 nm, MGL532, CNI) and red (639 nm, Power Tech. Inc.) lasers were spatially filtered and overlaid using a single-mode fibre (PMC-460, Schäfter and Kirchhoff), collimated, and coupled into an objective (60x Water Immersion, Olympus) via a dichroic mirror (XF94 535/635TBDR, Omega). Fluorescence was collected and filtered by the same objective-dichroic pair; split into its green and red components (XF2021 630DRLP, Omega); further bandpass filtered (HQ 585BP70 and HQ 725BP110, Chroma), and focused onto avalanche photodiodes (SPCM-AQR-14 Photon Counting Module, Perkin Elmer), the finite APD detector area acting as the pinhole component. ALEX spectroscopy at 10 kHz and measurement of

photon arrival times were digitally controlled (PCI-6602 board, National Instruments). A piezo electric scanning stage (P-733.2CD), with a E-761 Digital Piezo Controller PCI board (Physik Instrumente) and custom written LabVIEW software (National Instruments) was used to automatically locate single immobilised molecules.

CAP experiments

Production, expression, and purification of CAP(His6,17C,178S) protein was performed by Javier Periz, according to Kapanidis et al. (2001). Briefly, plasmid pAKCRP-(His6,17C,178S) was generated from pAKCRP-His6 using Quick Change II Site-directed Mutagenesis Kit (Stratagene). The sequence of the plasmid was confirmed by DNA sequencing. CAP(His6,17C,178S) was labelled with 10 fold molar excess of Alexa647 maleimide in 40 mM HEPES-NaOH pH 7, 200 mM KCl, 1 mM EDTA, 0.1 M TCEP, 0.2 mM cAMP. The reaction sample was quenched with 10 mM DTT for 15 min and purified using a NAP5 gel filtration column (Amersham) equilibrated with storage buffer (20 mM HEPES-NaOH pH 7, 200 mM NaCl, 1 mM DTT, 1 mM MEA and 5% glycerol). The DNA-binding activity of labelled CAP was evaluated using a shift assay described by Heilemann et al. (2009a).

A 39 bp dsDNA construct containing the CAP recognition sequence (ICAP DNA) was prepared similar to the other DNA constructs (Fig. 2.16). ICAP DNA was incubated with tenfold molar excess of CAP on PEG-passivated and neutravidin-coated coverslides. Switchable FRET experiments were performed in switching buffer (i) with additional 0.2 mM cAMP.

Holliday junction experiments

The Holliday junction construct of four separate DNA strands (X,B,R, and H strands; Fig. 2.16) was based on previous designs (McKinney et al., 2003; Hohng et al., 2004). Control experiments using Holliday junctions labelled with Cy3B and e-Cy5 only were performed with unlabelled B-strand. Hybrids were formed using 1 : 0.8 molar ratios of strands X,B,R with H. Experiments were performed in switching buffer (ii) with additional 200 mM MgCl₂.

Cy5-Cy3B-Cy5 construct	
Top Strand	5'-CAG ACG GCT GGC ACG ACT T <u>X</u> T GCA CGA TCA GCC CTG GGC GAG CGC ATT GTT GCG C-Biotin, modified at the 3'-end with Biotin and at position 20 (<u>X</u>) with Cy3B.
Bottom Strand	5'-GCG CAA CAA TGC GCT CGC CCA GGG C <u>Y</u> G ATC GTG CAA AAG TCG TGC CAG CCG TCT G-Cy5, modified at the 3'-end and at position 30 (<u>Y</u>) with Cy5.
Cy5-Cy3B Control: Identical to Cy5-Cy3B-Cy5 construct, except that the bottom strand is only 3'-modified with Cy5.	
Cy3B-Cy5 Control: Identical to Cy5-Cy3B-Cy5 construct, except that the bottom strand is only modified at position 30 with Cy5.	
ATTO655-Cy3B-ATTO655 construct	
Top Strand	5'-AAG ACG GCT GGC ACG ACT T <u>X</u> T GCA CGG TAA ATG CTG GGC GAG CGC ATT GTT GCG C-Biotin, modified at the 3'-end with Biotin and at position 20 (<u>X</u>) with Cy3B.
Bottom Strand	5'-GCG CAA CAA TGC GCT CGC CCA GCA T <u>Y</u> T ACC GTG CAA AAG TCG TGC CAG CCG TCT T-ATTO655, modified at the 3'-end and at position 30 (<u>Y</u>) with ATTO655.
ICAP DNA construct	
Top Strand	5'-Cy3B-AAC GCA ATA AAT GTG AAG TAG ATC ACA TTT TAG GCA CCA, modified at the 5'-end with Cy3B.
Bottom Strand	5'-Biotin-TGG TGC CTA AAA TGT GAT CTA CTT CAC ATT TAT TGC GTT, modified with Biotin at the 5'-end.
Holliday junction constructs	
X Strand	Cy3B-5'-CCC AGT TGA GAG CTT GAT AGG G, modified at the 5'-end with Cy3B.
B Strand	5'-CCC TA <u>X</u> CAA GCC GCT GCT ACG G, modified at position 6 (<u>X</u>) with Cy5 (internal "i-Cy5").
R Strand	Cy5-5'-CCC ACC GCT CTT CTC AAC TGG G, modified at the 5'-end with Cy5 (end "e-Cy5").
H Strand	Biotin-5'-CCG TAG CAG CGA GAG CGG TGG G, modified at the 5'-end with Biotin.
DNA polymerase primer-template construct for TIRF microscopy	
Primer Strand	Biotin-5'-CCT CAT TCT TCG TCC CAT <u>X</u> AC CAT ACA TCC-H, modified at position 19 with Atto647N; 3'-dideoxy terminated.
Template Strand	5'-ATG CTA GGA TGT ATG GTA ATG GGA CGA AGA ATG AGG
DNA polymerase primer-template construct for confocal microscopy	
Hairpin	5' GGT AGA GAC CTG TTG ACT CGA AAG AGT CAA CAG GTC-H, 3'-dideoxy terminated.

Figure 2.16: DNA sequences (5' to 3') and modifications for switchable FRET (chapter 2) and DNA polymerase experiments (chapter 3).

Chapter 3

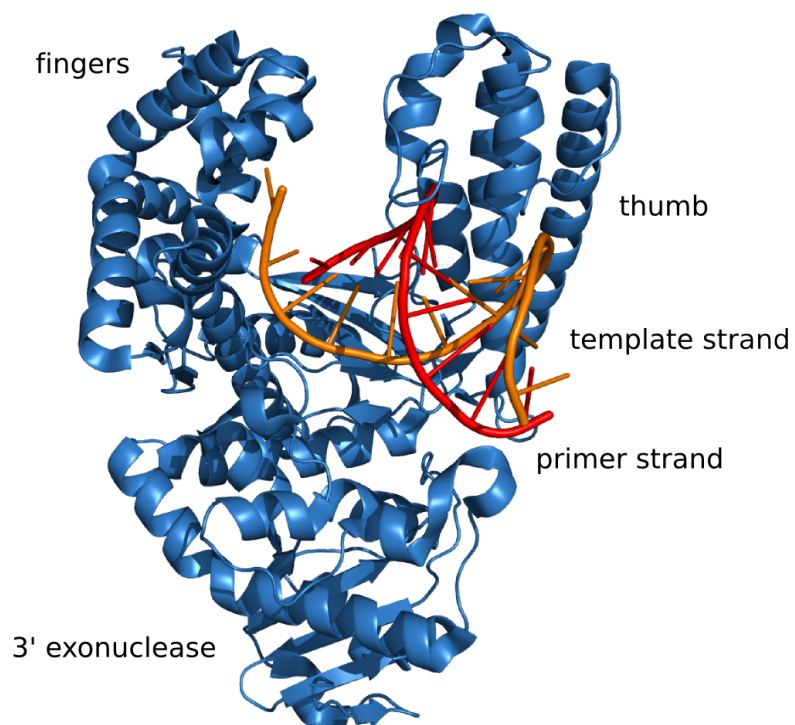
Conformational states of DNA polymerase I

3.1 Introduction

In this chapter, I discuss application of single-molecule FRET to study structure and conformational dynamics of DNA polymerase I. This work provides an initial experimental characterisation towards future application of switchable FRET.

DNA polymerase I (Pol), encoded by the *polA* gene in *E.coli*, was the first DNA polymerase to be discovered (Lehman et al., 1958) and remains the primary model enzyme for studies of DNA synthesis (reviewed by Kornberg (2005); Patel et al. (2001); Steitz and Yin (2004)). Moreover, it serves as a general prototype for a complex molecular machine with multiple domains and cellular functions. At 400 copies per cell, Pol is the most abundant DNA polymerase in *E.coli* with a primary role in replacing short RNA primers generated during lagging-strand replication (chapter 6). Moreover, Pol is the main enzyme for DNA repair synthesis, filling ssDNA gaps in base excision and nucleotide excision repair processes (chapter 5). Mutant *E.coli* deficient in Pol are hyper-sensitive to DNA damage

A



B

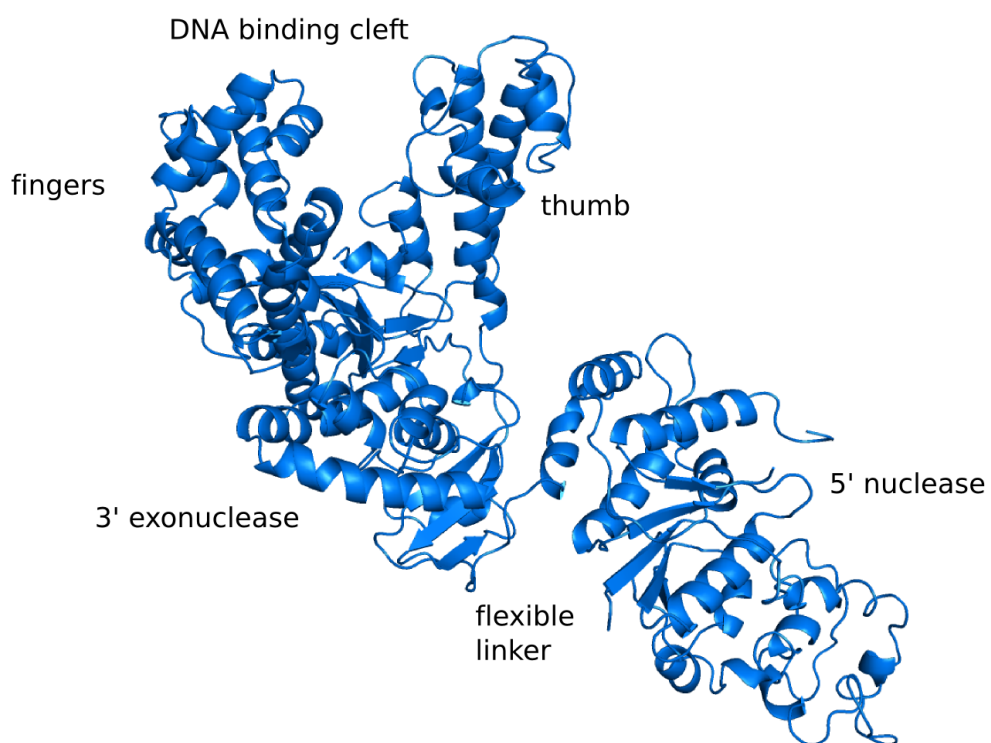


Figure 3.1: Crystal structures of DNA polymerase I homologues (A) Klenow fragment from *Geobacillus stearothermophilus* (Johnson et al., 2003) (PDB file 1L3U)(B) Full-length protein from *Thermus aquaticus* (Kim et al., 1995) (PDB file 1TAQ) .

(Lucia and Cairns, 1969; Ljungquist et al., 1976). Pol is a 109 kDa enzyme of 928 amino acids that catalyses DNA synthesis from a 3' primer terminus in 5' to 3' direction according to the sequence of the template strand. The main steps of the general polymerisation reaction are: (i) Pol locating the primer site; (ii) binding of a deoxyribonucleoside triphosphate (dNTP); (iii) rejection of a mismatching dNTP or positioning of the complementary dNTP for catalysis; (iv) phosphoryl transfer reaction catalysis; (v) pyrophosphate release. Translocation to the next template position after nucleotide incorporation makes Pol processive for tens of nucleotides at a modest synthesis rate of about 10 nt/s (Schwartz and Quake, 2009).

Proteolysis of Pol generates the large Klenow fragment (KF) of 605 residues containing the polymerase and 3' exonuclease activities, and a small fragment with the 5' nuclease activity. X-ray crystallography revealed the structure of the polymerase domain, broadly resembling a right hand, with thumb, fingers, and palm subdomains (Fig. 3.1A) (Ollis et al., 1985; Johnson et al., 2003). The thumb subdomain binds the upstream DNA duplex while the fingers bind the incoming dNTP and the downstream ssDNA template strand. The palm subdomain hosts the polymerase active site in the cleft between fingers and thumb. Two divalent magnesium ions are positioned by carboxylate side-chains in the catalytic center and interact with the primer terminus OH group and the incoming nucleotide phosphate group to enable phosphodiester bond synthesis.

The extraordinary DNA synthesis fidelity, with error rates as low as $10^{-5} - 10^{-8}$, can not be accounted for by the energetic differences between correct and incorrect base-pairing alone (Kunkel and Bebenek, 2000). It is rather a result of the active site geometry which sterically restricts catalysis to correctly base-paired nucleotides. To this end, a reversible conformational change closes the fingers subdomain upon complementary nucleotide binding (Li et al., 1998; San-

toso et al., 2010). Fingers closing promotes catalysis of the phosphoryl transfer reaction by forming a tight binding pocket for the correct nucleotide, while mispaired nucleotides or ribonucleotides appear to be rejected earlier in the reaction pathway by a partially closed fingers conformation (Santoso et al., 2010; Wu and Beese, 2011). Further reduction of the error rate is achieved by the 3' to 5' exonuclease proofreading function in the palm subdomain. Misincorporation of a nucleotide stalls Pol synthesis and destabilises the duplex primer terminus which then binds the 3' exonuclease site for removal of the mismatched terminal nucleotide.

The N-terminal fragment of full-length Pol contains the 5' nuclease domain which degrades downstream duplex DNA or RNA. It has a separate fold connected to the Klenow fragment via a flexible linker (Kim et al., 1995) (Fig. 3.1B). However, the overall structure and conformational changes of Pol in solution, with nucleotides, and in complex with different DNA substrates remain unknown. Downstream strand displacement by the polymerase domain produces a 5' flap for endonucleolytic cleavage. It has been proposed that the conformational states of the 5' nuclease domain and the polymerase domain might be coupled to function in a coordinated manner for efficient nick translation activity (Lundquist and Olivera, 1982; Xu et al., 1997, 2000).

Single-molecule FRET and switchable FRET might be able to test this hypothesis and reveal the domain organisation of Pol by measuring several distances between the two protein domains and relative to different DNA substrates. Previous confocal ALEX experiments on diffusing KF by Santoso et al. (2010) showed the relative populations of the open and closed fingers states but the millisecond observation times per molecule prohibited prolonged observation of the conformational dynamics. Here, I used TIRF microscopy to monitor KF bound to surface-immobilised DNA. I describe single-molecule FRET experiments for an

initial characterisation of the DNA binding characteristics and conformational changes of KF and full-length Pol.

3.2 Binding of Klenow fragment to immobilised DNA

Following pioneering studies on HIV reverse transcriptase (Liu et al., 2008) and KF (Christian et al., 2009; Schwartz and Quake, 2009), I studied transient complex formation of KF with immobilised DNA. The particular assay I used was established by Johannes Hohlbein and Ludovic Le Reste (Le Reste et al., 2012). KF was mutated, expressed, purified, and site-specifically labelled with donor fluorophore Cy3B at position 744 on the fingers subdomain by Catherine Joyce (Yale Biochemistry), as described by Santoso et al. (2010). A D424A mutation abolishes the 3' exonuclease activity to avoid DNA degradation. I used DNA with a 30 bp duplex portion upstream of a 6 nt overhang of the template strand. The non-template strand was labelled with acceptor fluorophore Atto647N at a distance of 12 nt from the 3' terminus. Primer extension was prevented by a dideoxy (3'-H) modification of the terminus, which prohibits the phosphoryl transfer reaction but does not interfere with nucleotide binding. Biotinylated DNA constructs were immobilised on PEG-passivated and neutravidin-coated microscope slides following the protocol by Selvin and Ha (2008) and experiments were performed under oxygen scavenging conditions with Trolox according to Rasnik et al. (2006). TIRF excitation limited the background fluorescence resulting from 10 nM labelled KF in solution.

The repeated appearance of Cy3B fluorescence at positions marked by constant Atto-647N fluorescence showed transient binding of KF to immobilised DNA

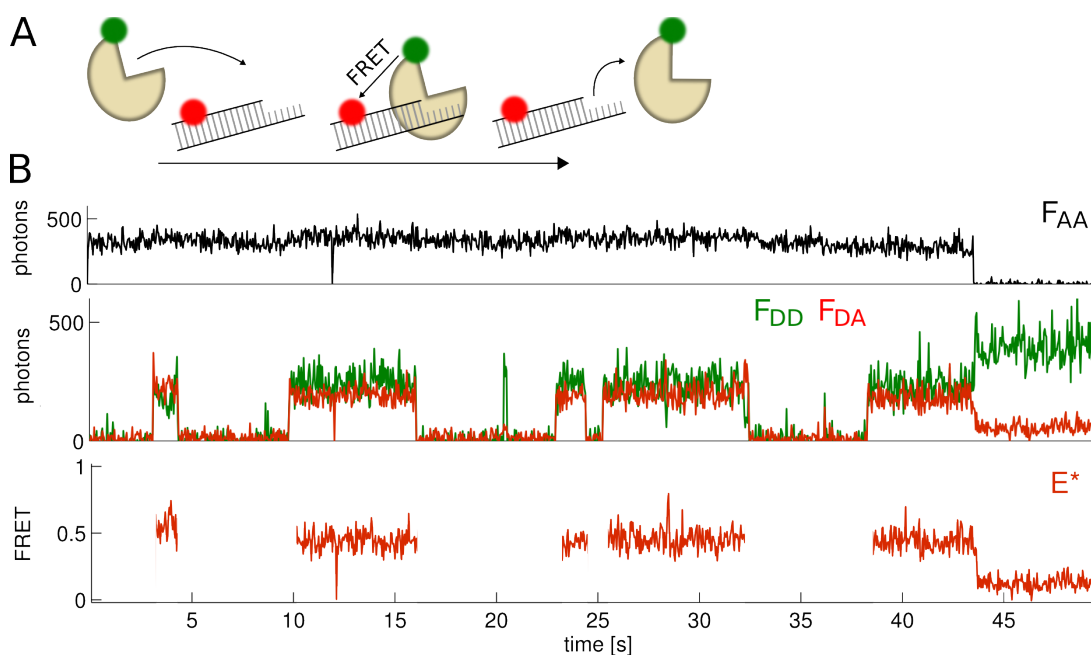


Figure 3.2: Direct observation of KF binding to a DNA primer. **(A)** Schematic of Cy3B-labelled KF binding to Atto647N-labelled primer-template DNA. **(B)** Fluorescence and FRET time-traces showing repeated binding of KF (intervals with simultaneous F_{DD} and F_{DA} fluorescence) to a single immobilised primer-template DNA construct (constant F_{AA} fluorescence). The Atto647N acceptor on the DNA bleaches after 43 s. The FRET time-trace is only shown during KF binding events where $F_{DD} + F_{DA} > 100$ photons.

(Fig. 3.2). Single KF binding events lasted for several seconds, in agreement with Markiewicz et al. (2012) who used a similar assay. The crystal structure of KF binary complex (Johnson et al., 2003) with manual extension of the upstream duplex DNA suggests a distance of 6.3 nm between the C-5 carbon of the DNA and the C- α carbon of the protein labelling positions. Consistent with this and the Förster Radius of 6.7 nm for Cy3B-Atto647N (Santoso et al., 2010), KF binding events in Fig. 3.2 and the combined data of 50 time-traces in Fig. 3.3A exhibited one main FRET state at $E^* = 0.48$. The static crystal structure hence appears to reflect the major conformation of the KF-DNA binary complex. More detailed structural insight from these data could be obtained by modeling the fluorophore linkers (Woźniak et al., 2008; Muschielok et al., 2008) and calculating accurate FRET efficiencies to recover donor-acceptor distances (Lee et al., 2005).

3.3 Conformational dynamics of Klenow fragment

To investigate the conformational change of the fingers subdomain with nucleotide binding, I added 1 μM dTTP complementary to the adenosine template base. This resulted in the appearance of an additional higher FRET state at $E^* = 0.68$ (Fig. 3.3B). A corresponding single-molecule time-trace shows reversible FRET changes between the two states with dwell times around 100-200 ms in either state (Fig. 3.3D). Both states are equally populated, in broad agreement with a dissociation constant of $k_D = 0.83 \mu\text{M}$ for the complementary nucleotide dTTP (Hohlbein et al., 2012). Saturating 100 μM dTTP shifted the population almost completely to the high FRET state (Fig. 3.3C). The increased FRET efficiency for the ternary complex matches the distance change in crystal structures from 6.3 nm in the open fingers configuration to 5.6 nm in the closed fingers conformation with the complementary nucleotide (Johnson et al., 2003). In the absence of nucleotides, KF binding events also show occasional increased FRET excursions that might reflect conformational sampling of the fingers subdomain in the absence of nucleotides. This is reflected in the spread of the E^* distribution towards higher FRET values (Fig. 3.3A).

3.4 Comparing Klenow fragment and full-length DNA polymerase I

Mechanistic studies of DNA polymerases typically employ Klenow fragment. The full-length enzyme, however, includes the 5' nuclease domain at the N-terminus. To test whether the conformational states of the polymerase domain are affected

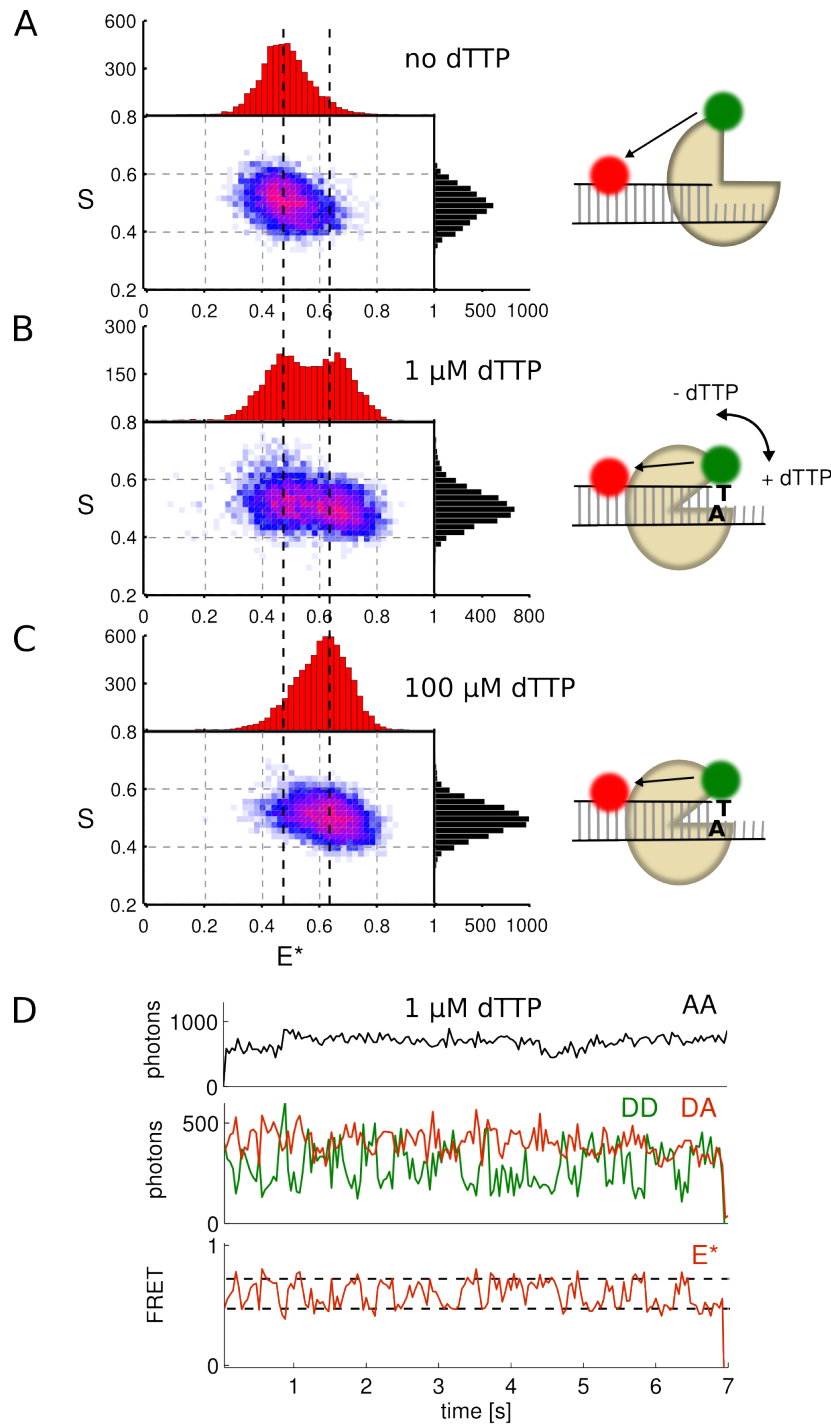


Figure 3.3: Reversible conformational changes of the KF fingers subdomain upon binding of the complementary nucleotide. **(A)** E^* - S histograms of Cy3B-labelled KF bound to immobilised primer-template DNA labelled with Atto647N, recorded at 40 ms ALEX period without nucleotides (top, 50 molecules), with 1 μM complementary nucleotide dTTP (middle, 30 molecules), with 100 μM complementary nucleotide dTTP (bottom, 31 molecules). KF binding events were identified by thresholding $F_{\text{DD}} + F_{\text{DA}} > 100$ photons. Cartoons show label positions and protein conformation. **(B)** Fluorescence and FRET time-traces of a single Cy3B-labelled KF molecule bound to immobilised primer-template DNA labelled with Atto647N in the presence 1 μM complementary nucleotide dTTP.

by the absence or presence of the 5' nuclease domain, I compared FRET measurements of KF and Pol labelled at positions 550 (thumb) with acceptor Atto647N and at 744 (fingers) with donor Cy3B¹. Santoso et al. (2010) used this labelling configuration previously to show conformational changes of KF upon DNA binding and nucleotide selection. To prevent DNA degradation by the 5' nuclease domain, I replaced MgCl₂ with CaCl₂ in the imaging buffer. Using confocal ALEX spectroscopy (Kapanidis et al., 2004), I measured fluorescence bursts of freely diffusing molecules as they transit the excitation volume². Distributions of E^* and S represent the combined data of thousands of bursts.

Unliganded KF and Pol display broad distributions of conformations between $E^* = 0.4 - 0.7$ (Fig. 3.4A). Torella et al. (2011) showed that the mobile fingers subdomain dynamically samples conformations in the unliganded state. The equilibrium between open and closed fingers conformations appears slightly shifted to the open state for unliganded Pol compared to KF. In the presence of saturating 100 nM primer-template DNA (3'-dideoxy terminated; dissociation constant $k_D \approx 7.9$ nM for KF (Guest et al., 1991); sequence in Fig. 2.16), KF and Pol both primarily adopt a low FRET state corresponding to the open fingers conformation (Fig. 3.4B). With saturating 1 mM complementary nucleotide dTTP, KF and Pol data shift to the high FRET state reflecting closed fingers subdomains (Fig. 3.4C). Because the fingers-thumb conformations are very similar for KF and Pol and match previous data for KF (Santoso et al., 2010), I conclude that the structure of the polymerase domain is largely unaffected by the presence or absence of the 5' nuclease domain. This validates the interpolation of results obtained using KF to represent the function of the full-length enzyme.

¹Doubly-labelled Pol was produced by Timothy Craggs and Catherine Joyce.

²I used the confocal microscope constructed by Geraint Evans and Johannes Hohlbein, described in chapter 2.6.

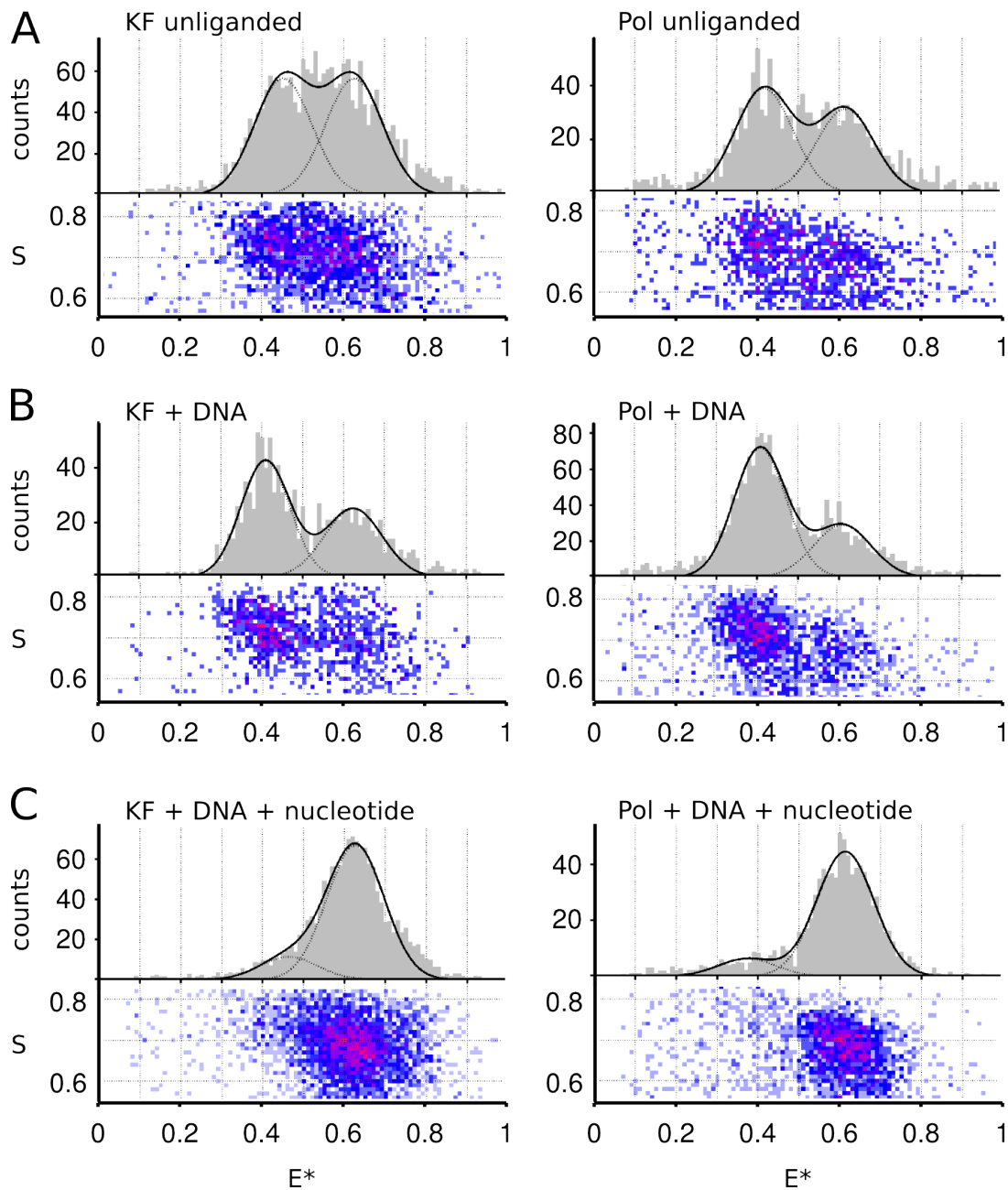


Figure 3.4: Comparing conformational states of the polymerase domain of Klenow fragment (KF, left column) and full-length DNA Polymerase I (Pol, right column). Both enzymes were labelled at positions 550 (thumb) with Atto647N and at 744 (fingers) with Cy3B. Fitted double-Gaussian distributions and the single-Gaussian components are shown in black and gray, respectively. **(A)** Unliganded enzymes. **(B)** Additional 100 nM DNA with dideoxy primer terminus. **(C)** Additional 100 nM DNA with dideoxy primer terminus and 1 mM complementary nucleotide dTTP.

3.5 Conclusion

Direct detection of single polymerases binding to immobilised DNA constructs has been used to measure their dissociation constants (Christian et al., 2009; Le Reste et al., 2012; Markiewicz et al., 2012), synthesis speed (Schwartz and Quake, 2009), and time-traces of conformational changes (Liu et al., 2008). Here, I applied single-molecule FRET to observe DNA binding, conformational states, and structural dynamics of DNA polymerase I. I showed dynamic closing and opening of the polymerase fingers subdomain in the presence of a complementary nucleotide. This technique could be used to quantify the steps in the process of nucleotide selection and incorporation using complementary and mismatching nucleotides and ribonucleotides.

Berezhna et al. (2012) recently employed a similar FRET labeling scheme to show the partially closed state of KF in the presence of incorrect nucleotides. However, the temporal resolution of their experiments was not sufficient to resolve the state transitions and dwell times in the open, partially closed, and closed states. Instead, most binding events exhibited one averaged FRET level which might explain the inconsistency in the relative population of the states with my data and previous results (Santoso et al., 2010). Berezhna et al. also interpreted an additional fourth FRET state as binding of a fully paired primer terminus to the 3' exonuclease site. On the other hand, Markiewicz et al. (2012) showed that an intact primer exclusively binds the polymerase site, which is in agreement with my observation of two major FRET states of KF relative to DNA.

I replicated previous KF measurements using full-length Pol and found that the polymerase domain conformation is largely unaffected by the absence or presence of the 5' nuclease domain. Timothy Craggs complemented these results using doubly-labelled Pol at position 550 on the thumb and at different positions on

the 5' nuclease domain towards generating a model of the domain arrangement in solution and in complex with different DNA substrates. The initial results presented here pave the way for application of switchable FRET to investigate coordination between the polymerase and nuclease domains during strand displacement synthesis and nick translation. Measuring multiple nick displacement reactions on the same substrate could be achieved using a rolling-circle DNA replication system (Tanner et al., 2009). While single-pair FRET measurements can only monitor the position of one domain relative to the DNA substrate, switchable FRET should be able to measure the organisation of both domains in the same complex.

3.6 Materials and methods

TIRF microscopy experiments

KF(N-His6,D424A,L744C,C907S) was produced and site-specifically labelled with Cy3B at position 744 by Catherine Joyce as described by Santoso et al. (2010). The primer-template DNA, labelled with Atto647N, was prepared as above (sequences and modifications in Fig. 2.16) and immobilised on PEG-passivated slides as described above (chapter 2.6). Imaging was performed with 10 nM KF in 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 5% glycerol, 100 µg/ml BSA, 4% (w/v) glucose, 1 mg/ml glucose oxidase, 40 µg/ml catalase, and 1 mM Trolox. TIRF microscopy data acquisition and analysis were as described in the previous chapter.

Confocal microscopy experiments

Pol(N-His6,C262S,D424A,K550C,L744C,C907S) was produced and site-specifically labelled with Cy3B at position 744 and Atto647N at position 550 by Timothy

Craggs and Catherine Joyce. Doubly-labelled KF and primer-template DNA, described previously (Santoso et al., 2010), were also provided by Catherine Joyce. ALEX spectroscopy on diffusing molecules was performed using the confocal microscopy setup described above. I measured 100 pM KF or Pol in 50 mM Tris-HCl pH 7.5, 10 mM CaCl_2 , 100 $\mu\text{g/ml}$ BSA, 5% glycerol; and additional 100 nM primer-template DNA (sequence in Fig. 2.16) and 1 mM dTTP as mentioned in the text. Data were analysed to detect fluorescence bursts and extract E^* and S distributions using LabView and MATLAB algorithms developed previously (Kapanidis et al., 2004; Selvin and Ha, 2008).

Chapter 4

Super-resolution fluorescence microscopy

4.1 Introduction

The previous chapters focused on single-molecule FRET spectroscopy to probe structure and dynamics of DNA molecules and protein-DNA complexes with sub-nanometre resolution. However, the accessible length-scale of this method is limited by the steep distance-dependence of the FRET interaction to within 10 nm. Investigating the molecular organization of larger complexes, biopolymers, up to organelles, cells, and tissues requires microscopy methods that achieve molecular resolution on larger scales. Optical diffraction broadens a fluorescence point source into a ~ 125 nm wide point-spread function (PSF) in the microscope image. PSFs of dense fluorophore labels on cellular structures consequently overlap and produce a blurred representation of the sample. Abbe's formula limits the resolution of overlapping emitters in the sample, yet, it does not limit the localisation precision of a single isolated point-source. In principle, the center of an isolated PSF can be determined with arbitrary precision,

limited only by the number of sampled photons (Thompson et al., 2002). This concept allowed direct imaging of myosin motor steps with nanometre precision (Yildiz et al., 2003). Imaging densely labelled specimen by the same method requires spatial isolation of PSFs for their individual localisation. This problem was overcome by the discovery of photoswitching in fluorophores (Bates et al., 2005; Heilemann et al., 2005; Vogelsang et al., 2009) and development of photoactivatable fluorescent proteins (Patterson and Lippincott-Schwartz, 2002; Subach et al., 2009b). Stochastic Optical Reconstruction Microscopy (STORM) forms a super-resolution image by sequential imaging and localisation of small subsets of stochastically activated fluorophores while the majority of labels reside in the non-fluorescent state (Rust et al., 2006). Photoactivated Localization Microscopy (PALM) employs the same concept using particular photoactivatable fluorescent proteins (Betzig et al., 2006). Complementary super-resolution microscopy approaches including Structured Illumination Microscopy (SIM) and Stimulated Emission Depletion Microscopy (STED) use particular excitation patterns for super-resolution imaging of cellular structures (Hell, 2007). Here, I focus on photoswitching and photoactivation techniques that allow both super-resolution representation of structures in the cell as well as tracking of single molecules at high intracellular concentrations. This chapter describes the implementation and further development of these methods, followed by applications to study protein-DNA interactions in live bacteria (chapters 5 and 6).

4.2 Microscope design and construction

In cooperation with Justin Pinkney (Achillefs Kapanidis group), I designed, built, and maintained a custom multi-colour TIRF microscope for super-resolution localisation and single-molecule FRET applications. Our design principles followed optimisation of several factors: (i) maximising photon detection efficiency by minimising the number of optical components, (ii) flexibility to use the setup for super-resolution microscopy and single-molecule FRET and allow for future extensions, (iii) mechanical stability.

Based on these requirements I generated the design shown in Fig. 4.1¹. We constructed the setup on an actively damped optical table from individual opto-mechanical components (Thorlabs). We also engineered custom mounts for several components in the mechanical workshop, namely for the objective, fibre coupling unit, stage, and five lasers. To allow imaging a range of fluorophores spanning the visible spectrum, we used the following lasers: (i) 405 nm laser (MLL 100 mW, OptoEngine) for photoactivation of e.g. PAmCherry, Alexa532, Alexa647; (ii) 475 nm laser (Stradus 70 mW, Vortran) for imaging e.g. YPet, Alexa488; (iii) 543 nm laser (MGL 100 mW, Laser2000) for imaging e.g. Cy3B, Alexa532, Alexa546; (iv) 561 nm laser (SLIM-561 200 mW, Oxixus) for imaging e.g. PAmCherry, Alexa568; (v) 640 nm laser (Cube 100 mW, Coherent) for imaging e.g. Cy5, Alexa647, Atto647N, Atto655. All laser beams apart from 405 nm were combined using dichroic mirrors (z561bcm, z488bcm, Chroma) and directed into an acousto optical tunable filter (AOTF, MSD040-150-0.2ADS-A5H-8X1, Gooch & Housego) to adjust excitation intensities and drive alternating-laser excitation schemes. We used National Instruments Compact DAQ analog and TTL digital output modules (NI 9263 and NI 9401) to control the AOTF and trigger

¹The final setup configuration is shown. The initial design featured only three lasers and was optimised for FRET applications.

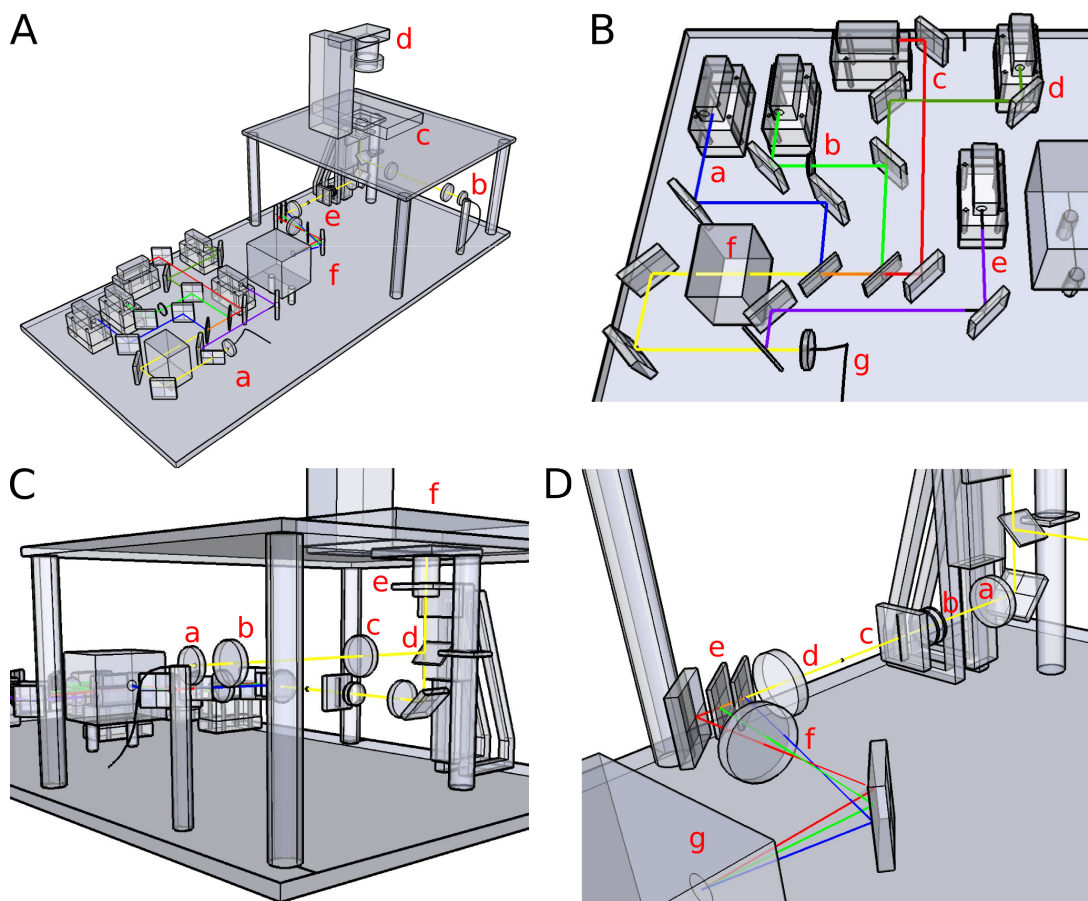


Figure 4.1: 3D design of the multi-colour single-molecule TIRF microscope generated in Google Sketchup. **(A)** Overview of the setup: (a) multi-laser excitation source, (b) TIRF excitation, (c) sample stage, (d) LED light source, (e) three-colour fluorescence emission path, (f) EMCCD camera. **(B)** Multi-laser excitation source: (a) 473 nm laser, (b) 543 nm laser, (c) 640 nm laser, (d) 561 nm laser, (e) 405 nm laser, (f) AOTF, (g) optical fibre. **(C)** TIRF excitation: (a) optical fibre output, (b) collimating lens, (c) TIRF lens, (d) polychroic mirror, (e) objective, (f) sample stage. **(D)** Three-colour fluorescence emission path: (a) lens, (b) polychroic emission filters, (c) slit, (d) lens, (e) dichroic mirrors for three emission channels, (f) tube lens, (g) EMCCD camera.

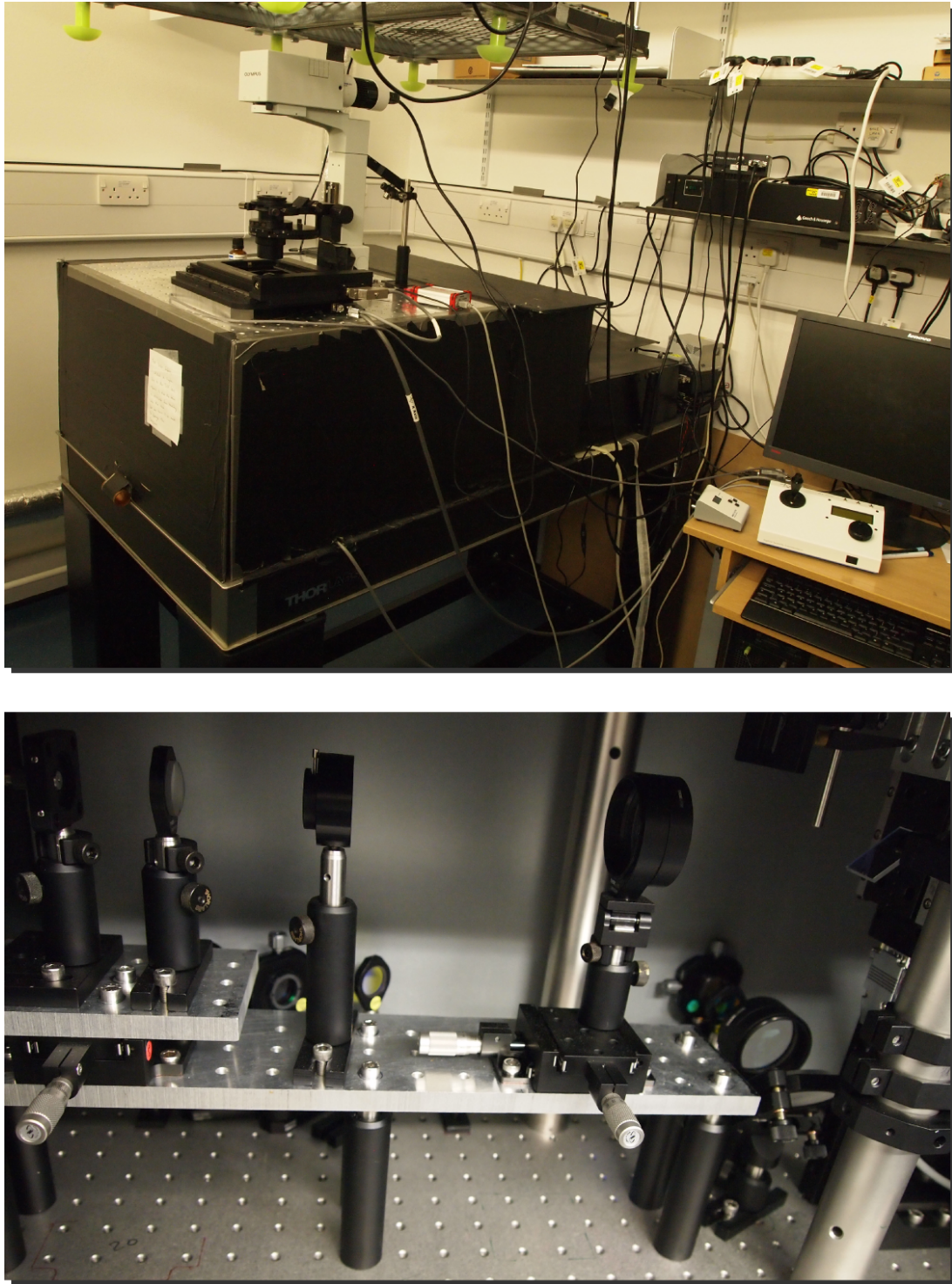


Figure 4.2: Photos of the custom TIRF microscope. Top: Whole setup with the sample stage in front and the excitation source at the back (compare to Fig. 4.1A). Bottom: TIRF excitation path with the optical fibre output at the left and the polychroic mirror at the right (compare to Fig. 4.1C). Part of the emission path can be seen in the lower right corner (compare to Fig. 4.1D).

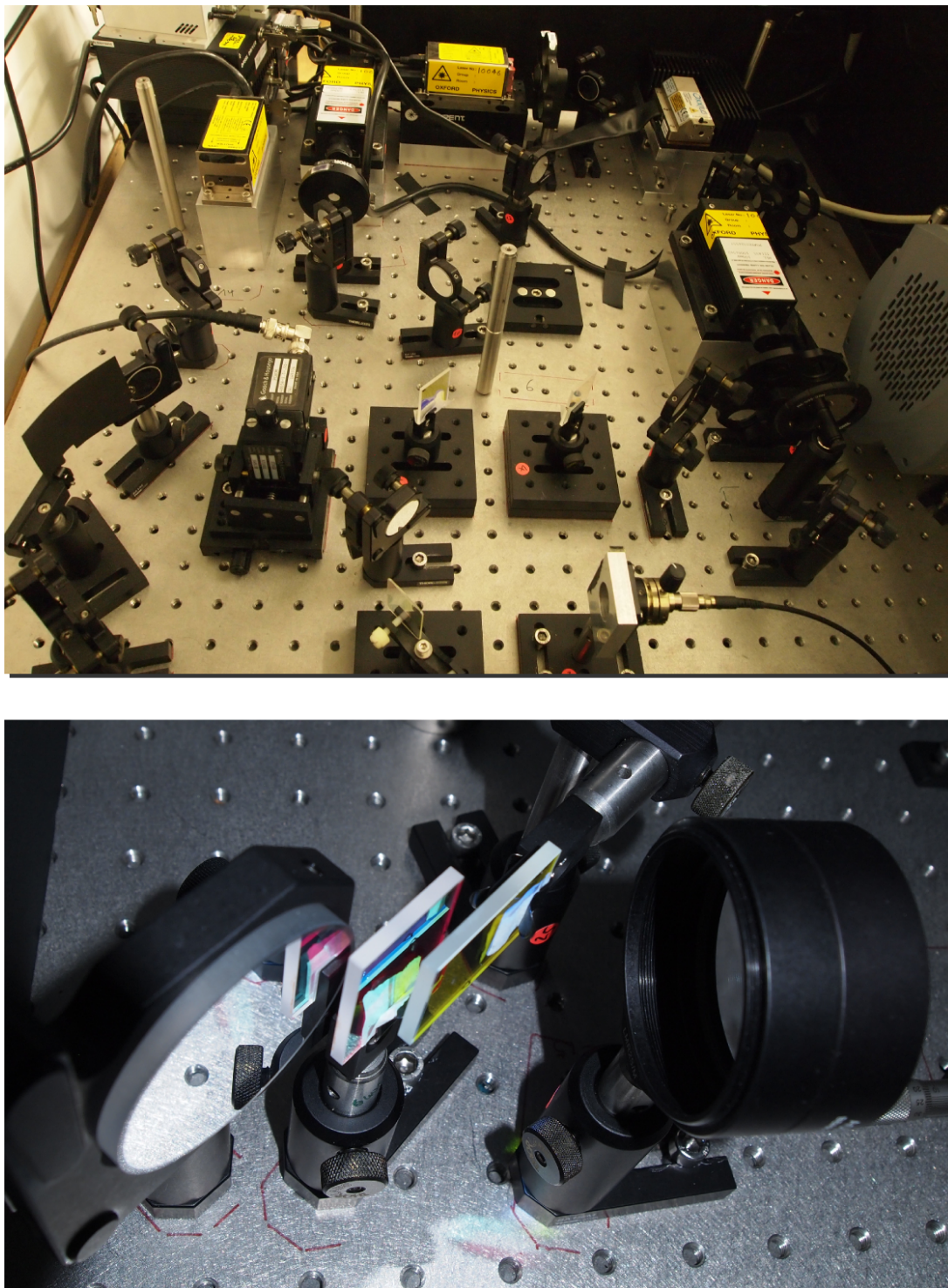


Figure 4.3: Photos of the custom TIRF microscope. Top: Multi-laser excitation source (compare to Fig. 4.1B). Bottom: Dichroic mirrors for three emission channels (compare to Fig. 4.1D).

the microscope camera (The LabView software was written by Justin Pinkney). The 405 nm laser was later added to the setup with a separate neutral density filter wheel for intensity control, bypassing the AOTF because the wavelength was not compatible. All beams were coupled into the same single-mode optical fibre to separate alignment of the excitation path from the rest of the microscope and achieve a Gaussian excitation profile.

At the fibre output, the beams were collimated (50 mm EFL aspherized achromatic lens, Edmund Optics) and focused (250 mm visible achromatic lens, Thorlabs) in the back focal plane of the objective (100x oil immersion objective, NA 1.4, focal length 1.8 mm, Olympus) such that the angle of excitation can be controlled by translating the lens for TIRF excitation (Fig. 4.1C). Reflected TIRF excitation was blocked by a beam stop underneath the objective. Fluorescence emission was collected by the same objective, separated from excitation light by a custom polychroic mirror and quad-band emission filter (ZT405/473/561/640rpc, ZET405/473/561/640NF, Chroma)². The microscope image was focused using two lenses (150 mm visible achromatic lenses, Thorlabs) on either side of a slit (Thorlabs) to create a rectangular shape for simultaneous imaging of three spectral regions side by side on the same camera. To this end, we constructed a triple-view component that sequentially separates blue emission up to 540 nm (540lpxr dichroic, Chroma), followed by green-yellow emission up to 640 nm (640lpxr dichroic, Chroma), leaving red emission beyond 640 nm. A single tube lens (250 mm visible achromatic lens, Thorlabs) was used to focus all emission channels onto a back-illuminated electron-multiplying CCD camera with 512x512 pixels and -80°C air cooling (iXon 897, Andor). Our combination of 1.8 mm

²The original design had two separate polychroic mirrors and filter sets for FRET or PALM applications.

focal length objective and 250 mm tube lens results in an overall magnification of $250 \text{ mm}/1.8 \text{ mm} = 138.9$. This corresponds to a relative pixel size on the camera of $16 \text{ } \mu\text{m}/138.9 = 0.1152 \text{ } \mu\text{m}$ ($0.1145 \text{ } \mu\text{m}$ by microruler measurement).

To record transmitted light images of cells, we included an LED light source (pE-100, coolLED) in an Olympus condensor (IX-2, Olympus). Sample position and focus were controlled using a motorized stage with z-piezo and a separate z-motor for coarse focusing (MS2000 and LS50, ASI). We also installed an objective heater for experiments at 37°C (Biopetechs). The emission path was completely shielded for experiments under ambient lighting. Photos of the microscope are shown in Fig. 4.2 and 4.3. My following results demonstrate the super-resolution capabilities of this microscope; other group members performed dual- and three-colour single-molecule FRET experiments.

4.3 Stochastic Optical Reconstruction Microscopy (STORM)

Localisation precision

Super-resolution localisation microscopy is based on accurate detection and localisation of fluorophores. In the case of isotropic rotational freedom of fluorophores, the fluorescence PSF can be approximated by a circular two-dimensional Gaussian function (Thompson et al., 2002; Mortensen et al., 2010). The standard analysis procedure involves initial identification of candidate PSFs using a fixed intensity threshold on each frame of a super-resolution movie. Band-pass filtering removes spurious high-frequency noise and low frequency variation in intensity across the image. Candidate positions serve as initial guesses for precise fitting of the PSF centre. Here, I used the Levenberg-Marquardt C++ least-squares optimisation algorithms that was incorporated into the custom MATLAB (MathWorks) software for single-molecule FRET analysis by Seamus Holden (Holden et al., 2010). In the future, improved resolution could be obtained by maximum likelihood parameter estimation which accurately models the non-uniform Poisson noise in the signal (Mortensen et al., 2010; Smith et al., 2010). To account for optical aberrations and motion blurring by moving fluorophores that distort the PSF, I used an elliptical Gaussian function to fit the intensity I_{ij} per pixel with coordinates (i, j)

$$I_{ij} = \frac{I}{2\pi \cdot s_x s_y} \cdot \exp \left(-\frac{x'^2}{2s_x^2} - \frac{y'^2}{2s_y^2} \right) + c . \quad (4.1)$$

Here, I is the total number of photons and the parameter c accounts for the camera offset and background signal. To allow for arbitrary rotation (θ) and position (x_0, y_0) of the elliptical PSF, x' and y' coordinates were expressed within

the coordinate system aligned to the minor and major elliptical axes, such that $x' = (i - x_0) \cdot \cos(\theta) - (j - y_0) \cdot \sin(\theta)$ and $y' = (i - x_0) \cdot \sin(\theta) + (j - y_0) \cdot \cos(\theta)$. The precision of localising the centre of a symmetrical PSF by least-squares Gaussian fitting is given by (Mortensen et al., 2010):

$$\sigma_{\text{loc}} = \sqrt{2} \cdot \sqrt{\frac{\sigma_{\text{PSF}}^2 + \frac{a^2}{12}}{I} \cdot \left(\frac{16}{9} + \frac{8\pi \cdot b^2 (\sigma_{\text{PSF}}^2 + \frac{a^2}{12})}{I \cdot a^2} \right)}, \quad (4.2)$$

with PSF width σ_{PSF} (std), pixel size a , and background noise per pixel b (std). The factor $\sqrt{2}$ accounts for electron-multiplying noise of the camera gain. Typical experimental parameters for the microscope described in section 4.2 and relatively dim fluorescent proteins used for PALM are $\sigma_{\text{PSF}} = 125$ nm, $a = 114.5$ nm, $I = 100$ photons, $b = 3$ photons. This gives a localisation precision of $\sigma_{\text{loc}} = 39.5$ nm, in agreement with my experimental estimate of 40 nm (chapter 4.4). Bright synthetic fluorophores with $I > 1000$ photons used for STORM provide a better localisation precision of $\sigma_{\text{loc}} < 8.3$ nm

STORM imaging of microtubule

Microtubule cytoskeleton filaments, immunostained with photoswitchable fluorophores, served as a standard to demonstrate super-resolution microscopy (Bates et al., 2007; Heilemann et al., 2008). Microtubule are protein polymers with a diameter of 25 nm that form a dense filamentous network of tracks for motor proteins. Polymerisation and decay of microtubule drives chromosome separation during mitosis and provides structural integrity to eukaryotic cells. Cos-7 cells were kindly provided by Argyris Papantonis (Dunn School of Pathology, Oxford). I fixed cells in LabTek 8-well chambered cover-glass (Nunc) with 4% paraformaldehyde in PBS for 20 min at 20°C, and permeabilised with 0.2% Triton X-100 in PBS for 3 min. Non-specific binding sites were blocked by incu-

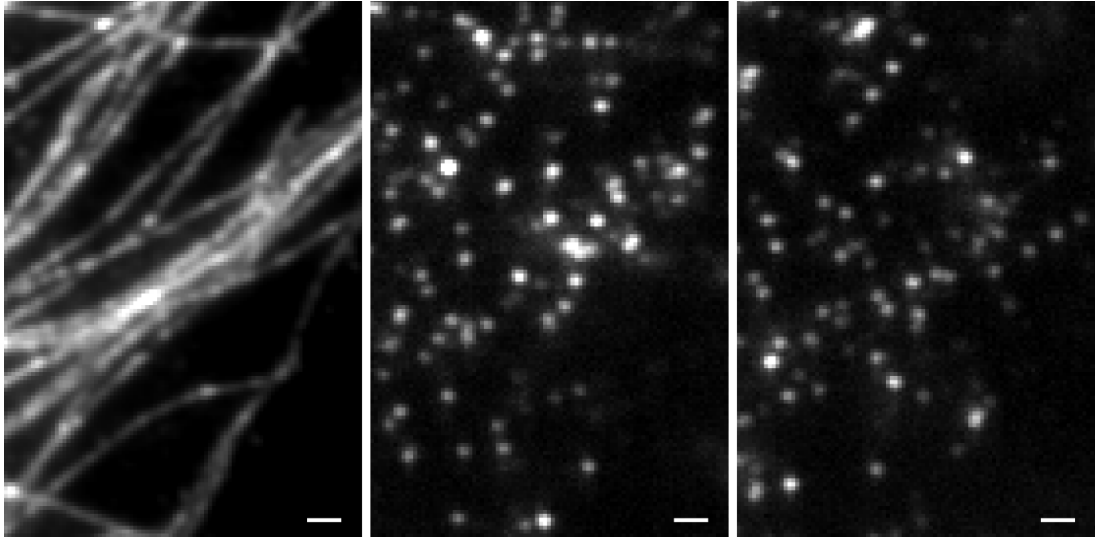


Figure 4.4: Diffraction-limited TIRF microscopy image of Alexa647-labelled microtubule in Cos-7 cells and two example frames of the corresponding STORM movie with photoswitching Alexa647 fluorophores. The TIRF image was generated by accumulating fluorescence of the whole STORM movie. Scale bars: 1 μm .

bating cells with 5% BSA for 20 min. Microtubule were specifically labelled with mouse monoclonal anti-beta-tubulin (Sigma T9026) for 1 h at 37°C, followed by incubation with secondary Alexa647-labelled antibody (goat anti-mouse F(ab')₂ fragment, Invitrogen A-21237) for 1 h at 20°C.

Photoswitching of Alexa647 fluorophores was induced in freshly prepared imaging buffer with 50 mM mercaptoethylamine (MEA) in PBS, supplemented with the standard oxygen scavenging system (10% w/v glucose, 1 mg/ml glucose oxidase, and 40 $\mu\text{g}/\text{ml}$ catalase). I acquired STORM movies at 40 ms exposure time for 4000 frames using 20 mW 640 nm TIRF excitation. Alexa647 photoactivation was promoted by 100 μW 543 nm irradiation. Figure 4.4 shows a dense microtubule network represented by diffraction-limited TIRF microscopy and example frames of the corresponding STORM movie with clear PSFs of photoswitching Alexa647 fluorophores. To detect fluorophore localisations even in the case of dense overlapping PSFs, I used automated image analysis software de-

veloped by Seamus Holden (Achillefs Kapanidis group), described in section 4.3. I reconstructed the STORM image as a colour-coded density map from $9.2 \cdot 10^5$ individual localisation coordinates (Fig. 4.5). Increased STORM resolution over diffraction-limited TIRF microscopy is apparent from the significantly decreased width of the filaments. This is particularly clear in regions with closely spaced and intersecting filaments. A cross section of STORM localisations demonstrates clear resolution of two filaments separated by ~ 200 nm that remain unresolved in the same cross section of the TIRF image (Fig. 4.6). The low number of background localisations prove that labelling and detection were specific to the target structure.

Dual-colour STORM imaging

Dual-colour STORM is an important extension to image two different proteins or structures in the same cell (Bates et al., 2007). High precision colocalisation analysis promises a direct view on molecular interactions in the cell. Our STORM microscope features multiple excitation lasers and three emission channels for multi-colour FRET and super-resolution imaging. I tested dual-colour STORM by imaging two different transcription factors in the nucleus of mouse embryonic fibroblast cells. Sox2 and Oct4 serve as master regulators of the pluripotent state of stem cells (reviewed by Rizzino (2009)). Expression of Sox2 and Oct4 is used as a marker for pluripotency. The ability to reprogram differentiated cells into an induced pluripotent state by expressing a combination of transcription factors including Sox2 and Oct4 has much potential in the field of regenerative medicine. Close proximity of their binding sites and chromatin immunoprecipitation experiments showed synergistic activation of several hundred genes by Oct4 and Sox2 (Boyer et al., 2005), potentially acting as heterodimers (Lam et al., 2012). STORM imaging of Oct4 and Sox2 was initiated by Chen Sok Lam (Neural

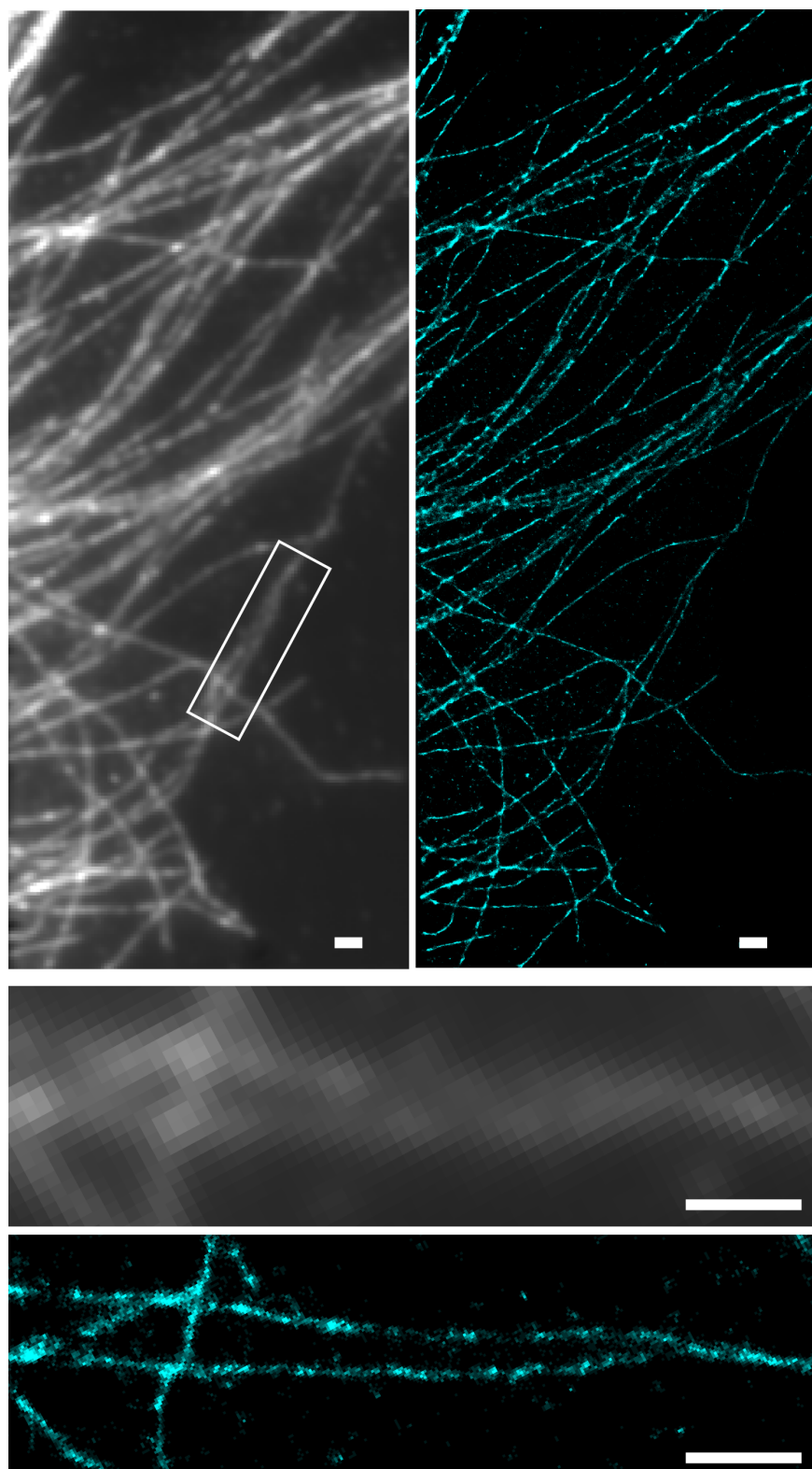


Figure 4.5: Comparison of diffraction-limited TIRF microscopy and STORM. STORM images show density maps of Alexa647 localisations. Scale bars: 1 μm . The boxed region is magnified below.

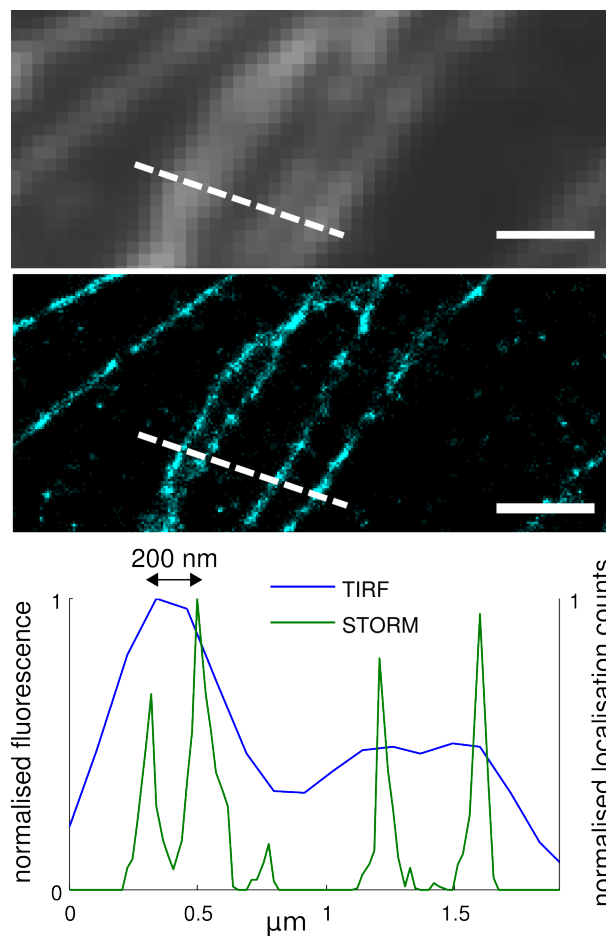


Figure 4.6: The cross sections (dashed lines) of microtubule filaments show the resolution improvement of STORM over diffraction-limited TIRF microscopy. Scale bars: $1\ \mu\text{m}$.

Stem Cell Laboratory, Singapore) who sent samples of fixed mouse embryonic fibroblast cells on chambered cover-glass (Nunc). I performed immunolabelling as described above. Heilemann et al. (2009b) and Dempsey et al. (2011) showed that a large range of synthetic fluorophores can potentially be used for dual-colour STORM imaging. According to these studies and my own experience, Alexa647 has the best properties for STORM imaging, namely a long-lived non-fluorescent state and high brightness in the fluorescent state. For the second colour, I tested the green fluorophores Alexa568, Alexa546, and Alexa532. The latter showed the best photoswitching characteristics and little fluorescence leakage into the red emission channel. I labelled Sox2 using mouse-anti Sox2 (Santa Cruz) with anti-mouse Alexa532 (Invitrogen) antibodies. For Oct4, I used goat anti-Oct4 (Santa Cruz) with anti-goat Alexa647 (Invitrogen) antibodies.

To excite fluorophores inside the nucleus several micrometre above the surface of the coverslide, I adjusted the excitation angle into highly inclined thin illumination mode (Tokunaga et al., 2008). Diffraction-limited fluorescence of Sox2-Alexa532 was approximately evenly distributed within the nucleus (Fig. 4.7A). Using the photoswitching buffer described above, single Oct4-Alexa647 molecules could be detected although the signal-to-noise ratio was lower than for TIRF excitation and there was significant background fluorescence from fluorophores above and below the focal plane (Fig. 4.7B). I acquired dual-colour STORM data sequentially because of Alexa532 fluorescence leakage into the red emission channel and the complication that green excitation activates and photobleaches Alexa647. For this reason, I first recorded Alexa647 fluorescence in the red emission channel under 640 nm excitation with 543 nm activation (as for microtubule), and subsequently imaged Alexa532 in the green emission channel under 543 nm excitation.

I analysed STORM movies using Quick-PALM real-time localisation software by

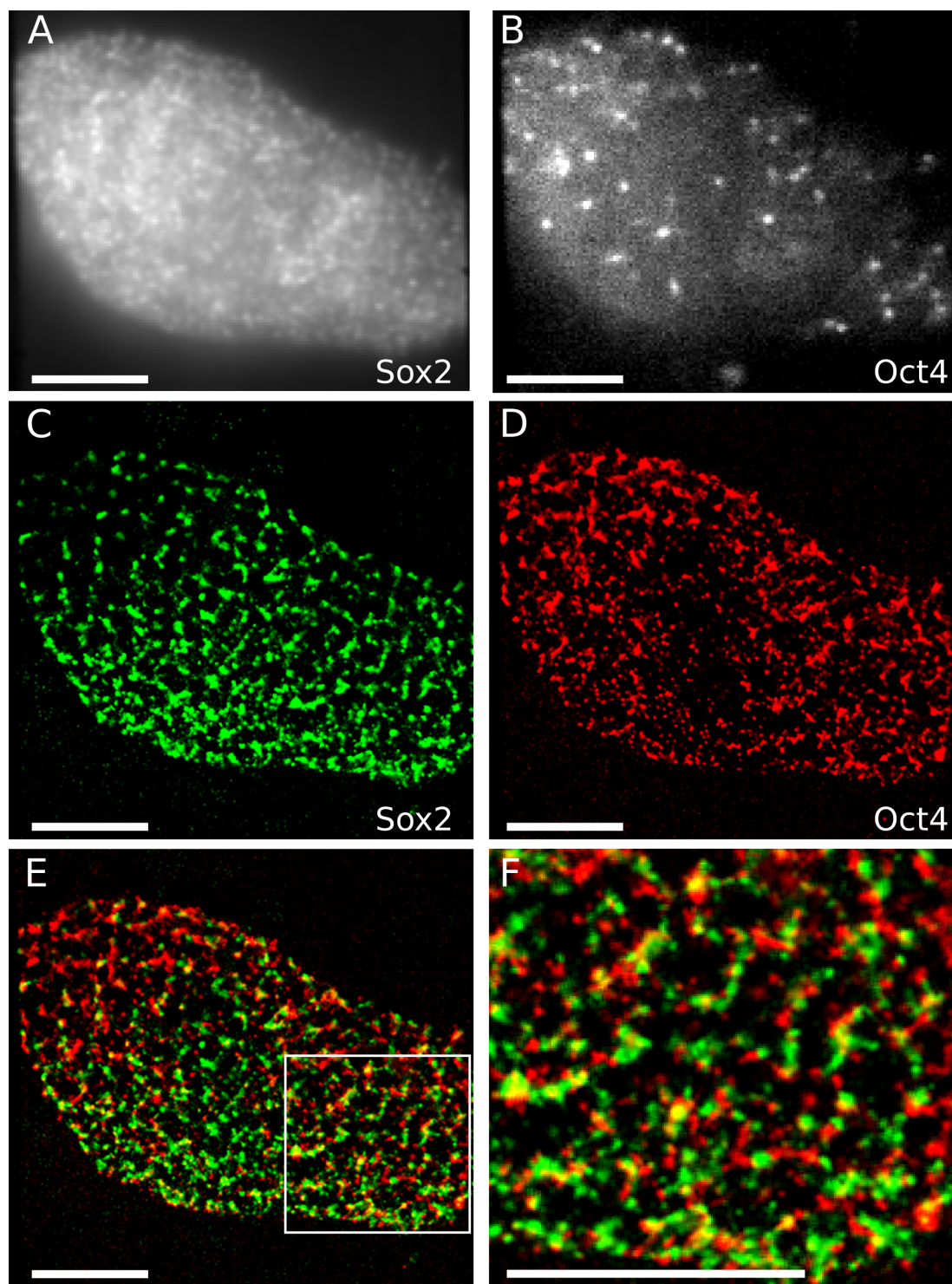


Figure 4.7: Dual-colour STORM imaging of transcription factors Sox2 and Oct4 in mouse embryonic stem cells. Scale bars: 5 μm . **(A)** Diffraction-limited fluorescence of Sox2-Alexa532. **(B)** Example frame of a STORM movie of Oct4-Alexa647. **(C)** Reconstructed STORM image of Sox2-Alexa532. **(D)** Reconstructed STORM image of Oct4-Alexa647. **(E)** Overlaid dual-colour STORM image of Sox2-Alexa532 (green) and Oct4-Alexa647 (red). **(F)** Magnified view of the boxed region in (E).

Henriques et al. (2010). Reconstructed STORM images show clustered localisations of Sox2 and Oct4 within the nucleus (Fig. 4.7C,D), reminiscent of STORM images of chromatin fibres (Benke and Manley, 2012). Interpreting individual images of either transcription factor is difficult without being able to relate their spatial distributions to other structures or proteins in the cell. The dual-colour STORM image, however, reveals partial colocalisation of Sox2 and Oct4, evident from green and red localisations in close proximity (Fig. 4.7E,F).

This brief project served as an initial demonstration of the dual-colour STORM capabilities of our microscope without further optimisation or method development. I used Quick-PALM software for fast and convenient data analysis at the sacrifice of localisation precision. Identification of optimal fluorophore combinations and photoswitching buffer conditions will help to further improve STORM image quality. While the qualitative dual-colour STORM images give evidence for colocalisation of Sox2 and Oct4 in the cell, stronger conclusions require quantitative colocalisation analysis with precise alignment of the two emission channels. The image registration algorithm of the single-molecule FRET software (Holden et al., 2010) can be adapted for this purpose.

High-density super-resolution localisation (DAOSTORM)

The overall resolution of localisation microscopy is defined by the localisation precision and the labelling density, according to the Nyquist sampling theorem (Heilemann et al., 2009b). However, isolating single fluorophores for localisation becomes very challenging at high labelling densities, due to lack of control over photoswitching rates and limited acquisition times, especially in living cells. This calls for optimised localisation algorithms that can deal with high densities of active fluorophores per frame. Seamus Holden (Achillefs Kapanidis group) realised that PSFs of fluorophores in microscopy images and of stars in astronomy

images are closely related. He developed the “DAOSTORM” software for high-density super-resolution localisation by adapting a method previously developed for analysis of dense stellar fields (Holden et al., 2011). The algorithm detects localisations of multiple overlapping PSFs using an iterative fitting procedure, thereby achieving reliable localisation at densities up to $10 \mu\text{m}^{-2}$. My work on STORM, sample preparation, and data acquisition contributed to this project. Here, I used DAOSTORM to generate the super-resolution images displayed in Figs. 4.5 and 4.6. DAOSTORM helps to improve spatial resolution and data acquisition times as demonstrated in a recent study by Shim et al. (2012), and served as a benchmark for further development of super-resolution image analysis algorithms (Zhu et al., 2012; Babcock et al., 2012; Wang et al., 2012).

4.4 Photoactivated Localization Microscopy (PALM) and single-particle tracking

Both PALM and STORM rely on photoswitching or activation and localisation. STORM achieves unrivaled image resolution but the use of synthetic fluorophores restricts the technique to immunostained fixed cells, although specialised live-cell labelling strategies are under development (van de Linde et al., 2012). On the other hand, PALM employs expression of genetically encoded photoactivatable fluorescent proteins in live cells (reviewed in Lippincott-Schwartz and Patterson (2009)). This approach effectively allows complete replacement of any protein with the fluorescent version with minimal perturbation to the cell. Beyond super-resolution microscopy of cellular structures, PALM can be extended to follow the movement of individual molecules in live cells (Manley et al., 2008). Single-particle tracking techniques that connect localisations in a series of images to form particle trajectories where previously restricted to low particle densities (Biteen et al., 2008). PALM allows extension of this approach to track almost arbitrary numbers of molecules. Measuring relatively slow diffusion of molecules within the cell membrane has been the main target of single-particle tracking techniques, benefiting from the two-dimensional structure of the membrane and its accessibility for labelling (Saxton and Jacobson, 1997). Increased sensitivity and temporal resolution of fluorescence microscopes recently allowed tracking individual diffusing fluorescent proteins in the bacterial cytoplasm (English et al., 2011), hence opening the technique for application to any bacterial fusion protein. The combination of photoswitching and single-particle tracking promises fundamental insights into the molecular organisation of cellular processes.

Although I focused on STORM applications in my original microscope design (section 4.2), extension to PALM simply required addition of a 405 nm laser for

photoactivation. The combination of 405 nm and 561 nm is ideal for common photoactivatable fluorescent proteins. I chose PAmCherry (Subach et al., 2009b) for several reasons that are further discussed below: (i) monomeric fluorescent protein, such that each detection reports on a single fusion protein; (ii) no aggregation in cells³; (iii) irreversible bleaching (no reversible switching was observed by Subach et al. (2009b)), such that the number of activated PAmCherry directly reports on the protein abundance (section 4.5); (iv) acceptable brightness and photostability.

***E.coli* fusion strains**

Rodrigo Reyes-Lamothe (Oxford Biochemistry) generated C-terminal PAmCherry fusions of DNA binding-proteins, expressed from their native promoter in *E.coli* AB1157, by replacing the genes using lambda-Red recombination (Reyes-Lamothe et al., 2008a). To this end, he amplified PAmCherry1 (including an 11 amino acids linker) and Kanamycin resistance genes from the plasmid pROD85 by PCR. The primers had 40 - 50 nt homologous sequence to the chromosome insertion site. Remaining plasmid DNA was digested with Dpn1 restriction endonuclease before electroporating 1 μ g of DNA into AB1157 cells that carry the plasmid pKD46 for overexpressing lambda-Red insertion proteins as described by Datsenko and Wanner (2000). He validated successful recombination using colony PCR with primers flanking the insertion site and confirmed the gel shift corresponding to the additional DNA fragment length. The cell growth rates were similar for AB1157 wild type and fusion strains in LB medium and minimal medium (M9) (Table 4.1), indicating full functionality of the fusion proteins. Fusion strains of DNA repair proteins DNA polymerase I (Pol) and ligase (Lig)

³mEos2 and other common fluorescent proteins have a tendency to aggregate in *E.coli* (Endesfelder et al., 2012; Landgraf et al., 2012).

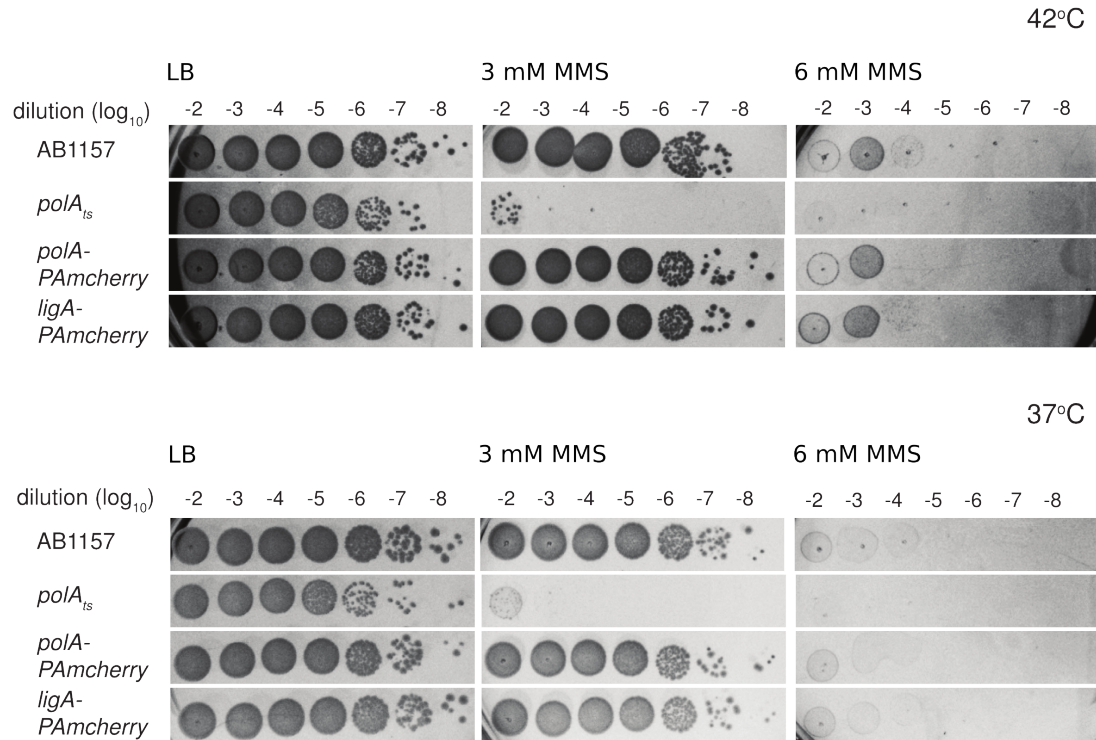


Figure 4.8: MMS sensitivity assay performed together with Rodrigo Reyes-Lamothe. Serial dilutions of cell strains were grown overnight on LB agar plates at 42°C (top), and 37°C (bottom) without MMS (left), 3 mM MMS (middle), 6 mM MMS (right). Wild type AB1157 shows the same MMS sensitivity as fusion strains *polA-PAmCherry* and *ligA-PAmCherry*. The temperature sensitive DNA polymerase I mutant *polA_{ts}* is hyper-sensitive to MMS.

showed the same sensitivity and growth rate in the presence of DNA damaging agent methyl methanesulfonate (MMS) as wild-type cells, judged by the number and size of colonies in a plating assay (Fig. 4.8). A temperature sensitive Pol mutant strain exhibited hyper-sensitivity to MMS. Integrity of the fusion proteins is evident from the following microscopy studies.

	wild type	<i>polA-PAmCherry</i>	<i>ligA-PAmCherry</i>	<i>fis-PAmCherry</i>
LB (min)	28.7	29.7	28.8	28.4
M9 (min)	86.6	87.7	88.8	84.5

Table 4.1: *E.coli* AB1157 growth rates at 37°C for wild-type, and PAmCherry fusions DNA polymerase I (Pol, *polA-PAmCherry*), DNA ligase (Lig, *ligA-PAmCherry*), and Factor for inversion stimulation (Fis, *fis-PAmCherry*).

Cell preparation

To prepare cells for microscopy, I streaked strains onto LB plates with 25 $\mu\text{g/ml}$ kanamycin. Single colonies were inoculated into LB and grown at 37°C for 4-5 h, then diluted 1/10.000 into M9 medium (M9 salts, MEM amino acids + proline, MEM vitamins, 0.2% glycerol) and grown over night at 37°C to OD 0.4 - 0.6. The following morning, cultures were diluted to OD 0.025 in M9 medium and grown for 2 h at 37°C to early exponential phase (OD $\sim$ 0.1). Cells were centrifuged and sandwiched on agarose pads between two coverslides (No. 1.5 thickness, burned to remove fluorescent background particles). 1% agarose pads were prepared by mixing low fluorescence 2% agarose in dH₂O 1:1 with 2x M9 medium. For fixation, centrifuged cells prepared as above were resuspended into 2.5 % paraformaldehyde in PBS and fixed for 45 min shaking at room temperature. Fixed cells were washed with PBS and immobilised on agarose pads as above.

Data acquisition

I recorded PALM movies under continuous 3.5 mW 561 nm excitation ($\sim$ 0.6 kW $\cdot$ μm^{-2} for $25 \times 25 \mu\text{m}^{-2}$ illumination area) at 15.26 ms/frame for 7.500 to 10.000 frames. The total fluorescence of Pol-PAmCherry gave a broad diffraction-limited distribution inside live *E.coli* cells (Fig. 4.9A). By adjusting 405 nm excitation from 0 to 10 μW over the duration of the movie, I controlled photoactivation to image less than one PAmCherry fluorophore per cell at any time, giving clear isolated PSFs (Fig. 4.9B). Conveniently, autofluorescent background of the cells could be bleached before starting data acquisition and photoactivating PAmCherry from the dark state. I applied the least-squares super-resolution localisation algorithm described above (chapter 4.3). To follow the movement of single

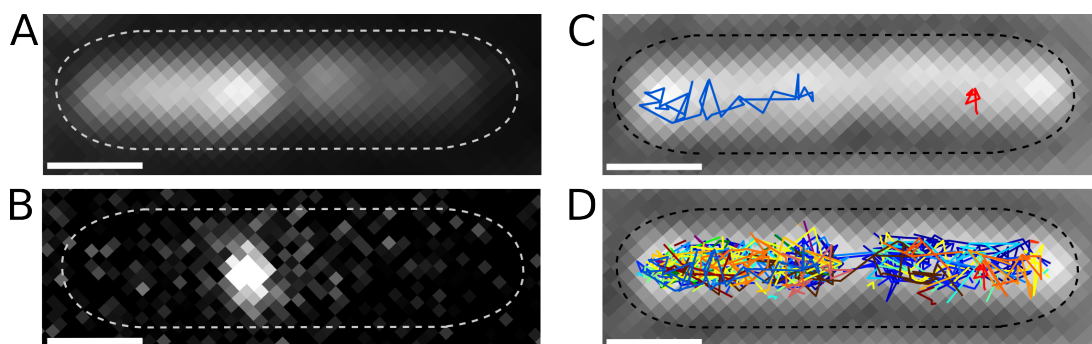


Figure 4.9: (A) Accumulated Pol-PAmCherry fluorescence in a live E.coli cell. Scale bars: $1\ \mu\text{m}$. (B) Example PSF at 15 ms exposure time. (C) Example tracks of single Pol molecules on transmitted light image of the cell. (D) All measured Pol tracks in the cell.

molecules over multiple frames, I adapted the MATLAB implementation of the single-particle tracking algorithm described by Crocker and Grier (1996). Positions were linked to a track if they appeared in consecutive frames within a window of 5 pixels ($0.57\ \mu\text{m}$); this window size ensures that 98% of steps are correctly linked for an apparent diffusion coefficient of $1\ \mu\text{m}^2\text{s}^{-1}$. If multiple localisations fell within the tracking radius, tracks were linked such that the sum of distances was minimised. I used a memory parameter of 1 frame to allow for transient disappearance of the PSF within a track due to missed localisation. I plotted the resulting tracks on top of light microscopy images showing cell outlines to visualise the spatial distribution of protein movement within cells (Fig. 4.9C-D).

I confirmed cell viability after PALM imaging (Fig. 4.10). Using agarose pads with a rich non-fluorescent medium (Teknova), I followed normal cell growth and division on the microscope slide for 45 min. I subsequently imaged Pol-PAmCherry with standard acquisition and excitation conditions. After PALM imaging, cells continued normal growth, but I noticed a 15 - 25 min time lag before growth continued. 67 out of 70 imaged cells (95.7%) recovered growth during the following 150 min observation, showing no sign of substantial light-

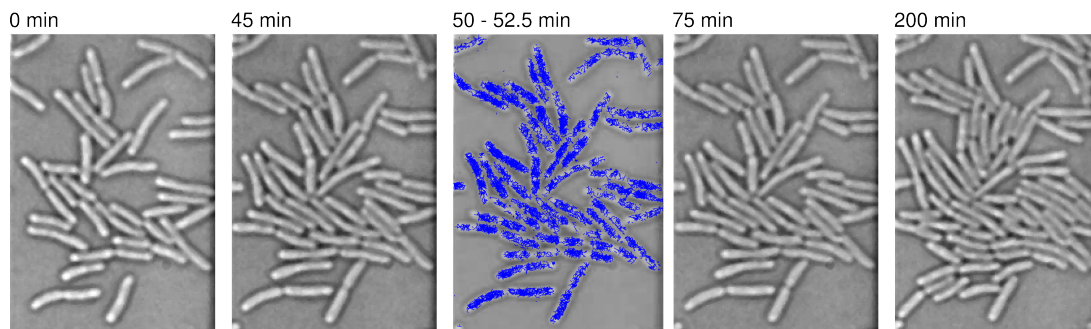


Figure 4.10: *E.coli* cells are viable after super-resolution imaging. 0 - 45 min: Cells grow on the microscope slide. 50 - 53 min: Super-resolution imaging of Pol-PAmCherry under standard illumination conditions. 53 - 75 min: Cells halt growth after imaging for ~ 20 min. 75 - 200 min: Cells continue normal growth.

induced DNA damage.

Diffusion analysis

In the introduction, I discussed the linear relation between the mean-squared displacement (MSD) of a molecule and the time interval $n\Delta t$ between observations:

$$\text{MSD}(n\Delta t) = 4 \cdot D \cdot n\Delta t . \quad (4.3)$$

Single-particle tracking data give a direct estimate of the diffusion coefficient D by calculating the MSD between localisations. For a two-dimensional trajectory of N steps, the MSD reads:

$$\text{MSD}(n\Delta t) = \frac{1}{N-n} \sum_{i=1}^{N-n} (x_{i+n} - x_i)^2 + (y_{i+n} - y_i)^2 . \quad (4.4)$$

I calculated MSD curves from Pol, Lig, and Fis tracks in live cells (Fig. 4.11A). These curves show an approximately linear rise but saturate with increasing lag time. This reflects the confined diffusion of molecules within the cell. The flat MSD curve for immobile Pol in fixed cells confirmed that the observed diffusion in live cells was not an artifact of the experiment or analysis. The cumulative

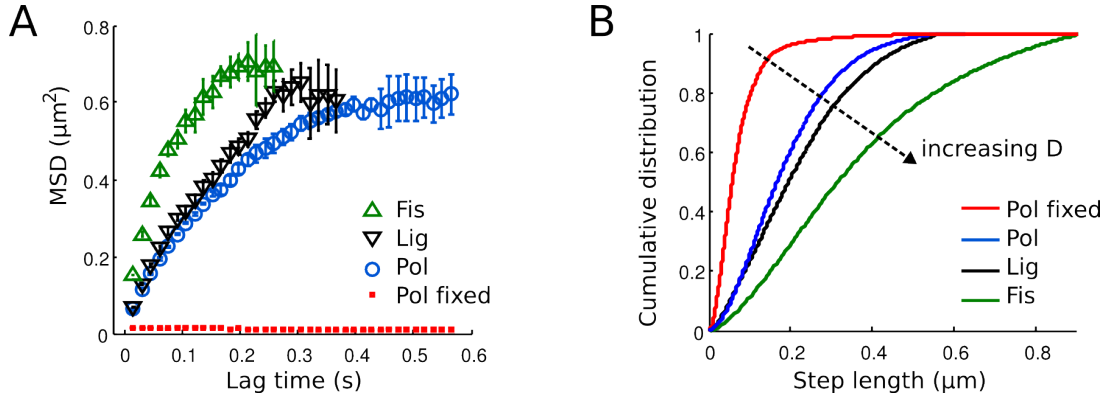


Figure 4.11: (A) Mean-squared displacement, averaged over all observed tracks ($\pm$ std) versus lag time for Fis (green triangles), Lig (black triangles), Pol (blue circles) in live cells, and Pol in fixed cells (red squares). (B) Cumulative distribution curves of the diffusion step length between localisations in subsequent frames.

distribution of the measured step lengths shows slower diffusion with increasing molecular weight (Fig. 4.11B): 76 kDa Fis dimers diffuse fastest, followed by 102 kDa Lig, and 136 kDa Pol (molecular weights include the 27 kDa of PAmCherry). In fixed cells, Pol steps are shortest (red curve), representing the localisation precision of ~ 40 nm. The scaling of the diffusion steps with the expected molecular weight confirms the correct identity and integrity of the fusion proteins.

To classify movement of individual molecules, I computed an apparent diffusion coefficient D^* from the MSD for each track, correcting for the localisation error of σ_{loc} (Michalet and Berglund, 2012):

$$D^* = \frac{\text{MSD}}{4 \cdot \Delta t} - \frac{\sigma_{\text{loc}}^2}{\Delta t} \quad (4.5)$$

Only tracks with a minimum of four steps were included in this analysis to lower the statistical error. Note that the apparent diffusion coefficient is biased by cell confinement and motion blurring and therefore does not equal the accurate microscopic diffusion coefficient D (Michalet and Berglund, 2012). However,

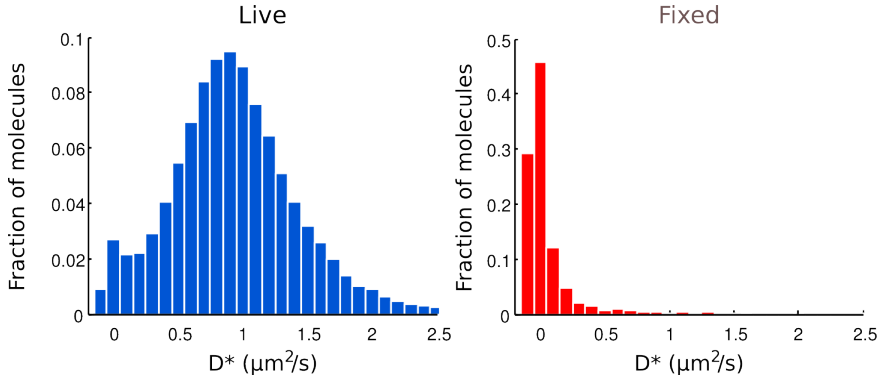


Figure 4.12: D^* distributions for Pol in live and fixed cells.

distributions of D^* distinguish well between bound and diffusing molecules (Fig. 4.12). The distribution for immobile Pol in fixed cells peaks at $D^* = 0 \mu\text{m}^2\text{s}^{-1}$ for $\sigma_{\text{loc}} = 40 \text{ nm}$, while live cells show a broad distribution centred around $D^* \approx 0.9 \mu\text{m}^2\text{s}^{-1}$. In the following chapters 5 and 6, I describe application of MSD analysis to identify protein-DNA interactions by the change in the diffusion coefficient of single proteins upon binding DNA.

Landgraf et al. (2012) reported a tendency of fluorescent proteins to cause artificial aggregation of fusion proteins in *E.coli*. Several characteristics of the tracking data shown in this chapter and the following chapter 5 exclude significant aggregation of Pol, Lig, and Fis fusion proteins in live cells: (i) The majority ($> 95\%$) of Pol, Lig, and Fis showed rapid diffusion in live cells (Figs. 4.12, 5.4, 5.5); (ii) The MSD and CDF slopes scale with the expected individual molecular weights of the fusion proteins (Fig. 4.11); (iii) Protein localisations were distributed over a large part of the cell without any sign of clustering (Figs. 5.2, 5.6); (iv) Simulations showed that the observed diffusion is consistent with ordinary Brownian motion and the extracted diffusion coefficients for Pol and Lig match those of other proteins with similar molecular weight (Fig. 5.3). Furthermore, Endesfelder et al. (2012) found a random spatial distribution of PAmCherry in *E.coli* but clusters of mEos2 under the same conditions.

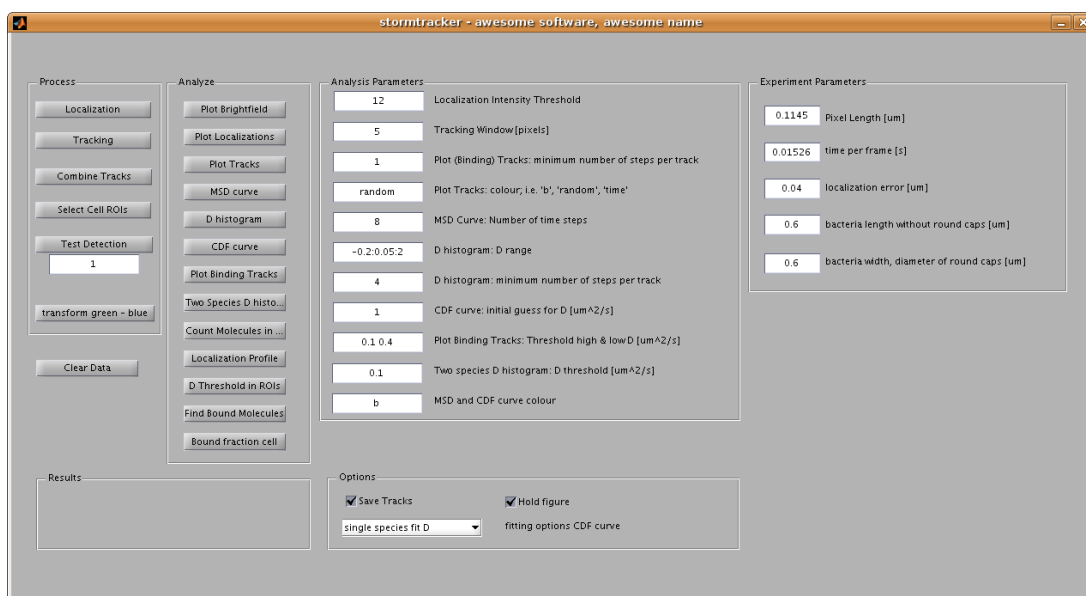


Figure 4.13: Screenshot of the graphical user interface of the super-resolution microscopy and single-particle tracking data analysis software.

Software

I compiled my data analysis code into an extensive user friendly software package for automated super-resolution localisation, tracking, and further data analysis in MATLAB (Fig. 4.13). This program also incorporates previous software developments by Holden et al. (2010, 2011), and freely available code by Crocker and Grier (1996).

4.5 Counting protein copy numbers

Apart from protein localisation, cell function is largely dictated by protein abundance. Cell behaviour follows dynamic regulation of protein production and degradation. Gene expression noise, however, sets fundamental limits on the cellular control over protein copy numbers (Lestas et al., 2010), and can lead to large phenotypical differences across a population of clonal cells (Choi et al., 2008). Single-cell fluorescence studies have shown dynamic changes in gene ex-

pression (Locke et al., 2011), and quantified protein copy numbers (Huang et al., 2007; Reyes-Lamothe et al., 2010) by integrating the signal of fluorescent proteins over the cell area. These methods, however, rely on the precise calibration factor between the fluorescence intensity and copy number which depends on the microscope alignment and excitation conditions.

I considered direct counting of fluorescent proteins by single-molecule localisation as a promising alternative. PALM imaging with PAmCherry allows simultaneous measurement of the protein localisation and copy number per cell, because each PAmCherry fusion is activated only once (Subach et al., 2009b), giving one track per protein. In practice, I ensured activation of essentially all PAmCherry molecules by acquiring long movies (up to $2 \cdot 10^4$ frames) with a gradual increase of 405 nm intensity. I segmented cell outlines in transmitted light microscopy images to determine the number of tracks within each cell. Cells were split if a septum was clearly visible. I estimated a detection efficiency of $\sim 80\%$, with 23 min PAmCherry maturation time (Subach et al., 2009b) and ~ 120 min steady state generation time in M9 glycerol medium at room temperature. The tracking algorithm with a memory parameter of 1 frame accounted for potential transient blinking. While TIRF excitation gives the brightest signal of fluorophores closely attached to the coverslide, imaging the majority of molecules within $\sim 0.8 \mu\text{m}$ thick *E.coli* cells required highly inclined thin illumination (Tokunaga et al., 2008).

I counted 479 Pol, 226 Lig, and 1560 Fis copies per cell (median values). These numbers are in agreement with previous reports of 400 Pol (Kornberg, 2005), 200 Lig (Friedberg, 2006), and thousands of Fis in early exponential phase (Ball et al., 1992). Ishihama et al. (2008) reported 321 Pol and 98 Lig per cell. In fixed cells, Pol counts were 20% lower than in live cells, suggesting partial destruction of PAmCherry by paraformaldehyde treatment. The protein copy number dis-

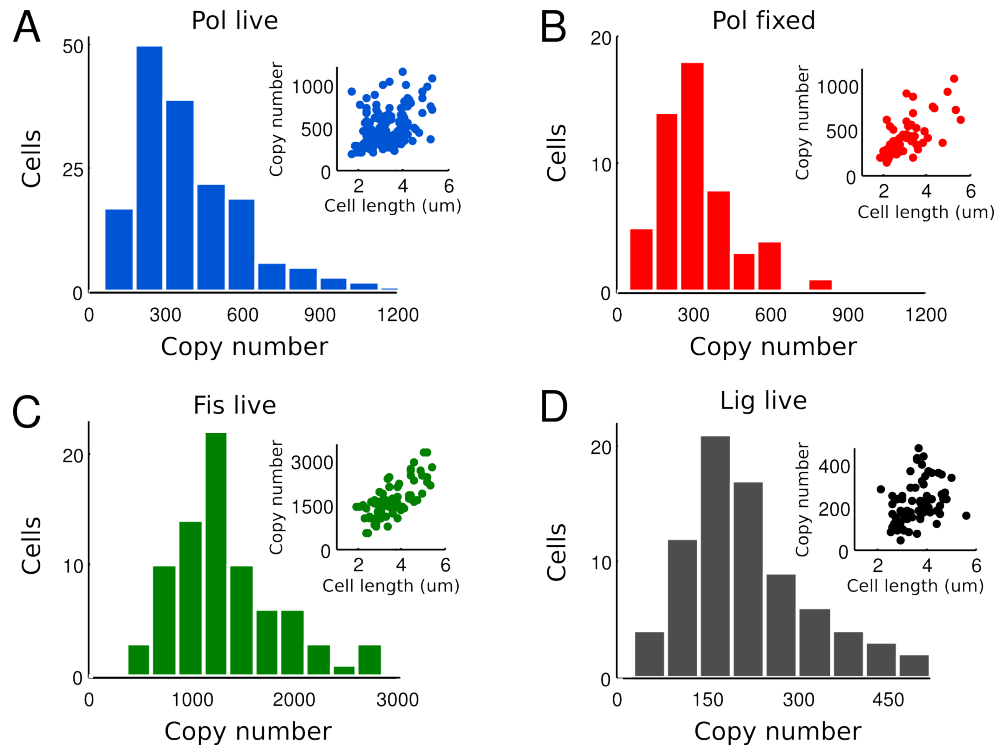


Figure 4.14: Protein copy numbers per cell. **(A)** Pol-PAMCherry in live *E.coli* (mean 553, median 479). **(B)** Pol-PAMCherry in fixed *E.coli* (mean 440, median 421). **(C)** Fis-PAMCherry in live *E.coli* (mean 1649, median 1560). **(D)** Lig-PAMCherry in live *E.coli* (mean 308, median 226).

tributions show expression variation across cells and the expected correlation of protein abundance with the *E.coli* cell length (Fig. 4.14). Lee et al. (2012) also recently evaluated molecule counting by PALM and found that persistent blinking of mEos2 lead to significant overcounting errors, while the relatively stable emission of Dendra2 resulted in more reliable results. Similar to my approach, they also proposed uniform activation rates by increasing the 405 nm intensity during the experiment. In chapter 5, I used the total Pol and Lig copy numbers to calculate absolute DNA repair rates per cell. Other applications of molecule counting in this thesis rely on relative abundances that are independent of the overall detection efficiency.

4.6 Conclusion

This chapter showed my implementation of super-resolution fluorescence microscopy techniques that rely on photoswitching and single-molecule localisation. I discussed the design and construction of a TIRF microscope and demonstrated single- and dual-colour STORM imaging of microtubule and transcription factors in mammalian cells. For live-cell applications, I focused on PALM imaging, tracking, and counting of fluorescent fusion proteins. This general method allowed single-molecule imaging at almost arbitrary concentrations in live bacterial cells. The following chapters 5 and 6 utilise this technique to study protein-DNA interactions in *E.coli*.

In the future, our flexible microscope design can be extended to allow three-dimensional super-resolution imaging. The simplest approach involves insertion of a cylindrical lens into the emission path such that the elliptical shape of the PSF reports on the precise z-position of the emitter (Huang et al., 2008). Our z-piezo stage and elliptical Gaussian fitting function are well suited for the required calibration scans and data analysis. Alternatively, a spatial light modulator can be used to engineer a double-helical PSF such that the rotation of the helix in the imaging plane reports on the z-position of the emitter. This more complicated approach provides almost uniform z-localisation precision over a range of 2 μm (Pavani et al., 2009). Three-dimensional localisation methods will be particularly useful for colocalisation analysis because two-dimensional projections can give the impression of colocalising molecules that are in fact separated in the z-dimension.

Chapter 5

Single-molecule DNA repair in live bacteria

5.1 Introduction

Cellular DNA damage is reversed by balanced DNA repair pathways that minimise accumulation of toxic intermediates. Despite their importance, the *in vivo* DNA repair mechanisms have remained unclear due to the inability to time individual reaction steps and search processes in single cells and link them to the abundance of repair proteins. Fluorescence microscopy is well suited to study DNA repair *in vivo* (Houtsmuller et al., 1999; Feuerhahn and Egly, 2008; Dion et al., 2012; Miné-Hattab and Rothstein, 2012), but is limited by its current spatial and temporal resolution. For instance, although individual fluorescent proteins can be detected in live bacteria, they typically report on the position and movement of a few fluorescent molecules per cell. Attempts by Reyes-Lamothe et al. (2010) to visualise lagging-strand replication by fluorescent DNA polymerase I (Pol) molecules (~ 400 copies per cell) only showed a broad diffraction-limited distribution inside *E.coli* (as in Fig. 4.9A). To detect single reactions

directly, I imaged proteins sequentially using the single-particle tracking PALM method described in the previous chapter. I measured individual Pol and Lig proteins searching and repairing DNA gaps in live cells (Uphoff et al., 2012). I found that single Pol and Lig reactions last 2.1 and 2.5 s, respectively, and their cellular repair rates appear to be limited by upstream pathway steps. Excess Pol and Lig saturated substrates, minimising the number and lifetime of gapped and nicked intermediates. My general approach provides a quantitative, systems-level overview of a model DNA repair pathway.

Bacterial base excision repair

DNA damage by exogenous or endogenous factors is a major cause of mutations that lead to cancer and other pathological conditions in humans (Wilson and Bohr, 2007). Normal metabolism causes more than 10,000 DNA lesions in a human cell per day (Lindahl, 1993), and up to 5,000 lesions in a single *E.coli* generation (Cox, 1998). All cellular organisms rely on DNA repair mechanisms to reverse DNA damage for faithful chromosome replication and maintenance of the genome integrity (Friedberg, 2006). The variety of DNA lesions requires modular repair pathways that carry out damage recognition, damage removal, repair synthesis, and ligation in sequential steps by a series of enzymes. The conserved base-excision repair (BER) pathway deals with the most common DNA lesions, including oxidised and alkylated bases, anomalous bases (e.g. uracil), and base losses (Friedberg, 2006). DNA alkylation is caused by naturally occurring endogenous metabolites; humans might also be exposed to alkylating agents in tobacco smoke, food, and chemotherapy (reviewed in Sedgwick (2004)). Methyl methanesulfonate (MMS) is a DNA methylation agent commonly used in research that preferentially creates 7-methylguanine and 3-methyladenine lesions in dsDNA. Efficient repair of 3-methyladenine lesions that block DNA replica-

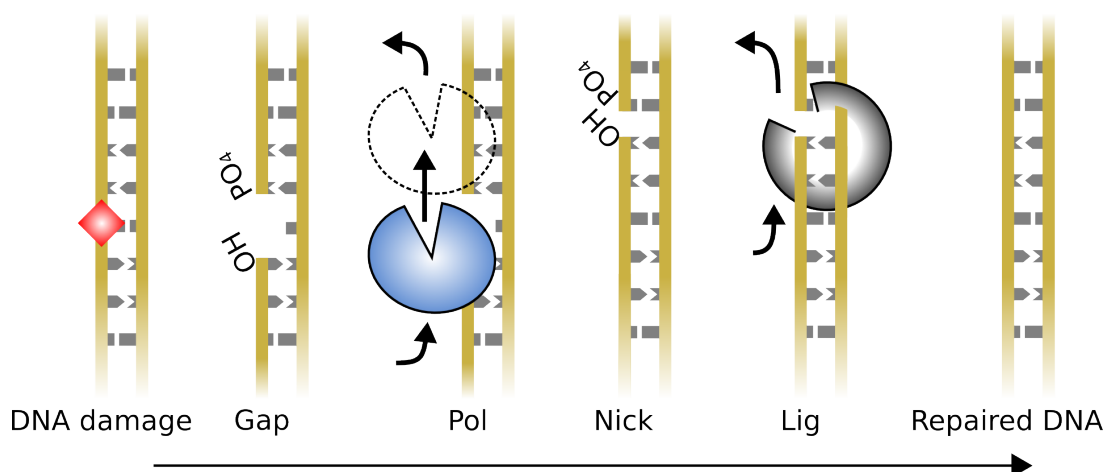


Figure 5.1: Schematic of a general DNA repair pathway involving a DNA Polymerase (Pol) for repair synthesis and a DNA ligase (Lig) to seal the DNA backbone.

tion is paramount to cell survival. In *E.coli*, BER involves at least five different enzymes. Damage specific DNA glycosylases (e.g. *tag*, *alkA*, *fpg*, *ung*) recognise and remove damaged bases, leaving apurinic or apyrimidinic (AP) sites. Subsequently, AP lyases (e.g. *xth*, *nfo*, *nth*) and deoxyribosephosphodiesterases (e.g. *recJ*) create a short DNA gap that is filled by Pol (*polA*), and the resulting nick is sealed by Lig (*ligA*). Strand displacement and 5' nuclease activities of Pol appear to be responsible for the alternative long patch BER modality (Dianov and Lindahl, 1994). DNA repair pathways are finely balanced to avoid accumulation of toxic species (Fu et al., 2012), particularly gapped and nicked intermediates (Senejani et al., 2012) that can cause lethal double-stranded breaks upon encountering replication forks (Kuzminov, 2001); a single such break can lead to chromosome loss and cell death.

Here, I focus on the last two steps of DNA synthesis and ligation in BER, which are conserved from bacteria to humans and across repair pathways (Fig. 5.1) and also link adjacent Okazaki fragments in lagging-strand DNA replication (Kornberg, 2005). The DNA synthesis mechanism and fidelity checkpoints of Pol have been discussed above (chapter 3). *E.coli* DNA ligase (Lig), encoded

by the *ligA* gene, is a 77 kDa enzyme that joins DNA nicks with adjacent 5'-phosphoryl and 3'-hydroxyl groups (Lehman, 1974; Tomkinson et al., 2006). In a series of three reaction steps, Lig acquires an AMP group from an NAD⁺ cofactor, associates with a DNA nick by a pyrophosphate bond between AMP and the phosphorylated 5'-terminus, thus creating the target for AMP displacement by the 3'-hydroxyl group to form a phosphodiester bond. Lig is present at approximately 200 copies per cell. A temperature-sensitive *E.coli* Lig mutant accumulated unliganded Okazaki fragments when grown above the restrictive temperature (Lehman, 1974). *E.coli* mutants deficient in Pol and Lig activity are hypersensitive to MMS (Ljungquist et al., 1976; Konrad et al., 1973). However, Pol activity could be depressed to about 1% of wild type levels without impairment in growth (Lucia and Cairns, 1969). This suggests that only a small fraction of cellular Pol is required for basal replication and repair, while the excess copies are important for severe exogenous DNA damage conditions.

E.coli cells acquire resistance against the toxic and mutagenic effects of DNA methylation damage after treatment with a non-lethal dose (Samson and Cairns, 1977). The adaptive response is regulated by the direct repair protein Ada, which is initially present at very low numbers in undamaged cells. An irreversible methyltransfer reaction from DNA onto Ada turns it into a transcription activator of genes involved in DNA alkylation repair: *ada* itself, *alkA* (glycosylase), *alkB* (oxidative demethylase), and *aidB* (function unclear) (Evensen and Seeberg, 1982; Volkert, 1988; Kleibl, 2002). This induction creates a thousand-fold increase of Ada levels and a ten-fold increase for AlkA. Furthermore, the global SOS response upregulates nucleotide excision repair (NER) proteins (Janion, 2008), as well as mutagenic lesion bypass DNA polymerases IV and V that compete with high fidelity NER and BER (Goodman, 2002). The SOS response recognises DNA damage by the persistent presence of single-stranded DNA at a

stalled replication fork which activates RecA for degradation of the LexA gene repressor.

Experimental challenges in studying bacterial DNA repair

Despite extensive genetic, biochemical, and biophysical studies (Friedberg, 2006), the molecular mechanisms of DNA repair *in vivo* remain unclear. Because of back-up, competing, and branching DNA repair pathways, it is hard to predict which enzymes are actually performing a particular reaction in the cell and at what rate. The toxic and mutagenic effects of imbalanced or saturated repair were previously investigated by removing pathway components or altering their copy numbers (Konrad et al., 1973; Ljungquist et al., 1976; Glassner et al., 1998; Senejani et al., 2012). Complementing cell extract with additional enzyme copies has been used to identify the rate limiting step in repair pathways (Sung and Mosbaugh, 2003; Allinson et al., 2004). However, reaction pathways are not defined by copy numbers, but by reaction rates as a function of copy numbers, substrate concentrations, and enzymatic rates in the cell. Early *in vivo* work focused on the mean behavior of large cellular populations; it was thus unable to examine functionally important heterogeneity, such as the variation in the protein copy numbers between cells and over time (Elowitz et al., 2002; Huh and Paulsson, 2011), which may lead to variable repair rates across genetically identical cells. Combinations of low copy number effects and feedback amplification can produce complex gene expression patterns (Friedman et al., 2005; Locke et al., 2011), potentially causing large variation and loss of coordination between the different pathway components. Single-cell fluorescence studies can address some of these issues, but their limited resolution and sensitivity have left the detailed spatial distribution of the repair machinery under debate (Goodman, 2002; Jackson et al., 1994; Stuart et al., 2005). While many molecular processes

in eukaryotes show a high degree of spatial organization, such as “factories” of DNA replication, transcription, and repair (Cook, 1999), the existence and functional role of spatial organisation in bacteria is yet controversial. Fluorescence recovery after photobleaching (FRAP) experiments have shown recruitment of DNA repair proteins upon DNA damage (reviewed in Dinant et al. (2009)), but this model-dependent approach can not easily resolve molecular subpopulations to extract binding and diffusion constants.

Further, much of our mechanistic understanding relies on *in vitro* ensemble studies, which cannot replicate the complex cellular environment and the stochastic nature of repair processes, and are complicated by the presence of molecular subpopulations and dynamics. As a result, modeling repair pathways quantitatively has proven difficult, because the reaction rates are either unknown or come from a range of different *in vitro* experiments often performed under different conditions (Sokhansanj et al., 2002; Ni et al., 2007; Karschau et al., 2011). Consequently, there is need to measure the overall repair rates at the single-cell level and relate them to the concentrations and spatial distribution of substrates and repair proteins, as well as to the rates of individual reactions and search processes.

5.2 DNA polymerase I and ligase in undamaged cells

Spatial distribution

Using single-particle tracking PALM as described in chapter 4.4, I investigated the spatial distribution of Pol and Lig in undamaged *E.coli*. Figure 5.2 shows that their tracks reflect the organisation of the nucleoid in replicating *E.coli* (Reyes-Lamothe et al., 2008b). In newborn cells ($\sim 2.5 \mu\text{m}$ length) with a single chromosome, the tracks occupy the central area of the cell, avoiding the cell poles (Fig. 5.2A). Chromosome separation is evident from separated areas of high track density in cells of intermediate length ($\sim 3.5 \mu\text{m}$; Fig. 5.2B). In the longest cells ($\sim 4.5 \mu\text{m}$), the two areas of high track density separated entirely even before the appearance of a constricting septum that marks cell division (Fig. 5.2C). These results suggest that the confinement of Pol and Lig diffusion (evident in the MSD curves in Fig. 4.11; chapter 4.4) is due to association with the chromosome rather than the cytoplasm. The average size of a chromosome was $2\mu\text{m} \times 0.65\mu\text{m}$ (Fig. 5.2D).

Computer simulation of Brownian motion and PALM

I implemented stochastic diffusion simulations in MATLAB that accurately model the raw fluorescence data. This approach promises a better understanding of experimental biases that might effect diffusion measurements and can be used to evaluate the performance of the data analysis algorithms. Comparing experimental data with simulation results can also reveal to what extent a simple diffusion model describes the molecular motion in the cell. The model was based on ordinary three-dimensional Brownian motion of molecules with diffusion coef-

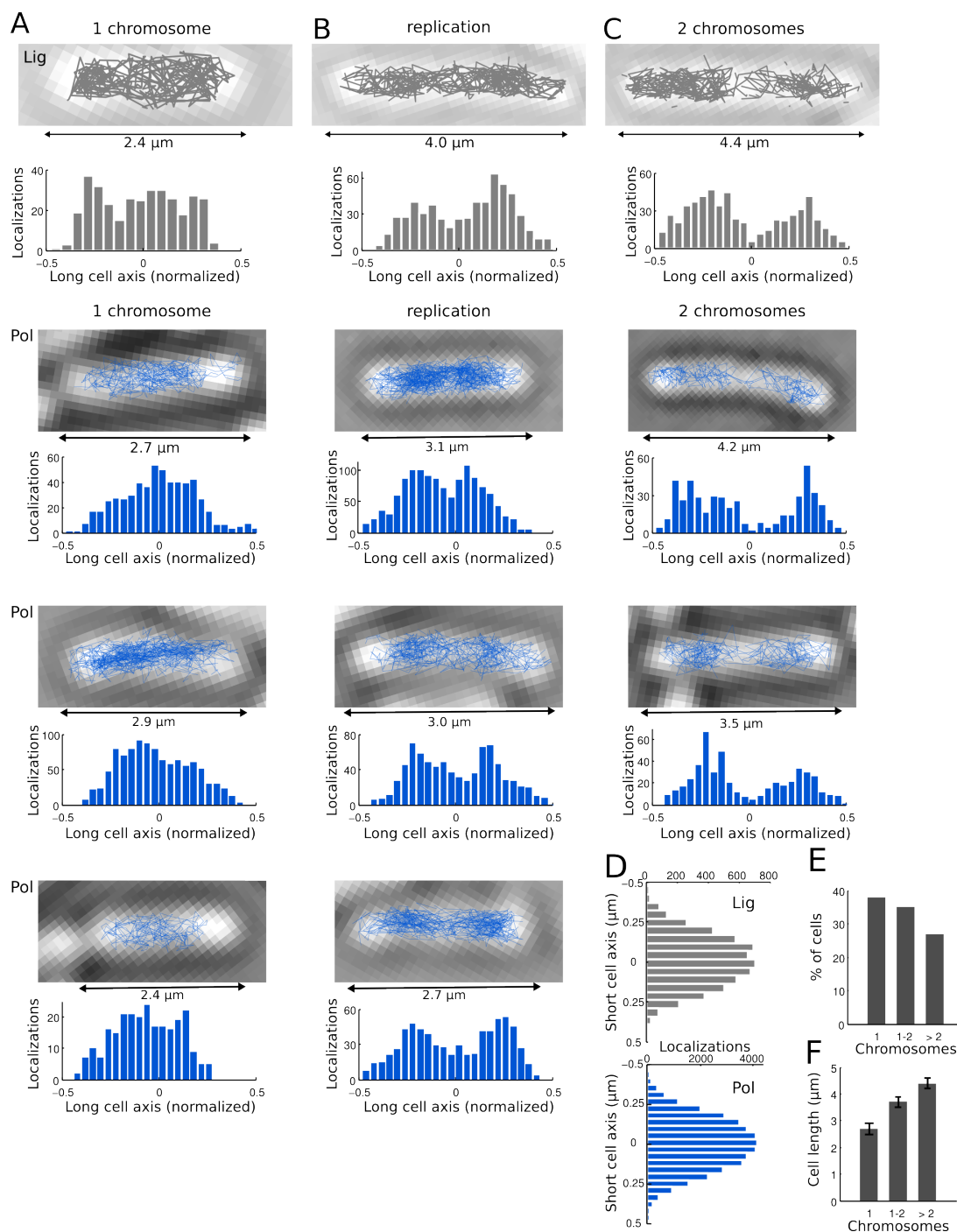


Figure 5.2: The spatial distributions of Lig and Pol tracks reflect the organisation of the nucleoid in replicating *E. coli*. Tracks are shown in blue/grey, histograms of localizations projected onto the normalized long cell axis are shown underneath together with the cell length. Like the nucleoid, Pol and Lig tracks occupied the central area of the cell, avoiding the cell poles. Chromosome separation is evident from discrete areas of high track density in longer cells. **(A)** Short cells with a single chromosome. **(B)** Medium length cells during replication in the process of chromosome separation. **(C)** Long cells with two separate chromosomes. **(D)** Lig localizations from 20 cells and Pol localizations from 47 cells projected onto the short cell axis, reflecting the $\sim 0.6 \mu\text{m}$ width of the nucleoid. **(E)** Based on Pol localization distributions, cells were manually sorted by the number of chromosomes and counted (134 cells). **(F)** Average cell length versus the number of chromosomes ($\pm\text{sem}$, 134 cells).

ficient D . Stochastic steps in x -, y , and z -directions were drawn from a Gaussian distribution with zero mean and variance $2D \cdot \Delta t$, with a simulation time step $\Delta t = 0.1526$ ms, according to equation 1.7 in chapter 1. Consistent with the experimental results described below in section 5.3, a subpopulation of immobile tracks was generated by randomly assigning individual tracks with $D = 0 \mu\text{m}^2\text{s}^{-1}$ for the whole duration of the track, such that the average percentage of immobile tracks matched experimental results of 2.7% bound Pol and 3.8% bound Lig in undamaged cells. The rod-shaped nucleoid confinement was approximated by a cylinder with round caps ($2 \mu\text{m}$ length, $0.65 \mu\text{m}$ diameter). Molecules were initially placed at uniform random positions within the confinement. Steps that would lead to a position outside the cell boundary were rejected. For stochastic photo-activation and bleaching, single tracks started at random times uniformly distributed over the total number of time steps and track durations before photobleaching were drawn from an exponential distribution with a mean of 54 ms. I simulated 400 tracks per cell in 152 s long movies (10^6 time steps) such that the density of activated particles was similar to experiments.

To model the raw experimental fluorescence data, PSFs were centred on each position of the simulated Brownian motion tracks that were projected into the x - y plane. Matching the experimental frame rate, every 100 simulation time steps were integrated to form a frame of a movie with 10.000 frames in total. This procedure generated PSF motion blurring as observed for rapidly diffusing molecules in experimental data. Simulation parameters for PSF intensity, PSF width, background noise per pixel, number of frames per movie, number of molecules per cell, and mean number of frames per track were chosen according to experimental values. Simulated movies were saved for super-resolution localisation, tracking, and further data analysis, performed exactly the same as for experimental data.

Comparing Pol and Lig diffusion in experiment and simulation

To evaluate whether a simple Brownian motion model matches experimental observations and to extract accurate diffusion coefficients, I performed simulations for a range of diffusion coefficients $D = 1.5 - 5.5 \mu\text{m}^2\text{s}^{-1}$ with increments of $0.1 \mu\text{m}^2\text{s}^{-1}$. All other simulation input parameters were directly obtained from experimental data and fixed. I evaluated simulation results by calculating the squared deviation between simulated and experimental data based on MSD curves, cumulative distribution curves, and D^* distributions. The best parameter estimate for the accurate diffusion coefficient D was obtained from the minimum of the squared deviation values (least squares estimate). I found diffusion coefficients $D_{\text{Pol}} = 2.7 \pm 0.4 \mu\text{m}^2\text{s}^{-1}$ and $D_{\text{Lig}} = 3.5 \pm 0.3 \mu\text{m}^2\text{s}^{-1}$ from simulations that precisely matched experimental MSD curves and cumulative distribution curves (Fig. 5.3). The diffusion coefficients for Pol and Lig are in the range of reported values for proteins of similar size: $D = 2.5 \mu\text{m}^2\text{s}^{-1}$ for a 72 kDa GFP fusion (Elowitz et al., 1999) and $D \sim 2.8 \mu\text{m}^2\text{s}^{-1}$ for a 138 kDa GFP pentamer (Nenninger et al., 2010). Due to confined diffusion and motion blurring, D^* distributions peak at lower values $D^*_{\text{Pol}} \sim 0.8 \mu\text{m}^2\text{s}^{-1}$ and $D^*_{\text{Lig}} \sim 1 \mu\text{m}^2\text{s}^{-1}$, as for experimental data. Because the particular shape, chemical composition, and interactions govern the movement of a protein in the cell, diffusion coefficients of different proteins can not be directly associated with their molecular weights. For instance, addition of a hexa-histidine tag reduced the diffusion coefficient of a GFP fusion by almost 50% (Elowitz et al., 2002). Rather, the close agreement between simulation and experiment shows that Pol and Lig diffusion are generally consistent with Brownian motion inside the nucleoid. While the nucleoid association of Pol and Lig indicates non-specific interactions with DNA, the significant displacement between subsequent localisations limits transient DNA binding and

sliding to < 15 ms. This finding excludes models of substrate search processes wherein Pol and Lig processively scan long chromosomal segments. The observed Brownian motion (i.e. Gaussian distribution of displacement) likely results from a sum over multiple displacements and non-specific binding events per frame. Regarding interactions of Pol and Lig with other proteins, the scaling of D with the individual protein size suggests that Pol and Lig do not stably form larger replication or repair complexes in the cytoplasm.

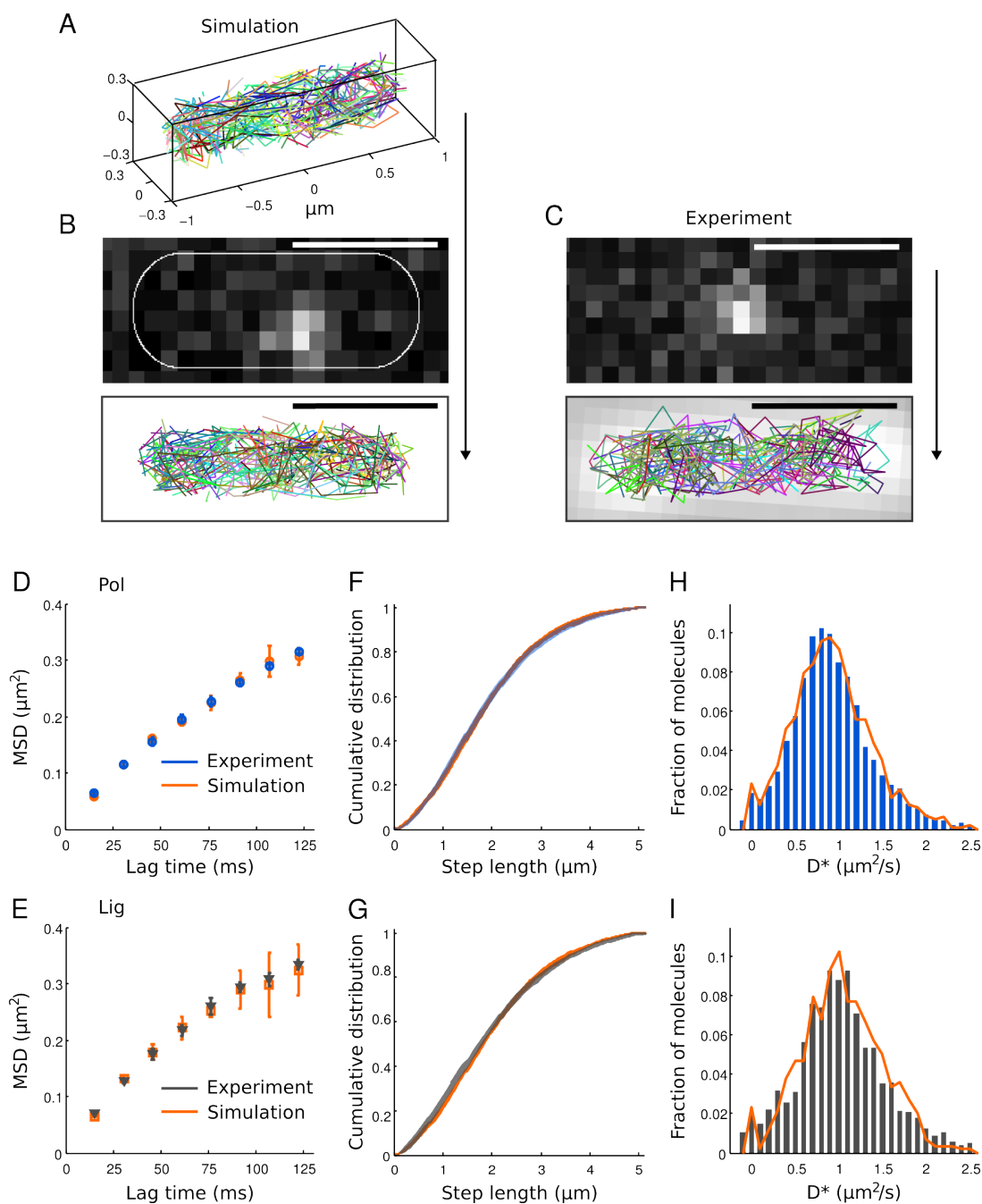


Figure 5.3: Pol and Lig diffusion match simulations of Brownian motion with diffusion coefficients $D_{\text{Pol}} = 2.7 \pm 0.4 \mu\text{m}^2\text{s}^{-1}$ and $D_{\text{Lig}} = 3.5 \pm 0.3 \mu\text{m}^2\text{s}^{-1}$. **(A)** Simulated 3D Brownian motion tracks. **(B)** Simulated tracks were used to model PALM movies (example frame shown) for localization and tracking analysis performed identical to analysis of experimental data. **(C)** Example frame from an experimental PALM movie and resulting tracks. **(D-I)** Comparison between simulation (orange) and experiment for Pol (blue, top) and Lig (grey, bottom) in undamaged cells: **(D-E)** Mean-squared displacement (MSD) versus lag time ($\pm\text{std}$). **(F-G)** Cumulative distributions of the diffusion step length (curve width: $\pm\text{std}$). **(H-I)** Apparent diffusion coefficient (D^*) distributions (normalized to maximum values).

5.3 Direct observation of DNA repair

To probe DNA repair activities of Pol and Lig, I generated single-nucleotide gap substrates in live *E.coli*, using BER of DNA-methylation damage by MMS. Cells were immobilised and imaged on agarose pads containing 100 mM MMS (Sigma-Aldrich) and imaged within 15 - 25 minutes of treatment, unless stated otherwise. I hypothesised that molecules performing repair while bound to DNA might have a much lower D^* than freely diffusing molecules, because DNA loci in the chromosome have a diffusion coefficient on the order of $10^{-5} \mu\text{m}^2\text{s}^{-1}$ (Elmore et al., 2005). Indeed, I measured an increased population of bound Pol and Lig in MMS treated cells (Fig. 5.4A).

To quantify the fraction of bound molecules, I set two thresholds: $D^* < 0.15 \mu\text{m}^2\text{s}^{-1}$ for $\text{MSD}(\Delta t = 15\text{ms})$ and $D^* < 0.075 \mu\text{m}^2\text{s}^{-1}$ for $\text{MSD}(\Delta t = 30\text{ms})$. The percentage of bound molecules was calculated from the number of bound tracks (passing both MSD thresholds) divided by the total number of tracks with four or more steps; this ratio is independent of the absolute detection efficiency of PAm-Cherry in live cells. Using simulated data of a single diffusing species matching experimental data (section 5.2), I estimated false positive identification of bound tracks (erroneous bound identification of diffusing molecules) to 0.7%. Using simulations of a mixed population of bound and diffusing tracks (containing 15% of immobile tracks), I also estimated a false negative identification (missed identification of a bound track) of 1.1%.

Undamaged cells showed primarily diffusing Pol and Lig with distinct subpopulations of $2.7 \pm 0.2\%$ and $3.8 \pm 0.4\%$ bound Pol and Lig with $D^* = 0 \mu\text{m}^2\text{s}^{-1}$, respectively (Fig. 5.4A, red bars). This population of bound Pol and Lig is also evident from the deviation between simulated and experimental CDF curves at small step lengths (Fig. 5.3). Pol in fixed cells was $> 95\%$ bound. MMS treat-

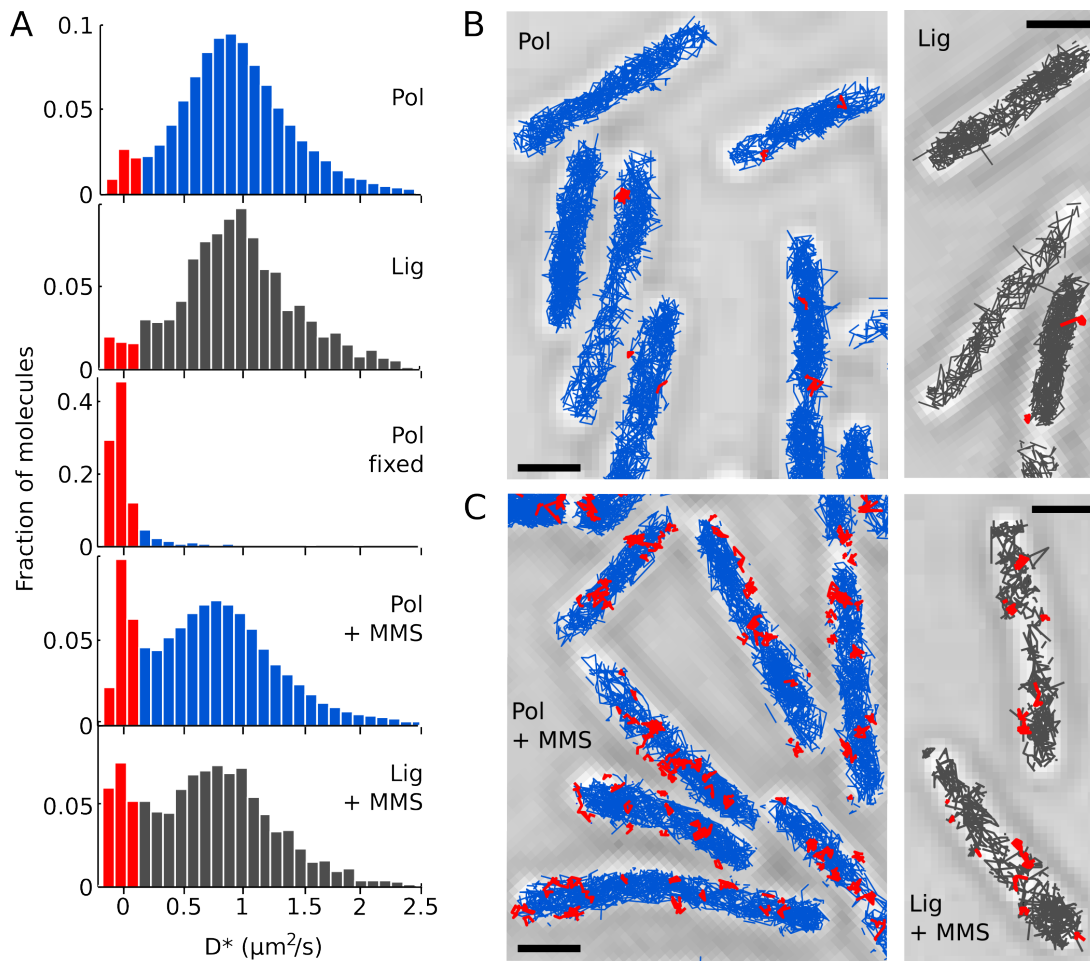


Figure 5.4: Direct observation of DNA repair in live *E. coli*. **(A)** Distributions of the apparent diffusion coefficient (D^*) for Pol and Lig in undamaged cells; Pol in fixed cells; Pol and Lig in live cells under constant 100 mM MMS treatment (populations of bound molecules in red). **(B)** Pol and Lig tracks in undamaged cells; **(C)** with 100 mM MMS. Diffusing tracks in blue/grey, bound tracks in red. Scale bars: 1 μm .

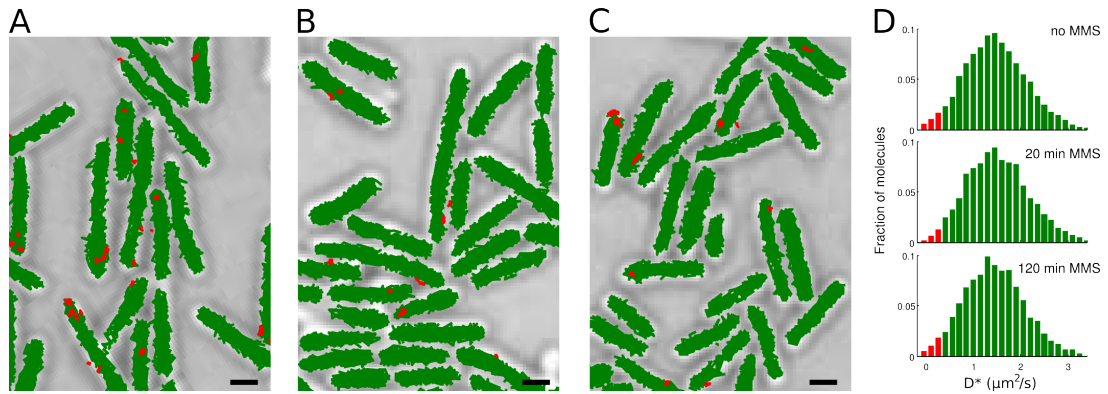


Figure 5.5: Tracking Fis-PAmCherry in undamaged cells and with 100 mM MMS treatment. (A–C) Representative images of Fis tracks in undamaged cells (A), 20 min MMS treated cells (B), and 120 min MMS treated cells. Tracks of diffusing and bound molecules are plotted in green and red, respectively. Scale bars: 1 μm . (D) Distributions of the apparent diffusion coefficient (D^*) for Fis in undamaged cells (top), 20 min MMS treated cells (middle), and 120 min MMS treated cells (bottom). Populations of bound molecules are shown in red.

ment of live cells resulted in fivefold increased populations of $13.0 \pm 0.2\%$ bound Pol and $17.9 \pm 0.3\%$ bound Lig (Fig. 5.4A bottom). The nucleoid-associated protein Fis (Ball et al., 1992), a control without DNA repair activity, showed no increase in binding even after 2 h MMS treatment (Fig. 5.5). Since the binding response to MMS was specific to the two DNA-repair proteins, I interpret the tracks of bound Pol and Lig as direct reports of DNA synthesis and ligation reactions during DNA repair (Fig. 5.4B, red tracks). Notably, diffusion of unbound Pol and Lig with MMS was unchanged, suggesting that their global diffusion and search processes remain identical upon DNA damage.

By examining the localisations of bound Pol and Lig tracks projected onto the long cell axis, I found that the spatial distribution of repair sites was spread throughout the cell, matching the distribution of unbound molecules (Fig. 5.6). This result supports a distributive search model of individual Pol and Lig molecules for damage sites on the chromosome, instead of these sites being directed to spatially organised “repair factories” (Goodman, 2002), as suggested for human cells and mitochondria (Jackson et al., 1994; Stuart et al., 2005).

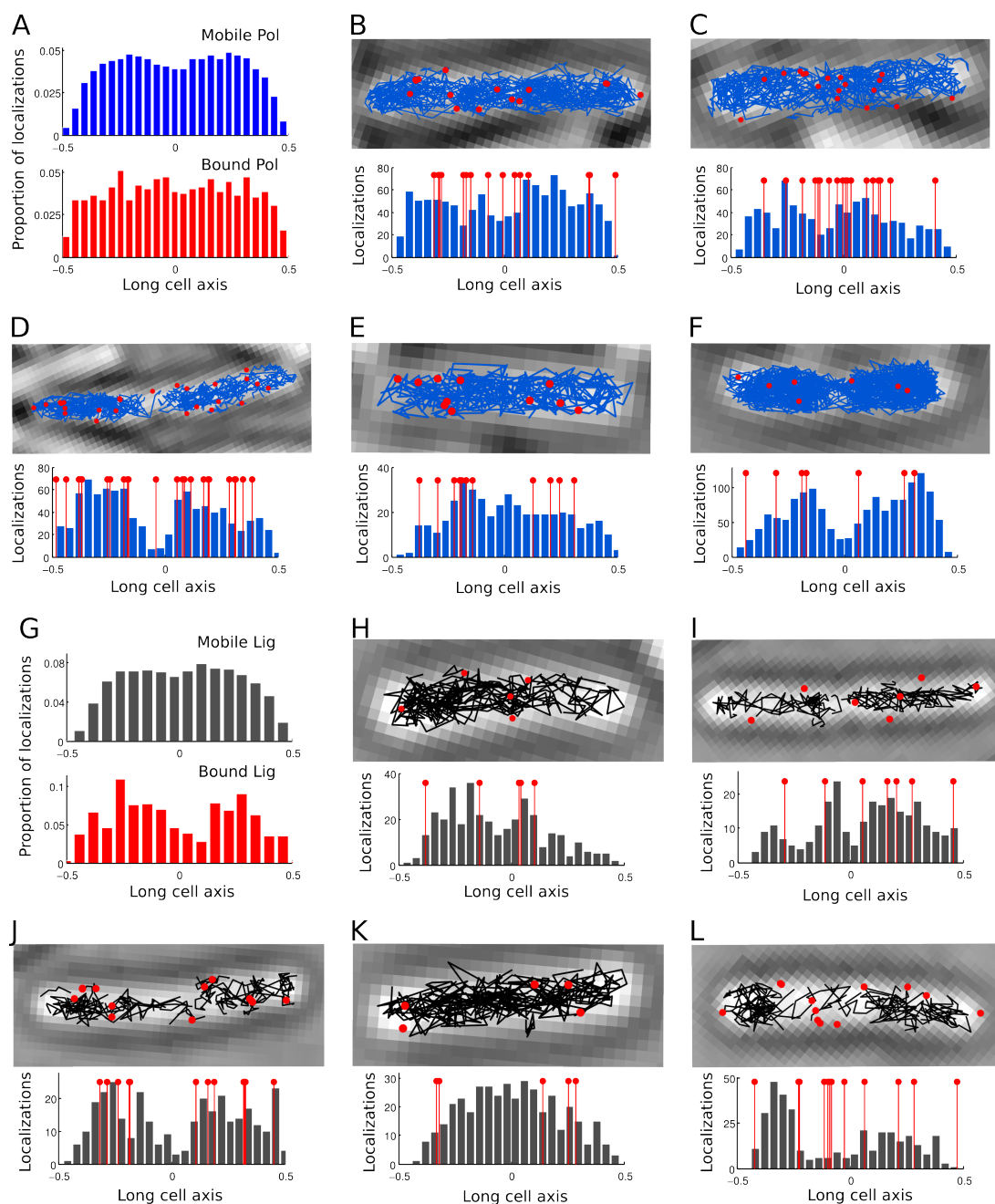


Figure 5.6: Spatial distribution of Pol (top) and Lig (bottom) repair sites. **(A, G)** Distribution of localisations from mobile tracks (top, blue/grey) and bound tracks (bottom, red) across the normalised long cell axis (combined data from 95 cells for Pol and 125 cells for Lig). **(B-F, H-L)** Example cells with mobile Pol/Lig (blue/grey tracks) and mean positions of bound Pol/Lig (red dots). The distribution of localisations over the normalised long cell axis are shown below each cell image. Mean localisations of bound tracks are shown in red.

5.4 DNA repair rates

Binding times

Next, I set out to measure single-molecule reaction rates *in vivo*. To directly visualize complete substrate search and repair cycles, I identified individual tracks that showed free diffusion leading to a binding event, followed by dissociation and continued diffusion (Fig. 5.7). Corresponding time-traces of the stochastic diffusion steps displayed characteristic binding intervals with $D^* \sim 0 \mu\text{m}^2\text{s}^{-1}$. The bound populations are hence in dynamic exchange with the pool of diffusing molecules. 1% of all traces with more than 12 steps showed complete binding events in MMS damaged cells; 0.2% in undamaged cells, respectively. This reflects the fivefold increase of Pol and Lig binding with MMS. However, most bound molecules remained immobile for the entire track, limited by photobleaching. The lifetime of PAmCherry before photobleaching followed an exponential distribution with a mean lifetime of 3.7 frames (54 ms).

To minimise the effect of PAmCherry photobleaching on the apparent repair-time distributions, I used long exposure times (500 ms, 750 ms, 1000 ms), low 561 nm excitation intensities (700 μW , 525 μW , 350 μW), and very sparse photoactivation. Under these conditions, PSFs of diffusing molecules were dim and blurred (due to motion during exposure) but PSFs of bound molecules remained narrow (Fig. 5.8). Hence, bound and diffusing molecules could be distinguished by the mean fitted PSF width (std) over each track, with thresholds x-width < 160 nm and y-width < 200 nm for bound molecules (Fig. 5.9). Note that the elliptical Gaussian fit includes two independent width parameters, and an angle to rotate the PSF such that x-width corresponds to the short elliptical axis by convention. I extracted on-time distributions from the duration of tracks that were classified as bound according to the PSF width thresholds (Table 5.1). The proba-

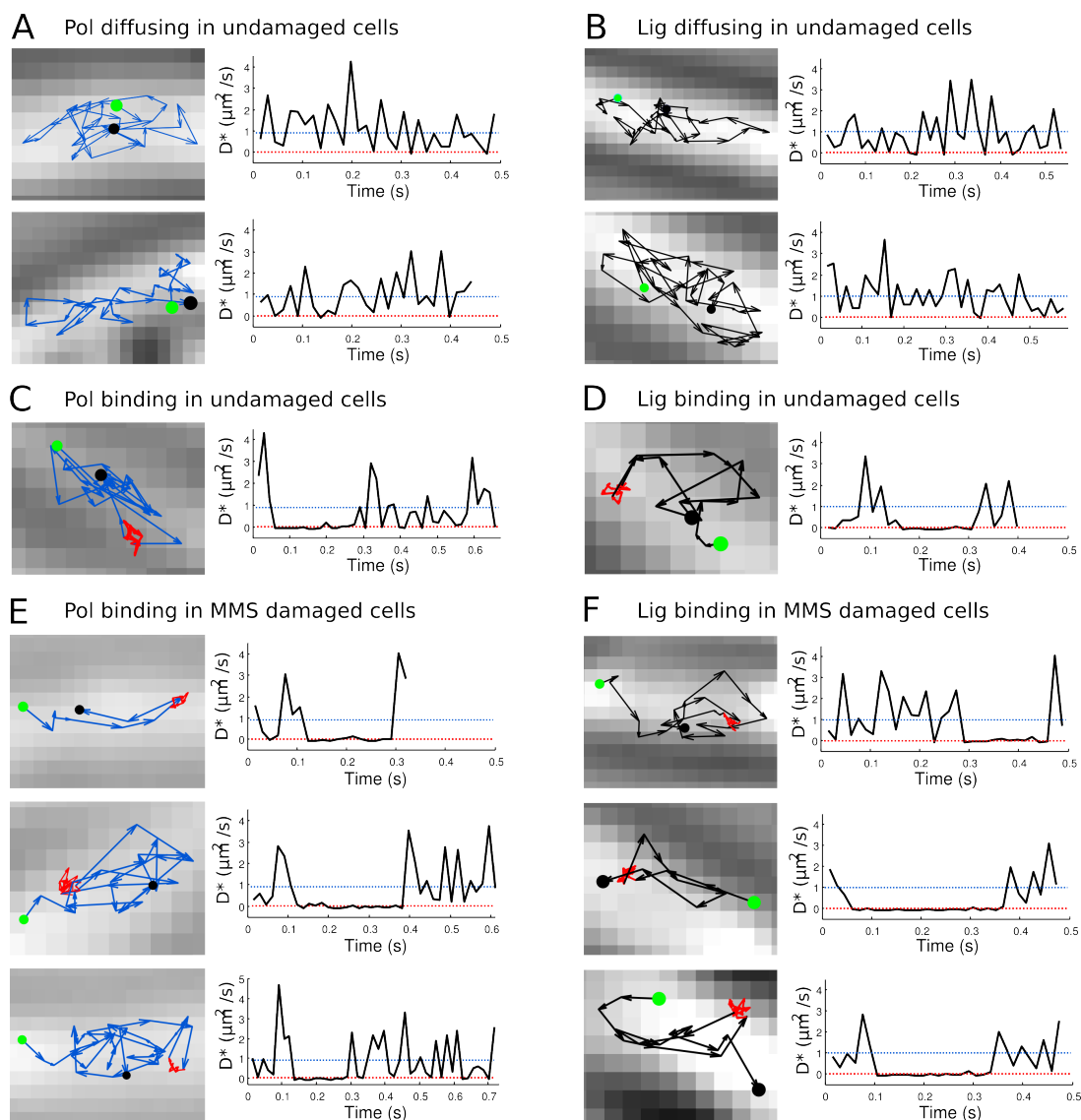


Figure 5.7: Complete binding events of Pol (left) and Lig (right) in undamaged and 100 mM MMS damaged cells. Length-scale: 114.5 nm per pixel. **(A, B)** Example tracks (starting at green dot, finishing at black dot) and time-traces of the diffusion steps showing Pol and Lig movement in undamaged cells. Dotted blue and red lines indicate the average D^* for mobile and bound molecules. **(C, D)** Example tracks and time-traces showing full binding events (red arrows) flanked by diffusion (blue/grey arrows) in undamaged cells. **(E, F)** Example binding tracks with MMS treatment.

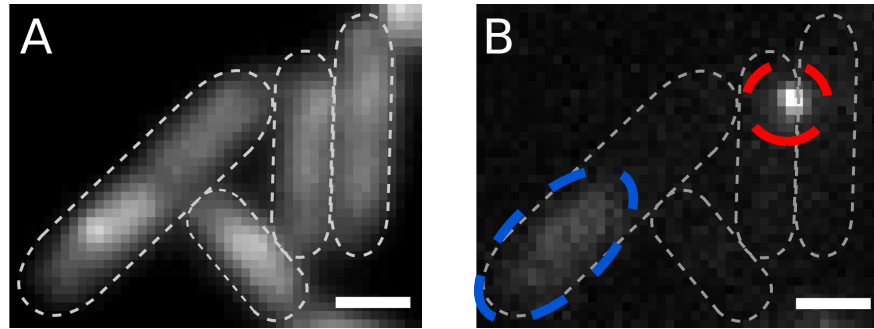


Figure 5.8: Fluorescence images of Pol-PAmCherry in live cells at 750 ms exposure time. Scale bars: 1 μm . (A) Accumulated fluorescence over a whole PALM movie. (B) Example frame showing the broad PSF of a diffusing molecule (blue outline) and the sharp PSF of a bound molecule (red outline).

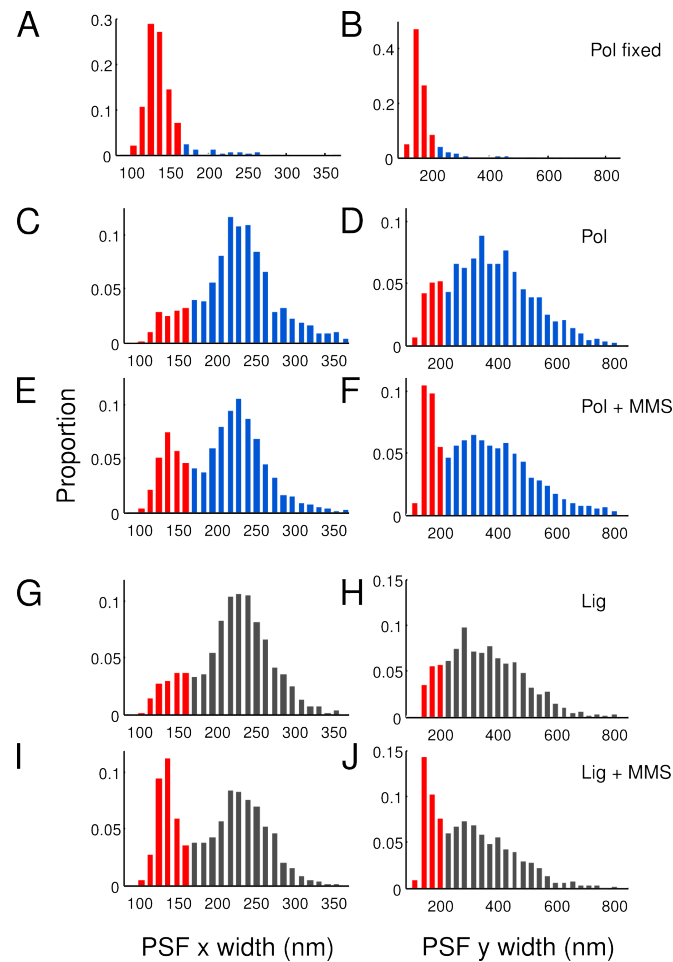


Figure 5.9: Fitted PSF width (std) for 750 ms exposure time movies. The elliptical Gaussian fit gives two PSF width parameters, x-width and y-width, and a rotation angle. The x-width parameter is arbitrarily assigned to the smaller of the two width parameters for each fit. (A, B) Pol in fixed cells. (C, D) Pol in undamaged cells. (E, F) Pol in cells with MMS treatment. (G, H) Lig in undamaged cells. (I, J) Lig in cells with MMS treatment.

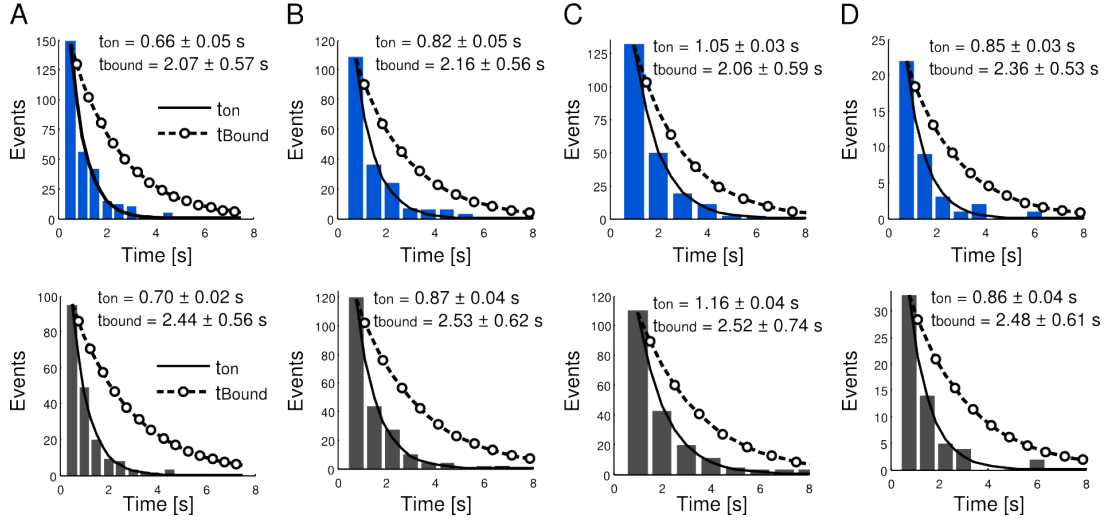


Figure 5.10: On-time distributions for Pol (top), and Lig (bottom) with MMS treatment measured at different exposure times and excitation intensities. Solid lines represent exponential fits to the measured on-time distributions ($t_{\text{on}} \pm \text{std}$) for bound molecules. The binding-time distribution (dashed line with circles, $t_{\text{bound}} \pm \text{std}$) was recovered by correcting the measured on-time for photobleaching of PAmCherry. Curves were normalised by their maximum value. (A) 500 ms exposure time, 700 μW excitation intensity. (B) 750 ms, 525 μW . (C) 1000 ms, 350 μW . (D) In cells without MMS treatment. 750 ms, 525 μW .

bility of observing a particular on-time ($\propto \exp(-t/t_{\text{on}})$) is the product of the underlying binding-time probability ($\propto \exp(-t/t_{\text{bound}})$) and the bleaching-time probability ($\propto \exp(-t/t_{\text{bleach}})$). The bleaching-time distributions were measured independently using Pol-PAmCherry in fixed cells with the same acquisition and excitation conditions. On-time and bleaching-time distributions were fitted with single-exponential functions to extract exponential time constants, and the binding-time constant was calculated by

$$t_{\text{bound}} = t_{\text{on}} \cdot t_{\text{bleach}} / (t_{\text{bleach}} - t_{\text{on}}) . \quad (5.1)$$

Stochastic photoactivation of molecules before or during binding events does not influence this measurement, because the observed binding times follow an exponential distribution and are therefore memoryless.

Since the population of narrow PSFs increased with MMS damage (Fig. 5.9,

	500 ms 700 μ W	750 ms 525 μ W	1000 ms 350 μ W	no MMS 750 ms, 525 μ W
t_{bleach} (s)	0.97 ± 0.04	1.33 ± 0.06	2.15 ± 0.15	1.33 ± 0.06
Pol t_{on} (s)	0.66 ± 0.05	0.82 ± 0.05	1.05 ± 0.03	0.85 ± 0.03
Pol t_{bound} (s)	2.07 ± 0.57	2.16 ± 0.56	2.06 ± 0.59	2.36 ± 0.53
Lig t_{on} (s)	0.70 ± 0.02	0.87 ± 0.04	1.16 ± 0.04	0.86 ± 0.04
Lig t_{bound} (s)	2.44 ± 0.56	2.53 ± 0.62	2.52 ± 0.74	2.48 ± 0.61

Table 5.1: Exponential time constants for PAmCherry photobleaching ($t_{\text{bleach}} \pm \text{std}$), measured on times ($t_{\text{on}} \pm \text{std}$), and corrected binding times ($t_{\text{bound}} \pm \text{std}$) for Pol and Lig with 100 mM MMS treatment (columns 1-3) and without MMS treatment (column 4).

red bars), I interpret the exponentially distributed Pol and Lig binding times as repair times $t_{\text{Pol}} = 2.1 \pm 0.3$ s and $t_{\text{Lig}} = 2.5 \pm 0.4$ s (average $\pm$ sem from three exposure times and excitation intensity conditions; Fig. 5.10, Tables 5.1 and 5.2). I measured binding times of Pol and Lig in undamaged cells that were similar to those under MMS treatment, suggesting that the rate-limiting steps are similar for DNA repair and lagging-strand replication reactions. The total Pol repair time includes filling the single nucleotide gap and potentially further strand displacement synthesis at a rate of > 10 nt/s (Schwartz and Quake, 2009), followed by nuclease cleavage of the generated 5' flap and product release; the latter being the likely single rate-limiting step. Accordingly, the Lig repair time corresponds to the ligation reaction and product release, with an overall reaction rate of $(1/2.5) \text{ s}^{-1} = 0.4 \text{ s}^{-1}$, in agreement with the *in vitro* reaction rate of 0.36 s^{-1} (Frey and Hegeman, 2007). However, it is generally not straightforward to compare *in vivo* results to previous *in vitro* studies because of the strong dependence of the DNA-binding affinity on salt and substrate concentrations, as well as macromolecular crowding effects. Markiewicz et al. (2012) measured *in vitro* binding times of KF to primer-template DNA of ~ 2 s, with a strong dependence on the concentration of correct or incorrect nucleotides and NaCl.

Reaction rates

To convert binding times into reaction rates, I used the measured fraction of bound molecules:

$$N_{\text{bound}} / (N_{\text{bound}} + N_{\text{free}}) = t_{\text{bound}} / (t_{\text{bound}} + t_{\text{search}}) \quad (5.2)$$

This gives the reaction rate per molecule: $1 / (t_{\text{bound}} + t_{\text{search}})$. For undamaged cells, I obtained basal reaction rates of $\sim 0.7 \text{ min}^{-1}$ per single Pol and $\sim 0.9 \text{ min}^{-1}$ per single Lig. To estimate the total reaction rate per cell, I used the Pol and Lig copy number distributions with median values of 479 Pol, 226 Lig per cell (Fig. 4.14), corresponding to cellular reaction rates of $\sim 330 \text{ min}^{-1}$ for Pol and $\sim 210 \text{ min}^{-1}$ for Lig. The basal activities match expectations: DNA replication accounts for 30-50% of the observed activity, considering two replication forks per cell generating Okazaki fragments at a rate of $\sim 120 \text{ min}^{-1}$ (Kornberg, 2005). Additionally, Pol and Lig are involved in the repair of replication errors and damage by reactive metabolites, including endogenous methylating agents. Oxidative DNA damage alone produces 3000-5000 lesions during normal metabolism per *E.coli* generation, requiring a repair rate of $30 - 60 \text{ min}^{-1}$ (Cox, 1998).

With MMS, the reaction rates increase to $\sim 3.7 \text{ min}^{-1}$ and $\sim 5.1 \text{ min}^{-1}$ per single Pol and Lig, leading to cellular repair rates of $\sim 1780 \text{ min}^{-1}$ and $\sim 1150 \text{ min}^{-1}$ for Pol and Lig; the 50% higher Pol activity over Lig likely reflects the wider range of Pol substrates and its nick-translation activity, which competes with the ligation reaction (Sung and Mosbaugh, 2003). As the final step in the pathway, the Lig activity corresponds to the overall BER rate. Pol and Lig activities are summarised in table 5.2.

	Pol	Lig	Pol + MMS	Lig + MMS
% bound molecules	2.7 ± 0.2	3.8 ± 0.4	13.0 ± 0.2	17.9 ± 0.3
Bound molecules per cell	13 ± 1	9 ± 1	62 ± 1	40 ± 1
Binding time (s)	2.4 ± 0.5	2.5 ± 0.6	2.1 ± 0.3	2.5 ± 0.4
Rate per molecule (min^{-1})	0.7 ± 0.2	0.9 ± 0.2	3.7 ± 0.5	5.1 ± 0.7
Rate per cell (min^{-1})	329 ± 76	207 ± 54	1782 ± 248	1152 ± 167
Search time (s)	85 ± 20	63 ± 17	14 ± 2	9 ± 2

Table 5.2: Summary of Pol and Lig activities in undamaged cells and with constant 100 mM MMS treatment. Bound molecules per cell and reaction rates per cell are stated for a cell with exact copy numbers of 479 Pol and 226 Lig (median values of the copy number distributions in Fig 4.14). Errors are $\pm$ sem for binding times and % bound molecules; the other errors were estimated by error propagation. Differences in copy numbers cause additional natural variation in the number of bound molecules and reaction rates across cells.

Substrate-limited repair rates

I examined four factors that may influence the observed DNA repair rates: Pol and Lig copy numbers, MMS concentration, upstream reaction rates in the BER pathway, and MMS treatment time. To test the influence of copy numbers, I linked the natural copy number variation between cells (Fig. 4.14) with the bound-fraction distribution across cells (Fig. 5.11 A-B); Pol and Lig binding saturated in cells with higher copy numbers but similar size (Fig. 5.11 C-D). Excess copies joined the diffusing pool of molecules (Fig. 5.11 E-F), demonstrating that the DNA repair activities of Pol and Lig are limited by the number of substrate sites available for binding.

Pol and Lig binding also saturated with increasing MMS concentration (at $\sim 13\%$ and $\sim 18\%$ bound fraction, Fig. 5.12), showing that the repair rates are not limited by the number of initial damage sites, but instead by upstream BER steps, such as the glycosylase and endonuclease reactions. Similarly, Lig activity is limited by the rate that Pol creates nicked substrates. Cells benefit from high Pol and Lig rates relative to repair initiation, because gapped and nicked intermediates are typically more toxic than the initial DNA lesions (Fu et al., 2012). Overabundance of high-fidelity Pol relative to error-prone DNA

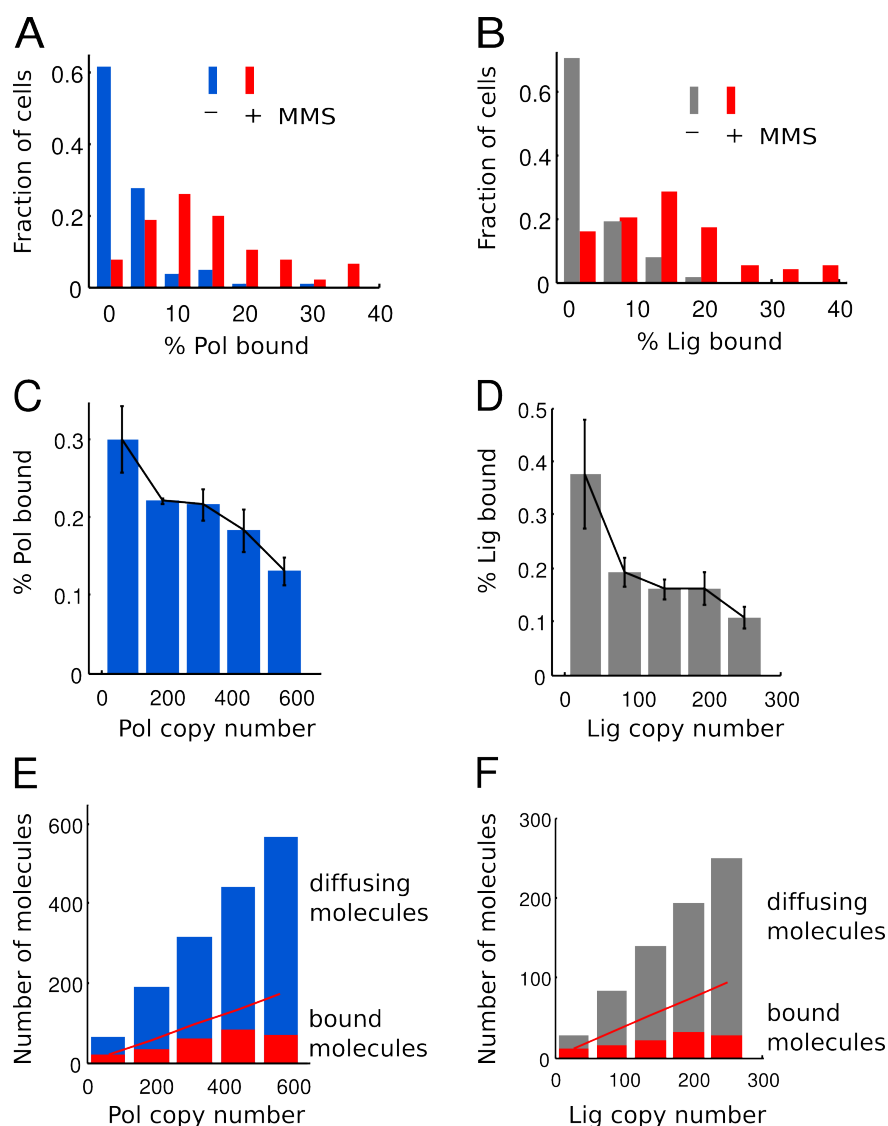


Figure 5.11: Single cell analysis shows that Pol (left) and Lig (right) saturate repair substrates. **(A, B)** Distributions of the percentage of bound Pol and Lig across cells with and without 100 mM MMS treatment. **(C, D)** Percentages of bound Pol and Lig as a function of the copy number per cell with 100 mM MMS treatment ($\pm$ sem), only including cells of similar size (length $> 3.5 \mu\text{m}$). **(E, F)** Alternative representation of the data in (C, D), showing the number of diffusing and the number of bound molecules as a function of the copy number per cell. The number of bound molecules saturates with increasing copy number. Binding is hence saturated by limiting amounts of substrates, while excess copies increase the diffusing pool of molecules. Red lines indicate the hypothetical case of unlimited substrates, when the percentage of bound molecules is independent of the copy number and binding does not saturate.

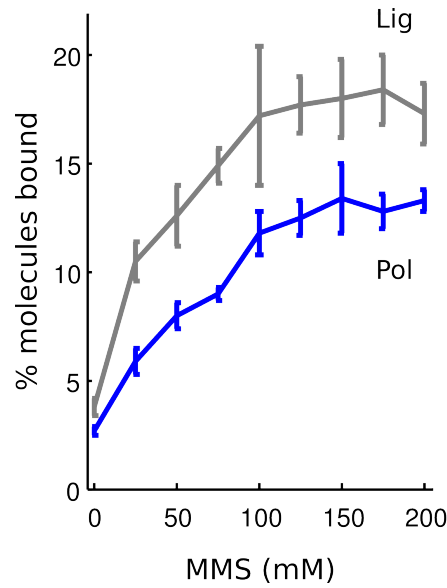


Figure 5.12: Pol and Lig dose response to MMS damage. Average percentage of bound Pol and Lig for different MMS concentrations ($\pm$ sem; $n = 4$).

polymerase IV also contributes towards minimising mutations rates associated with DNA repair (Goodman, 2002; Hastings et al., 2010). I measured a similar trend of saturated Pol activity with MMS treatment over time, with the bound Pol fraction increasing significantly within 5 min and reaching $\sim 13\%$ binding within 15 min of treatment (Fig. 5.13A).

To explore the Pol response to DNA damage and recovery, I immobilised cells using 1% polyethylenimine on glass slides with silicon gaskets. This allowed imaging cells on the same slide in M9 medium first, then incubating with 100 mM MMS for 15 min, followed by removal of MMS and washing with M9 medium for DNA damage recovery. I observed rapid reduction of Pol binding within 5 min and return to basal activity within 15 min of recovery (Fig. 5.13B); I estimated that Pol catalyses an integrated number of ~ 18000 repair events per cell during the damage and recovery experiment (Fig. 5.13B, orange shaded area).

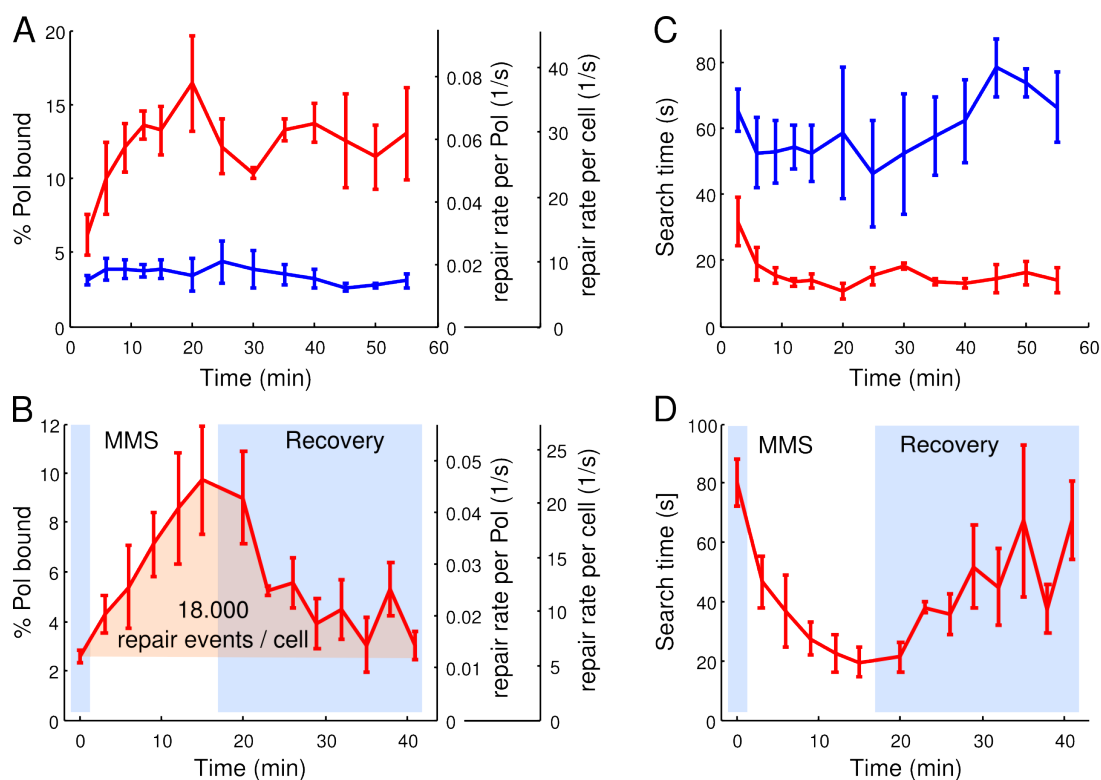


Figure 5.13: Timing of Pol response to DNA damage and recovery. **(A)** Percentage of bound Pol (left axis), repair rates per Pol molecule and per cell (right axes), measured over time during constant 100 mM MMS treatment (red), and control (blue). ($\pm$ sem; $n = 3$) **(B)** Pol DNA damage response during 100 mM MMS treatment for 15 min; followed by recovery after removing MMS (blue background). ($\pm$ sem; $n = 3$). Integrated Pol repair activity above the basal level is highlighted in orange. **(C, D)** Pol search times corresponding to data in (A, B).

Substrate search times

The search time between two reactions (t_{search}) decreased from ~ 85 s for Pol in undamaged cells to ~ 14 s with saturating MMS damage, and increased during recovery (Fig. 5.13C, D); similar times were obtained for Lig (Table 5.2). Decreasing search times reflect increasing substrate concentrations. Using the measured search times, I calculated the concentration of free gaps and nicks in the cell. The diffusion-limited search time of a single enzyme with diffusion coefficient, D , for free substrates with concentration, $C_{\text{substrate}}$, and association radius, a , is given by Halford and Marko (2004):

$$t_{\text{search}} = 1 / (4\pi \cdot D \cdot a \cdot C_{\text{substrate}}) . \quad (5.3)$$

Using $C_{\text{substrate}} = N_{\text{substrate}} / V_{E.coli}$, I calculated the number of free substrates:

$$N_{\text{substrate}} = V_{E.coli} / (4\pi \cdot D \cdot a \cdot t_{\text{search}}) . \quad (5.4)$$

I used parameter values $V_{E.coli} = 10^{-18} \text{ m}^3$; $D_{\text{Pol}} = 2.7 \mu\text{m}^2\text{s}^{-1}$ and $D_{\text{Lig}} = 3.5 \mu\text{m}^2\text{s}^{-1}$; $a = 0.5 \text{ nm}$; $t_{\text{search}} = 14 \text{ s}$ for Pol and $t_{\text{search}} = 9 \text{ s}$ for Lig (with saturating MMS). This gives $N_{\text{substrate}} = 4.2$ free binding sites for Pol, and $N_{\text{substrate}} = 5.1$ free binding sites for Lig.

I further calculated the lifetime of a free substrate before enzyme binding:

$$t_{\text{free}} = V_{E.coli} / (4\pi \cdot D \cdot a \cdot N_{\text{protein}}) . \quad (5.5)$$

Importantly, N_{protein} refers to the number of free enzymes that are available for substrate binding, accounting for the fraction of enzymes that are already bound. I obtained this fraction from the measurements (Table 5.2): $N_{\text{protein}} =$

$479 - 62 = 417$ for Pol, and $N_{\text{protein}} = 22640 = 186$ for Lig. The lifetime of gapped substrates, considering only Pol binding, gives $t_{\text{free}} = 140$ ms. With Pol and Lig both competing for nicked substrates, I find $t_{\text{free}} = 90$ ms. Alternatively, I can use the fact that the turnover rate of free substrates equals the enzyme repair rate per cell (Table 5.2) at equilibrium: $r_{\text{cell}} = N_{\text{substrate}}/t_{\text{free}}$. Using $r_{\text{cell}} = 29.7 \text{ s}^{-1}$ for Pol, and $r_{\text{cell}} = 19.2 \text{ s}^{-1}$ for Lig gives the same results for t_{free} (within rounding error).

In summary, I estimated there are about five free gaps and nicks with lifetimes of 140 ms and 90 ms, respectively, to be present with saturating MMS at any time; this contrasts to 62 Pol-bound and 40 Lig-bound substrates. The short lifetime of free gaps and nicks, a consequence of the distributive search by excess numbers of Pol and Lig, minimises the accumulation of these repair intermediates and safeguards genomic stability.

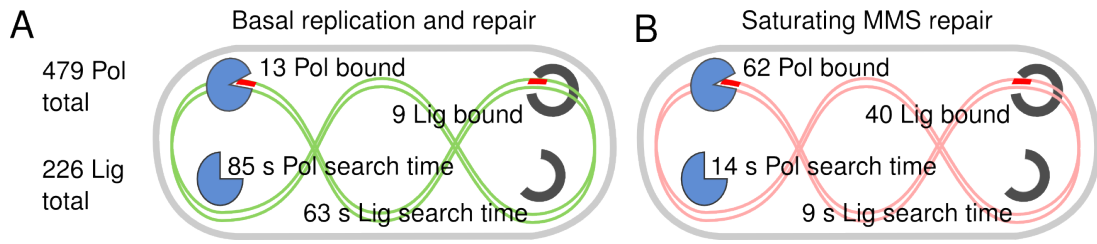


Figure 5.14: Model for Pol and Lig activity in undamaged *E. coli* (A), and under saturating MMS treatment (B).

5.5 Conclusion

This chapter gave a quantitative account of the ubiquitous DNA synthesis and ligation steps of the bacterial DNA base excision repair pathway, summarised in Fig. 5.14. Pol and Lig in *E. coli* show similar behaviour to mammalian repair enzymes (Houtsmuller et al., 1999; Dinant et al., 2009): Individual enzymes diffuse freely with little affinity for undamaged DNA but become immobilised for DNA repair on damage sites that are distributed throughout the nucleoid/nucleus. While mammalian repair complexes assemble on damage sites for several minutes, Pol and Lig binding events last only few seconds which indicates that these prokaryotic enzymes act individually instead of joining multi-protein repair complexes.

The spatial distribution and movement of DNA binding proteins report on the underlying structure of the chromosome. Further experiments and analysis might inform about potential reorganisation and compaction of the bacterial nucleoid in response to DNA damage (Levin-Zaidman et al., 2000; Smith et al., 2002). Future work can also explore the role of dynamic DNA damage responses. For instance, increased DNA glycosylase activity upon induction of the adaptive response by DNA alkylation damage is expected to change the rate of BER and should thus affect Pol and Lig activities. The SOS response, on the other hand, induces alternative DNA polymerases that might compete with Pol for gapped

DNA substrates. It will be instructive to quantify the DNA repair activity and binding time of a Pol mutant with a lower DNA-binding affinity (Kelly and Grindley, 1976). Furthermore, observing the repair activity of Lig-PAmCherry in the defective Pol mutant strain *polA_{ts}* could directly test the requirement of Pol to fill gaps and generate ligatable nicks.

Cancer, neurodegeneration, and ageing diseases have all been linked to corrupted DNA repair that causes DNA damage accumulation and mutagenesis (Wilson and Bohr, 2007). A major goal towards understanding and solving these conditions is the generation of comprehensive quantitative models that predict an organism's response and sensitivity to DNA damage. The general experimental approach described here provides a starting point for quantitative modeling of entire repair pathways. Such models might eventually have a clinical impact in cancer therapy by predicting the optimal treatment type and dose of medication (Kinsella, 2009).

Chapter 6

Imaging DNA replication and chromosome organising proteins in live bacteria

6.1 Introduction

The previous chapter has focused on organising principles of DNA repair. In this chapter, I will present two further applications of single-molecule imaging to investigate protein-DNA interactions in live *E.coli* and discuss the results within the wider context of bacterial chromosome biology. These projects were initiated by Rodrigo Reyes-Lamothe and Anjana Badrinarayanan (Oxford Biochemistry). The organisation of the chromosome is key to the function of all DNA-binding proteins *in vivo*. The single circular 4.6 Mb *E.coli* chromosome folds into a nucleoid with the left and right arms occupying either side of the origin of replication (*oriC*) at mid-cell, while the terminus region (*ter*) spans between the polar ends of the nucleoid (Reyes-Lamothe et al., 2008b). Upon initiation, two replication forks commence bi-directional DNA synthesis at mid-cell, then separate

and translocate to quarter-cell positions for most of the replication cycle, and return close to mid-cell for termination (Reyes-Lamothe et al., 2008a). Here, I exploit the faithful positioning of replication forks to test the ability of my tracking method to distinguish molecules bound to DNA from the diffusing pool of molecules. Single-molecule imaging has previously been used to determine the stoichiometry of the multi-protein replisome *in vivo* (Reyes-Lamothe et al., 2010). Extending these findings towards an understanding of the dynamic composition of the replisome, I used PALM imaging to measure residence times of the ϵ subunit of DNA polymerase III and the δ subunit of the clamp loader complex within the replisome.

Concurrently with replication, nascent DNA folds and segregates into sister chromosomes (Wang et al., 2006; Berlatzky et al., 2008). Contrary to the original view, the bacterial chromosome is highly organised throughout the cell cycle, with genes more distant from *oriC* on the genetic map also being physically positioned further from mid-cell. Replication and segregation result in a (left-right)-(left-right) organisation of the chromosome arms in the two sister nucleoids. My data for the spatial distribution of Pol tracks showed the ellipsoidal shape of the *E.coli* nucleoid and its segregation at different stages of the cell cycle (Fig. 5.2). However, much remains to be learned about the mechanisms that achieve organisation, compaction, and segregation of the nucleoid. Non-specific chromosome condensation arises from DNA supercoiling and phase-separation between DNA and cytoplasmic proteins (Woldringh and Nanninga, 2006). Furthermore, a palette of highly expressed nucleoid-associated proteins (NAPs), including Fis (Ball et al., 1992), compact chromosomes by bridging, bending, and wrapping DNA, akin to the function of histones in eukaryotes (Dillon and Dorman, 2010; Stavans and Oppenheim, 2006). My results in chapter 4.5 confirmed that Fis is

present at very high copy numbers during early exponential growth (Fig. 4.14). In addition to small NAPs that coat the chromosome globally, large Structural Maintenance of Chromosomes (SMC) complexes appear to clamp distant DNA segments at specific cellular positions to aid chromosome compaction and segregation (Nasmyth and Haering, 2005; Hirano, 2002; Danilova et al., 2007; Badrinarayanan et al., 2012b). In the second part of this chapter, I apply PALM imaging to show that the *E.coli* SMC complex, MukBEF, forms stoichiometric complexes in solution, while immobile MukBEF are located within clusters that are further organised into densely packed sub-clusters.

6.2 Organisation of the *E.coli* replisome

The replisome is a multi-protein complex that duplicates DNA into two identical copies (Kornberg, 2005; Johnson and O'Donnell, 2005; Benkovic et al., 2001). The semiconservative nature of DNA replication requires each strand of the parental duplex DNA to serve as a replication template. Because the template strands have opposite polarity while DNA polymerisation functions exclusively in 5' to 3' direction, only one of the strands, termed the leading strand, can be copied in a continuous fashion. The lagging strand, on the other hand, is replicated in discontinuous stretches, so-called Okazaki fragments. The replisome couples synthesis on both strands within the same complex to act as a single functional unit.

The *E.coli* replisome is composed of at least 10 proteins. At the front of the progressing replication fork, the hexameric helicase (*dnaB*) unwinds duplex DNA in 5' to 3' direction. The actual DNA synthesis is performed by DNA polymerase III (PolIII) core enzyme, consisting of the catalytic α subunit (*dnaE*), the 3'-exonuclease ϵ subunit (*dnaQ*) for proofreading, and the non-essential θ subunit

(*holE*). The β sliding clamp (*dnaN*) encircles dsDNA as a dimeric ring and binds PolIII core, rendering DNA synthesis highly processive (> 50 kb) and fast (750 nt/s).

The clamp loader complex engages the β clamp and acts as a scaffold for the replication fork by binding the helicase and α of the PolIII core. Three τ subunits (*dnaX*) attach a PolIII core each. Alternatively, γ , encoded via a ribosomal frameshift by the same gene (*dnaX*), can replace τ but does not bind a PolIII core. ATP binding causes a conformational change in τ for the recruitment of β clamp. Tanner et al. (2011) showed that β clamp can be recycled from a completed Okazaki fragment on the lagging strand. The single δ subunit (*holA*) opens the β dimer for loading onto a primer terminus with help of the δ' subunit (*holB*). The sliding clamp closes upon ATP hydrolysis in the presence of DNA and the clamp loader complex is recycled after ADP release. In addition, heterodimers of χ and ψ subunits (*holC* and *holD*) form part of the clamp loader complex and interact with SSB.

Efficient and robust replication requires coordinated leading and lagging strand synthesis. Repeated cycles of discontinuous lagging strand synthesis are initiated by primase (*dnaG*) creating short 10 - 12 nt RNA primers. β clamp and PolIII core load onto a new primer and fill in the Okazaki fragment, while the helicase continues to generate single-stranded DNA (ssDNA) forming a loop for the next Okazaki fragment. Either completion of an Okazaki fragment or synthesis of a new primer can trigger the release of the loop (Hamdan et al., 2009). The lagging strand ssDNA is coated by ssDNA binding protein SSB (*ssb*) homotetramers that prevent DNA secondary structure formation and protect against nucleases. SSB also aids DNA recombination and triggers the SOS DNA damage response that rescues stalled replication forks by replacing PolIII with lesion bypass DNA polymerase V (Goodman, 2002). Removal of RNA primers by DNA polymerase

I (*polA*) and ligation (*ligA*) complete synthesis of each Okazaki fragment (section 5.1).

Conventional ensemble experiments have struggled to dissect the dynamic function of the replisome due to the stochastic timing of enzymatic reactions and transient protein interactions. This challenge has been overcome by single-molecule techniques that measure isolated components of the replisome or within reconstituted systems *in vitro* (van Oijen and Loparo, 2010; Perumal et al., 2010). Live-cell fluorescence imaging of replisome components with single-molecule sensitivity has been possible because the replication fork is a very stable and slowly moving complex in the cell (Reyes-Lamothe et al., 2008a). Reyes-Lamothe et al. (2010) (Oxford Biochemistry) determined the stoichiometry of replisome components by counting the number of fluorescent proteins within replication fork foci in live *E.coli*. Extension of this work focused on the dynamic exchange of proteins in the complex using fluorescence recovery after photobleaching (FRAP) experiments. To directly visualise the diffusing molecules and those that form part of the replication fork, Rodrigo Reyes-Lamothe prepared *E.coli* AB1157 strains with PAmCherry fusions to ϵ and δ as described in chapter 4.4. I contributed to single-particle tracking PALM data acquisition and performed the data analysis.

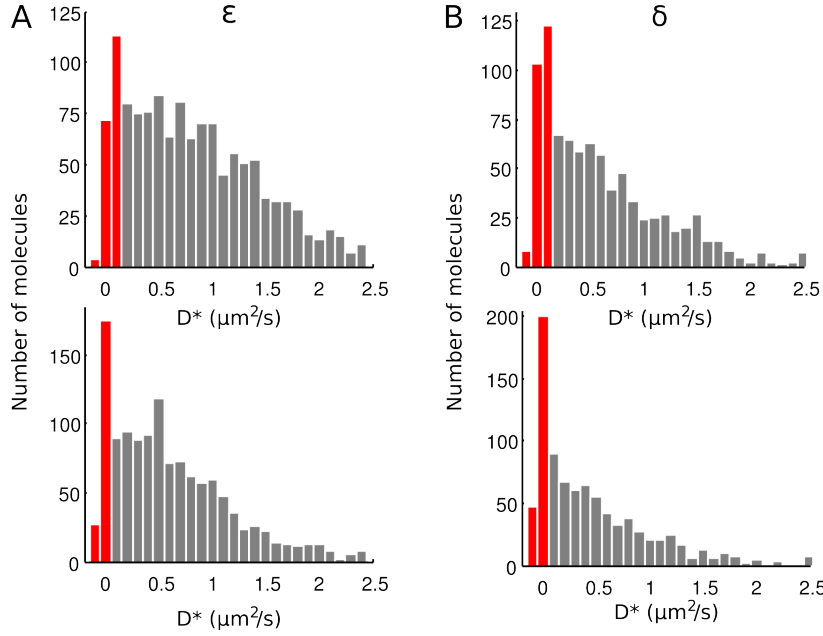


Figure 6.1: Distributions of the apparent diffusion coefficient (D^*) calculated from MSD($\Delta t = 15\text{ms}$) (top graphs) and MSD($\Delta t = 30\text{ms}$) (bottom graphs). The bound population of molecules is shown in red. **(A)** ϵ subunit of PolIII. **(B)** δ subunit of the clamp loader complex.

6.3 Imaging replisome components

Using cell preparation and PALM imaging conditions as described (chapter 4.4), I tracked the movement of ϵ -PAmCherry and δ -PAmCherry in live *E.coli*. Based on MSD analysis described in chapter 5.3, 13% of ϵ and 25% of δ molecules were classified as bound in cells. D^* distributions for ϵ show the bound population and a major slowly diffusing population, likely corresponding to ϵ complexed within the 166 kDa PolIII core complex in solution (Fig. 6.1A). The spread of the distribution towards higher D^* suggests the presence of free 27.5 kDa ϵ . Similarly, the slowly diffusing population of δ appears to be part of the large 297 kDa clamp loader complex in solution, with a minor population of free 38.7 kDa δ at higher D^* (Fig. 6.1B).

Using both the single-step and double-step displacements (MSD for $\Delta t = 15\text{ms}$ and 30ms) facilitated identification of bound molecules with good specificity de-

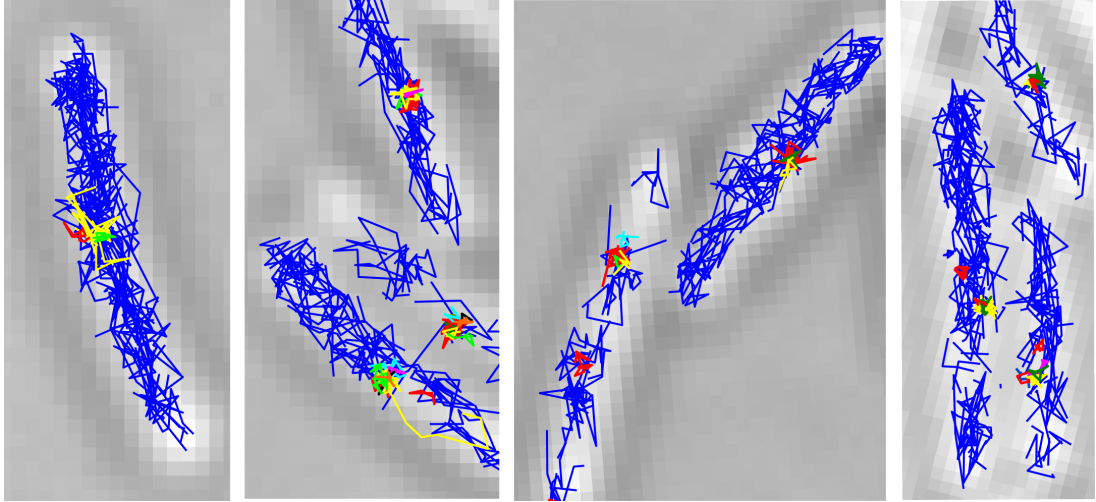


Figure 6.2: Tracking the ϵ subunit of PolIII in live *E.coli*. Tracks of diffusing molecules are plotted in blue; tracks of bound molecules in different colours.

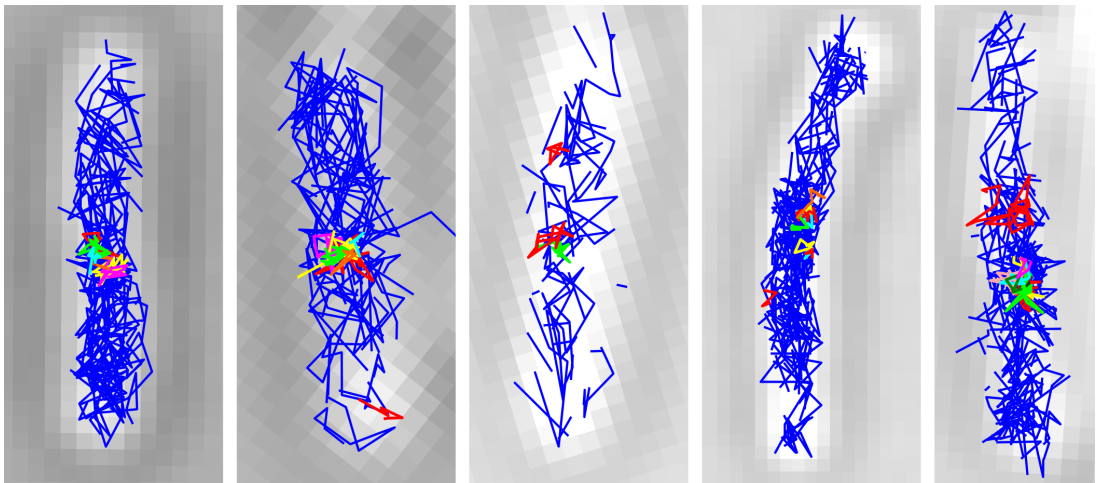


Figure 6.3: Tracking the δ subunit of the clamp loader complex in live *E.coli*. Tracks of diffusing molecules are plotted in blue; tracks of bound molecules in different colours.

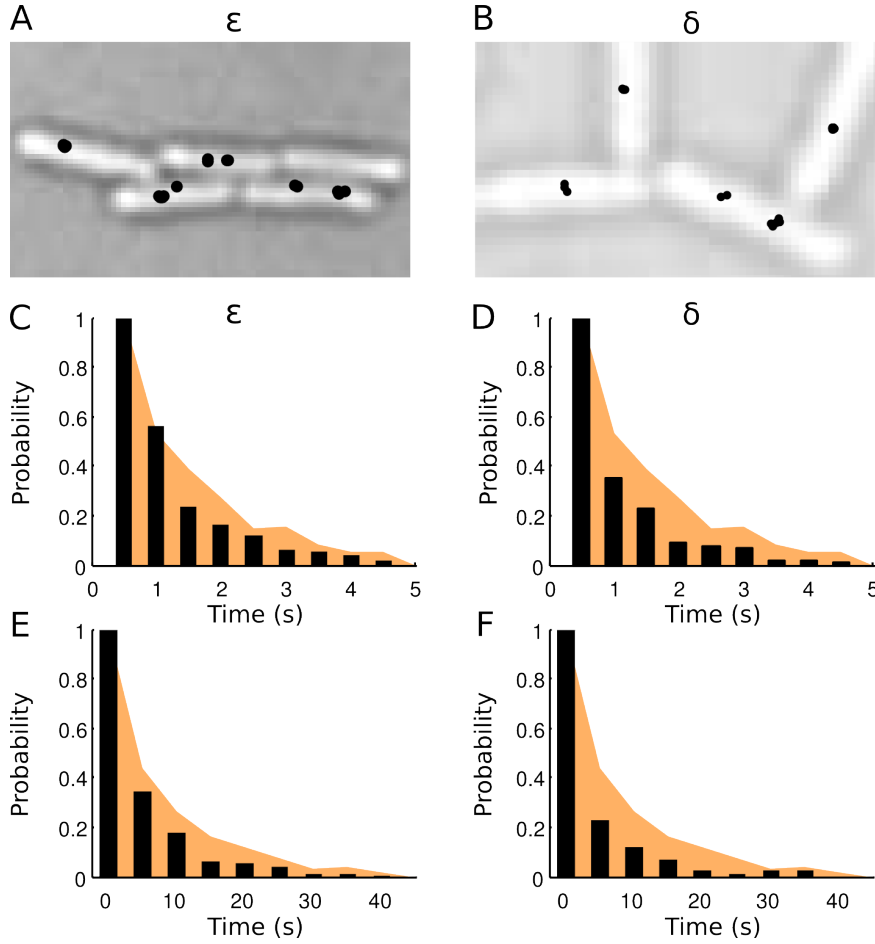


Figure 6.4: Binding times of ϵ and δ at the replication fork recorded at 500 ms exposure time. (A, B) Localisations of bound ϵ and δ (black circles). (C, D) Distribution of the observed on-times for bound ϵ and δ recorded at 500 ms exposure time (black bars), and corresponding photobleaching distribution (orange background). (E, F) Distribution of the observed on-times for bound ϵ and δ recorded at 500 ms exposure time with 5 s time-lapse (black bars), and corresponding photobleaching distribution (orange background).

spite the broad D^* distributions. Bound ϵ and δ proteins were almost exclusively localised at the expected positions of the replication fork, namely at mid-cell or at quarter-cell positions (Figs. 6.2 and 6.3). Diffusing molecules were distributed randomly throughout the cell. These results prove the ability of the tracking experiments and data analysis to specifically identify single proteins that are bound to DNA or form part of large immobile complexes and distinguish them from the diffusing pool of molecules in live cells.

To investigate the residence time of ϵ and δ components within the replisome, I

used long 500 ms exposure time and low 700 μW excitation intensity, as described in chapter 5.4. Bound molecules could be distinguished from diffusing molecules by the PSF width, as demonstrated by localisation of bound ϵ and δ at the expected replication fork positions (Fig. 6.4A, B). Measuring the duration of bound tracks provided on-time distributions (Fig. 6.4C, D). The distribution for δ exhibited a steeper decay compared to ϵ , indicating a shorter residence time at the replication fork. Observed on-times were limited to a large extent by photobleaching (orange background in Fig. 6.4C, D), which I measured under the same conditions using ϵ -PAmCherry in fixed cells. To gain access to a larger range of residence times I applied time-lapse microscopy with 500 ms exposure time at a period of 5 s. This gave a distribution of on-times for ϵ of 5 s and longer (Fig. 6.4E). Time-lapse imaging of δ confirmed its shorter residence time compared to ϵ (Fig. 6.4F).

Discussion

The recently revealed presence of three replicative DNA polymerases in the replisome (McInerney et al., 2007; Reyes-Lamothe et al., 2010) has been shown to enhance lagging strand replication processivity and avoid ssDNA gaps that are generated by replication forks containing only two polymerases (Georgescu et al., 2012). The third PolIII may be required to avoid early termination of an Okazaki fragment upon primer synthesis (“signal release”). My preliminary results complement FRAP experiments that showed exchange of bleached ϵ and δ in the replication fork against fluorescent molecules from solution within several seconds¹. It is possible that the binding-time distribution of ϵ contains slow and fast exponential components because the leading-strand PolIII core is expected to be more stably bound than the other two PolIII, which have been reported

¹FRAP experiments by Rodrigo Reyes-Lamothe, personal communication

to exchange dynamically on the lagging strand (Lia et al., 2012). However, my data show that a large part of ϵ molecules bind the replication fork for several seconds, indicating that lagging-strand PolIII copies stay associated with the replication fork for the synthesis of more than one Okazaki fragment. Further time-lapse experiments will explore longer binding times and test for the presence of ϵ components with different binding kinetics.

6.4 Structural Maintenance of Chromosome proteins

Controlling the organisation of highly compacted chromosomes for efficient transcription, replication, repair, and segregation constitutes a remarkable task that requires the concerted action of a range of DNA-binding proteins. Amongst these, Structural Maintenance of Chromosome (SMC) complexes, such as condensins and cohesins serve a central role in organising the higher-order architecture and segregation of chromosomes (Nasmyth and Haering, 2005; Hirano, 2002). The eukaryotic cell cycle allows for chromosome replication, segregation, and cell division to take place in separate stages with checkpoints ensuring completion of each stage. It remains puzzling how bacteria manage to simultaneously perform these tasks, especially at high growth rates, when replication initiates from nascent origins before completion of the previous replication round. Bacterial SMC proteins appear to be particularly important under these conditions: *E.coli* condensin mutants are unable to grow above 30°C and lose chromosome separation fidelity as shown by the frequent occurrence of anucleate daughter cells (Niki et al., 1991).

The central component of the *E.coli* condensin complex is MukB (Niki et al., 1991), a 177 kDa ATPase and DNA-binding protein that shares structure and function but little sequence similarity with other members of the SMC protein family. MukB twists into an antiparallel coiled coil of ~50 nm length such that the N- and C-termini jointly form a globular head domain. On the opposite side of the coiled coil, the central hinge domain provides the interface for MukB dimerisation. Two smaller accessory proteins, 27 kDa MukE and 51 kDa MukF, bind MukB (Yamazoe et al., 1999) via interaction with the head domains of MukB homodimers (Woo et al., 2009). In relation to the kleisin subunit of

eukaryotic condensins (Fennell-Fezzie et al., 2005), MukF dimers link a MukB pair to form a V-shaped closed ring (Woo et al., 2009) that can clamp DNA, potentially bringing distant chromosomal segments together (Cui et al., 2008). Intriguingly, this interaction could also join multiple MukBEF complexes into larger assemblies. The ATP-dependent interaction between MukB and F (Woo et al., 2009) might regulate reversible opening and closing of the ring entry gate for capture and release of DNA strands (Hirano and Hirano, 2004).

Although structural similarities with eukaryotic SMC complexes suggest related function of the bacterial relatives, the precise role and composition of the SMC complex in *E.coli* remains unclear. Fluorescence microscopy showed that MukBEF complexes form clusters *in vivo* (Ohsumi et al., 2001), that colocalise with *oriC* during replication (Danilova et al., 2007). Moreover, *mukB* mutants lost faithful positioning of *oriC* at mid-cell (Danilova et al., 2007). By counting fluorescent MukB, E, F, Anjana Badrinarayanan and Rodrigo Reyes-Lamothe (Oxford Biochemistry) established a minimal stoichiometric unit of 4:4:2 MukB:E:F molecules, and that 8-10 of these units accumulate within elongated spots (Badrinarayanan et al., 2012b). Diffraction-limited microscopy, however, could not resolve the underlying organisation of MukBEF within these spots. To address this problem using single-particle tracking PALM, they prepared *E.coli* AB1157 strains with PAmCherry fusions to MukB, E, and F as described in chapter 4.4. I contributed to single-particle tracking PALM data acquisition and performed the data analysis.

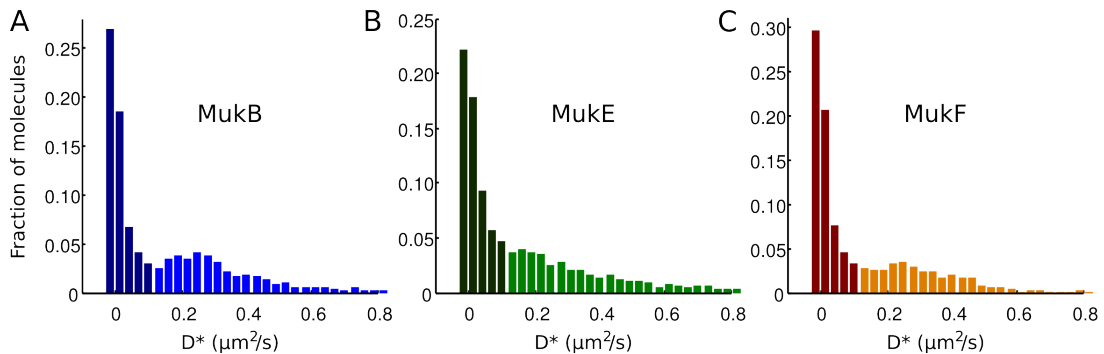


Figure 6.5: Distributions of the apparent diffusion coefficient (D^*) for (A) MukB; (B) MukE; (C) MukF. Populations of bound molecules are shown in darker colour.

6.5 Imaging MukBEF proteins

PALM imaging of MukB, E, and F proteins fused to PAmCherry in live *E. coli* showed slowly diffusing as well as bound molecules at 50 ms temporal resolution². All three components gave very similar D^* distributions that classified $\sim 50\%$ of the molecules as bound (Fig. 6.5). The diffusing populations had comparatively low apparent diffusion coefficients $D^* \sim 0.25 \mu\text{m}^2\text{s}^{-1}$. MSD curves for the diffusing populations of MukB, E, and F had a linear slope that indicates ordinary Brownian motion (Fig. 6.6). These curves were essentially identical despite the large differences in the molecular weights of MukB, E, and F. Individual MukE and MukF components with predicted higher diffusion coefficients were not observed in the D^* distributions (Fig. 6.5). These results suggest that stoichiometric MukBEF complexes exist fully assembled in solution.

Next, I analysed the spatial distribution of MukBEF within cells. Accumulated PAmCherry fluorescence showed elongated spots consistent with previous observations. While tracks of diffusing MukB, E, and F were randomly spread in cells (Fig. 6.7A), bound molecules localised within the diffraction-

²Slower diffusion of MukB, E, F permitted longer exposure time. Apart from this, sample preparation and imaging conditions were as described in chapter 4.4.

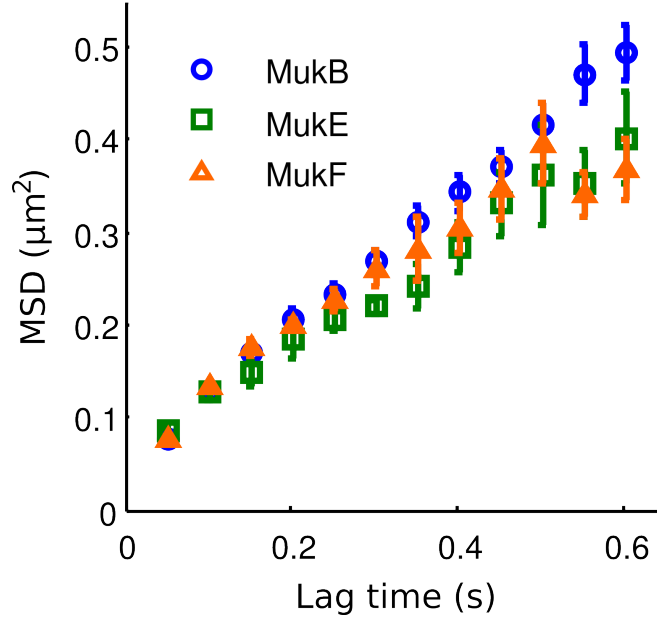


Figure 6.6: Mean-squared displacement (MSD) versus lag time for diffusing MukB (blue circles), MukE (green squares), MukF (orange triangles). The curves were averaged over all tracks classified as “unbound” according to the D^* distributions in Fig. 6.5; ($\pm$ sem).

limited spots (Fig. 6.7B). To evaluate MukBEF cluster formation, I implemented an algorithm that assigned the mean localisations of bound molecules to the same cluster if their separation is less than a distance threshold (here: $5 \times \sigma_{\text{loc}} = 200$ nm). There were typically 1-3 main clusters per cell at approximately mid-cell or quarter-cell positions and additional small clusters or individual immobile molecules in their periphery. The increased PALM resolution revealed that the diffraction-limited elongated spots were composed of several isolated sub-clusters, separated by > 100 nm (Fig. 6.7C). Clustering analysis (distance threshold $2 \cdot \sigma_{\text{loc}} = 80$ nm) showed that these sub-clusters contained tightly packed MukB, E, and F molecules. The distribution of localisations in sub-clusters was within the super-resolution limit, i.e. entirely consistent with scattering solely due to the localisation error for all three components (Fig. 6.8). MukBEF sub-clusters were hence compacted to < 40 nm.

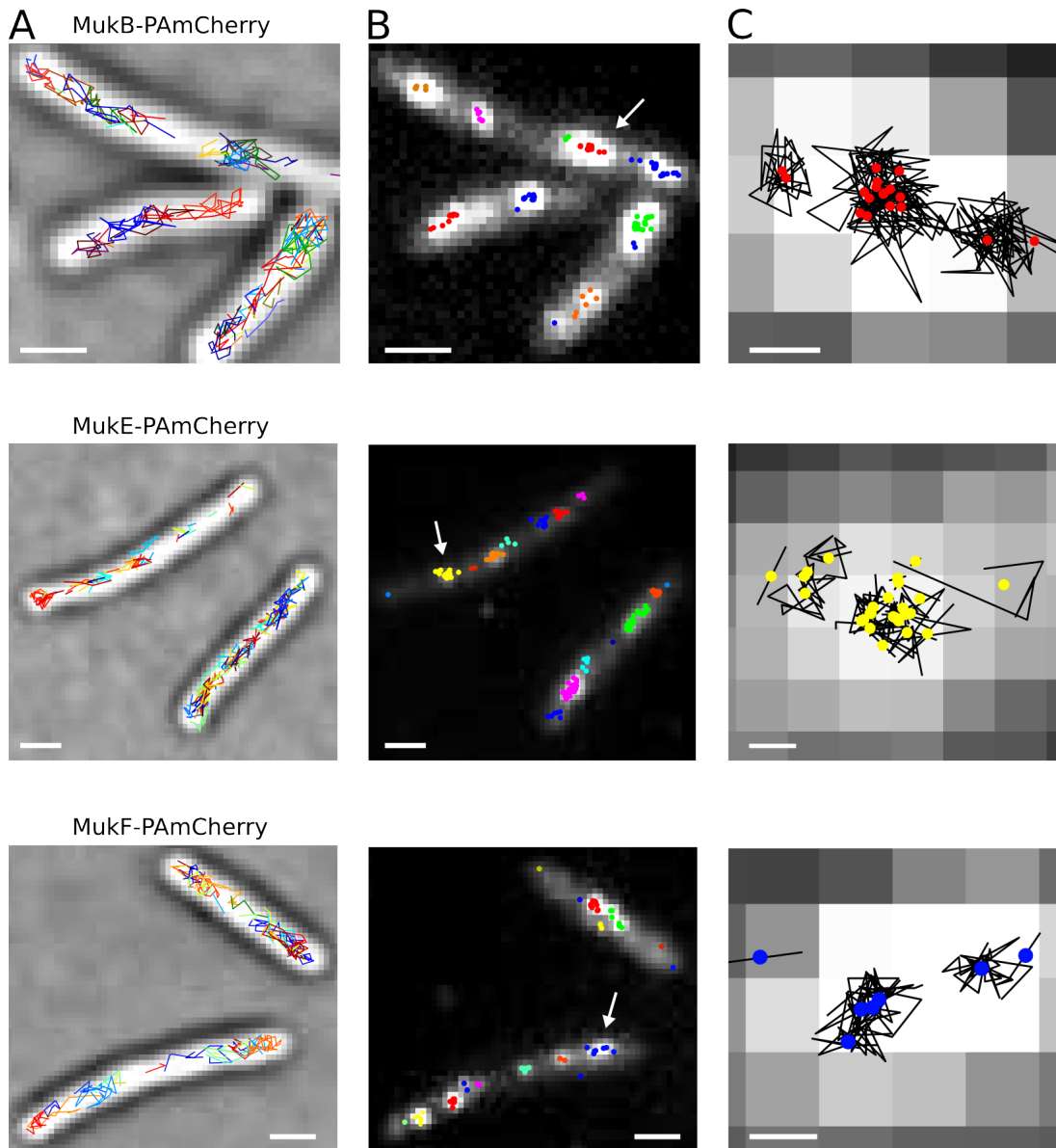


Figure 6.7: Imaging MukB, E, F in live *E. coli*. (A) Tracks of diffusing MukB, E, F in random colours on transmitted light images. (B) Mean localisations of bound MukB, E, F on accumulated fluorescence images. Localisations belonging to the same cluster share colours. Scale bars: 1 μm . (C) Magnified views showing MukB, E, F subclusters inside the clusters that are marked by an arrow in (B). Scale bars: 100 nm.

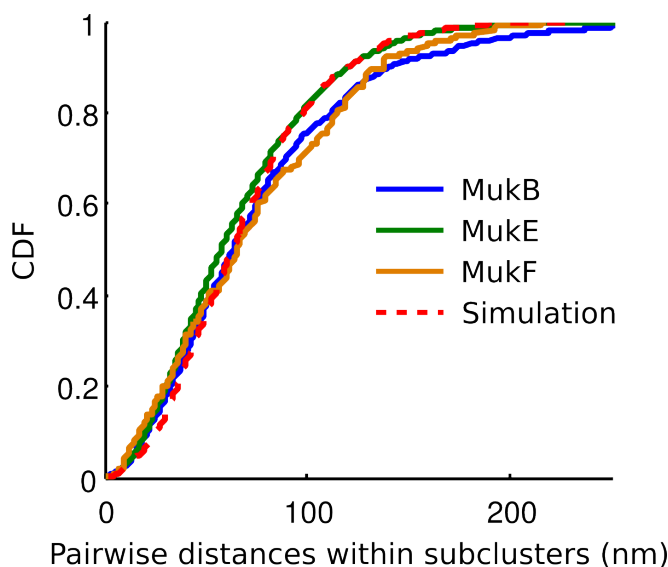


Figure 6.8: Cumulative distribution function of the distances between localisations within subclusters of MukB (blue), MukE (green), MukF (orange). The simulated distribution of distances that arises solely from scattering of localisations due to the localisation error is also shown (red dotted line).

FRAP analysis showed that single bound MukBEF complexes have dwell times of about 50 s within clusters (Badrinarayanan et al., 2012b). I measured an on-time of 2.16 s for bound MukE (Fig. 6.9), using 1 s exposure time and 350 μ W excitation intensity and identifying bound molecules by their PSF width as described in chapter 5.4. Because MukE unbinding occurs on a much longer time scale according to the FRAP analysis, this on-time reflects the photobleaching-time of PAmCherry in agreement with the previously measured value for Pol-PAmCherry in fixed cells (Table 5.1).

Discussion

My findings open the question about the functional role of MukBEF sub-clusters. Are they intermediates in the dynamic assembly or disassembly of larger MukBEF complexes that drag together different segments of the chromosome? Or are they static features of a larger complex with a structural connection yet to be discovered? Sub-clusters potentially correspond to multiple binding sites within

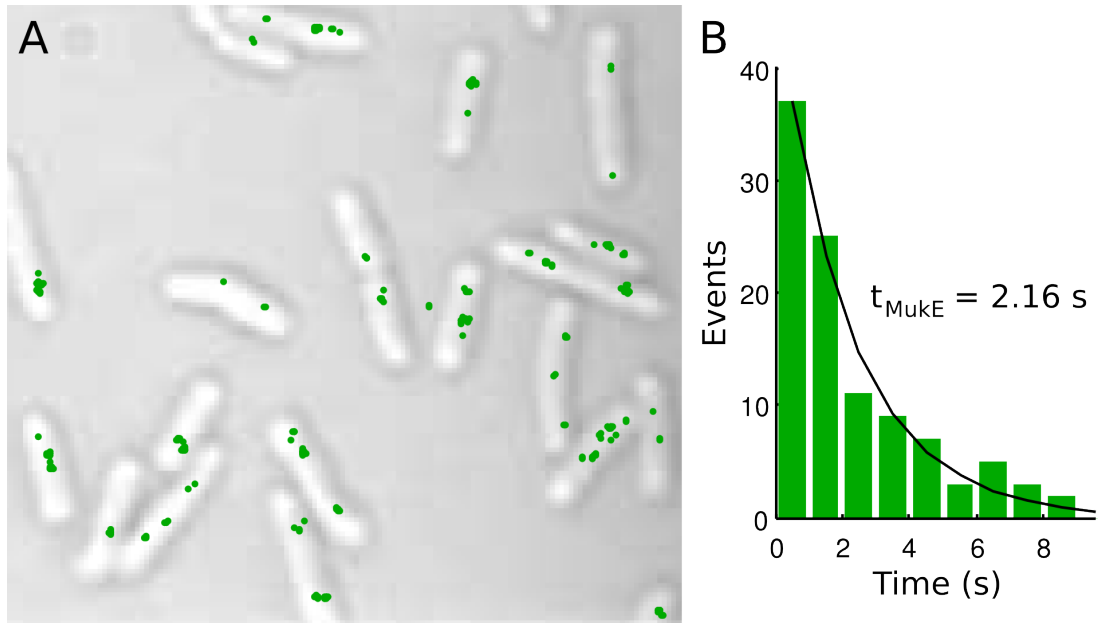


Figure 6.9: MukE binding times at 1 s exposure time and $350 \mu\text{W}$ excitation intensity. **(A)** Localisations of bound MukE (green circles). **(B)** Distribution of the observed on-times for bound MukE (green bars) and single exponential fit with time constant $t_{\text{MukE}} = 2.16 \text{ s}$.

the broad *ori* region, analogous to recruitment of the *Bacillus subtilis* SMC complex to eight *ori*-proximal sites by the chromosome partitioning protein SpoOJ (Sullivan et al., 2009). Two models for the function of MukBEF in folding and separating sister chromosomes have been discussed (Graumann and Knust, 2009). Because of the association with *oriC*, it seems plausible that MukBEF might associate with newly replicated DNA at replication forks and connect DNA loops by forming larger MukBEF complexes. However, MukBEF foci formation was shown to be independent of active DNA replication (Badrinarayanan et al., 2012a). Alternatively, MukBEF might load onto DNA at various positions on the chromosome and connect DNA loops by random collisions of MukBEF. The presence of sub-clusters within larger spots as well as isolated clusters outside the main spots disfavour models that involve restriction of MukBEF to a single chromosomal locus or cellular position. Instead, stoichiometric MukBEF complexes perform Brownian motion in cells and appear to be binding DNA initially

at different locations. These clustering “seeds” could capture and release DNA strands in a stepping manner via sequential opening and closing of the two rings of the double MukB dimer (Badrinarayanan et al., 2012b). These minimal functional units would then join into larger assemblies at roughly defined positions, rather than DNA being recruited into pre-formed large complexes. Time-lapse super-resolution imaging will help to elucidate whether sub-clusters are a consequence of dynamic complex assembly. Tracking analysis of individual diffusing MukBEF complexes should further inform about the mechanics of the search process for potential target sites and cluster formation. While ATP binding is essential for SMC complex formation (Arumugam et al., 2003), MukB mutants impaired in ATP hydrolysis still form clusters in cells but do not allow DNA unbinding (Badrinarayanan et al., 2012b). In light of these results, it will be informative to investigate the effect of these mutations on the diffusing complexes to test whether the search and cluster assembly processes involve DNA binding and require ATP turnover. This might give a clue as to whether the excess diffusing molecules merely increase the rate of cluster formation, or if their transient DNA interactions contribute to general chromosome organisation.

Chapter 7

Conclusion and Outlook

The development of experimental methods that allow observation and manipulation of individual molecules has lead to a more direct and intuitive understanding of molecular biology. In this thesis, I discussed further developments and applications of single-molecule fluorescence microscopy techniques that probe the structure, localisation, movement, and interactions of biomolecules. I worked on two complementary methods. Single-molecule FRET measurements probe nanometre distances within the molecular frame of reference, providing a direct readout of biomolecular structure and dynamics. Super-resolution microscopy, on the other hand, records the precise localisations of molecules within cells.

I used photoswitching to measure different FRET pairs sequentially, thereby extending single-molecule FRET for the measurement of multiple distances per molecule. Switchable FRET should be generally applicable to any surface immobilised-molecule that can be specifically labelled with fluorophores. Protein-DNA interactions are particularly suited because labeling and immobilisation techniques for synthetic DNA constructs are well established.

Photoswitching is also key for super-resolution microscopy to localise and track non-overlapping emitters in a densely labelled sample. I described a custom

TIRF microscope for STORM and PALM imaging. I implemented data analysis algorithms and computer simulations to move from qualitative observations towards more rigorous and quantitative analysis of microscopy data. To this end, I measured diffusion coefficients, counted protein copy numbers, and classified localisation clusters. These efforts might help to generate quantitative models of biological processes that uncover the underlying fundamental laws of nature.

The final chapters of this thesis focused on three applications that demonstrated the performance and versatility of super-resolution microscopy and gave new insight into the molecular organisation of reaction pathways and multi-protein complexes. Protein-DNA interactions are ideal applications because DNA binding results in a clear change of the protein diffusion coefficient. Generally, it should be possible to quantify *in vivo* reaction rates for all enzymes that significantly change diffusion characteristics upon substrate binding and product release.

The applications also pointed out technical and conceptional limitations. Foremost, photobleaching ultimately limits the localisation accuracy, diffusion coefficient estimation, and hence distinction between diffusing and bound molecules. Measurement of binding times that last longer than a few seconds was achieved by time-lapse imaging only at the sacrifice of temporal resolution. Mechanistic insight into the dynamic composition of multi-protein complexes like the replisome could be obtained by dual-colour tracking of two protein species and visualising their simultaneous or sequential turnover. However, the sequential observation concept of PALM means that only one or few molecules per cell are observed at any time, so that colocalising two particular molecules while they interact or exchange becomes statistically highly unlikely.

There is hence a need for further development of *in vivo* single-molecule imaging techniques that simultaneously monitor multiple molecular species over extended

periods of time. Engineering novel fluorophores with better photostability and optimised photoswitching characteristics is a key task towards this goal. Current microscopy techniques have been designed for the purpose of observation with minimal perturbation to the system; yet, understanding molecular mechanisms *in vivo* beyond mere observation requires measuring responses upon specific genetic or biochemical alterations. Much potential lies in new methods that link biochemical approaches with super-resolution microscopy to perturb and measure cellular responses simultaneously. These efforts can build on achievements in optogenetics (Pastrana, 2011) and *in vivo* nano-manipulation technologies (Oddershede, 2012).

Future combination of single-molecule FRET, super-resolution microscopy, and single-particle tracking has potential to reveal the structure and dynamics of molecules inside cells. Fluorescent proteins are not very useful for this task because of their large size and low intensity and it has been challenging to internalise proteins labelled with synthetic fluorophores or quantum dots. For this reason, Sako et al. (2000) measured FRET between Cy3 and Cy5 labelled epidermal growth factor molecules on the cell surface. Single-molecule FRET measurements inside cells have been achieved by microinjecting labelled proteins (Sakon and Weninger, 2010). My work on the combination of photoswitching and single-molecule FRET might prove useful for super-resolution FRET microscopy to resolve high concentrations of molecules and their FRET states by photoswitching (Grecco and Verveer, 2011).

Single-molecule technologies have current and potential future applications beyond fundamental research, as in DNA sequencing (Eid et al., 2009), lab-on-a-chip devices (Craighead, 2006), biosensing (Lymperopoulos et al., 2010), and drug screening (Skinner and Visscher, 2004), which underlines the timeliness and relevance of this research project.

Bibliography

- Aitken, C., Marshall, R., and Puglisi, J. (2008). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.*, 94:1826–1835.
- Allinson, S. L., Sleeth, K. M., Matthewman, G. E., and Dianov, G. L. (2004). Orchestration of base excision repair by controlling the rates of enzymatic activities. *DNA Repair*, 3:23–31.
- Altman, R. B., Terry, D. S., Zhou, Z., Zheng, Q., Geggier, P., Kolster, R. A., Zhao, Y., Javitch, J. A., Warren, J. D., and Blanchard, S. C. (2011). Cyanine fluorophore derivatives with enhanced photostability. *Nature Meth.*, 9:68–71.
- Andrecka, J., Lewis, R., Brückner, F., Lehmann, E., Cramer, P., and Michaelis, J. (2008). Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, 105:135–140.
- Arumugam, P., Gruber, S., Tanaka, K., Haering, C. H., Mechtler, K., and Nasmyth, K. (2003). ATP hydrolysis is required for cohesin’s association with chromosomes. *Curr. Biol.*, 13:1941–1953.
- Axelrod, D. (2001). Total internal reflection fluorescence microscopy in cell biology. *Traffic*, 2:764–774.
- Babcock, H., Sigal, Y. M., and Zhuang, X. (2012). A high-density 3D localization algorithm for stochastic optical reconstruction microscopy. *Opt. Nanoscopy*, 1:1–6.
- Badrinarayanan, A., Lesterlin, C., Reyes-Lamothe, R., and Sherratt, D. J. (2012a). The Escherichia coli SMC complex, MukBEF, shapes nucleoid organization independently of DNA replication. *J. Bacteriol.*, 194:4669–4676.
- Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M. C., and Sherratt, D. J. (2012b). In vivo architecture and action of bacterial structural maintenance of chromosome proteins. *Science (in print)*.
- Balci, H., Arslan, S., Myong, S., Lohman, T. M., and Ha, T. (2011). Single-molecule nanopositioning: structural transitions of a helicase-DNA complex during ATP hydrolysis. *Biophys. J.*, 101:976–984.

- Ball, C. A., Osuna, R., Ferguson, K. C., and Johnson, R. C. (1992). Dramatic changes in fis levels upon nutrient upshift in *Escherichia coli*. *J. Bacteriol.*, 174:8043–8056.
- Bates, M., Blosser, T., and Zhuang, X. (2005). Short-range spectroscopic ruler based on a single-molecule optical switch. *Phys. Rev. Lett.*, 94:101–108.
- Bates, M., Huang, B., Dempsey, G., and Zhuang, X. (2007). Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*, 317:1749–1753.
- Benke, A. and Manley, S. (2012). Live-cell dSTORM of cellular DNA based on direct DNA labeling. *ChemBioChem*, 13:298–301.
- Benkovic, S. J., Valentine, A. M., and Salinas, F. (2001). Replisome-mediated DNA replication. *Annu. Rev. Biochem.*, 70:181–208.
- Berezhna, S. Y., Gill, J. P., Lamichhane, R., and Millar, D. P. (2012). Single-molecule förster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase I. *J. Am. Chem. Soc.*, 134:11261–11268.
- Berlatzky, I. A., Rouvinski, A., and Ben-Yehuda, S. (2008). Spatial organization of a replicating bacterial chromosome. *Proc. Natl. Acad. Sci. U.S.A.*, 105:14136–14140.
- Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313:1642–1645. 090.
- Biteen, J. S., Thompson, M. A., Tselentis, N. K., Bowman, G. R., Shapiro, L., and Moerner, W. E. (2008). Super-resolution imaging in live *Caulobacter crescentus* cells using photoswitchable EYFP. *Nature Meth.*, 5:947–949.
- Bowers, P. M., Strauss, C. E., and Baker, D. (2000). De novo protein structure determination using sparse NMR data. *J. Biomol. NMR*, 18:311–318.
- Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122:947–956.
- Bronson, J. E., Fei, J., Hofman, J. M., Gonzalez, R. L., and Wiggins, C. H. (2009). Learning rates and states from biophysical time series: a bayesian approach to model selection and single-molecule FRET data. *Biophys. J.*, 97:3196–3205.

- Choi, P. J., Cai, L., Frieda, K., and Xie, X. S. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*, 322:442–446.
- Christian, T. D., Romano, L. J., and Rueda, D. (2009). Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc. Natl. Acad. Sci. U.S.A.*, 106:21109–21114.
- Chung, H., Louis, J., and Eaton, W. (2009). Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc. Natl. Acad. Sci. U.S.A.*, 106:11837–11844.
- Clamme, J.-P. and Deniz, A. (2005). Three-color single-molecule fluorescence resonance energy transfer. *ChemPhysChem*, 6:74–77.
- Clegg, R. (1992). Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.*, 211:353–388.
- Cook, P. R. (1999). The organization of replication and transcription. *Science*, 284:1790–1795.
- Cordes, T., Vogelsang, J., and Tinnefeld, P. (2009). On the mechanism of Trolox as antiblinking and antibleaching reagent. *J. Am. Chem. Soc.*, 131:5018–5019.
- Cox, M. M. (1998). A broadening view of recombinational DNA repair in bacteria. *Genes Cells*, 3(2):65–78.
- Craighead, H. (2006). Future lab-on-a-chip technologies for interrogating individual molecules. *Nature*, 442:387–393.
- Crocker, J. and Grier, D. (1996). Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.*, 179:298–310.
- Cui, Y., Petrushenko, Z. M., and Rybenkov, V. V. (2008). MukB acts as a macromolecular clamp in DNA condensation. *Nature Struct. Mol. Biol.*, 15:411–418.
- Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D. J., and Possoz, C. (2007). MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. *Mol. Microbiol.*, 65:1485–1492.
- Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.*, 97:6640–6645.
- Deich, J., Judd, E. M., McAdams, H. H., and Moerner, W. E. (2004). Visualization of the movement of single histidine kinase molecules in live Caulobacter cells. *Proc. Natl. Acad. Sci. U.S.A.*, 101:15921–15926.

- Dempsey, G. T., Bates, M., Kowtoniuk, W. E., Liu, D. R., Tsien, R. Y., and Zhuang, X. (2009). Photoswitching mechanism of cyanine dyes. *J. Am. Chem. Soc.*, 131:18192–18193.
- Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M., and Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Meth.*, 8:1027–1036.
- Deniz, A., Dahan, M., Grunwell, J., Ha, T., Faulhaber, A., Chemla, D., Weiss, S., and Schultz, P. (1999). Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci. U.S.A.*, 96:3670–3675.
- Dianov, G. and Lindahl, T. (1994). Reconstitution of the DNA base excision-repair pathway. *Curr. Biol.*, 4:1069–1076.
- Dillon, S. C. and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Rev. Mol. Cell Biol.*, 8(3):185–195.
- Dinant, C., Luijsterburg, M. S., Höfer, T., von Bornstaedt, G., Vermeulen, W., Houtsmuller, A. B., and van Driel, R. (2009). Assembly of multiprotein complexes that control genome function. *J. Cell Biol.*, 185:21–26.
- Dion, V., Kalck, V., Horigome, C., Towbin, B. D., and Gasser, S. M. (2012). Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nature Cell Biol.*, 14:502–509.
- Eggeling, C., Widengren, J., Brand, L., Schaffer, J., Felekyan, S., and Seidel, C. (2006). Analysis of photobleaching in single-molecule multicolor excitation and Förster resonance energy transfer measurements. *J. Phys. Chem. A*, 110(9):2979–2995.
- Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., A., B., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., deWinter, A., Dixon, A., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., and Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, 323:133–138.
- Elf, J., Li, G.-W., and Xie, X. S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, 316(5828):1191–1194.

- Elmore, S., Müller, M., Vischer, N., Odijk, T., and Woldringh, C. L. (2005). Single-particle tracking of oriC-GFP fluorescent spots during chromosome segregation in *Escherichia coli*. *J. Struct. Biol.*, 151:275–287.
- Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297:1183–1186.
- Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B., and Leibler, S. (1999). Protein mobility in the cytoplasm of *Escherichia coli*. *J. Bacteriol.*, 181:197–203.
- Endesfelder, U., Finan, K., Holden, S. J., Cook, P. R., Kapanidis, A. N., and Heilemann, M. (2012). Multi-scale spatial organization of RNA polymerase in bacteria. *in review*.
- English, B. P., Hauryliuk, V., Sanamrad, A., Tankov, S., Dekker, N. H., and Elf, J. (2011). Single-molecule investigations of the stringent response machinery in living bacterial cells. *Proc. Natl. Acad. Sci. U.S.A.*, 108:365–373.
- Evensen, G. and Seeberg, E. (1982). Adaptation to alkylation resistance involves the induction of a DNA glycosylase. *Nature*, 296:773–775.
- Fennell-Fezzie, R., Gradia, S. D., Akey, D., and Berger, J. M. (2005). The MukF subunit of *Escherichia coli* condensin: architecture and functional relationship to kleinsins. *EMBO J.*, 24:1921–1930.
- Feuerhahn, S. and Egly, J.-M. (2008). Tools to study DNA repair: what’s in the box? *Trends Genet.*, 24:467–474.
- Förster, T. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen d. Phys.*, 437:55–75.
- Frey, P. A. and Hegeman, A. D. (2007). *Enzymatic Reaction Mechanisms*. Oxford University Press, 1st edition.
- Friedberg, E. C. (2006). *DNA Repair and Mutagenesis*. American Society for Microbiology Press, 2nd edition.
- Friedman, N., Vardi, S., Ronen, M., Alon, U., and Stavans, J. (2005). Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biol.*, 3:e238.
- Fu, D., Calvo, J. A., and Samson, L. D. (2012). Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nature Rev. Cancer*, 12:104–120.
- Georgescu, R. E., Kurth, I., and O’Donnell, M. E. (2012). Single-molecule studies reveal the function of a third polymerase in the replisome. *Nature Struct. Mol. Biol.*, 19:113–116.

- Glassner, B. J., Rasmussen, L. J., Najarian, M. T., Posnick, L. M., and Samson, L. D. (1998). Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proc. Natl. Acad. Sci. U.S.A.*, 95:9997–10002.
- Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.*, 71:17–50.
- Graumann, P. and Knust, T. (2009). Dynamics of the bacterial SMC complex and SMC-like proteins involved in DNA repair. *Chromosome Res.*, 17:265–275.
- Grecco, H. E. and Verveer, P. J. (2011). FRET in cell biology: still shining in the age of super-resolution? *ChemPhysChem*, 12:484–490.
- Guest, C. R., Hochstrasser, R. A., Dupuy, C. G., Allen, D. J., Benkovic, S. J., and Millar, D. P. (1991). Interaction of DNA with the Klenow fragment of DNA polymerase I studied by time-resolved fluorescence spectroscopy. *Biochemistry*, 30:8759–8770.
- Ha, T., Enderle, T., Ogletree, D., Chemla, D., Selvin, P., and Weiss, S. (1996). Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. U.S.A.*, 93:6264–6268.
- Halford, S. E. and Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.*, 32(10):3040–3052.
- Hamdan, S. M., Loparo, J. J., Takahashi, M., Richardson, C. C., and van Oijen, A. M. (2009). Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature*, 457:336–339.
- Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E. G., Berg, O. G., and Elf, J. (2012). The lac repressor displays facilitated diffusion in living cells. *Science*, 336:1595–1598.
- Hastings, P. J., Hersh, M. N., Thornton, P. C., Fonville, N. C., Slack, A., Frisch, R. L., Ray, M. P., Harris, R. S., Leal, S. M., and Rosenberg, S. M. (2010). Competition of escherichia coli DNA polymerases I, II and III with DNA pol IV in stressed cells. *PLoS ONE*, 5:e10862.
- Heilemann, M., Hwang, L., Lymperopoulos, K., and Kapanidis, A. (2009a). Single-molecule FRET analysis of protein-DNA complexes. *Methods Mol. Biol.*, 543:503–521.
- Heilemann, M., Margeat, E., Kasper, R., Sauer, M., and Tinnefeld, P. (2005). Carbocyanine dyes as efficient reversible single-molecule optical switch. *J. Am. Chem. Soc.*, 127:3801–3806.

- Heilemann, M., van de Linde, S., Schüttelpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., and Sauer, M. (2008). Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew. Chem. Int. Ed.*, 47:6172–6176.
- Heilemann, M., vandeLinde, S., Mukherjee, A., and Sauer, M. (2009b). Super-resolution imaging with small organic fluorophores. *Angew. Chem. Int. Ed.*, 48:6903–6908.
- Hell, S. (2007). Far-field optical nanoscopy. *Science*, 316:1153–1158.
- Henriques, R., Lelek, M., Fornasiero, E. F., Valtorta, F., Zimmer, C., and Mhlanga, M. M. (2010). QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nature Meth.*, 7:339–340.
- Hirano, M. and Hirano, T. (2004). Positive and negative regulation of SMC-DNA interactions by ATP and accessory proteins. *EMBO J.*, 23:2664–2673.
- Hirano, T. (2002). The ABCs of SMC proteins: two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes Dev.*, 16:399–414.
- Hohlbein, J., Aigrain, L., Craggs, T. D., Bermek, O., Potapova, O., Shoolizadeh, P., Grindley, N. D. F., Joyce, C. M., and Kapanidis, A. N. (2012). Conformational landscapes of DNA polymerase ϵ and its mutator derivatives establish critical checkpoints for nucleotide insertion fidelity. *in preparation*.
- Hohlbein, J., Gryte, K., Heilemann, M., and Kapanidis, A. N. (2010). Surfing on a new wave of single-molecule fluorescence methods. *Phys. Biol.*, 7:031001.
- Hohng, S., Joo, C., and Ha, T. (2004). Single-molecule three-color FRET. *Biophys. J.*, 87:1328–1337.
- 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. *Biophys. J.*, 99:3102–3111.
- Holden, S. J., Uphoff, S., and Kapanidis, A. N. (2011). DAOSTORM: an algorithm for high-density super-resolution microscopy. *Nature Meth.*, 8:279–280.
- Houtsmuller, A. B., Rademakers, S., Nigg, A. L., Hoogstraten, D., Hoeijmakers, J. H., and Vermeulen, W. (1999). Action of DNA repair endonuclease ERCC1/XPF in living cells. *Science*, 284:958–961.
- Huang, B., Bates, M., and Zhuang, X. (2009). Super-resolution fluorescence microscopy. *Annu. Rev. Biochem.*, 78:993–1016.
- Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319:810–813.

- Huang, B., Wu, H., Bhaya, D., Grossman, A., Granier, S., Kobilka, B. K., and Zare, R. N. (2007). Counting low-copy number proteins in a single cell. *Science*, 315:81–84.
- Huh, D. and Paulsson, J. (2011). Non-genetic heterogeneity from stochastic partitioning at cell division. *Nature Genet.*, 43:95–100.
- Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J., and Frishman, D. (2008). Protein abundance profiling of the Escherichia coli cytosol. *BMC Genomics*, 9:102.
- Jackson, D. A., Balajee, A. S., Mullenders, L., and Cook, P. R. (1994). Sites in human nuclei where DNA damaged by ultraviolet light is repaired: visualization and localization relative to the nucleoskeleton. *J. Cell Sci.*, 107:1745–1752.
- Janion, C. (2008). Inducible SOS response system of DNA repair and mutagenesis in Escherichia coli. *Int. J. Biol. Sci.*, 4:338–344.
- Johnson, A. and O'Donnell, M. (2005). Cellular DNA replicases: components and dynamics at the replication fork. *Annu. Rev. Biochem.*, 74:283–315.
- Johnson, S. J., Taylor, J. S., and Beese, L. S. (2003). Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. U.S.A.*, 100:3895–3900.
- Kapanidis, A., Lee, N., Laurence, T., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U.S.A.*, 101:8936–8941.
- Kapanidis, A. N., Ebright, Y. W., Ludescher, R. D., Chan, S., and Ebright, R. H. (2001). Mean DNA bend angle and distribution of DNA bend angles in the CAP-DNA complex in solution. *J. Mol. Biol.*, 312:453–468.
- Kapanidis, A. N. and Strick, T. (2009). Biology, one molecule at a time. *Trends Biochem. Sci.*, 34:234–243.
- Karschau, J., de Almeida, C., Richard, M. C., Miller, S., Booth, I. R., Grebogi, C., and de Moura, A. P. S. (2011). A matter of life or death: modeling DNA damage and repair in bacteria. *Biophys. J.*, 100:814–821.
- Karymov, M., Daniel, D., Sankey, O. F., and Lyubchenko, Y. L. (2005). Holliday junction dynamics and branch migration: single-molecule analysis. *Proc. Natl. Acad. Sci. U.S.A.*, 102:8186–8191.
- Kelly, W. S. and Grindley, N. D. (1976). polA6, a mutation affecting the DNA binding capacity of DNA polymerase I. *Nucleic Acids Res.*, 3:2971–2984.

- Kim, Y., Eom, S. H., Wang, J., Lee, D.-S., Suh, S. W., and Steitz, T. A. (1995). Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature*, 376:612–616.
- Kinsella, T. J. (2009). Coordination of DNA mismatch repair and base excision repair processing of chemotherapy and radiation damage for targeting resistant cancers. *Clin. Cancer Res.*, 15:1853–1859.
- Kleibl, K. (2002). Molecular mechanisms of adaptive response to alkylating agents in *Escherichia coli* and some remarks on O(6)-methylguanine DNA-methyltransferase in other organisms. *Mutat. Res.*, 512:67–84.
- Konrad, E. B., Modrich, P., and Lehman, I. R. (1973). Genetic and enzymatic characterization of a conditional lethal mutant of *Escherichia coli* K12 with a temperature-sensitive DNA ligase. *J. Mol. Biol.*, 77:519–529.
- Kornberg, A. (2005). *DNA Replication*. University Science Books, 2nd edition.
- Kunkel, T. A. and Bebenek, K. (2000). DNA replication fidelity. *Annu. Rev. Biochem.*, 69:497–529.
- Kuzminov, A. (2001). Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc. Natl. Acad. Sci. U.S.A.*, 98:8241–8246.
- Lakowicz, J. (2006). *Principles of fluorescence spectroscopy*. Springer, 3rd edition.
- Lam, C., Mistri, T., Foo, Y., Sudhaharan, T., Gan, H., Rodda, D., Lim, L., Chou, C., Robson, P., Wohland, T., and Ahmed, S. (2012). DNA dependent Oct4-Sox2 interaction and diffusion properties characteristic of the pluripotent cell state revealed by fluorescence spectroscopy. *Biochem. J.* (*in print*).
- Landgraf, D., Okumus, B., Chien, P., Baker, T. A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nature Meth.*, 9:480–482.
- Lawson, C. L., Swigon, D., Murakami, K. S., Darst, S. A., Berman, H. M., and Ebright, R. H. (2004). Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.*, 14:10–20.
- Le Reste, L., Hohlbein, J., Gryte, K., and Kapanidis, A. (2012). Characterization of dark quencher chromophores as nonfluorescent acceptors for single-molecule FRET. *Biophys. J.*, 102:2658–2668.
- Lee, J., Lee, S., Ragunathan, K., Joo, C., Ha, T., and Hohng, S. (2010). Single-molecule four-color FRET. *Angew. Chem. Int. Ed.*, 49:9922–9925.

- Lee, N., Kapanidis, A., Koh, H., Korlann, Y., Ho, S., Kim, Y., Gassman, N., Kim, S., and Weiss, S. (2007). Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances. *Biophys. J.*, 92:303–312.
- Lee, N., Kapanidis, A., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebricht, R., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.*, 88:2939–2953.
- Lee, S.-H., Shin, J. Y., Lee, A., and Bustamante, C. (2012). Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc. Natl. Acad. Sci. U.S.A.*, page early edition.
- Lehman, I. R. (1974). DNA ligase: structure, mechanism, and function. *Science*, 186:790–797.
- Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958). Enzymatic synthesis of deoxyribonucleic acid. i. preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J. Biol. Chem.*, 233:163–170.
- Lestas, I., Vinnicombe, G., and Paulsson, J. (2010). Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467:174–178.
- Levin-Zaidman, S., Frenkiel-Krispin, D., Shimoni, E., Sabanay, I., Wolf, S. G., and Minsky, A. (2000). Ordered intracellular RecA-DNA assemblies: a potential site of in vivo RecA-mediated activities. *Proc. Natl. Acad. Sci. U.S.A.*, 97:6791–6796.
- Li, Y., Korolev, S., and Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.*, 17:7514–7525.
- Lia, G., Michel, B., and Allemand, J.-F. (2012). Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science*, 335:328–331.
- Lilley, D. M. J. (1997). All change at Holliday junction. *Proc. Natl. Acad. Sci. U.S.A.*, 94:9513–9515.
- Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, 362:709–715.
- Lippincott-Schwartz, J. and Patterson, G. H. (2009). Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging. *Trends Cell Biol.*, 19:555–565.

- Little, M. and Jones, N. (2010). Sparse bayesian step-filtering for high-throughput analysis of molecular machine dynamics. In *IEEE Acoust. Speech Signal Process.*
- Liu, S., Abbondanzieri, E., Rausch, J., Le Grice, S., and Zhuang, X. (2008). Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science*, 322:1092–1097.
- Ljungquist, S., Lindahl, T., and Howard-Flanders, P. (1976). Methyl methane sulfonate-sensitive mutant of *Escherichia coli* deficient in an endonuclease specific for apurinic sites in deoxyribonucleic acid. *J. Bacteriol.*, 126:646–653.
- Locke, J. C. W., Young, J. W., Fontes, M., Jiménez, M. J. H., and Elowitz, M. B. (2011). Stochastic pulse regulation in bacterial stress response. *Science*, 334:366–369.
- Loveland, A. B., Habuchi, S., Walter, J. C., and Oijen, A. M. v. (2012). A general approach to break the concentration barrier in single-molecule imaging. *Nature Meth.*, 9:987–992.
- Lucia, P. D. and Cairns, J. (1969). Isolation of an *E. coli* strain with a mutation affecting DNA polymerase. *Nature*, 224:1164–1166.
- Lundquist, R. C. and Olivera, B. M. (1982). Transient generation of displaced single-stranded DNA during nick translation. *Cell*, 31:53–60.
- Lymperopoulos, K., Crawford, R., Torella, J. P., Heilemann, M., Hwang, L. C., Holden, S. J., and Kapanidis, A. N. (2010). Single-molecule DNA biosensors for protein and ligand detection. *Angew. Chem. Int. Ed.*, 49:1316–1320.
- Manley, S., Gillette, J. M., Patterson, G. H., Shroff, H., Hess, H. F., Betzig, E., and Lippincott-Schwartz, J. (2008). High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nature Meth.*, 5:155–157.
- Mao, S., Benninger, R. K. P., Yan, Y., Petchprayoon, C., Jackson, D., Easley, C. J., Piston, D. W., and Marriott, G. (2008). Optical lock-in detection of FRET using synthetic and genetically encoded optical switches. *Biophys. J.*, 94:4515–4524.
- Margeat, E., Kapanidis, A., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R., and Weiss, S. (2006). Direct observation of abortive initiation and promoter escape withing single immobilized transcription complexes. *Biophys. J.*, 90:1419–1431.
- Markiewicz, R. P., Vrtis, K. B., Rueda, D., and Romano, L. J. (2012). Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucl. Acids Res.*, 40:7975–7984.

- McInerney, P., Johnson, A., Katz, F., and O'Donnell, M. (2007). Characterization of a triple DNA polymerase replisome. *Mol. Cell*, 27:527–538.
- McKinney, S. A., Declais, A.-C., Lilley, D. M., and Ha, T. (2003). Structural dynamics of individual holliday junctions. *Nature Struct. Mol. Biol.*, 10:93–97.
- McKinney, S. A., Joo, C., and Ha, T. (2006). Analysis of single-molecule FRET trajectories using hidden markov modeling. *Biophys. J.*, 91:1941–1951.
- Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A., Niu, W., Ebright, Y., Levy, R., and Ebright, R. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, 108:599–614.
- Michalet, X. and Berglund, A. J. (2012). Optimal diffusion coefficient estimation in single-particle tracking. *Phys. Rev. E*, 85:061916.
- Miné-Hattab, J. and Rothstein, R. (2012). Increased chromosome mobility facilitates homology search during recombination. *Nature Cell Biol.*, 14:510–517.
- 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 Meth.*, 7:377–381.
- Muschielok, A., Andrecka, J., Jawhari, A., Brückner, F., Cramer, P., and Michaelis, J. (2008). A nano-positioning system for macromolecular structural analysis. *Nature Meth.*, 5:965–971.
- Napoli, A. A., Lawson, C. L., Ebright, R. H., and Berman, H. M. (2006). Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: recognition of pyrimidine-purine and purine-purine steps. *J. Mol. Biol.*, 357:173–183.
- Nasmyth, K. and Haering, C. H. (2005). The structure and function of SMC and kleisin complexes. *Annu. Rev. Biochem.*, 74:595–648.
- Neher, E. and Sakmann, B. (1976). Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature*, 260:799–802.
- Nenninger, A., Mastroianni, G., and Mullineaux, C. W. (2010). Size dependence of protein diffusion in the cytoplasm of Escherichia coli. *J. Bacteriol.*, 192:4535–4540.
- Ni, M., Wang, S.-Y., Li, J.-K., and Ouyang, Q. (2007). Simulating the temporal modulation of inducible DNA damage response in Escherichia coli. *Biophys. J.*, 93:62–73.

- Niki, H., Jaffé, A., Imamura, R., Ogura, T., and Hiraga, S. (1991). The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. *EMBO J.*, 10:183–193.
- Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K. J. (1997). Direct observation of the rotation of F1-ATPase. *Nature*, 386:299–302.
- Oddershede, L. B. (2012). Force probing of individual molecules inside the living cell is now a reality. *Nature Chem. Biol.*, 8:879–886.
- Ohsumi, K., Yamazoe, M., and Hiraga, S. (2001). Different localization of SeqA-bound nascent DNA clusters and MukF-MukE-MukB complex in *Escherichia coli* cells. *Mol. Microbiol.*, 40:835–845.
- Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature*, 313:762–766.
- Pastrana, E. (2011). Optogenetics: controlling cell function with light. *Nature Meth.*, 8:24–25.
- Patel, P. H., Suzuki, M., Adman, E., Shinkai, A., and Loeb, L. A. (2001). Prokaryotic DNA polymerase I: evolution, structure, and "base flipping" mechanism for nucleotide selection. *J. Mol. Biol.*, 308:823–837.
- Patterson, G. H. and Lippincott-Schwartz, J. (2002). A photoactivatable GFP for selective photolabeling of proteins and cells. *Science*, 297:1873–1877.
- Pavani, S., Thompson, M., Biteen, J., Lord, S., Liu, N., Twieg, R., Piestun, R., and Moerner, W. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. U.S.A.*, 106:2995–2999.
- Perumal, S. K., Yue, H., Hu, Z., Spiering, M. M., and Benkovic, S. J. (2010). Single-molecule studies of DNA replisome function. *Biochim. Biophys. Acta*, 1804:1094–1112.
- Rasnik, I., McKinney, S., and Ha, T. (2006). Nonblinking and long-lasting single-molecule fluorescence imaging. *Nature Meth.*, 3:891–893.
- Reyes-Lamothe, R., Possoz, C., Danilova, O., and Sherratt, D. J. (2008a). Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell*, 133:90–102.
- Reyes-Lamothe, R., Sherratt, D. J., and Leake, M. C. (2010). Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science*, 328:498–501.

- Reyes-Lamothe, R., Wang, X., and Sherratt, D. J. (2008b). *Escherichia coli* and its chromosome. *Trends Microbiol.*, 16:238–245.
- Rizzino, A. (2009). Sox2 and Oct-3/4: a versatile pair of master regulators that orchestrate the self-renewal and pluripotency of embryonic stem cells by functioning as molecular rheostats. *Wiley Interdiscip. Rev. Syst. Biol. Med.*, 1:228–236.
- Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. *Nature Meth.*, 5:507–516.
- Rust, M., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Meth.*, 3:793–796.
- Sako, Y., Minoghchi, S., and Yanagida, T. (2000). Single-molecule imaging of EGFR signalling on the surface of living cells. *Nature Cell Biol.*, 2:168–172.
- Sakon, J. J. and Weninger, K. R. (2010). Detecting the conformation of individual proteins in live cells. *Nature Meth.*, 7:203–205.
- Samson, L. and Cairns, J. (1977). A new pathway for DNA repair in *Escherichia coli*. *Nature*, 267:281–283.
- 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. *Proc. Natl. Acad. Sci. U.S.A.*, 107:715–720.
- Saxton, M. J. and Jacobson, K. (1997). Single-particle tracking: applications to membrane dynamics. *Annu. Rev. Biophys. Biomol. Struct.*, 26:373–399.
- Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science*, 253:1001–1007.
- Schwartz, J. J. and Quake, S. R. (2009). Single molecule measurement of the "speed limit" of DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, 106:20294–20299.
- Sedgwick, B. (2004). Repairing DNA-methylation damage. *Nature Rev. Mol. Cell Biol.*, 5:148–157.
- Selvin, P. and Ha, T. (2008). *Single molecule techniques: A laboratory manual*. Cold Spring Harbor Laboratory Press, 1st edition.
- Selvin, P. R. (2000). The renaissance of fluorescence resonance energy transfer. *Nature Struct. Mol. Biol.*, 7:730–734.

- Senejani, A. G., Dalal, S., Liu, Y., Nottoli, T. P., McGrath, J. M., Clairmont, C. S., and Sweasy, J. B. (2012). Y265C DNA polymerase beta knockin mice survive past birth and accumulate base excision repair intermediate substrates. *Proc. Natl. Acad. Sci. U.S.A.*, 109:6632–6637.
- Shim, S.-H., Xia, C., Zhong, G., Babcock, H. P., Vaughan, J. C., Huang, B., Wang, X., Xu, C., Bi, G.-Q., and Zhuang, X. (2012). Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes. *Proc. Natl. Acad. Sci. U.S.A.*, 109:13978–13983.
- Skinner, G. M. and Visscher, K. (2004). Single-molecule techniques for drug discovery. *Assay Drug Dev. Tech.*, 2:397–405.
- Smith, B. T., Grossman, A. D., and Walker, G. C. (2002). Localization of UvrA and effect of DNA damage on the chromosome of *Bacillus subtilis*. *J. Bacteriol.*, 184:488–493.
- Smith, C. S., Joseph, N., Rieger, B., and Lidke, K. A. (2010). Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature Meth.*, 7:373–375.
- Sokhansanj, B. A., Rodrigue, G. R., Fitch, J. P., and Wilson, D. r. (2002). A quantitative model of human DNA base excision repair. i. mechanistic insights. *Nucleic Acids Res.*, 30:1817–1825.
- Stavans, J. and Oppenheim, A. (2006). DNA-protein interactions and bacterial chromosome architecture. *Phys. Biol.*, 3:R1–10.
- Stein, I. H., Steinhauer, C., and Tinnefeld, P. (2011). Single-molecule four-color FRET visualizes energy-transfer paths on DNA origami. *J. Am. Chem. Soc.*, 133:4193–4195.
- Steitz, T. A. and Yin, Y. W. (2004). Accuracy, lesion bypass, strand displacement and translocation by DNA polymerases. *Philos. Trans. R. Soc. London Ser. B*, 359:17–23.
- Stuart, J. A., Mayard, S., Hashiguchi, K., Souza-Pinto, N. C., and Bohr, V. A. (2005). Localization of mitochondrial DNA base excision repair to an inner membrane-associated particulate fraction. *Nucleic Acids Res.*, 33:3722–3732.
- Subach, F. V., Malashkevich, V. N., Zencheck, W. D., Xiao, H., Filonov, G. S., Almo, S. C., and Verkhusha, V. V. (2009a). Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. *Proc. Natl. Acad. Sci. U.S.A.*, 106:21097–21102.
- Subach, F. V., Patterson, G. H., Manley, S., Gillette, J. M., Lippincott-Schwartz, J., and Verkhusha, V. V. (2009b). Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature Meth.*, 6:153–159.

- Sullivan, N. L., Marquis, K. A., and Rudner, D. Z. (2009). Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell*, 137:697–707.
- Sung, J.-S. and Mosbaugh, D. W. (2003). Escherichia coli uracil- and ethencytosine-initiated base excision DNA repair: rate-limiting step and patch size distribution. *Biochemistry*, 42:4613–4625.
- Svoboda, K., Schmidt, C. F., Schnapp, B. J., and Block, S. M. (1993). Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, 365:721–727.
- Tafvizi, A., Huang, F., Fersht, A. R., Mirny, L. A., and van Oijen, A. M. (2011). A single-molecule characterization of p53 search on DNA. *Proc. Natl. Acad. Sci. U.S.A.*, 108(2):563–568.
- Tanner, N. A., Loparo, J. J., Hamdan, S. M., Jergic, S., Dixon, N. E., and van Oijen, A. M. (2009). Real-time single-molecule observation of rolling-circle DNA replication. *Nucl Acids Res*, 37:e27.
- Tanner, N. A., Tolun, G., Loparo, J. J., Jergic, S., Griffith, J. D., Dixon, N. E., and van Oijen, A. M. (2011). E. coli DNA replication in the absence of free beta clamps. *EMBO J.*, 30:1830–1840.
- Tawfik, D. S. (2010). Messy biology and the origins of evolutionary innovations. *Nature Chem. Biol.*, 6:692–696.
- Thompson, R., Larson, D., and Webb, W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.*, 82:2775–2783.
- Tokunaga, M., Imamoto, N., and Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Meth.*, 5:159–161.
- Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., and Ellenberger, T. (2006). DNA ligases: structure, reaction mechanism, and function. *Chem. Rev.*, 106:687–699.
- Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J., and Kapanidis, A. N. (2011). Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis. *Biophys. J.*, 100:1568–1577.
- Tsien, R. Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.*, 67:509–544.
- Uphoff, S. (2010). Probing molecular and cellular structure using single molecule photoswitching. Master’s thesis, University of Oxford.

- Uphoff, S., Gryte, K., Evans, G., and Kapanidis, A. N. (2011). Improved temporal resolution and linked hidden markov modeling for switchable single-molecule FRET. *ChemPhysChem*, 12:571–579.
- Uphoff, S., Holden, S. J., Le Reste, L., Periz, J., van de Linde, S., Heilemann, M., and Kapanidis, A. N. (2010). Monitoring multiple distances within a single molecule using switchable FRET. *Nature Meth.*, 7:831–836.
- Uphoff, S., Reyes-Lamothe, R., Garza de Leon, F., Sherratt, D. J., and Kapanidis, A. N. (2012). Single-molecule DNA repair in live bacteria. *in preparation*.
- Vale, R. D. (2008). Microscopes for fluorimeters: the era of single molecule measurements. *Cell*, 135(5):779–85.
- Vale, R. D., Funatsu, T., Pierce, D. W., Romberg, L., Harada, Y., and Yanagida, T. (1996). Direct observation of single kinesin molecules moving along microtubules. *Nature*, 380:451–453.
- van de Linde, S., Heilemann, M., and Sauer, M. (2012). Live-cell super-resolution imaging with synthetic fluorophores. *Annu. Rev. Phys. Chem.*, 63:519–540.
- van Oijen, A. M. and Loparo, J. J. (2010). Single-molecule studies of the replisome. *Annu. Rev. Biophys.*, 39:429–448.
- Vogelsang, J., Cordes, T., Forthmann, C., Steinhauer, C., and Tinnefeld, P. (2009). Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 106:8107–8112.
- 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. *Angew. Chem. Int. Ed.*, 47:5465–5469.
- Volkert, M. R. (1988). Adaptive response of Escherichia coli to alkylation damage. *Environ. Mol. Mutagen.*, 11:241–255.
- Wang, X., Liu, X., Possoz, C., and Sherratt, D. J. (2006). The two Escherichia coli chromosome arms locate to separate cell halves. *Genes Dev.*, 20:1727–1731.
- Wang, Y., Quan, T., Zeng, S., and Huang, Z.-L. (2012). PALMER: a method capable of parallel localization of multiple emitters for high-density localization microscopy. *Opt. Express*, 20:16039–16049.
- Watson, J. and Crick, F. (1953). A structure for deoxyribose nucleic acid. *Nature*, 171:737–738.
- Wilson, David M, r. and Bohr, V. A. (2007). The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair*, 6:544–559.

- Woldringh, C. L. and Nanninga, N. (2006). Structural and physical aspects of bacterial chromosome segregation. *J. Struct. Biol.*, 156:273–283.
- Woo, J.-S., Lim, J.-H., Shin, H.-C., Suh, M.-K., Ku, B., Lee, K.-H., Joo, K., Robinson, H., Lee, J., Park, S.-Y., Ha, N.-C., and Oh, B.-H. (2009). Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions. *Cell*, 136:85–96.
- Woźniak, A., Schröder, G., Grubmüller, H., Seidel, C., and Oesterhelt, F. (2008). Single-molecule FRET measures bends and kinks in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, 105:18337–18342.
- Wu, E. Y. and Beese, L. S. (2011). The structure of a high fidelity DNA polymerase bound to a mismatched nucleotide reveals an ajar intermediate conformation in the nucleotide selection mechanism. *J. Biol. Chem.*, 286:19758–19767.
- Xie, S. and Lu, P. (1999). Single-molecule enzymology. *J. Biol. Chem.*, 274:15967–15970.
- Xie, X. S., Choi, P. J., Li, G.-W., Lee, N. K., and Lia, G. (2008). Single-molecule approach to molecular biology in living bacterial cells. *Annu. Rev. Biophys.*, 37:417–444.
- Xu, Y., Derbyshire, V., Ng, K., Sun, X. C., Grindley, N. D., and Joyce, C. M. (1997). Biochemical and mutational studies of the 5′-3′ exonuclease of DNA polymerase I of *Escherichia coli*. *J. Mol. Biol.*, 268:284–302.
- Xu, Y., Grindley, N. D., and Joyce, C. M. (2000). Coordination between the polymerase and 5′-nuclease components of DNA polymerase I of *Escherichia coli*. *J. Biol. Chem.*, 275:20949–20955.
- Yamazoe, M., Onogi, T., Sunako, Y., Niki, H., Yamanaka, K., Ichimura, T., and Hiraga, S. (1999). Complex formation of MukB, MukE and MukF proteins involved in chromosome partitioning in *Escherichia coli*. *EMBO J.*, 18:5873–5884.
- Yildiz, A., Forkey, J., McKinney, S., Ha, T., Goldman, Y., and Selvin, P. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, 300:2061–2065.
- Yu, J., Xiao, J., Ren, X., Lao, K., and Xie, X. S. (2006). Probing gene expression in live cells, one protein molecule at a time. *Science*, 311:1600–1603.
- Zhu, L., Zhang, W., Elnatan, D., and Huang, B. (2012). Faster STORM using compressed sensing. *Nature Meth.*, 9:721–723.