Real-time single-molecule observations of conformational changes in DNA Polymerase

Geraint Wyn Evans
University College, Oxford

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

Trinity Term, 2013

Abstract

Genetic information is encoded in the long sequence of bases which form DNA, which is replicated during cell division by enzymes known as DNA Polymerases. Polymerases replicate DNA extremely accurately to avoid errors which can cause cell death and diseases such as cancer, although the mechanisms behind these extraordinary fidelities are not well understood.

A large conformational change in the protein, in which the "fingers" subdomain closes around an incoming nucleotide, is thought to be implicated in these fidelity mechanisms. Here we present an assay to monitor this conformational change in single polymerase molecules, in real-time. We achieve this using total-internal-reflection-fluorescence (TIRF) microscopy to monitor the fluorescence resonance energy transfer (FRET) of an intra-protein dye labeled DNA Polymerase I (KF) as it binds to surface-immobilised DNA.

Initially, we investigated the polymerase fingers-conformations during the pre-chemistry polymerisation reaction, resolving forward and backward rates which would be challenging to observe using ensemble techniques. These observations confirmed that KF closes rapidly around complementary nucleotide, but we discovered that the reverse step, fingers-opening, is particularly slow relative to chemistry. These finger kinetics act to remove the influence of the reaction rate-limiting step on fidelity, surprising given decades of investigations have focused on the rate-limiting step as the key determinant of fidelity. We also use our kinetic measurements to quantify contributions of different reaction steps to the macroscopic error rate of the polymerase.

Subsequently, we developed our assay to investigate the fingers-conformations across the entire DNA polymerisation reaction. We observed single-nucleotide incorporations, and processive DNA polymerisation at high and low nucleotide concentrations, which suggested heterogeneous nucleotide incorporation rates. The observations demonstrated that the post-chemistry slow step that limits processive polymerisation occurs before post-chemistry fingers-opening, or is accounted for by post-chemistry fingers-opening. We observe a correlation in turn-over kinetics and binary complex kinetics, suggesting that turn-over rates could be limited by the intrinsic dynamics of the binary complex, as seen in other protein systems, although more work is needed on this.

Disclaimer

All of the work in this thesis is my own, unless declared otherwise, with other contributions described below. The pronoun "we" is used in the place of "I" in this thesis for stylistic purposes, although the latter would be more accurate.

- Figures in Chapters 1 and 2 are adapted from lab commons, or from previously published images.
- The polymerase (KF) construct with an intra-protein FRET pair was synthesised and characterised by researchers at Yale University, a significant achievement without which none of this work could have been performed.
- The "KF6" buffer formulation used in these experiments was developed by Johannes Hohlbein (50mM Tris, pH 7.5; 100μg/ml BSA; 1mM DTT; 5% glycerol; 1mM Trolox (UV treated); 1.9mg/ml glucose oxidase;0.075mg/ml catalase; 1.5%w/v D-glucose; 10mM $MgCl_2$).
- The single-molecule TIRF microscope used in these investigations "Halfdome" was built by other members of the Kapanidis lab.
- The single-molecule confocal microscope used in these investigations "Snowdon" was built by Johannes Hohlbein and I. Subsequently I programmed the microscope to realise its scanning confocal capability. Although not documented in this thesis, the scanning confocal capability was used in the following publication:

  Uphoff, S., Gryte, K., Evans, G., & Kapanidis, A. N. (2011). Improved Temporal Resolution and Linked Hidden Markov Modeling for Switchable Single-Molecule FRET. *ChemPhysChem*, 12(3), 571-579.
- The surface passivation protocol was developed by previous group members and refined by Louise Aigrain.

Geraint Evans
October 2013

Acknowledgements

My thanks go to my parents and family, without whom I would never have reached a position to investigate the fascinating interface of physics and biology. Thanks to Prof. Achillefs Kapanidis, for allowing me to work in his lab, for providing ever insightful contributions and distilling the core essence of a problem, and for his invaluable tutorials in scientific communication.

Thanks to members of the "downstairs" office for moral support, particularly those who were there from the beginning, Justin, Stephan, Rob, Chetan and Lejla. Thanks to Dr. Johannes Hohlbein for the microscope building collaboration, providing the tips and tricks of practical optics and microscopy. Thanks to Dr. Timothy Craggs for his enthusiasm for all things polymerase, reading multiple thesis chapters and consistently providing interesting and useful discussions.

Many thanks to Alexandra Tomescu and Louise Aigrain for proof-reading. Additional thanks to Prof. Robin Nicholas, to all those whose prior work formed a basis for my own, and to anyone I've left out!

Contents

1 Introduction **1**

2 Single-molecule fluorescence **5**
2.1 Introduction . . . 5
2.2 Fluorescence . . . 5
2.3 FRET . . . 8
2.4 Single-molecule techniques . . . 10
2.5 Single-molecule fluorescence . . . 10
2.5.1 Single-molecule confocal microscopy . . . 12
2.5.2 TIRF microscopy . . . 15
2.5.3 Photophysics . . . 17

3 A primer on DNA polymerase **19**
3.1 Introduction . . . 19
3.2 A brief history of investigations into Klenow Fragment (KF) . . . 22
3.2.1 Kinetic investigations . . . 23
3.2.2 Crystal structures . . . 28
3.2.3 Bulk fluorescence investigations . . . 31
3.2.4 Single-molecule fluorescence investigations of KF . . . 34
3.3 Theories of DNA polymerase fidelity . . . 38
3.3.1 "A new paradigm for DNA polymerase specificity" - T7 DNA polymerase 41

4 An assay to monitor the conformational dynamics of DNA polymerase I (KF) **43**
4.1 Introduction . . . 43
4.2 Results and discussion . . . 46
4.2.1 An assay to monitor the conformational dynamics of Klenow Fragment 46
4.2.2 Transient immobilisation of KF . . . 49
4.2.3 Evidence for KF dimerisation on DNA . . . 52
4.2.4 Influence of Atto647N photo-physics on FRET efficiency . . . 56
4.2.5 Validation using a diffusion based assay . . . 59
4.3 Conclusions . . . 64
4.4 Materials and Methods . . . 64

5 Binary and ternary complex dynamics in DNA Polymerase I (KF) 67
5.1 Introduction . . . 67
5.2 Results and discussion . . . 70
5.2.1 Binary complex dynamics . . . 70
5.2.2 Ternary complex dynamics . . . 74
5.2.3 Mismatching nucleotide . . . 83
5.3 Materials and Methods . . . 86

6 An assay to monitor KF conformation during DNA polymerisation 88
6.1 Introduction . . . 88
6.2 Initiating DNA polymerisation on a TIRF microscope without microfluidics . 91
6.2.1 Challenges . . . 91
6.2.2 Method . . . 92
6.2.3 Results . . . 93
6.3 Diversity of single-molecule behaviours during DNA polymerisation . . . 96
6.3.1 Data representation challenges in the single-molecule field . . . 96
6.3.2 A novel visualisation representing hundreds of single-molecules simultaneously . . . 97
6.3.3 Classifying single-molecules into behavioural groups . . . 99
6.3.4 Behaviour of KF molecules before, during and after DNA polymerisation 101
6.3.5 KF fingers conformation during DNA polymerisation . . . 102
6.4 A complementary method for initiating DNA polymerisation without the use of microfluidics. . . . 104
6.4.1 Initiating DNA polymerisation using a low KF concentration . . . 105
6.5 Discussion . . . 107

7 KF fingers conformations during DNA polymerisation 110
7.1 Introduction . . . 110
7.2 Results: Processive DNA polymerisation at high nucleotide concentrations . 114
7.2.1 A partially extendable DNA substrate . . . 114
7.2.2 Partial DNA extension: Single-molecules . . . 115
7.2.3 Partial DNA extension: Single-molecule ensemble . . . 118
7.2.4 Rates of polymerisation by KF . . . 120
7.3 Results: Single-nucleotide incorporations . . . 125
7.4 Results: Processive DNA polymerisation at low nucleotide concentrations . . 130
7.5 Discussion . . . 136
7.6 Materials and Methods . . . 139

8 Concluding remarks 141
8.0.1 Pre-chemistry investigations . . . 141
8.0.2 On the role of large scale conformational changes more generally . . . 142
8.0.3 Investigating the entire DNA polymerisation pathway . . . 144

A Deriving an expression for the fingers-closing rate 148

B Single KF molecules translocating along DNA during polymerisation 150

Bibliography **156**

Chapter 1

Introduction

Enzymes are crucial for cellular activity, catalysing reactions which would rarely occur in their absence, enhancing reaction rates by factors of 10^{10} to 10^{29}.[1,2] Enzymes are thought to enhance reaction rates by binding and stabilising the reaction transition state, lowering the energy barrier which defines the reaction rate according to transition state theory. However, beyond this relatively broad statement, it remains unclear how rate enhancement is mechanistically achieved in enzymes.[3,4] Although there are a multitude of propositions, quantifying the precise contribution of various components is challenging.

Some argue that electrostatics dominate rate enhancement, with the dielectric constant of protein interior allowing charges to exert more significant forces than in solution. Other proposals centre around the contribution of entropic effects, in which substrate binding removes the entropic contribution required to align and proximate the substrates.[5,6] There have also been discussions on the contribution of hydrophilic effects, and quantum tunnelling.[7] Of recent, the dynamic aspects of proteins have received much attention as techniques allowing their observation have matured.

Crystal structures have demonstrated that proteins can adopt a variety of structures to achieve their function, and the noisy thermal environment allows proteins, which are inherently flexible, to interconvert between a range of conformations. These structures can be abstracted in an energy landscape, in which the energy levels of the conformations define their equilibrium populations via thermodynamics, and the energy barriers between the states

define their interconversion rates (Fig 1.1). Energy landscapes of proteins were first discussed when conformational substates were inferred by the binding characteristics of myoglobin to carbon monoxide and oxygen under different temperature conditions, which elucidated different energy barriers in the binding process.[8,9]

The great successes of crystallography in providing the static structures of enzymes have been followed by the development of techniques which can add a time component to these structures. Protein dynamics have been observed at fast time scales (ps-ns) using NMR, demonstrating how the structural flexibility of calmodulin relates to its ability to bind many different molecules.[10] In addition, rapid correlated side-chain motions during binding of carbon monoxide to myoglobin have observed using Laue X-ray diffraction[11], and investigations of the temperature dependence of kinetic isotope exchange experiments have suggested that molecular oscillations could enhance quantum tunnelling reaction rates.[3] Protein dynamics can also exist at much slower μs-ms time-scales, such as observations that the intrinsic dynamics of cyclophilin A corresponds with it's turnover rate[12], and similarly, that the turn over rate of adelynate kinase is limited by its lid opening rate.[13] While there has been significant interest recently in the role that dynamical fluctuations play in enzyme function, the use of the phrase "protein dynamics" has led to confusions as the phrase has been used to describe dynamics over a broad range of time-scales and behaviours.

The time evolution of enzymatic reactions can be observed using computer simulations to create atomistic "movies" based on crystal structures, but these are generally limited to very short time-scales. While short time-scales are relevant to the motions and fluctuations of atoms, electrons, rotations of protein residues and the chemical event, enzymatic turnovers frequently occur on longer time-scales, up to seconds and milliseconds, which are difficult to simulate computationally. In this thesis, we are interested in conformational dynamics at these slower millisecond time-scales, of interest because the time-scale of these dynamics often coincide with the enzymatic turn-over rate, whilst the speed of the chemical reaction its self occurs at a sub-picosecond time-scale.[14] Slow fluctuations have been termed tier-0 dynamics, defined by conformational changes which are limited by energy barriers of a few kT.

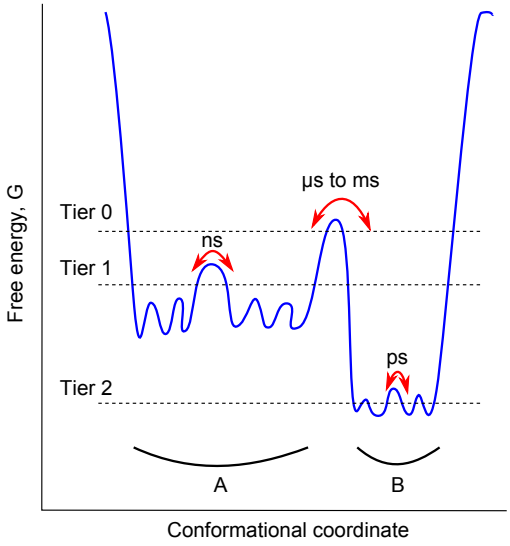


Figure 1.1: Schematic illustrating the different energies of a protein's structures along a particular conformational coordinate. The occupancies of state A and B are defined by the difference in energy between the two states, while the interconversion rate between the states is determined by the energy barriers between them, defining the time-scales of interconversions.

Single-molecule techniques have presented an opportunity to observe the behaviour of enzymes at the time-scale of reaction turn-over, inaccessible to computer simulations. Single-molecule enzymology also allow observation of phenomena which would otherwise be shrouded in ensemble averages, such as molecular sub-populations, dynamical fluctuations, and observations of individual steps in a reaction, as well as transient intermediates which would otherwise not be observed, such as observations of the 120° rotations steps in F_1-ATPase[15], or the transient folding intermediates in a ribozyme.[16] Single-molecule investigations have revealed phenomena such as dynamic disorder in cholesterol oxidase, in which reaction rates demonstrate a "memory effect", fluctuating over time. These fluctuations were attributed to slow interconversions of the enzyme conformation, thus highlighting the importance of slow time-scale conformational fluctuations in enzymes.[17,18,14]

In this thesis we investigate the enzyme DNA polymerase I, and how the motion of its sub-domains contributes to its ability to replicate and repair DNA extremely accurately, minimizing the replication errors and DNA mutations which can lead to diseases. Specifically we investigate how a large scale structural change in DNA polymerase ("fingers-closing") contributes kinetically to its reaction mechanism, which we observe using single-molecule

techniques, the principles of which we describe in Chapter 2. We describe DNA polymerase and the experiments which form the basis of its current understanding in Chapter 3.

Brief observations of the conformation of single polymerase molecules were achieved in previous work, using a dye pair on the "fingers" and "thumb" of the polymerase, which inferred protein conformation using fluorescence resonance energy transfer (FRET) and diffusion-based single-molecule techniques.[19,20] Using this protein construct, we devise a single-molecule assay based on Total Internal Reflection Fluorescence (TIRF) microscopy, which allows extended ($> 1s$) observations of the polymerase's conformation. We present and characterise this assay in Chapter 4. In Chapter 5 we observe the polymerase fingers-conformations of binary (KF-DNA) and ternary (KF-DNA-dNTP) complexes, with inhibition of the polymerisation reaction allowing investigation of contribution of the fingers-dynamics to the pre-chemistry reaction. We then analyse how these fingers-closing kinetics contribute to the polymerase reaction mechanism and fidelity.

In Chapter 6 we present and characterise an extension to our assay which allows observation the fingers-conformation across the rapid DNA polymerisation reaction, providing observation of the polymerase's conformational dynamics in the presence of chemistry. In Chapter 7 we investigate the fingers-conformations during DNA polymerisation in more detail, and follow with concluding remarks in Chapter 8.

This thesis presents the application of single-molecule fluorescence techniques to the investigation of the kinetic contribution of a specific structural change to a protein's functioning.

Chapter 2

Single-molecule fluorescence

2.1 Introduction

This thesis is concerned with how structural changes in the enzyme DNA polymerase I contribute to its DNA replication mechanisms. We use fluorescence techniques both to localise single polymerase molecules and to infer conformational changes via energy transfer between two dye molecules attached to the protein. In this chapter, we outline the concepts of fluorescence, fluorescence/Förster resonance energy transfer (FRET), and the techniques used to observe single-molecules.

2.2 Fluorescence

Fluorescence is the photon emission which occurs when a molecule relaxes radiatively from a higher energy state to a lower energy state. The higher energy state is typically occupied as a result of photon absorption, whose wavelength is usually shorter than the emitted radiation. The wavelength distinction between absorbed and emitted radiation allows fluorescence to be used for highly sensitive detection.[21]

The phenomenon of fluorescence is described well by a Jablonski diagram, which illustrates the electronic states of a molecule at distinct energy levels, which are further split into vibrational sub-levels (Fig. 2.1a, energy levels at S_0 and S_1). For instance, absorption of a photon can excite an electron from the electronic state S_0 to a vibrational energy level

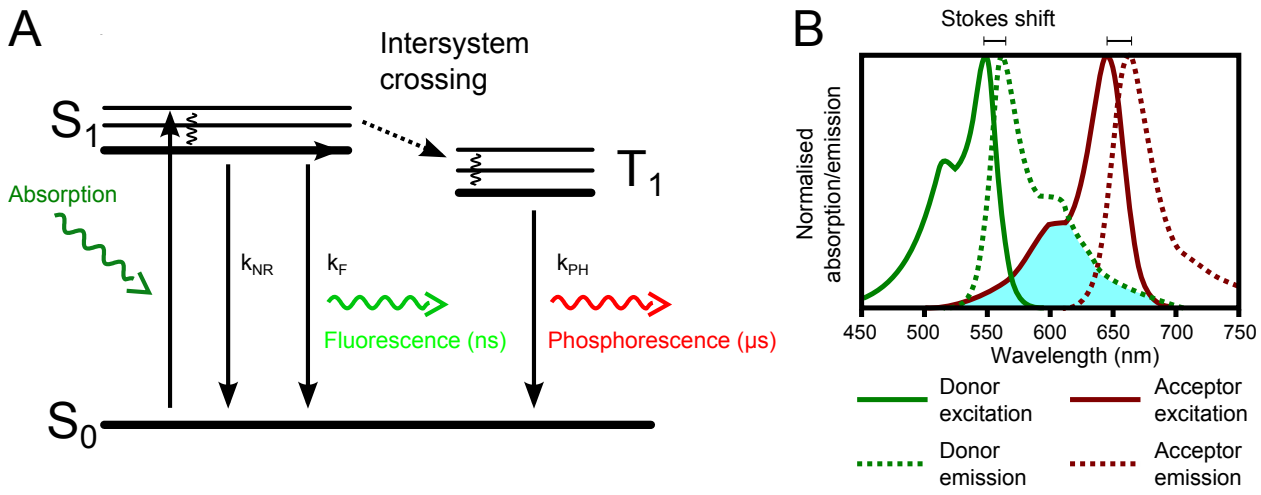


Figure 2.1: (a) Jablonski diagram illustrating the electronic states of a molecule (S_0 and S_1) split into vibrational energy levels. Absorption of a photon can excite the molecule from S_0 to S_1, which then rapidly relaxes (10^{-12} s) non-radiatively to the lowest vibrational S_1 level. Subsequent relaxation to the S_0 level can occur through photon emission (k_F) or non-radiative relaxation (k_{NR}) (within $\sim 10^{-9}$ s). Additionally, intersystem crossing to the triplet state (T_1) can occur, which has a slower return to the ground state, accompanied by release of a photon (k_{PH}, phosphorescence). (b) Normalised absorption and emission spectrum of two dyes (lines and dotted lines respectively). Absorption and emission wavelengths are separated, known as Stokes shift. The emission and excitation spectra of these two dyes overlap (cyan), allowing these two dyes to transfer energy via fluorescence resonance energy transfer (FRET).

of state S_1, with the electron then rapidly relaxing to the lowest vibrational energy level of the S_1 band, within $\sim 10^{-12}$ s (internal conversion). Subsequent relaxation to the electronic ground state S_0, can be accompanied by release of a photon, which occurs within about 10 ns - this is the fluorescence (Fig. 2.1a, k_F). Relaxation to the ground state can also occur via collisional quenching with solvent molecules, and by other non-radiative relaxation pathways (Fig. 2.1a, k_{NR}). The emitted fluorescence photon is lower in energy than the absorbed photon due to the vibrational level relaxation after initial absorption, leading to a difference in wavelengths of the absorbed and emitted radiation known as the Stokes shift (Fig. 2.1b).

Fluorescence occurs from the singlet state, when the excited electron is paired with an electron of an opposite spin in the ground state. However, when in the excited state, a spin conversion event can occur in which the electron switches its spin to match the ground state electron. Subsequent transition to the ground state is forbidden, leading to much slower photon emission known as phosphorescence (timescale 0.001 - 1s), commonly used in "glow in the dark" toys. The triplet state is relevant as it can cause pauses or fluctuating fluorescence emission from individual molecules (Fig. 2.1a, T_1).

After fluorophore excitation, the timescale of relaxation is described by the lifetime τ, which we represent in the form of relaxation rates (the inverse of τ), which can be used to describe the probability per unit time of relaxing via a particular pathway. An excited molecule can relax via photon emission k_F, or by non-radiative pathways k_{NR}, with the relative rates of these pathways describing the "Quantum Yield" of a dye molecule (Q), defined as the fraction of emitted photons to absorbed photons. We quantify Q via the rate relaxation through photon emission k_F, to the relaxation rate through all other pathways:

$$Q = \frac{k_F}{k_F + k_{NR}}$$

Henceforth we refer to fluorescent dyes as fluorophores. (The fluorescence lifetime of a fluorophore is defined as: $\tau = (k_F + k_{NR})^{-1}$)

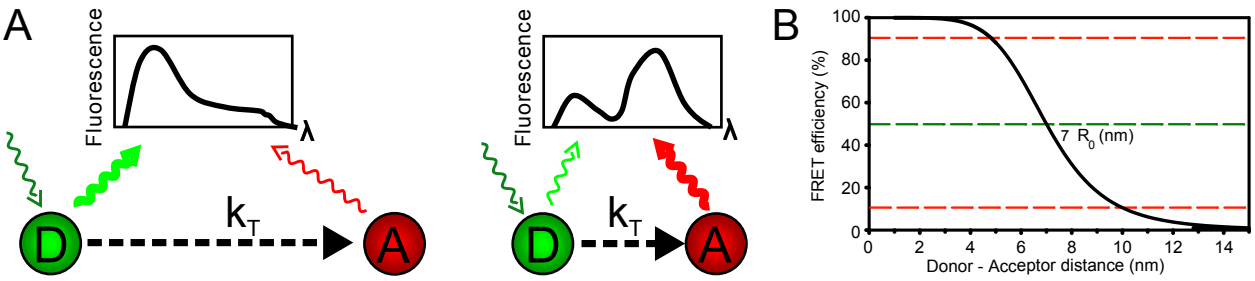


Figure 2.2: (a) Schematic illustrating the principle of FRET. As donor (D) and acceptor (A) fluorophores near each other, the excited donor dye transfers more energy to the acceptor dye, resulting in a decrease in green fluorescence emission, and an increase in red fluorescence emission. (b) Spatial dependence of the energy transfer efficiency between a donor and acceptor dye. The approximately linear range between 10% and 90% FRET is the most effective when using FRET as a spectroscopic ruler. Image adapted from lab commons.

2.3 FRET

Fluorescence resonance energy transfer (FRET) is the process by which an excited "donor" fluorophore (D) transfers its energy non-radiatively to an "acceptor" fluorophore (A), via a dipole-dipole interaction, which occurs if the emission spectrum of the donor fluorophore overlaps with the excitation spectrum of the acceptor fluorophore (Fig. 2.1b). This energy transfer is highly distance dependent over a range of 2-10nm, allowing FRET to be used as a "spectroscopic ruler" to monitor structural changes in proteins via attached fluorophores (Fig. 2.2a).[21]

The rate of energy transfer from donor to acceptor fluorophore has a distance dependence given by:

$$k_T(r) = k_D \left(\frac{R_0}{r}\right)^6$$

Where k_D is the decay rate of the excited donor in the absence of acceptor, r is the distance between the fluorophores, and R_0 is the Förster Radius, which defines the length scale over which FRET occurs, dependent on the spectral overlap between donor emission and acceptor absorption (J); the relative orientation of the fluorophore dipoles (κ); the refractive index of the medium n; and the donor quantum yield Q_D:

$$R_0^6 = \frac{9000 \ln 10 Q_D \kappa^2 J}{128 \pi^5 n^4 N_A}$$

$$J = \int f_D(\lambda)\epsilon_A(\lambda)\lambda^4 \, d\lambda$$

Where N_A is Avagadro's number; $f_D(\lambda)$ normalised donor emission intensity as a function of wavelength and $\epsilon_A(\lambda)$ is the extinction coefficient of the acceptor as a function of wavelength. R_0 defines the distance at which the rate of energy transfer is equal to the donor relaxation rate (i.e. 50% energy transfer efficiency). κ^2 describes the relative orientation of donor and acceptor dipoles, which averages to 2/3 for freely rotating dyes, but can vary between 0 and 4. Fluorophore orientation is usually investigated using fluorescence anisotropy experiments.

The efficiency of energy transfer between donor and acceptor is given by the ratio of energy transfer rate to total relaxation rate, essentially the probability by which a photon is transferred after excitation of a donor (note that k_D represents donor relaxation through photon or non-radiative relaxation) (Fig. 2.2b demonstrates the distance dependence of E):

$$E = \frac{k_T(r)}{k_D + k_T(r)} = \frac{1}{1 + (\frac{r}{R_0})^6}$$

Donor dye fluorescence emission decreases when an acceptor molecule nears, as it transfers more of its energy, leading to an increase in acceptor fluorescence. By measuring these intensities, we can infer the FRET efficiency via:

$$E = \frac{DA}{DA + \gamma DD}$$

Where DD and DA are the donor and acceptor fluorescence intensities under donor excitation, and $\gamma = \phi_a\eta_a/\phi_d\eta_d$ depends on the quantum yields ($\phi_{d,a}$) and collection efficiencies of the donor and acceptor fluorophores ($\eta_{d,a}$)[22]. In order to extract accurate FRET values, complications such as background fluorescence and cross-talk need to be accounted for.[23] E can also be inferred by changes in donor-lifetime, but equipment is more costly and sophisticated.[24] In situations in which absolute distances are not required, FRET is inferred via the proximity ratio (E*), the FRET approximation at which $\gamma = 1$:

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

Henceforth we extract FRET efficiencies using the measured intensities of donor and acceptor dyes under donor excitation (DA/DD). We use this notation throughout the thesis, where XY is the fluorescence emitted in the Y channel after excitation of the X channel, which can be donor (D) or acceptor (A) channels (for instance, DD represents the donor channel fluorescence after donor excitation, whilst DA represents the acceptor channel fluorescence after donor excitation, i.e. the energy transfer).

2.4 Single-molecule techniques

Single-molecule techniques have emerged as a powerful new tool for investigating biomolecular reactions, and although they remain technically challenging, are increasing in popularity.[25] Single-molecule techniques have distinct advantages over ensemble measurements, allowing the resolution of separate molecular populations within a heterogeneous sample which would otherwise be averaged (Fig. 2.3a, static heterogeneity); resolving the distinct states of dynamically interconverting single-molecules, rather than the averaged state measured using ensemble measurements (Fig. 2.3b, dynamic heterogeneity); and allowing observation of multi-step reactions without synchronisation, highlighting transient intermediates which would otherwise be missed, and allowing extraction of more detailed kinetics (Fig. 2.3c, multi-step reactions).

2.5 Single-molecule fluorescence

The high sensitivity of fluorescence makes it an obvious choice for single-molecule detection, with astonishing progress having been made in the last two decades. The first observations of single fluorescent molecules at room temperature were in 1990, via the use of flow and pulsed excitation to gate background photons.[26] The challenge in detecting single fluorescent molecules is in distinguishing fluorescence photons from background photons emitted from

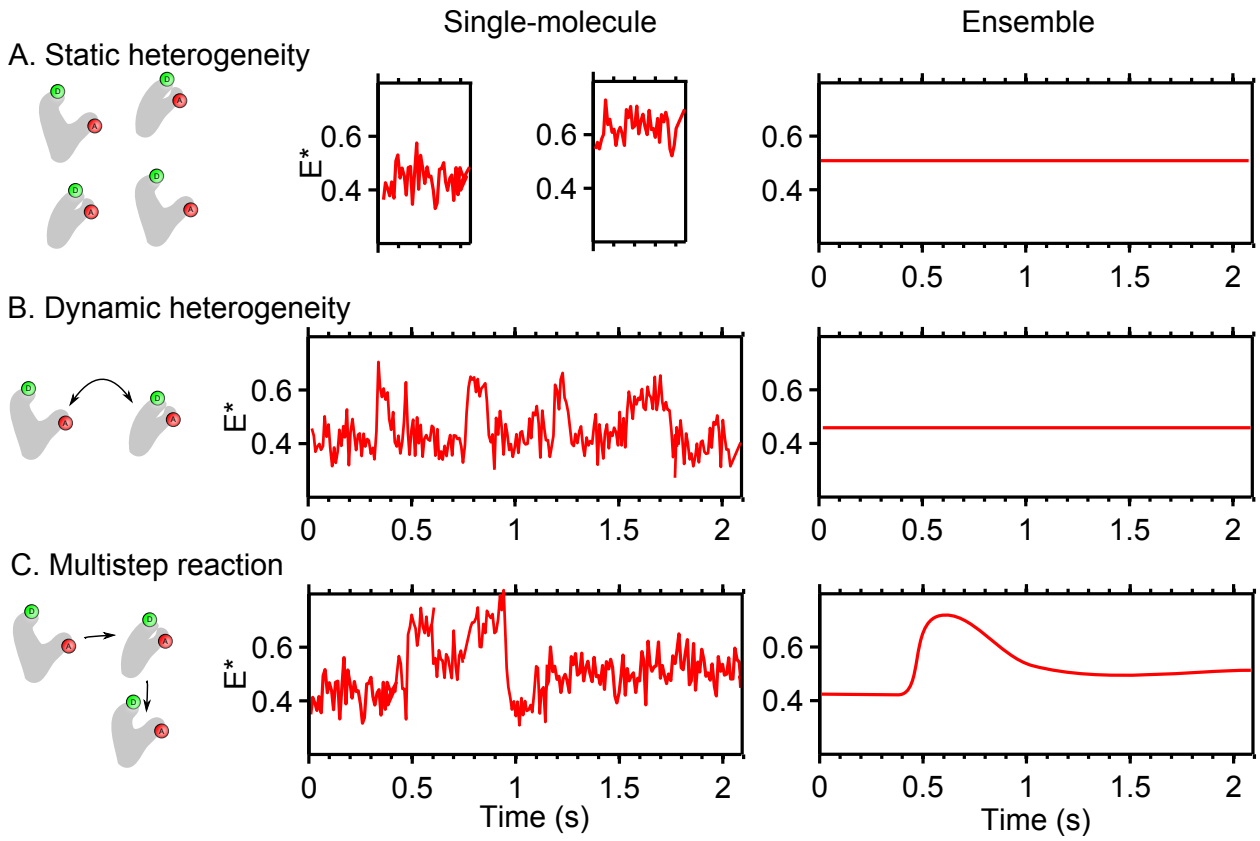


Figure 2.3: Schematic illustrating the advantages of single-molecule techniques over ensemble measurements, in which a FRET pair is used to infer a protein's conformation. (a) Distinct molecular populations can be observed using single-molecule techniques, which would otherwise exist as an averaged fluorescent signal in the ensemble. (b) Molecules interconverting between structural states can be observed using FRET variations of single-molecules, which exist as an averaged fluorescent signal in the ensemble. (c) A complicated multi-step reaction is illustrated in which a molecule interconverts between a series of FRET structural states during a reaction, all of which are observed individually. The FRET change of a synchronised ensemble reaction demonstrates the overall FRET change well, but the series of intermediate states are challenging to extract using the form of the FRET efficiency change.

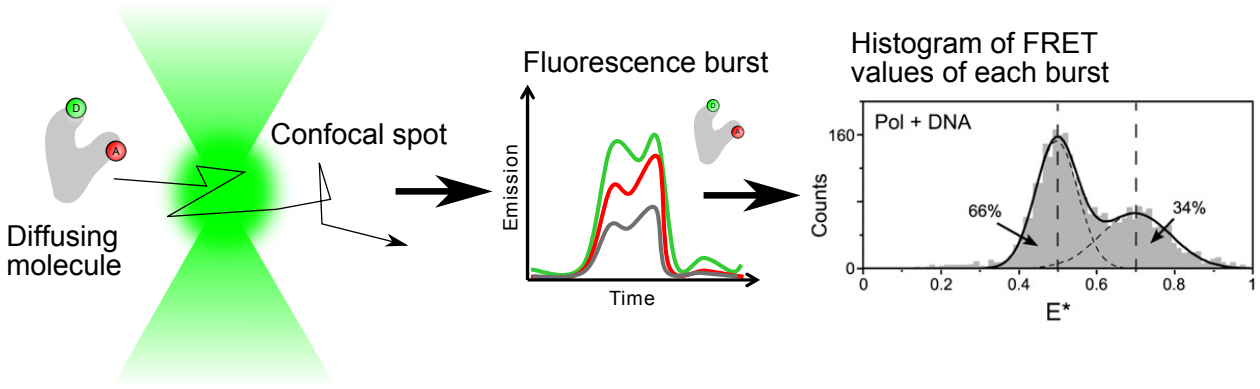


Figure 2.4: Schematic illustrating the principles of single-molecule confocal microscopy. A small illumination volume is formed in solution by a focused laser beam, through which fluorescent molecules can rapidly diffuse (in this case a donor-acceptor labelled polymerase). Bursts of fluorescence released in the DD (green) and DA (red) channels can be used to calculate a FRET value for each polymerase, which is plotted in a histogram. Adapted from Santoso 2010.[19]

a variety of sources, such as Raman scattering from solvent molecules; Rayleigh scattering of the excitation source; detector dark counts; and auto-fluorescence of optical components.

High optical density dichroic mirrors and filters can remove Rayleigh scattered excitation light from the fluorescence signal, and detectors with high quantum yields and low noise can be used to detect single photons. Raman scattering from individual solvent molecule is low, but the higher number of solvent molecules to fluorophore molecules leads to large Raman scattering signals, which can be reduced by using reduced illumination volumes. Considering the excitation and emission properties of Rhodamine 6G (R6G) and Raman scattering of the solvent ethylene glycol, it can be calculated an observation volume < 100 fl (4.9μm sided cube) is required for single molecule detection purely on the basis of Raman scattering.[21] Typical observations volumes used in SM experiments are 1 fl (1μm sided cube), readily produced using modern optical components.[27] Techniques such as spatial filtering are also used to remove the optical emission from out of focus illumination.

2.5.1 Single-molecule confocal microscopy

Single-molecule detection can be achieved using confocal microscopy, in which a small illumination volume is produced by tightly focusing a laser beam, producing a point-spread-

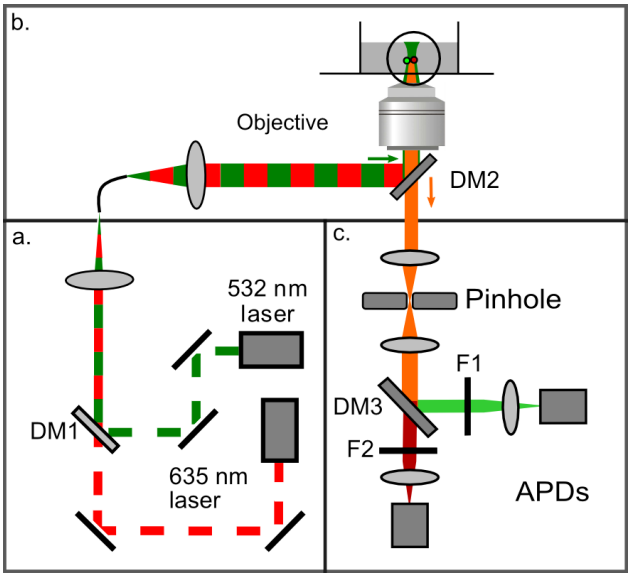


Figure 2.5: Schematic of a single-molecule confocal microscope. (a) Fluorescence excitation module. Green (532nm) and red (638nm) lasers are alternated at 20kHz and coupled into an optical fibre using a dichroic mirror (DM1) and mirrors. (b) The fibre output is collimated, reflected by a dichroic mirror (DM2) and focused in solution using an oil-immersion objective (60X, NA 1.35, UPLSAPO (Olympus)). Fluorescence light passes through the dichroic mirror (DM2) into (c) the fluorescence emission module. Collimated fluorescence is focused onto a pinhole for spatial filtering, recollimated, split into donor and acceptor emission components using a dichroic mirror (DM3), filtered (F1,F2) and focused onto avalanche photodiodes (APDs) for detection. Image adapted from lab commons.

function (PSF) of femtolitre volume. Fluorescent molecules in solution can diffuse rapidly through this PSF, emitting short ($\sim$ 1ms) bursts of fluorescence, which are collected and then detected using an avalanche photodiode (APD) (Fig 2.4). Using a low concentration of fluorescent molecules (pM) results in well separated fluorescence bursts, which can be extracted by applying a thresholding algorithm on the photon arrival times measured by the APD.[28]

Previous work in the Kapanidis group used single-molecule confocal microscopy to investigate conformational changes in the protein DNA Polymerase I, the enzyme which we study further in this thesis. An internal FRET pair reported on the conformation of the polymerase, via the ratio of intensities in the donor and acceptor channels under donor excitation (DD/DA). FRET values of each fluorescence burst were plotted in a histogram, demonstrating the conformational distribution of the polymerase sample (Fig. 2.4). Alternating laser excitation (ALEX) was used, which involves rapid alternation of the green and red lasers (20kHz), allowing quantification of both DD and DA fluorescence, but also the acceptor fluorescence under acceptor excitation (AA) for each fluorescence burst. This allowed probing for the presence of the acceptor dye in each burst, and also it's photophysical state. Practically, this allows distinction between donor-only molecules and low FRET molecules, which appear at similar low FRET values due to donor fluorescence cross-talk in the acceptor channel. ALEX also allows monitoring of acceptor dye quantum yield changes, permitting distinction between FRET changes due to structural changes or those due to acceptor quantum yield changes.[29] In this thesis, we use confocal microscopy to a limited degree, to validate surface based investigations, and use a confocal microscope described previously[20], outlined in Fig. 2.5.

Single-molecule confocal microscopy typically monitors diffusing molecules, and thus does not require the complicated immobilisation schemes required for surface-based experiments. However, diffusion confocal microscopy is limited to observations of $\sim$1ms at a time, making it difficult to observe continuous reactions or conformational changes over longer times scales, unless combined with scanning stages and surface-immobilisation techniques. Scanning confocal microscopy is, however, limited to observing a single molecule at a time, with

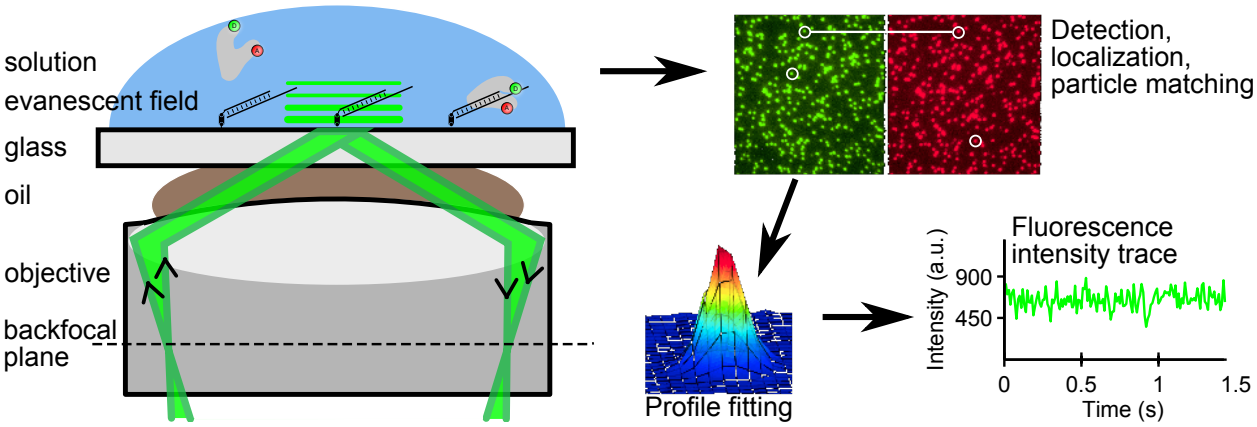


Figure 2.6: Schematic illustrating the principle of TIRF microscopy. A collimated laser beam is focused into the back-focal plane of an objective, off the optical axis such that the subsequently collimated beam is incident on the sample at an angle greater than the critical angle, resulting in the beam being totally internally reflected, producing an evanescent field in solution above the coverslide, on which there are immobilised single-molecules. The slide surface is imaged on a CCD camera, producing a "star-field" of single-molecule point-spread functions in the DD and DA channels. Software is used to localise these molecules, and associate them between the DD and DA channels. Gaussian profiles are fit to each PSF, to extract its intensity at each movie frame, which then produces a fluorescence intensity trace for each molecule.

the resultant low statistics making it challenging to sort and distinguish static and dynamic heterogeneity between inherent statistical variability.

2.5.2 TIRF microscopy

Observing the time evolution of a biomolecular reaction requires extended observations, which can be achieved by immobilising the molecules of interest. Total Internal Reflection Fluorescence (TIRF) microscopy allows observation of hundreds of surface-immobilised single-molecules simultaneously, whereas confocal microscopy is limited to observing individual immobilised molecules.[30]

TIRF microscopy produces a shallow illumination ($\sim 100nm$) across a surface via the evanescent field of a laser beam totally-internally-reflected at a glass/solution interface, reducing the background from the surrounding solution and allowing excitation of fluorescent molecules immobilised close to the surface. Total internal reflection occurs when light passes from a medium of higher refractive index (n_2), to a lower refractive index (n_1), providing the angle of incidence θ_{inc} is greater than the critical angle θ_C:

$$\theta_C = \sin^{-1}(\frac{n_1}{n_2})$$

The decay of the evanescent field is exponential, where z is the position above the surface:

$$I = I(0)\exp(-z/d)$$

The decay constant d is given by:

$$d = \frac{\lambda_0}{4\pi}(n_2^2 \sin\theta_{inc} - n_1^2)^{-1/2}$$

For microscopy, an incident beam with high enough angle for TIR can be achieved by using a prism, or by focusing a laser beam into the back-focal plane of an objective, off the optical axis, such that the collimated beam of light exits at a high angle, and reflects off the glass coverslide (Fig. 2.6). Fluorescence emission can then be filtered from excitation light using a dichroic mirror, spatially filtered if required, separated into donor and acceptor emission, with an image of the surface produced on a electron-multiplying charge-coupled device (EMCCD).

In this thesis, the majority of single-molecule experiments are performed on a custom built TIRF microscope (Fig. 2.7 schematic of a TIRF microscope). In the excitation path of this microscope, green (532nm Cobolt Samba) and red (635nm Cube Coherent) lasers are coupled into an optical fibre using a dichroic mirror (DM1), and alternated at the desired frequency (1000 - 10Hz). The fibre output is collimated, and focused into the back focal plane of a objective (100X oil-immersion, NA 1.4) off the optical axis. The collimated beam exits the objective at an angle, sufficient that the beam is totally internally reflected from the glass surface, creating the necessary evanescent field. Fluorescence emission from the focal plane is separated from excitation light by a dual pass dichroic mirror (DM2: 545nm/650nm, Semrock). The focal plane is imaged on a slit, re-collimated and passed through clean-up filters (F1: 545nm LP, Chroma; F2: 633/25nm red laser notch filter, Semrock), and split into donor and acceptor fluorescence components using a dichroic mirror (DM3: 630 nm DRLP; Omega), and subsequently imaged on a CCD camera (Andor iXon 897). A Continuous

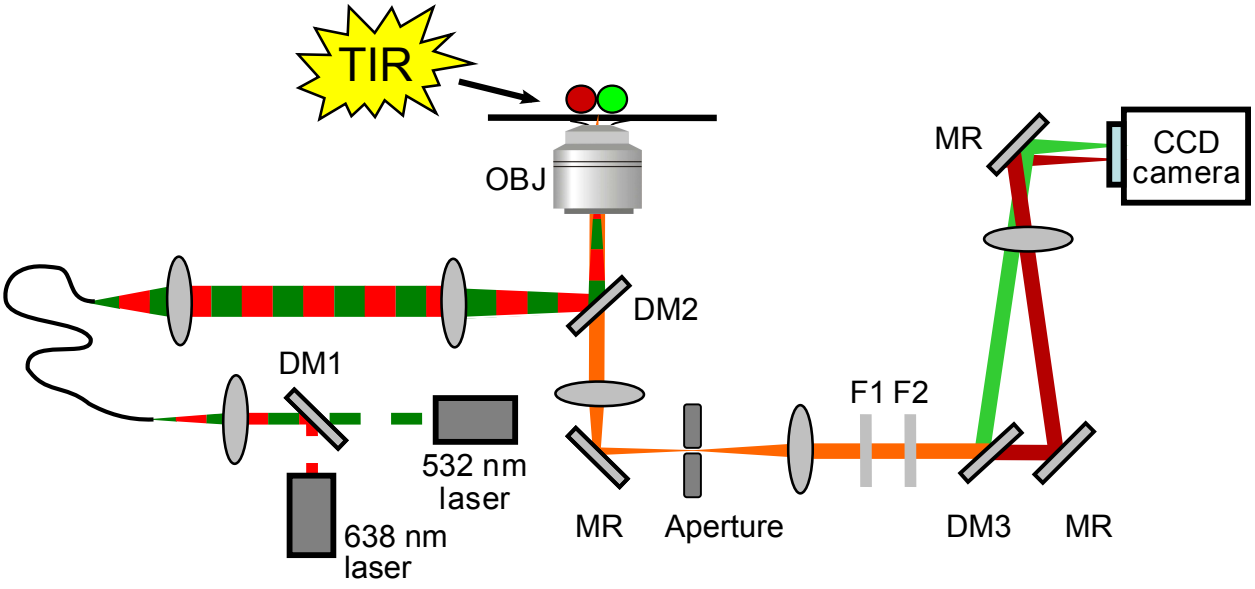


Figure 2.7: Schematic of a TIRF microscope. Green and red lasers are alternated and coupled into a fibre optic cable, whose output is collimated and focused in the back-focal plane of the objective, producing a totally internally reflected beam. The fluorescence emission is filtered using a dichroic mirror (DM2), imaged on a slit, recollimated, passed through clean-up filters (F1,F2), split into donor and acceptor emission channels, and imaged on an EMCCD. Image adapted from lab commons.

Reflective-Interface Feedback Focus System (CRIFF, ASI) is used to maintain the focal plane, and the microscope can be run using ALEX, or in continuous wave mode (CW) in which the green laser maintains a constant illumination.

Surface-immobilised single-molecules are imaged as a collection of point-spread functions (PSF) on the EMCCD (Fig. 2.6). The intensity emission at each PSF is quantified using TwoTone software[31], which localises the positions of the PSFs, matches the molecules in the green and red channels, and fits each PSF to a Gaussian function to extract intensity value in the DD, DA and AA channels (AA - acceptor emission under acceptor excitation).

2.5.3 Photophysics

There are practical difficulties associated with the observation of single-molecules, most notable are dye blinking and photobleaching. Photobleaching is the irreversible photo-chemical alteration of a dye which leads to cessation of fluorescence, thought to occur via reaction with oxygen radicals in solution, which exist because of excitation of the fluorophores.[32] Photobleaching can be reduced by using photostable dyes (we use Cy3B and Atto647N),

and by reducing solution oxygen levels by using an enzymatic oxygen scavenging system (glucose oxidase/catalase)[33]. Blinking is the fluctuating fluorescence emission which occurs as a result of transient entry to the triplet state or by long lived radical ion states (Fig. 2.1a, triplet state T_1). While oxygen induces photobleaching, it also quenches the triplet state[32], requiring the addition of a triplet state quencher such as Trolox (an analogue of vitamin E) if oxygen is removed. Trolox and it's oxidised form (TX-quinone) can act as a reducing and oxidising system (ROXS) which can depopulate triplet states and also remove radical ions from solution, leading to stable fluorescence signals.[34]

Chapter 3

A primer on DNA polymerase

In this chapter we present the enzyme DNA polymerase I (Klenow Fragment), which is responsible for cellular DNA replication and repair in *E. coli*, and is the subject of this thesis. We review the experiments used to understand the DNA polymerase reaction mechanism, and review the theories which attempt to explain how polymerases replicate DNA with such extraordinary accuracy.

3.1 Introduction

Cellular genetic material is encoded in DNA, a long heteropolymer whose base sequence encodes the "instructions" required to synthesise the wide range of proteins crucial to cellular function (See Fig. 3.1 for a description of DNA). The genome is replicated during cell division to provide a copy for each cell, however, errors in DNA replication or maintenance can result in cell death, changes in cell physiology, or diseases such as cancer. Conversely, DNA mutations may assist in driving evolution of a species.[35,36]

DNA polymerases are the enzymes which perform DNA replication during cell division, as well as associated functions, such as DNA damage and lesion repair, correcting nucleotide misincorporations, and repairing double stranded DNA breaks ($> 50,000$ damage events occur daily in each human cell, caused by UV radiation and other sources[37]). Diversity across biological domains and functions results in a wide range of polymerases, which are classified into families based on sequence homology (families A, B, C, D, X, Y and RT). Many demonstrate similar overall structures.[36]

DNA polymerisation is performed upon primer-template DNA - a DNA duplex with

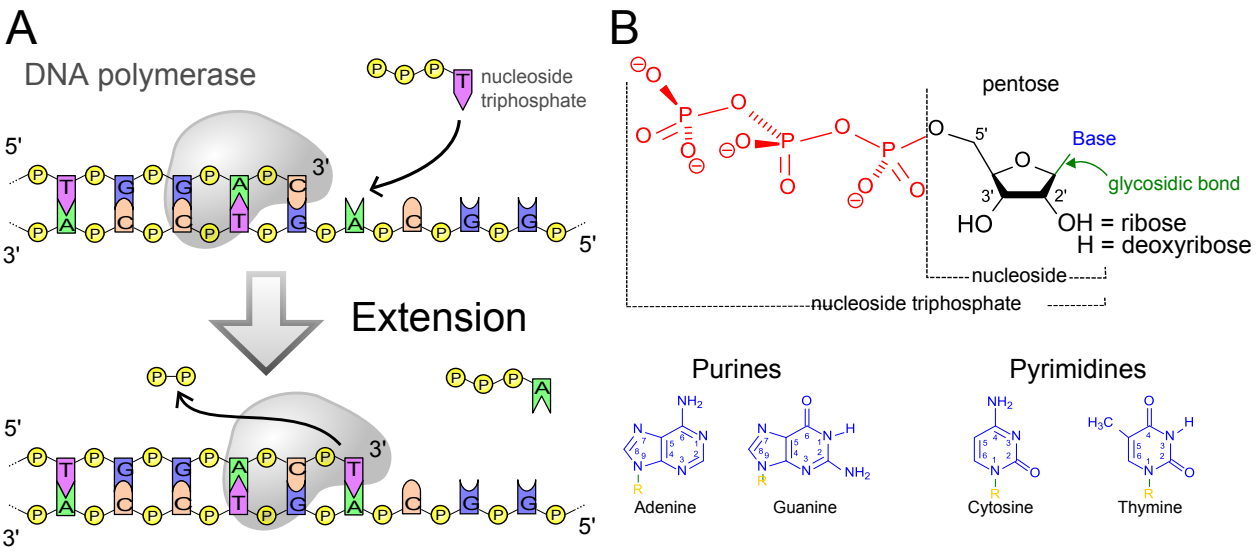


Figure 3.1: (a) Primer-template DNA with a duplex region, and a single-stranded template overhang. A dNMP is incorporated into the primer strand by a polymerase if the incoming dNTP base is complementary to the first templating base, with the reaction leading to a pyrophosphate product. (b) Structure of deoxyribonucleoside triphosphate. The sugar and phosphate form the backbone of the DNA, with each sugar connected to one of four possible bases (illustrated in blue). Two single-stranded DNA chains form a duplex structure if the bases of the two strands complement each other according to Watson-Crick base-pairing rules (A with T, C with G). The directionality of each DNA strand is defined by the 5' and 3' regions of the sugar. Image courtesy of Madeleine Price Ball and Wikimedia Commons.

a single-stranded overhang formed by the hybridisation of a single-stranded primer DNA molecule with a longer template DNA molecule (Fig. 3.1a). The template acts as a blueprint for the polymerase, which matches the template bases to incoming deoxyribonucleotide (dNTP) molecules according to Watson-Crick base-pairing rules (Fig. 3.1b, adenine (A) matches thymine (T), cytosine (C) with guanine (G)). When the dNTP base is complementary, the deoxyribonucleoside monophosphate (dNMP) of the dNTP is chemically incorporated onto the primer strand. Sequential incorporation then leads to primer extension, creating a fully complementary duplex. Chemical incorporation involves creating a phosphodiester bond between the terminal -OH of the 3' primer base and the alpha phosphate of the incoming nucleotide, catalysed by two Mg^{2+} ions (Fig. 3.1 shows dNTP structure)

DNA polymerases replicate DNA extremely accurately, incorporating an incorrect nucleotide as low as once in several hundred thousand incorporations (consider that the human genome contains 6 X 10^9 bases and that one error can be potentially fatal).[38] This accuracy cannot be accounted for by simple thermodynamic considerations such as the disfavourable

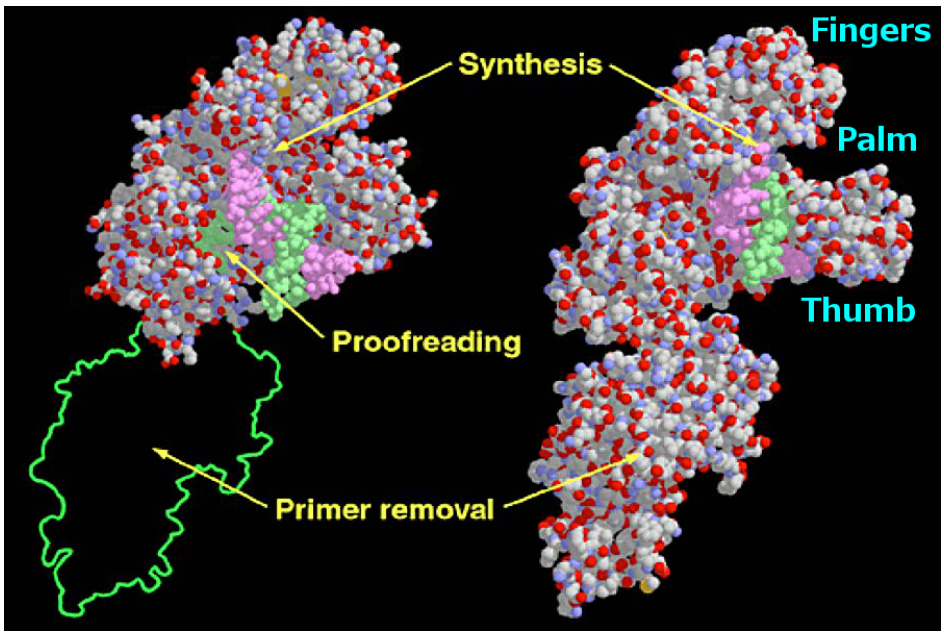


Figure 3.2: Crystal structures of DNA polymerase I (right, *Thermus aquaticus*, PDB:1tau) and Klenow fragment (left, *E.Coli*, PDB:1kln), in which the 5'→3' exonuclease is removed (green outline). DNA strands are seen in green and purple. Ignoring the 5'→3' nuclease, the right hand shape of the polymerase can be seen, with the fingers, palm and thumb labelled. Klenow fragment appears to be bound with the DNA in the 3'→5' exonuclease, rather than in the polymerisation site, as seen in the DNA polymerase I structure. Sourced from Wikimedia Commons.

energetics of mismatching base-pair hydrogen bonding, compared to the correct hydrogen bonding of Watson-Crick base-pairs.[39] A complete theory explaining the high fidelity of polymerases remains elusive.

In this study, we investigate the first polymerase to be discovered (by Kornberg in 1956), DNA polymerase I , a family A polymerase from *Escherichia coli* (*E. coli*) with low processivity (15-20 nucleotides).[40] The role of DNA polymerase I in *E. coli* is to perform lagging strand DNA replication and DNA repair. Its polymerisation domain can polymerise through DNA or RNA hybridised to the template strand, with the resulting flap removed using its 5'→3' nuclease, allowing the removal of the RNA primers on Okazaki fragments. DNA polymerase I can also fill the DNA gaps caused by the excision of nucleotides during DNA repair, and its 3'→5' exonuclease domain performs a proofreading function, excising nucleotides from the primer on misincorporation.

In this study we are mainly concerned with polymerisation function, therefore we use a fragment of DNA polymerase I in which the 5'→3' nuclease region is absent. The resultant

protein known as Klenow Fragment (KF) (Fig. 3.2, structures of DNA polymerase I and Klenow Fragment). Additionally, the 3'→5' exonuclease is rendered inactive via a point mutation (D424A), and a cysteine is removed (C907S) to allow site-specific dye labelling.[41,42] As for many polymerases, KF is shaped like a right hand, with fingers, thumb and palm subdomains (Fig. 3.2). The thumb binds the duplex DNA, whilst the palm contains the catalytic active site, and the fingers bind the incoming nucleotides and the template overhang, and undergoes conformational changes during the reaction.

In the following chapter we discuss the experimental evidence which has been used to construct a reaction mechanism for DNA polymerisation by KF. In brief, the reaction mechanism for correct nucleotide incorporation involves binding of KF to DNA, followed by binding of complementary nucleotide, followed by a series of subtle conformational changes in the fingers and DNA template overhang. These are followed by a large conformational change in which the fingers subdomain closes around the complementary nucleotide, forming a tight binding pocket around the nascent base-pair. A rate-limiting conformational change of unknown origin is proposed to occur, followed by activation of the catalytic domain and chemical incorporation of the nucleotide. Subsequently, the product pyrophosphate is released, there is another slow conformational change and translocation of KF to the next insertion point.

In addition, we discuss theories which have been proposed to explain polymerases high fidelity, thought to include selection via Watson-Crick base pairing, water exclusion, geometric selection and contributions from the fingers-closing conformational change.

3.2 A brief history of investigations into Klenow Fragment (KF)

Here we review the experimental basis for the known DNA polymerisation mechanism of KF. While we focus primarily on the KF polymerisation mechanism, many of the techniques and general results apply to closely related polymerases. The need to review the experimental methods which form the basis of our current understanding of KF will become clear in

later chapters, as our novel observations allow reinterpretation of previous results, requiring reconsideration of the conclusions drawn.

3.2.1 Kinetic investigations

Studying the rates of DNA replication (specifically extension of the primer) by KF under different conditions can be used to draw conclusions about the underlying reaction mechanisms by kinetic analysis. These sort of kinetic investigations were performed heavily in the late 1980s and early 1990s, resulting in a minimal mechanism for the polymerisation reaction, forming the basis for many further studies. The minimal mechanism consists of substrate binding steps, conformational changes and chemistry steps. The set of experiments leading to this minimal mechanism are described briefly below.

$$\mathrm{KF} + \mathrm{dATP} + 13/20\mathrm{mer} \rightleftharpoons \mathrm{KF} + 14/20\mathrm{mer} + \mathrm{PP}_i \tag{3.1}$$

The rate of DNA polymerisation by KF was analysed by measuring the production of a 14/20mer primer-template, formed via the incorporation of radioactively labelled nucleotide [^{32}P]dATP into a 13/20mer by KF (See Eq. 3.1). The time course of DNA formation was measured by halting the polymerisation reaction at time-points after initiation, and measuring the radioactivity of the sample. The high-speed of nucleotide incorporation necessitated the use of "rapid-quench" techniques to halt the reaction at these time points, typically using EDTA or acid. (See Fig. 3.3 for details on rapid quench apparatus, and an example kinetic experiment.)

For instance, a sample kinetic experiment is seen in Figure 3.3b, in which 14/20mer product is formed as DNA polymerase I (KF) catalyses the extension of a 13/20mer via incorporation of [^{32}P]dATP. An initial rapid increase in product occurs as the KF-DNA-dATP complexes incorporate a single nucleotide, with the plot slope defining the rate of nucleotide incorporation (Fig. 3.3b, the rapid increase between 0 - 0.15s is the pre-steady-state reaction) (Conditions are such that all KF molecules are initially bound to DNA, and only the nucleotide complementary to the first base is present, so that the pre-steady-state

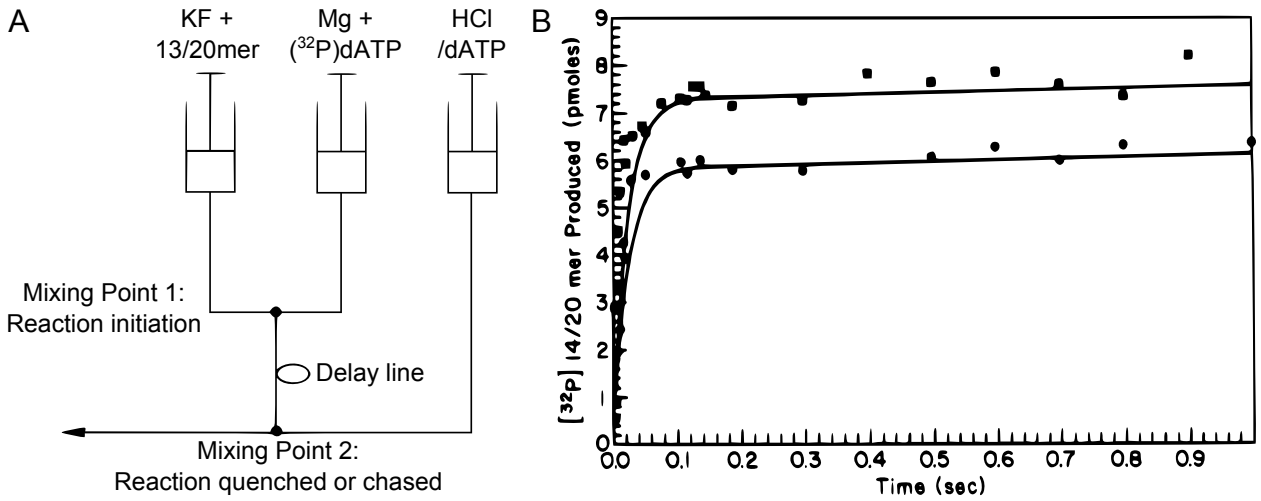


Figure 3.3: Use of rapid quench apparatus to measure the rate of DNA extension by KF. (a) Schematic of a rapid quench apparatus. Syringes are depressed, and polymerisation is initiated as the KF + DNA solution mixes with the Mg^{2+} + dNTP solution at Mixing Point 1. The reaction then flows along a delay line (which defines the time of the reaction, variable by line length) and it is halted (quenched) at Mixing Point 2 as it mixes with acid or EDTA from a third syringe. After quenching, the amount of radioactive elongated DNA is quantified and plotted (see b). Chased experiments are performed with unlabelled nucleotide (dATP) instead of acid or EDTA. (b) Quantity of elongated DNA ([^{32}P] 14/20-mer) measured at different reaction time-points (circles). An initial steep increase corresponds with single-nucleotide incorporations of the KF-DNA-dNTP bound complexes. Subsequent increase is slow and rate-limited by the time taken for KF to dissociate from the polymerised DNA, and bind unpolymerised DNA molecules. Squares show the elongated DNA quantity when chased with unlabelled dATP instead of being quenched. Fig. b adapted from Dahlberg 1991.[43]

region reflects a single-turnover).

In the subsequent slow phase, DNA extension is limited by the slow dissociation of KF from the 14/20mer product, required before the KF can bind to a new substrate and produce more DNA product (Fig. 3.3b, the slow increase between 0.15 - 2.1s is the steady-state reaction). Given it is limited by KF-DNA dissociation, this slow rate provides a measure of the KF-DNA dissociation rate ($k_{off} = 0.12s^{-1}$). Performing these observations under a range of dATP concentration can be used to quantify a catalytic rate $k_{pol} \sim 50s^{-1}$ and nucleotide dissociation constant $K_D = 5.5\mu M$. The mono-exponential increase also specifies polymerisation to be rate limited by a single-step.[44]

Sulphur-elemental effect

The chemical step during nucleotide incorporation involves creating a phosphodiester bond between the primer strand and the incoming nucleotide. During the 1980s, it was unclear if this step, or another, was rate-limiting for the polymerisation reaction, an important issue as the rate-limiting step was thought to contribute strongly towards fidelity.[45]

In order to determine if the chemical step was rate-limiting, chemically modified nucleotides were used to slow the polymerisation chemical step, to observe if this affected the overall DNA extension rate. If chemistry was rate-limiting, changing the rate of chemistry should influence the DNA extension rate strongly, whereas if another step was rate-limiting, then changing the chemistry rate would not alter the DNA extension rate significantly. Chemistry was slowed by using modified dNTPs (dTTPαS) containing a sulphur atom instead of an oxygen atom at the alpha phosphate, however, no significant change was found in the polymerisation rate between these and natural dTTPs, which suggested that the rate-limiting step occurred before the chemical step, but was not the chemical step its self. This pre-chemistry rate-limiting step was hypothesised to be a conformational change. (Work in the same paper used positional isotope exchange to infer that PP_i release after polymerisation was much faster than the reverse chemical step.)[44]

The sulphur elemental effect remains cited as evidence for or against rate-limiting chemistry for various polymerases, despite its use as a diagnostic being heavily criticised.[46,47,48]

The sulphur elemental effect was initially predicted to effect the rate of chemistry by a factor of $\approx 30-100$ based on solution observations of phosphoryl transfer reactions[49], however its effect was later revised as much smaller ($\approx 4-10$),[50] but larger observed effects have suggested that steric effects contribute as well as electronic effects.[51] Complications in predicting the effect of substituting sulphur atoms in the place of oxygen makes conclusions drawn based on the sulphur elemental effect questionable.[52]

Pulse-quench vs pulse-chase yields

Despite the ambiguities of the sulphur-elemental effect, the existence of a slow pre-chemistry step has been supported by further rapid-quench experiments. These experiments compared the product yield of a quenched polymerisation reaction with that of a "chased" polymerisation reaction. A "chased" polymerisation reaction involves using an excess of unlabelled substrate in place of the quenching acid. The excess of unlabelled substrate allows the reaction to proceed normally, with any nucleotide bound KF-DNA intermediate free to proceed with the reaction, or dissociate from the labelled nucleotide, which then becomes swamped by the excess of unlabelled nucleotide. Any product formed with the unlabelled substrates would not be measured in the product yield. Thus, the difference in yields between the "chased" and "quenched" reactions allow a measure of the approach to equilibrium reached between the unlabelled nucleotides and the KF-DNA nucleotide bound intermediates.[43]

Experiments on KF revealed the DNA product yield of quenched experiments to be 80% of the active enzyme concentration, whilst the yield of chased experiments was 100% of the active enzyme concentration (Fig. 3.3b, circles and squares respectively). This difference was used to infer the existence of a nucleotide bound intermediate, whose existence was used to infer slow steps both before *and* after the chemical reaction.[43] It was reasoned that fast steps before the reaction would allow labelled substrate to equilibrate rapidly with unlabelled substrate, resulting in no difference between the chase and quench yields, which suggested the existence of a slow step before chemistry to allow the nucleotide-bound intermediate population to form. In order to have sufficient build-up of this nucleotide-bound intermediate, a slow step after the reaction was also required in order that the intermediate

not immediately react, form product, and dissociate. These results were used as evidence supporting the existence of a slow pre-chemistry step, but also demonstrated the existence of a slow *post*-chemistry step, however, the interpretation was somewhat predicated upon the results of the sulphur-elemental effect experiments.[43]

The same study quantified the K_D for KF primer-template binding ($K_D = 5nM$); the K_D for nucleotide binding to the KF-DNA binary complex ($K_D = 5\mu M$); and the K_D for PP_i binding to the binary complex ($K_D = 100\mu M$), as well as forward and backward rates for the conformational change before chemistry ($k_3 = 50s^{-1}, k_{-3} = 3s^{-1}$); the forward and backwards conformational change after chemistry ($k_5 = 15s^{-1}, k_{-5} = 15s^{-1}$) and estimated the internal equilibrium constant of the chemical step to be ~ 4.[43,53] It was also demonstrated that KF-DNA binding was sequence dependent.[53]

Pyrophosphorolysis

The existence of a pre-chemistry rate-limiting step was also initially supported in work which investigated the reverse polymerisation reaction: pyrophosphorolysis.[53] Pyrophosphorolysis occurs when pyrophosphate (PP_i) reacts with the dNMP on the primer end, removing it from the DNA and forming a dNTP molecule (the reverse reaction seen in Eq .3.1). The pyrophosphorolysis reaction was studied by adding an excess of PP_i, to drive the polymerisation reaction in the reverse direction. This was monitored by observing the *decrease* in radioactive DNA as the [H^3]dNMP was removed from the DNA strand during phosphorolysis, to form [H^3]dNTP.

Initial experiments demonstrated a biphasic decrease in the 14/20-mer DNA population during pyrophosphorolysis, explained by a fast pyrophosphorolysis reaction followed by a slow conformational change before release of the dATP product, consistent with the slow pre-chemistry conformational change.[53] However, these experiments were found to be erroneous, with later work demonstrating pyrophosphorolysis to follow simple exponential kinetics. The sulphur elemental effect initially suggested that chemistry was rate-limiting for pyrophosphorolysis[44], but this was revised as the elemental effect was attributed to decreased KF binding caused by the phosophorothioate duplex, the revision demonstrating a

lack of elemental effect, and therefore supporting the existence of a post-chemistry slow-step.

The complexity of kinetic experiments, both when using altered chemistry, and the modelling used to fit the data, was reflected in the frequent revision and correction of results collected using these techniques. There is no consistent kinetic pattern across all polymerases, with the sulphur elemental effect providing a variety of results for or against rate-limiting conformational changes across various polymerases (consistent with the techniques unreliability). Pulse-chases vs. pulse-quench yields, however, suggest the existence of slow pre and post-chemistry steps for several polymerase: Klenow fragment; HIV reverse-transcriptase; T7 DNA polymerase; DNA polymerase η; and Dpo4 (*Sulfolobus solfataricus* DNA polymerase IV). Assertions about the rate-limiting step during misincorporations are based predominantly upon the sulphur elemental effect, and as such are variable between polymerases and of questionable reliability, whilst the experiments are generally more challenging because of the very slow rates of misincorporation.[46]

3.2.2 Crystal structures

Some of the most important insights into DNA polymerases' reaction mechanisms have arisen from their crystal structures, which over the years have developed from structures of the enzymes alone,[54,55,56] to structures of the enzyme-DNA binary complex[57], through to sequences of structures of the polymerase undergoing polymerisation.[58,59,60] Information on KF specifically has been derived from structures of KF and close homologues from *Bacillus stearothermophilus* and *Thermus aqaticus*.[61]

Crystal structures have demonstrated that many polymerases exist in a structure analogous to a "right hand" shape, in which the DNA binds along the cleft formed by the "thumb", "palm" and "fingers" subdomains (Figures 3.2, 3.5, 4.1)[58,61]. The thumb subdomain is particularly involved in DNA binding, with significant contacts to the DNA minor groove.[61,62], and has been implicated in processivity, most likely due to its role in DNA binding[63].

For A-family polymerases, such as KF, the palm domain consists of a β sheet structure which contains highly conserved residues involved in catalysis. The "fingers" subdomain is involved in binding of the incoming nucleotide, after which it transitions from the "fingers-

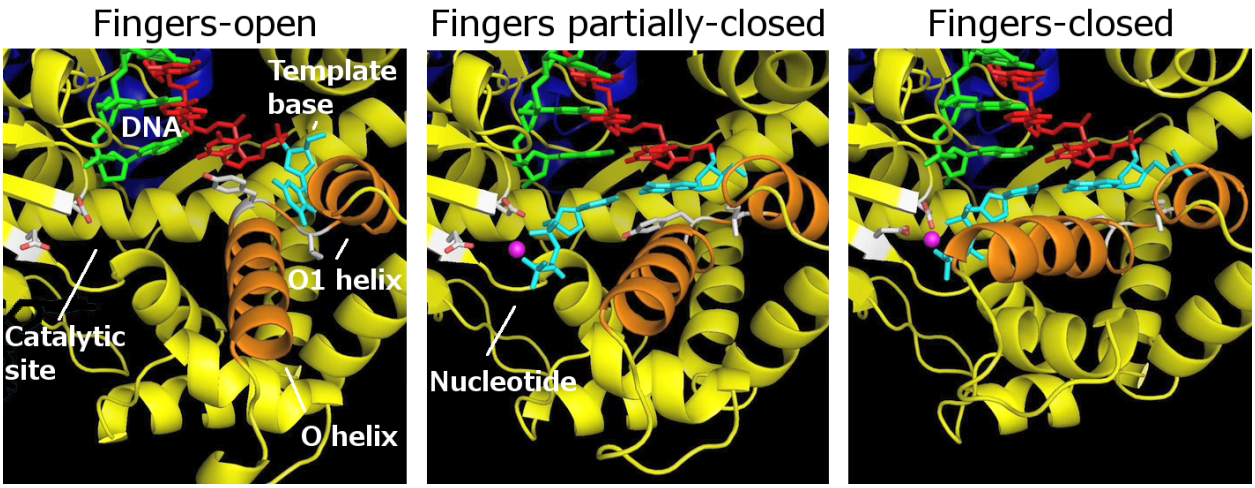


Figure 3.4: Crystal structures of the fingers-open, fingers-partially-closed and fingers-closed structures of the large DNA polymerase I fragment from *B. stearothermophilus*, from Wu 2011. Yellow shows the protein backbone, with the O and O1 helices representing the fingers of the polymerase (orange), with the O helix becoming increasingly "closed" throughout the structures. The DNA primer-template are in green and red respectively, with the first templating base (cyan) either stacking between the O helices or stacking on the end of the duplex, dependent on the rotation of the O1 helix and the position of the blocking tyrosine residue (grey). Catalytically active residues are seen in white, with magnesium ions in pink, in position to perform catalysis in the fingers-closed conformation. The incoming nucleotide is seen in cyan.[60]

open" conformation, into a "fingers-closed conformation" in the presence of complementary nucleotide, with the fingers closing around the nascent base-pair, leading to the catalytically active structure (See Fig. 3.4 for an active site view and Fig. 4.1 for fingers-closing in the context of the "right-hand" shape).[58]

In the fingers-open conformation, the first templating base is occluded from the insertion site via a tyrosine residue stacked on the end of the DNA duplex, and instead is flipped out and lying between the O and O1 helices (Fig. 3.4, fingers-open structure shows the templating base in blue between the orange O and O1 helices, with the grey tyrosine stacking on the red template strand). On fingers-closing the O helix of the fingers undergoes a rigid $\sim 40°$ rotation, displacing the blocking tyrosine from the duplex end, and simultaneously displacing the first templating base from its "pre-insertion site" between the O and O1 helices, allowing it to stack on the duplex end (Fig. 3.4, fingers-closed structure). The incoming nucleotide is positioned in line with magnesium ions and catalytic residues, ready for chemical incorporation.[58] The fingers closing conformational change was noted as a clear candidate for the

rate-limiting conformational change inferred in kinetic experiments described above.[64] The reader is strongly recommended to view a striking movie of this conformational change, as the important motions are highlighted far more clearly than in static images or descriptions. [i]

Recent structures have demonstrated how *B. stearothermophilus* polymerase responds to a mismatching nucleotide, via crystal structures of the ternary complex bound to a dG:dTTP mismatch. In this structure, the O-helix undergoes a slight rotation, although it remains in a predominantly fingers-open position. However, a kink in the O-helix causes the terminally stacked tyrosine to unstack from the DNA duplex, allowing the templating base to move from the pre-insertion site, to the insertion site, where it forms hydrogen bonds with the incoming mismatching nucleotide, forming a "wobble-pair" (Fig. 3.4, fingers-partially-closed structure). This "ajar state"/"partially-closed state" (in reference to the O-helix being somewhat closed, but not entirely) was hypothesised as allowing the templating base to sample the incoming nucleotide for complementarity, *before* progressing to the catalytically active closed state. The same work also hinted at the importance of removing water molecules from the active site for correct nucleotide incorporation.[60] An illustrative movie of this change can be seen online.[ii]

Non-homologues enzymes show similarities in the three-dimensional structure of the active site used to catalyse phosphodiester bond formation, reflecting the similarity of the catalytic mechanism across polymerases. Phosphodiester bond formation between the incoming nucleotide and the 3'-OH of the primer terminus is catalysed by a two metal-ion mechanism,[61] with the catalytically active polymerase bound to two metal ions (A and B) liganded by conserved carboxylate residues in the palm of the polymerase. Metal B is hypothesised to arrive with the incoming nucleotide, bound to its phosphate tail, while Metal A arrives when the polymerase is in the fingers-closed conformation. In the closed conformation, the nucleotide α-phosphate is positioned to react with the 3'-OH of the primer terminus, ready to be incorporated into the DNA strand. Metal A forms contacts with polymerase residues, the nucleotide α-phosphate and primer terminus 3'-OH; preparing the

[i] http://www.pnas.org/content/suppl/2003/03/12/0630532100.DC1/0532Movie1Legend.html
[ii] http://www.jbc.org/content/286/22/19758/suppl/DC1

3'-OH for nucleophilic attack by lowering its pK_a. Metal B could be involved in stabilising the negative charge on the β-phosphate, and assisting in stabilising the pentacoordinated intermediate of the reaction.[61,58]

More recent work has implicated a third metal ion in the reaction, and illustrated the role of water in the deprotonation of the 3'-OH, with an impressive series of crystal structures demonstrating the formation of a phosphodiester bond by DNA polymerase η,[59] again with a illustrative movie online.[iii] It has been suggested that the rate-limiting step might be caused by a conformational rearrangement of the residues in the active site before catalysis, associated with the binding of the second magnesium ion.[65,58]

While crystal structures have provided invaluable information, it must be noted that fidelity is purely a kinetic property of polymerases, which is very difficult to extract from structural information alone.[66]

3.2.3 Bulk fluorescence investigations

The rapid-quench kinetic techniques described above, were able to establish a minimal mechanism for DNA polymerisation by KF. However, kinetic techniques can only infer the existence of conformational changes which are rate-limiting. Fluorescence techniques, meanwhile, can investigate more subtle conformational changes. For instance, 2-aminopurine (2-AP) is a fluorescent base analogue which can be used in place of natural bases given its ability to base-pair with thymine and form a DNA duplex.[67] 2-AP fluorescence is environmentally dependent, and is strongly quenched by stacking with neighbouring bases in a DNA strand, allowing changes in interactions with neighbouring bases to be observed via rapid changes in 2-AP fluorescence, observed using a stopped-flow instrument (Fig. 3.5a describes a stopped-flow apparatus).[52,68] Non-extendable "dideoxy-terminated" DNA substrates and can be used to observe the fluorescence changes preceding the chemical step, whereas changes during the full polymerisation reaction can be observed using normal "deoxy" primers. (Dideoxy DNA substrates prevent nucleotide incorporation by lacking the

[iii] http://www.nature.com/nature/journal/v487/n7406/abs/nature11181.html#supplementary-information

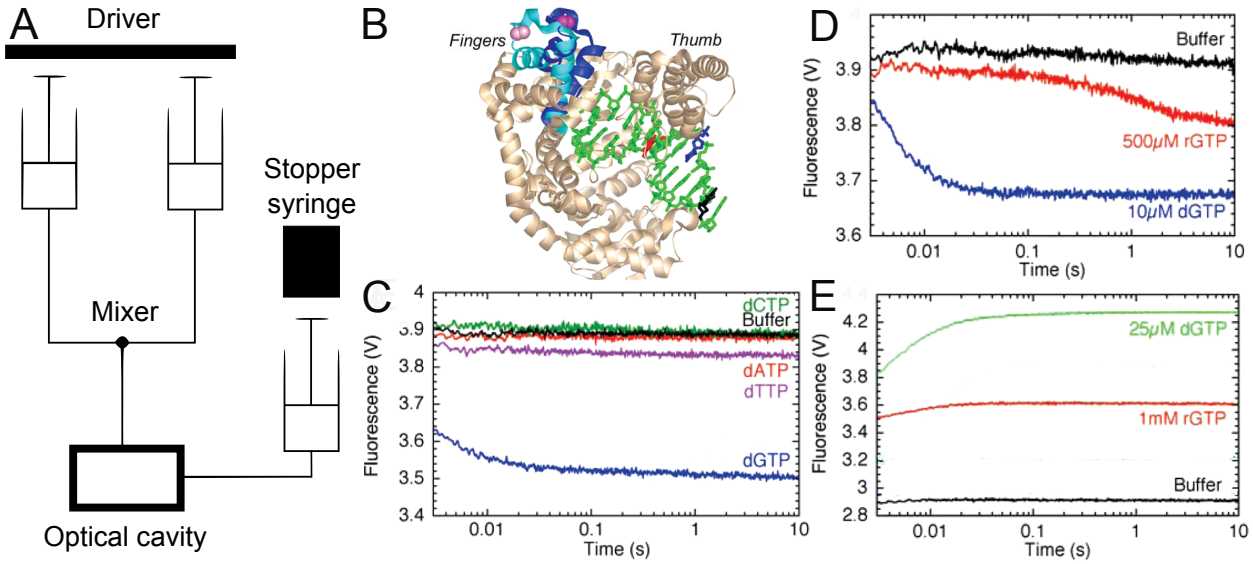


Figure 3.5: Inferring rapid conformational changes in KF via fluorescence. (a) Stopped flow apparatus used to observe rapid fluorescence changes. Syringe solutions are mixed using the driver, which then flow into the optical cavity after mixing. The delay between mixing and entering the optical cavity is the "dead-time" of the instrument. A stopper syringe halts the flow and begins the fluorescence observation. (b) FRET assay in which the KF fingers are donor labelled, while the DNA is acceptor (quencher) labelled, such that the donor fluorescence decreases on fingers-closing. (c) Donor fluorescence decreases rapidly, indicating fingers-closing on addition of complementary nucleotide (dGTP blue), but not in the presence of mismatching nucleotide (dCTP, dATP,dTTP) or buffer (black). Performed on a terminated DNA substrate, and allows extraction of the fast-fingers closing rate. (d) Fingers-closing is rapid on addition of complementary nucleotide, but change in fluorescence using complementary ribonucleotide is distinct from that seen using T(+1) 2-AP reporter. (e) Rapid fluorescence changes of a T(+1) 2-AP reporter on addition of KF and 25μM complementary nucleotide (dGTP - green); 1mM complementary ribo-nucleotide (rGTP - red); or buffer (black) when using a non-extendable primer, note that fluorescence changes occur for both ribo and deoxyribonucleotides. Adapted from Joyce 2008.[42]

3'-OH group on the primer terminus necessary for the chemical reaction).

Stopped-flow observations of 2-AP placed at the first templating base (T(0)) demonstrated a rapid fluorescence decrease during DNA polymerisation, which was ascribed to rapid nucleotide binding. (The change was faster than the time resolution of the instrument $> 1000s^{-1}$). A second fluorescence change had a rate ($= 40s^{-1}$) similar to the reaction rate and was ascribed to a large global transformation such as DNA translocation, which was determined by the rate-limiting step of the reaction.[68]

DNA polymerisation was also observed using a 2-AP molecule placed at the second templating base (T(+1)), which measured a rapid fluorescence increase on nucleotide addition,

although slower that the rapid change measured using T(0) 2-AP reporter (Fig. 3.5e shows a 2-AP fluorescence change in response to KF and complementary nucleotide on a non-extendable substrate).[68] A fluorescence increase was also observed whose rate ($= 59s^{-1}$) coincided with reaction rate, again limited by the rate-limiting step of the reaction ($= 46s^{-1}$).[68] Fluorescence changes on addition of mispairs were found to vary significantly dependent on the pair, suggesting a misalignment of the active site by the mispair.[68]

Linking the 2-AP fluorescence changes to specific structural changes was challenging, therefore the authors were unable to comment on the rate of fingers-closing, or isolate which conformational change in the reaction was rate-limiting. If the fluorescence changes are considered in the context of the recently discovered "partially-closed" structure (see above), they might be assigned to the observed rearrangement of the templating bases.

To address whether fingers-closing was rate-limiting for the polymerisation reaction, a FRET assay was devised to couple fluorescence changes to fingers-closing. A donor dye was attached to the fingers subdomain of KF, and an acceptor dye attached to the DNA, such that "fingers-closing" would quench the donor dye fluorescence via the increased acceptor dye proximity (although technically the placement of the dyes would also investigate DNA translocation) (Fig. 3.5b, FRET construct). A rapid decrease in fluorescence was seen on addition of complementary nucleotide, which was ascribed to the fingers-closing conformational change (Fig. 3.5c, biphasic fluorescence decrease (blue)). This demonstrated that the rate of "fingers-closing" ($=140s^{-1}$) was much faster than the polymerisation rate ($=47s^{-1}$), and hence not rate limiting for the reaction, as previously predicted.[42]

Both 2-AP T(+1) and FRET assays demonstrated rapid fluorescence changes of similar rates during polymerisation (rates of $> 220s^{-1}$ and $140s^{-1}$ respectively), however it was demonstrated that these fluorescence changes arose from different reaction steps. On complementary nucleotide addition, fluorescence changes occurred in both the 2-AP T(+1) and the FRET assays, while neither change occurred when using mismatching nucleotide.[42] However, the kinetics of the 2-AP T(+1) and FRET assays differed when using matching ribonucleotides (Fig. 3.5d shows slow fingers response to ribonucleotide (red) while Fig. 3.5d shows rapid response of 2-AP to ribonucleotide).[42] The inhibition of fingers-closing

by complementary ribonucleotide, but presence of DNA rearrangement step, suggested that matching ribonucleotides are discriminated against at the fingers-closing step, while mismatch discrimination occurs at the 2-AP T(+1) DNA rearrangement step. The authors suggested that base-discrimination is related to movement of the first templating base, sampling of the incoming nucleotide for complementarity (subsequently supported by discovery of the "partially-closed" structure). This extended previous minimal mechanisms, also demonstrating that fingers-closing rates are nucleotide dependent, and that mismatching nucleotide increased dissociation of the KF-DNA complex.[42]

The 2-AP T(0), 2-AP T(+1), and FRET assays demonstrated large fluorescence changes on nucleotide addition, however, additional low amplitude fluorescence changes were seen when using non-extendable DNA constructs. These low amplitude changes had rates similar to nucleotide incorporation[68,42], but were structurally difficult to interpret. The low amplitude changes were posited to relate to the rate-limiting step given the coincidence of the rates, and also that replacement of the Mg^{2+} metal ion cofactor with Ca^{2+}, which prohibits chemistry, led to a significant change in the low amplitude rate (from $\sim 20s^{-1}$ to $\sim 3s^{-1}$). This suggested that the rate-limiting step of the reaction may involve binding of the metal cofactor or some conformational rearrangement of the active site residues after metal ion binding, but with caveats due to the low amplitudes of the fluorescence changes.[42]

3.2.4 Single-molecule fluorescence investigations of KF

In Chapter 2 we discussed techniques used to observe single-molecules, and their advantages. Here we present some relevant single-molecule investigations of KF. Early work observing single-molecule DNA polymerisation used TIRF microscopy to observe FRET changes of a surface-immobilised DNA duplex.[69] A donor-quencher dye pair were attached to opposite strands of the duplex, such that the donor fluorescence increased as KF performed strand-displacement synthesis, separating the DNA strands and the dyes. Rapid donor fluorescence increases were seen, interpreted as rapid polymerisation, as well as fluorescence increase in "steps" which were interpreted as pausing events, lasting $\sim 13s$ for KF (Fig. 3.6a, assay and fluorescence changes). These pauses were found to be sequence related, and could account

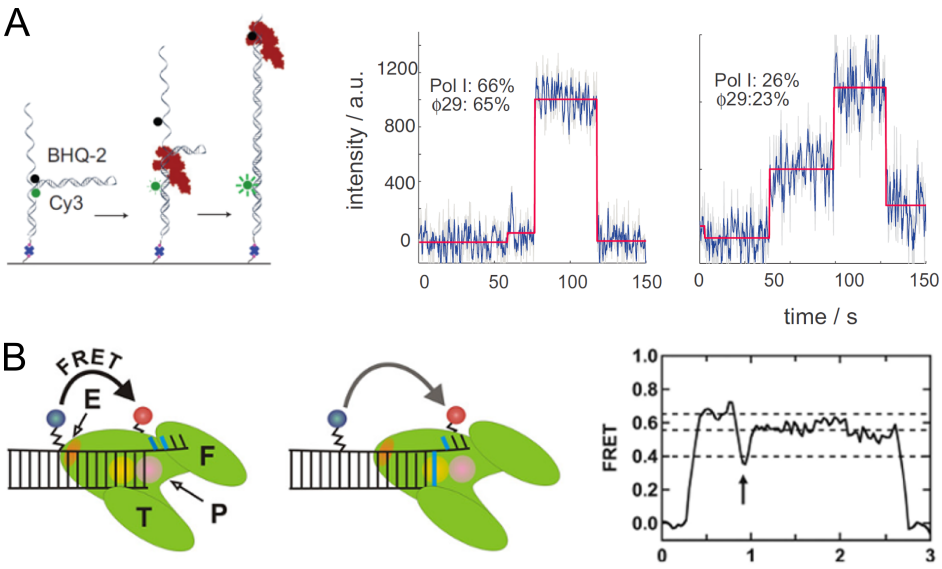


Figure 3.6: Single molecule fluorescence investigations of KF. (a) Donor-quencher (Cy3-BHQ-2) FRET pair on opposite sides of a DNA hairpin duplex are separated as KF performs strand displacement synthesis on a primer-template DNA. Fluorescence increases caused by the dye separation are rapid (middle panel) or occur in steps (right hand panel). (b) DNA-KF FRET pair monitors the position of KF on DNA as it performs polymerisation. FRET decreases from 0.65 to 0.55 as it translocates during a single nucleotide incorporation, however with an unexpected low FRET state inbetween (0.4). Adapted from Schwartz 2009 and Christian 2009.[69,72]

for the slow strand-displacement polymerisation rates observed in the literature ($1.2nt/s$)[70] Surprisingly, the corrected strand-displacement polymerisation rate ($14nt/s$) was similar to previously measured single-stranded primer extension rate, despite the expected difference in energetics based on the need to dehybridise the duplex DNA during strand-displacement synthesis.[70,43] On a cautionary note, unusual photophysics were observed in the paper but not explained, and the contribution of the known photophysics of Cy3[71] were not fully characterised (Fig. 3.6a, note the fluorescence intensity changes in the right hand panel).[69]

This study was followed shortly after by another surface-based investigation, which monitored the position of KF on DNA via a FRET pair on the DNA (template -10) and KF palm domain (Fig. 3.6b, assay). This work observed FRET decreasing in steps as KF incorporated nucleotides and translocated along the DNA. A relatively slow polymerisation rate was observed = $1.2s^{-1}$, and a novel state was observed during incorporation. For instance, on extension from a 15-mer ($FRET \sim 0.68$) primer to a 16-mer ($FRET \sim 0.62$), a low ($FRET \sim 0.47$) state was observed between these states, which was hypothesised

as a fidelity checking step, with the polymerase checking the base-pairing of the incorporated nucleotide, then shifting to the exonuclease domain to excise the base if necessary (Fig. 3.6b, sample FRET changes during single nucleotide incorporation). The state was relatively short lived ($\approx 270ms$), and did not contribute significantly to the incorporation time.[72] (Subsequent work on a similar construct demonstrated that KF-DNA binding is stabilised by the presence of complementary nucleotide,[73] and that KF binds carcinogenic adducts with distinct kinetics and conformations.[74])

Soon after, a single-molecule study of KF was performed by the Kapanidis lab, in which the fingers conformation of KF was directly monitored via an intra-protein FRET pair, with the mobile fingers sub-domain labelled with a donor dye (Cy3B), and the static thumb sub-domain labelled with an acceptor dye (Atto647N). The intra-protein dye pair allowed direct monitoring of fingers conformation, without the convolution of KF-DNA movements or translocations seen in KF-DNA dye pairs. Rather than TIRF, a single-molecule diffusion microscope was used to monitor the conformation of KF via the FRET values emitted from bursts of fluorescence as KF diffused rapidly ($\approx 1ms$) through a focused laser-beam (see Section 2.5.1 for a description of a confocal microscope). These FRET bursts reported the KF conformation and were plotted in a histogram. In the KF-DNA binary complex, KF was found predominantly in a low FRET state (E*=0.5) which was assigned to the "fingers-open" conformation, in correspondence with the crystal structures (Fig. 3.7b). Similarly, when in the KF-DNA-dTTP (complementary nucleotide) ternary complex, KF was found predominantly in a high FRET state (E*=0.7), assigned to the fingers-closed conformation (Fig. 3.7c). The unliganded KF existed in both open and closed conformations fairly equally (Fig. 3.7a).

Surprisingly, both binary and ternary complexes exhibited small population in the unexpected FRET state (Fig. 3.7b, binary complex has small fingers-closed population (E*=0.7). Fig. 3.7c, ternary complex has small fingers-open population (E*=0.5).). Whether KF was dynamically interconverting between two structures (dynamic heterogeneity), or whether the two peaks were formed from distinct species (static heterogeneity) was unclear, given it was difficult to directly resolve the interconversions between the open and closed struc-

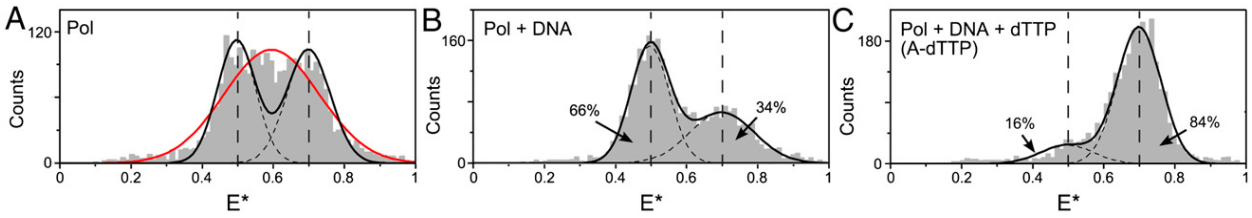


Figure 3.7: Histogram of E* values of KF complexes which have diffused rapidly through a focused laser beam under different conditions. (a) E* histogram of the unliganded KF, which demonstrates a broad distribution of E* values. (b) E* histogram of the KF-DNA binary complex, which has a large population around E*=0.5 (fingers-open), with a smaller population at E*=0.7 (fingers-closed) (c) E* histogram of the KF-DNA ternary complex, which has a large population E*=0.7 (fingers-closed), with a smaller population at E*=0.5 (fingers-open). Adapted from Santoso 2010.[19]

tures because of the short observation times. A variety of techniques were used to indirectly demonstrate that the unliganded KF was interconverting dynamically between the open and closed states, extending the static images presented by the crystal structures, demonstrating KF to exist in a dynamic equilibrium between the open and closed conformations, with ligands biasing it towards a certain conformation. The dynamics of binary and ternary conformations were estimated using the standard deviation of the FRET bursts (Burst Variance Analysis - BVA[75]) and by analysis of the shape of the FRET distribution histograms (Probability Distribution Analysis - PDA[76]).

This KF construct also demonstrated a small conformational change in the presence of mismatching nucleotides and matching ribonucleotides, with the fingers moving into a "partially closed" conformation, inferred via the small shift in the open state peak position from E*=0.5 to E*=0.55. Given it occurred with both mismatching nucleotide and ribonucleotide, it must occur before the DNA rearrangement step which is inhibited by mismatching nucleotide but allowed for matching ribonucleotide. The updated reaction scheme is seen in Figure 3.8. The existence of a "partially closed" structure was supported in other single-molecule experiments, as well as crystal structures.[77,60]

Single-molecule investigations of DNA polymerisation under tension have also been performed using magnetic tweezers, extracting a single-stranded polymerisation rate of ∼15bp/s for KF.[78] Recent investigations have used the variations in current through nanopores trapped with KF to infer DNA and nucleotide binding, sequence detection, nucleotide dis-

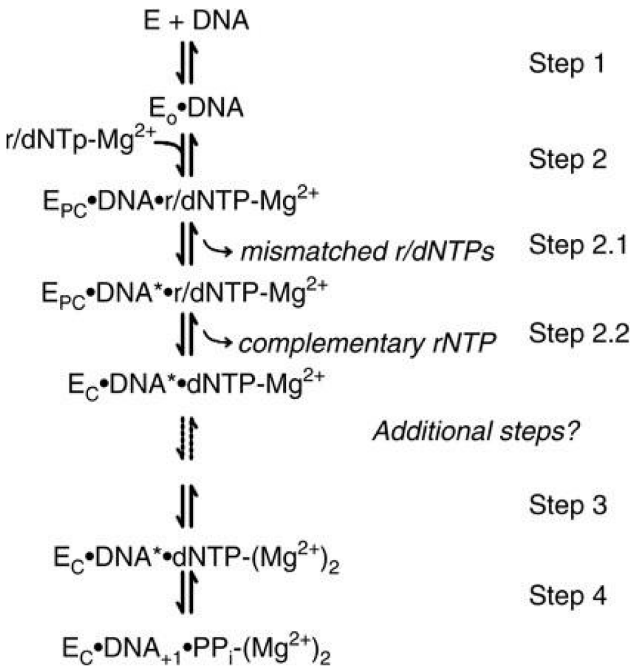


Figure 3.8: Reaction scheme for DNA polymerisation by KF. Step 1: KF (E) binds DNA. Step 2: The open binary complex (E_O.DNA) binds nucleotide, and enters the partially closed conformation (E_{PC}). Step 2.1: Mismatched nucleotides or ribonucleotides are rejected, but matching ribonucleotides or nucleotides progress to the DNA rearrangement step (DNA*). Step 2.2: Fingers-closing (E_C) occurs for matching nucleotide, but not matching ribonucleotide. Step 3 represents the unknown rate limiting step before catalysis, including binding of the second magnesium ion. Step 4: Catalysis occurs, leading to nucleotide incorporation, translocation and a PP_i product. Adapted from Hohlbein 2013.[20]

crimination and observations of replication,[79,80,81,82] and changes in KF conformation have been inferred using the change in current through a nanotube.[83] A single molecule investigation into T7 DNA polymerase demonstrated interconversions between conformations using an environmentally sensitive dye attached to the protein, although the structural origin of these changes was unclear.[84]

3.3 Theories of DNA polymerase fidelity

Although there exists no all-encapsulating theory describing the origin of DNA polymerases extraordinary fidelity, there are a number of hypothesis. For instance, it has been frequently stated that error rates as low as 1:10000 cannot simply be accounted for by thermodynamic energy differences between Watson-Crick base-pairs, and mismatching base-pairs, which could only contribute a factor of 2-100 to fidelity.[39,46,85,47,86]. (Despite the frequent use of this argument in the literature, recent work has contested this fact, suggesting that the energetic differences between correct and incorrect base-pairs are actually fairly significant, with confusion arising due to the method used to measure the ΔG of incorporation: directly, or indirectly by comparing melting energies of different DNA sequences.)[87,88,89] However, polymerases could amplify the enthalpic and entropic differences between correct

and incorrect nucleotides by excluding water molecules from their active sites.[87,89]

Crystal structures have demonstrated that the catalytic binding pockets of family A, B, X and RT polymerases tightly encapsulate the nascent base-pair, suggesting that size and shape are important in nucleotide selection, with steric clashes halting the polymerisation reaction for mismatches. This is supported by the structural similarity of A-T and G-C matches, and the distinct conformations of mismatched pairs.[39] Base-analogues which match the shape and size of natural bases, but lack their hydrogen bond ability, have been found to be efficiently incorporated by polymerases, suggesting a stronger role for size and shape than Watson-Crick base-pairing in nucleotide selection. Mispairs which are too small to sterically clash, such as pyrimidine-pyrimidine pairs, could demonstrate mis-shapen geometry by allowing water molecules into the active site. Polymerase variants, such as lesion bypass polymerases (family Y polymerases) impose less strict geometric requirements, and interestingly, also demonstrate lower fidelities, further supporting the role that geometry plays in fidelity.[89,90,91] Structural analysis using mutagenesis has suggested that residues which bind the DNA minor groove are particularly important in fidelity, (although residues distal to the nascent base-pair have also been found to influence fidelity).[91,39]

Fidelity is defined by the ability of the polymerase to incorporate complementary nucleotides over mismatches. While rates of correct incorporation span a wide range of values over different polymerases, mismatching incorporation rates do not vary so much. This suggests that the broad range of fidelities exhibited by polymerases are achieved largely via enhancing the rate of correct nucleotide incorporation, rather than slowing the rate of mismatch incorporation. This highlights the importance of understanding the polymerisation mechanism for correct nucleotide incorporation, in order to understand the origin of its error rate.[92]

It has been proposed that fidelity is caused by "induced fit" conformational changes which activate the polymerase, which occur on binding of complementary nucleotide, but not on mismatched nucleotide binding. However, transition state theory has demonstrated that k_{cat}/K_m (which defines specificity and fidelity) is a function of the rate-limiting step of the reaction, implying that conformational changes have to be rate-limiting to influence

fidelity.[45] While a pre-chemistry rate-limiting step has been inferred in KF for correct nucleotide incorporation, it is not the fingers-closing conformational change, and its origin remains unknown. Additionally, during mismatch incorporation, the chemistry step, not a conformational change, has been shown to be rate-limiting, resulting in a paradoxical situation in which the fidelity of the polymerase would be improved if the conformational change for correct nucleotide incorporation were removed.[47]

Another proposed fidelity mechanism is that the polymerase uses multiple steps, or "kinetic checkpoints", to screen for mismatching nucleotides during the reaction pathway, each of which contribute "partial fidelities".[46] Others have argued that fidelity is solely defined by the differences in the energies of the transition states for the incorporation and misincorporation complexes, and that kinetic checkpoints conceptually overcomplicate matters. They argue that kinetic checkpoints appear simply as a consequence of the different transition state energies being reflected in the early stages of the reaction - i.e. correct nucleotides may bind to the polymerase more efficiently than mismatches, but this does not constitute a checkpoint, as the increased binding is just reflecting the polymerases preference for substrates which are similar to the desired transition state.[46]

> *By this view, the fact that the structure of the enzymes active site, particularly in its closed conformation, is intrinsically complementary to the chemical transition state of a Watson-Crick base pair is the source of all other properties, including the appearance of partial fidelities at individual steps... we believe that it more accurately emphasizes the true source of polymerase fidelity.* - ***Showalter and Tsai (2002)***[47]

Proponents have countered that this is at odds with experimental results which suggest that phosphoryl transfer is not the rate-limiting step in the reaction, reducing the relevance of the chemical transition state. However, if this turns out to be incorrect, then even if kinetic checkpoints do not enhance fidelity, they do define the operational mechanism of fidelity, and assist reaction throughput by rejecting mismatches early in the reaction pathway.[46]

3.3.1 "A new paradigm for DNA polymerase specificity" - T7 DNA polymerase

Recent investigations into T7 DNA polymerase, combined with mathematical analysis, have provided a compelling alternative approach to analysing fidelity. These bulk fluorescence investigations used an environmentally sensitive dye to monitor a conformational change in T7 DNA polymerase, which suggested that the polymerase adopts distinct structures in response to complementary nucleotide or mismatch binding, inferred via the decreases or increases in fluorescence on their addition. This suggested that the binding energy of the mismatching nucleotide was being used to actively misalign the residues of the active site, slowing the rate incorporation of mismatches.

The structural interpretation of the conformational change on complementary nucleotide addition was unclear, however it was thought to be linked to the fingers-closing change, because of the position of the dye, so we refer to it hereafter as "fingers-closing".[93]

The kinetics of "fingers-closing" were investigated using stopped-flow apparatus, extracting a "fingers-closing" rate, and a "fingers-opening" rate (via a competition assay in which unlabelled DNA-Pol was added to sequester nucleotides released from the ternary complexes, with the fluorescence change measuring the rate at which the DNA-Pol-dNTP complex released complementary nucleotide). "Fingers-closing" was found to be rapid on complementary nucleotide addition, however the reverse conformational change, "fingers-opening" was slow. Although the structural interpretation was difficult, the investigation demonstrated that the polymerase bound nucleotide very tightly by "closing" rapidly around it, and remaining closed by having a slow "opening" rate.[93]

Intuitively, one can imagine that if reverse "opening" rate was much slower than the forward rate of catalysis, then each "closing" event would be followed by an incorporation, thus the rate of chemistry becomes irrelevant to specificity. This idea emerged from a rigorous mathematical analysis of a multi-step polymerisation reaction, demonstrating that the catalytic efficiency constant (k_{cat}/K_m) becomes independent of the rate-limiting step if this step is much faster than the reverse step, "fingers-opening". The ratio of the catalytic

efficiencies for correct and mismatching incorporation define the polymerases discrimination:

$$D = \frac{(k_{cat}/K_m)_{match}}{(k_{cat}/K_m)_{mismatch}}$$

The removal of the rate-limiting step from considerations of fidelity was surprising, given it was counter to decades of investigations which had been attempting to isolate the rate-limiting step in order to understand fidelity. Thus, provided the ratio of chemical incorporation to fingers-opening rates is high, fidelity becomes independent of the rate-limiting step of the reaction.

Additionally, the catalytic efficiency was found to be a complicated function of the various reaction rates. This demonstrated that analyses based on inspection of free energy barrier heights for incorporation and misincorporation,[46,47] in which fidelity was based on the single rate-limiting steps, were overly simplistic. Arguments that multiple-step mechanisms cannot enhance the fidelity were criticised as relying on over-simplifications such as an equilibrium between conformational changes and substrate binding.[45,93]

In this thesis, we argue that KF demonstrates a similar kinetic profile for fingers-closing to T7 DNA polymerase, allowing the same analysis to be applied, demonstrating that fidelity is not defined by the rate-limiting step of the reaction.

Chapter 4

An assay to monitor the conformational dynamics of DNA polymerase I (KF)

Here we present an assay to monitor the conformational dynamics of single DNA polymerase I (KF) molecules via an intra-protein FRET pair, as KF transiently binds to surface-immobilised DNA. In this chapter we characterise the assay, quantifying KF-DNA binding relative to non-specific KF-surface binding, and compare our assay with a diffusion-based assay, demonstrating that the surface environment has little adverse effect on the system. We also quantify the effect of acceptor dye photophysics on FRET efficiency values, such that the assay can be extended in the future to observe smaller FRET changes. Furthermore, we present observations suggesting that KF may bind DNA as a dimer. Fingers conformational dynamics are investigated in the next chapter.

4.1 Introduction

Crystal structures have demonstrated that DNA polymerase is shaped like a right-hand, and undergoes a large scale conformational change upon binding of complementary nucleotide, in which the fingers sub-domain closes around the incoming nucleotide, forming a binding pocket around the nascent base-pair (Fig. 4.1a, "fingers-open" conformation in the KF-DNA binary complex (orange), and the "fingers-closed" conformation in the KF-DNA-dTTP

ternary complex (blue)).[58] The contribution of fingers-closing to fidelity, which is a purely kinetic property, has been investigated by observing the kinetics of fingers-closing (other structural changes such as DNA rearrangement have also been investigated).

The kinetics of "fingers-closing" have been investigated in an ensemble assay using energy transfer between a donor dye on the KF fingers and an acceptor dye on the DNA, demonstrating fingers-closing to be rapid upon nucleotide binding ($k_{close} = 140s^{-1}$), and faster than the rate-limiting step of nucleotide incorporation ($k_{pol} = 47s^{-1}$) (see Section 3.2.3).[68,65] Fingers-closing does not occur in the presence of mismatching nucleotides or ribonucleotides, and is thought to act as a fidelity checkpoint for the rejection of matching ribonucleotides. Although fidelity checkpoints are important, fidelity is hypothesised to arise from enhanced complementary nucleotide incorporation rates rather than rejection of mispairs, making it important to quantify the kinetic contribution of fingers-closing to complementary nucleotide incorporation rates.[94]

"Finger-closing" was further investigated in a collaboration between the Kapanidis and Grindley (Yale) labs, which resulted in a KF construct in which the fingers and thumb subdomains were labelled with a donor-acceptor FRET pair (Fig. 4.1a, donor/acceptor dyes indicated by green/red spheres respectively). This intra-protein FRET was a more direct reporter of fingers-closing, removing convolutions of KF-DNA motions. In addition, this construct allowed observation of the unliganded KF conformation in the absence of DNA.[19]

This novel KF construct was investigated using a single-molecule diffusion techniques, allowing brief snapshots ($\sim 1ms$) of KF conformation via the FRET efficiency observed as KF diffused through a focused laser beam (for more details on these experiments, see Sections 2.5.1 and 3.2.4). These observations revealed KF to exist in two conformations, predominantly in the fingers-open conformation in binary complex (KF-DNA), and predominantly in the fingers-closed conformation in the ternary complex (KF-DNA-dTTP), consistent with crystal structures.[58] However, small occupancies were seen in the opposite state: the binary complex showed some closed KF, and ternary complex showed some fingers-open KF (Fig. 3.7). Such heterogeneity was previously undetectable, but given the short observations times, it was unclear whether KF was dynamically interconverting between open and closed

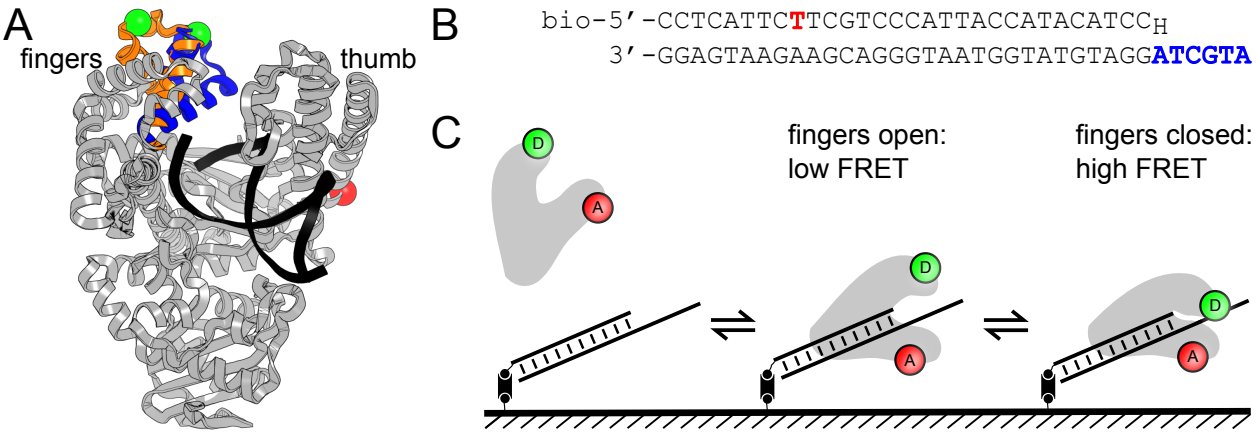


Figure 4.1: (a) The fingers closing conformational change in crystal structures of B. Stearothermophilus (Bst.) DNA polymerase, a close homologue of Klenow Fragment. The fingers sub-domain is in the open conformation (orange) in the Pol-DNA binary complex (PDB 1L3U), and in the closed conformation (blue) in the Pol-DNA-dNTP ternary complex (PDB 1LV5). The protein backbone is shown in grey, and the DNA strands in black. The Bst. structure is shown labelled at positions equivalent to the Pol I (KF) labelling positions, with Cβ atoms shown green on the mobile fingers portion and red on the static thumb, to represent Cy3B and Atto647N labelling positions respectively, which are separated by $\approx$ 52Å and $\approx$ 42Å in the open and closed structures respectively.[96] (Labelled residues are 744 and 550 respectively in KF). (b) Base-paired primer-template oligonucleotide used in these experiments. The -22 base on the primer strand is labelled with a Cy5 dye (red) for localisation, and the 5′ terminus is biotinylated. The 3′ terminal base is dideoxy modified (H) to prevent primer extension. The templating bases are shown in blue. (c) Schematic illustrating the single-molecule assay used to transiently immobilise KF. Dye labelled DNA primer-templates are immobilised via a biotin-neutravidin-biotin interaction onto a glass surface. TIRF microscopy is used to localise and measure subsequent binding of solution based KF. PEG passivation of the surface minimized non-specific adsorption.

conformations, or if sub-populations of molecules were stuck in either conformation (i.e. it was difficult to determine if the heterogeneity was static or dynamic).

The presence of dynamics was indirectly inferred using the standard deviation within the E* bursts[75], and the statistical distribution of E* values[76], and gave estimates on interconversion rates between the conformations. However these results were revised significantly after an optimised labelling procedure altered the static heterogeneity within the KF sample, by reducing the amount of reverse labelled KF present (it was desirable to minimise the presence of the fingers-acceptor and thumb-donor reverse labelled KF, as it demonstrated a low activity and low dye anisotropy)[19,20]. Further attempts to extract the interconversion rates of KF, extended single-molecule diffusion experiments into gels in order to slow diffusion.[95]

The lack of direct demonstration of fingers dynamics of KF, and the large variation in estimated interconversion rates, necessitated the development of an assay which would allow direct, extended observations of KF, in order to directly confirm the presence of fingers dynamics, and extract accurate interconversion rates if appropriate.

Here we present extended observations of KF fingers conformations through an assay in which KF transiently binds to surface-immobilised DNA molecules, and is observed using total internal reflection fluorescence microscopy (TIRF) (Fig. 4.1c). This assay allowed observations of KF interconverting between the fingers-open and fingers-closed states, which we investigate further in the next chapter. In this chapter we characterise this assay, by investigating the use of surface-immobilised DNA to transiently immobilise KF from solution, and complications caused by transient binding of KF to the surface. We also characterise the effect of red dye photo-physics on FRET values, and present evidence on the possibility of KF dimerisation. Furthermore, we compare our surface-based assay with a diffusion-based assay to demonstrate that the surface environment has little negative effect on the system.

4.2 Results and discussion

4.2.1 An assay to monitor the conformational dynamics of Klenow Fragment

A previous assay monitored equilibrium distribution of KF fingers conformations via intra-protein FRET and brief (1ms) KF observations (Fig. 4.1a, doubly-labelled KF construct).[19,20] We extended this assay to observe KF fingers-conformation for extended periods ($\sim 1s$), by using surface-immobilised DNA molecules to bind and transiently immobilise KF molecules from solution, thus permitting observation of the KF fingers-dynamics (Fig. 4.1c, assay schematic).

We used PEG passivated glass surfaces to minimise non-specific adsorption of KF, with 1:100 of the PEG molecules attached to a biotin molecule. Neutravidin was then bound and immobilised by these biotin molecules, which in turn bound and immobilised the biotinylated DNA molecules. The surface-immobilised dye-labelled DNAs were then localised using

TIRF microscopy, and were dideoxy terminated to prevent primer elongation. (Fig. 4.3a provides a rough example of what a TIRF field of view looks like). Fluorescence intensities at each DNA then reported on the binding of labelled-KF from solution, quantified by fitting Gaussian profiles to point-spread-functions (PSFs) of the DNAs imaged on the EMCCD. Fluorescence intensities at DNA molecules demonstrated stochastic increases/decreases across time, assigned to binding of KF molecules in the region of the DNA (Fig. 4.2a, donor and acceptor fluorescence under donor excitation (DD/DA respectively), and acceptor fluorescence under acceptor excitation (AA) show increases/decreases in intensity). (See Materials and Methods for details on surface-immobilisation and data extraction.)

The KF dye labelling protocol resulted in donor-acceptor, donor-only and acceptor-only labelled KF molecules[20], which we were able to distinguish as separate species via increases in the associated fluorescence channels. Binding of donor-acceptor labelled KF was observed via increases in DD, DA and AA (Fig. 4.2a, KF^{DA} at $\sim$ 10.5s); binding of donor-only labelled-KF via increases in DD only (Fig. 4.2a, KF^{D} at $\sim$ 26s); and binding of acceptor only labelled-KF via increases in AA only (Fig. 4.2a, KF^{A} at $\sim$ 40s). Doubly labelled polymerases accounted for $\sim 60\%$ of the binding events, with the singly labelled polymerases contributing the remaining $\sim 40\%$.

Dwell times of bound KF molecules were extracted by thresholding for intensity values, and demonstrated biexponential dissociation kinetics (Fig. 4.2b, histogram of binding dwell times) (see Materials and Methods). Biexponential fits extracted a fast and a slow KF dissociation rates ($k_{off-fast} = 3.3 \pm 0.3 s^{-1}$, $k_{off-slow} = 0.22 \pm 0.03 s^{-1}$). While these biexponential kinetics could have been accounted for distinct KF-DNA binding modes, investigations demonstrated these modes to be caused by KF-surface binding, and KF-DNA binding respectively (Section 4.2.2).

At each KF-DNA binding event, the conformation of the KF fingers was reported by the FRET efficiency at each frame:

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

E* was seen to interconvert between E*=0.4 and E*=0.6 in the presence of complementary nucleotide (Fig. 4.2c, example KF-DNA binding event and E* values). These E* values were

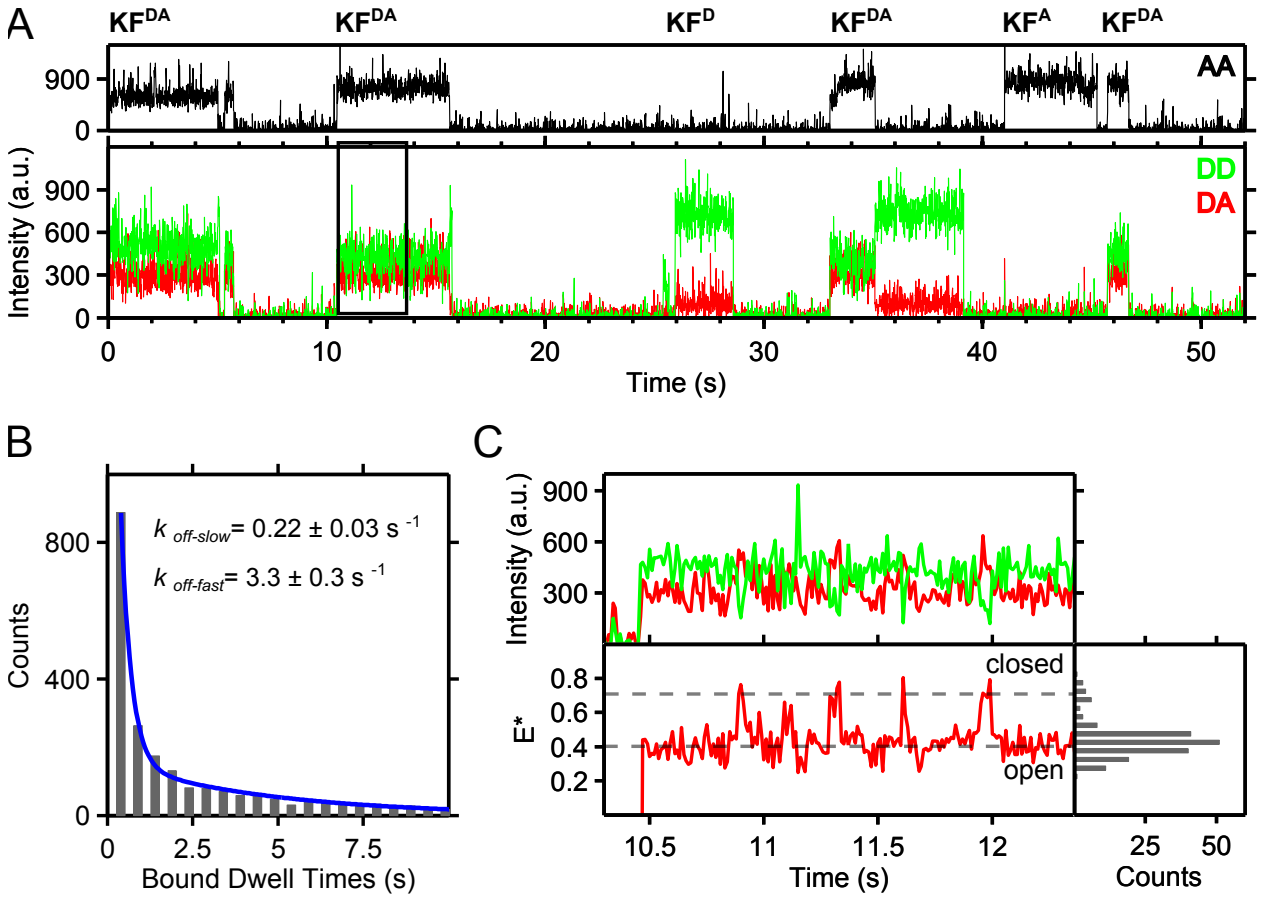


Figure 4.2: (a) Fluorescence intensities in the region of a localised DNA molecule demonstrate repeated polymerase binding and dissociation events. The upper panel shows acceptor fluorescence under acceptor excitation (AA - black trace), allowing monitoring of the acceptor dye fluorescence. The bottom panel shows donor and acceptor fluorescence under donor excitation (DD/DA respectively). Increases in fluorescence intensity are interpreted as KF binding events. Binding events are indicated by KF^X where X is a list of the dye labels on the molecule (D=donor, A=acceptor). Acceptor dye bleaching at the fifth binding event results in an increase in donor fluorescence, characteristic of a single molecule event. (b) Bound KF dwell times reveal bimodal dissociation kinetics. Binding events were defined by thresholding for fluorescence intensity. The dwell times fit to a double exponential, with fast and slow dissociation rates. (c) Expanding the binding event seen in the black box in a., with the fluorescence trace in the upper panel and corresponding FRET ratio (E*) in the lower panel. Clear interconversions between two FRET values at E* ≈ 0.4 and E* ≈ 0.6 are seen. Anti-correlation is seen in the DD and DA signals, corresponding with the FRET changes, whilst the AA signal shown in a. remains constant. A histogram of E* values for each frame is shown, with the KF observed predominantly in the low FRET state. Data for a and c were collected using Alternating Laser Excitation (ALEX)[29] at 10ms frame time, with laser powers of 1.5mW and 3mW for red and green lasers respectively, with 4nM KF and 0.125μM dTTP. Dwell times of KF-DNA complexes in b were collected using continuous wave (CW) excitation with a green laser power of 0.3mW and 100ms frame time, with a KF concentration of 4nM, the low green power used to minimize photo-bleaching.

assigned to the fingers-open and fingers-closed conformations respectively, in accordance with the open and closed states seen in the diffusion based work, which had values at E*=0.5 and E*=0.7 respectively. The difference in E* values was ascribed to the differing spectral characteristics of the microscopes, while ΔE* was consistent, supporting the surface assay as a valid environment for the investigation of the protein.[20] We investigate the fingers interconversion rates in the next chapter.

4.2.2 Transient immobilisation of KF

KF binding in the region of localised DNA molecules demonstrated biphasic dissociation kinetics (Fig. 4.2b), which could be accounted for either by two KF-DNA binding modes, or a single KF-DNA binding mode and non-specific adsorption (NSA) of KF to the surface. Our surface preparation protocol used poly-ethylene-glycol (PEG) to passivate the glass surfaces, which we expected to minimized KF NSA. Subsequent to PEG-ylation, the surface was primed to immobilise DNA via treatment with neutravidin, which then bound to biotinylated PEG molecules. The biotinylated DNA was then immobilised via binding to the immobilised neutravidin.

The effectiveness of our slide passivation was measured by taking 80 second images of the different surface environments during slide preparation, with KF diffusing in solution, free to bind the surface. The surfaces included: a PEG glass surface; a PEG glass surfaces treated with neutravidin; and a PEG glass surfaces treated with neutravidin and with immobilised DNAs. PEG surfaces were found to efficiently prevent KF NSA (Fig. 4.3a, left hand panel, surface image shows low number of point-spread functions (PSFs), indicating low KF binding). As expected, immobilised DNA enhanced KF binding significantly. (Fig. 4.3a, right hand panel, surface shows large number of PSFs). Surprisingly, the neutravidin treated slide, prior to DNA immobilisation, showed increased KF-surface binding, but at lower levels than when immobilised DNA was present (Fig. 4.3a Contrast PSFs in middle (neutravidin) and right (neutravidin +DNA) hand panels).

These results indicated that PEG passivation efficiently suppressed NSA, and that KF bound predominantly to DNA, however, there was also non-DNA binding enhanced by the

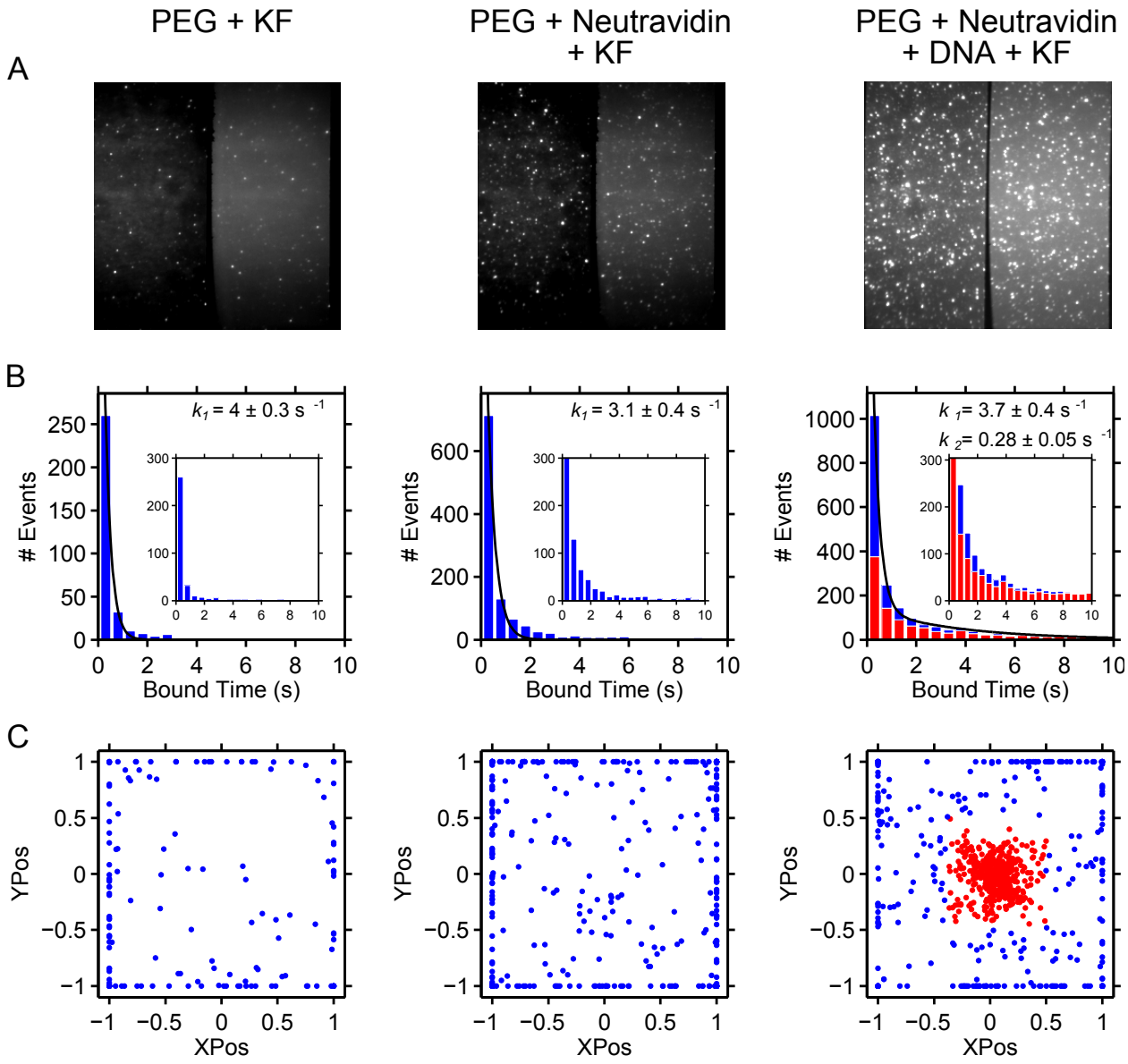


Figure 4.3: Transient surface-immobilisation of KF to surfaces treated with: PEG (left); PEG + neutravidin (middle); and PEG+neutravidin+DNA (right). (KF was diffusing in solution above the surfaces) (a) Regions of a glass slide imaged over 80s, demonstrating the transient immobilisation of KF to these surfaces, with low KF binding to PEG, higher binding in the presence of neutravidin, and higher in the presence of DNA. (b) Histograms of KF dwells at random locations on PEGylated (left) and neutravidin treated slide (middle), and KF dwells at localised DNAs (right), collated from three FOVs each. Blue and red correspond with the localised positions specified below. (c) Localised positions of the transient KF dwells relative to the randomly localised positions on a PEGylated (left) and neutravidin treated slides (middle). Localised positions of transient KF dwells relative to localised DNA positions (right), with red points are assigned as KF binding to DNA (right). These investigations were performed using continuous wave (CW) excitation with a green laser power of 0.5mW with an 40ms frame time, with a KF concentration of 4nM, and sampled from one FOV.

presence of neutravidin. Given that $\sim 1:100$ PEG molecules are biotinylated, it is possible for there to be much more immobilised-neutravidin than DNA. Thus a surface coated with neutravidin could compete with the DNA for binding KF, even if it had a relatively weak affinity for KF.

The kinetics of KF binding were characterised for each surface by extracting KF binding dwell times at random locations in a FOV, or at localised DNAs, across an 80 seconds experiment, repeated three times. This demonstrated that the fast dissociation mode (Figure 4.2, $k_{off-fast} = 3.3 \pm 0.3 s^{-1}$) was caused by KF-surface binding, predominantly induced by the presence of neutravidin (Fig. 4.3b, $k_{off-fast} = 3.1 \pm 0.4 s^{-1}$ in the absence of DNA but presence of neutravidin (middle panel), with a higher occupancy than in the absence of neutravidin and DNA (left hand panel)). The slow dissociation mode was only seen in the presence of DNA, which we attributed to KF-DNA binding (Fig. 4.3b, right hand panel inset, dwells reach higher values than dwells in the left and middle panel insets). These dwell time histograms demonstrated that KF-DNA binding has distinct although overlapping kinetics with KF-surface binding.

We characterised the relative contribution of KF-DNA binding to KF-surface binding, by quantifying the positions of binding events for each surface (see Materials and Methods). This revealed non-DNA binding to be randomly distributed across space (Fig. 4.3c, left and middle panels, blue dots indicating the position of localised KF in relation to the randomized initial locations in one field of view (FOV)), while KF-DNA binding events were clustered around the localised DNA (Fig. 4.3c, right hand panel, dots indicating the position of localised KF in relation to initial DNA locations, with binding events clustering around the centre, where the DNAs were initially localised).

The central cluster of localisations were ascribed KF-DNA binding (Fig. 4.3c, red dots delineated by eye), and the rest to KF-surface binding (Fig. 4.3c, blue dots). The bound dwell times were linked to their binding position using this colour notation, demonstrating that KF-DNA binding dominates over KF-surface binding if events shorter than 0.5s are ignored (Fig. 4.3c inset, red bars indicate KF-DNA binding dwells, which dominate over non-DNA binding dwells (blue bars), which are short in duration). In this case, 80% of

the time KF is bound to a DNA molecule, rather than surface bound. We used this time thresholding method to minimise contributions from KF-surface binding during subsequent investigations, as rigorous elimination of KF-surface binding via thresholding for the distance from the localised DNA was not reliably applicable in our studies, due to the low XY stability of the TIRF microscopy. This instability would lead to sample drift, and thus limited the temporal accuracy of DNA localisation.

4.2.3 Evidence for KF dimerisation on DNA

Single-molecule techniques allow the observation of rare events which might otherwise be shrouded in the ensemble average. Such rare events were observed in our system, in which a small number of frames demonstrate twice the fluorescence intensity of nearby frames (Fig. 4.4a, upper panel, total fluorescence under donor excitation (blue trace - D_{Ex}=DD+DA) demonstrates a binding event at 24-25s with twice the intensity ($\sim$ 1630 a.u.) of a subsequent binding event ($\sim$ 870 a.u.) at $\sim$ 25.5s. Similarly in Fig. 4.4b, the binding event begins at a high fluorescence intensity ($\sim$ 2640 a.u.) and decreases to half that value ($\sim$ 1260 a.u.)). Both DD and DA are seen to change in intensity, although DD more so than DA.

A second intensity state of the donor dye (Cy3B) could explain the increase in total fluorescence, however, Cy3B intensity has been characterised as predominantly stable in previous investigations, with rare intensity fluctuations differing by only 15% in intensity, unlike the 100% intensity change seen above.[97] While the FRET efficiency is weakly dependent on donor quantum yield, our approximate measure of FRET efficiency, the proximity ratio E*, should not change, as DD both and DA are linearly proportional to donor quantum yields. E* should thus not change if donor quantum yield changes, however, we observe E* changing between the high and low D_{Ex} regions (Fig. 4.4a, see E* values at the arrow, and the change in E* upon the decrease in D_{Ex} fluorescence).[98]

The distinct doubling of the total fluorescence signal suggests the presence of a second KF molecule, which could be binding to the DNA, or increasing the local fluorescence by binding to the surface nearby the DNA. We would expect a KF molecule binding nearby to a localised DNA to shift the mean position of the point-spread-function (PSF), used to

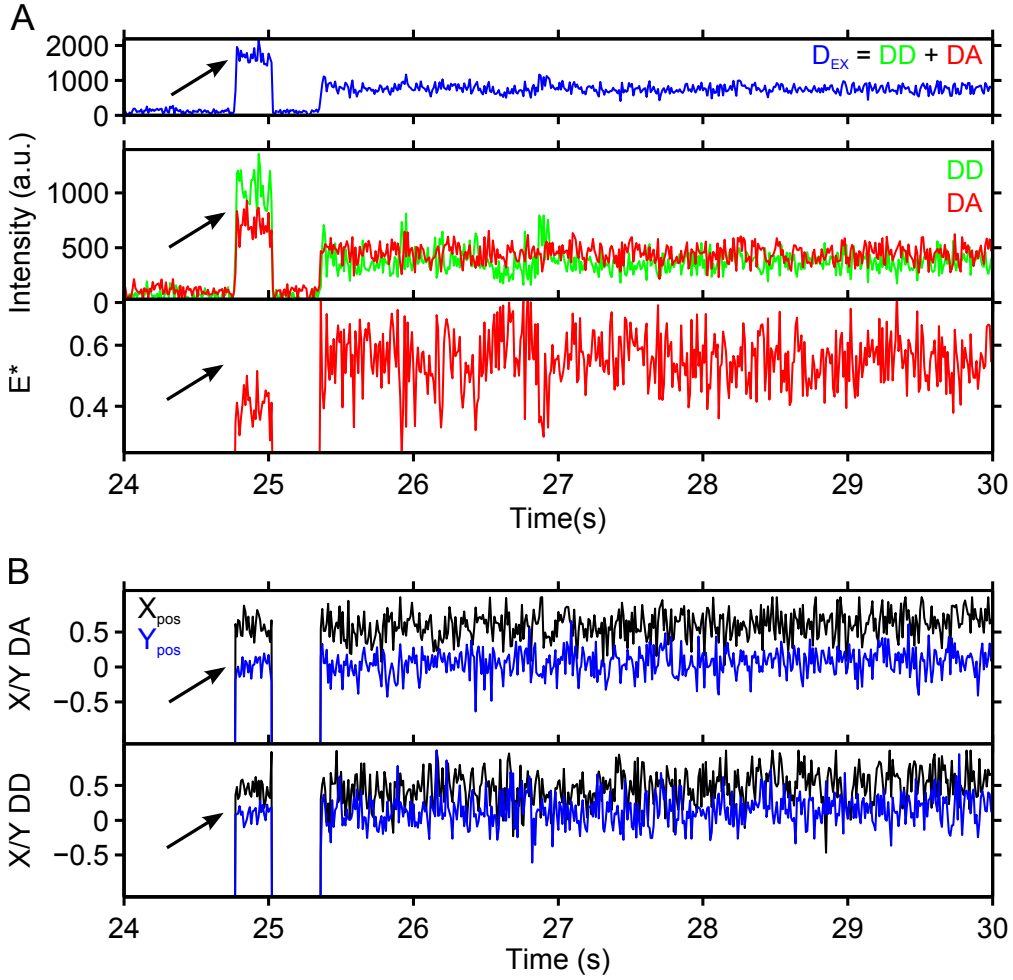


Figure 4.4: Variations in fluorescence intensities at KF-DNA binding events suggesting binding of two KF molecules. (a) Total fluorescence emission under donor-excitation at the first KF binding event is twice the intensity at the second binding event (D_{EX} - blue). Donor (DD) and acceptor (DA) emission under donor excitation, demonstrate similar differences in intensities. FRET efficiency differs between the first and the second binding event (E* - red). KF is predominantly closed because of the complementary nucleotide added in order to stabilise the KF-DNA complex. (b) X/Y positions of the localised PSFs (black/blue), from which the DD and DA values are extracted (upper/lower panel). Notice the lack of change in localised X/Y positions between the high and low fluorescence intensity regions. Y position is centred around the localised DNA position (zero), but X position is offset by 0.5 pixels, caused by drift in the microscope. Data was collected using continuous wave (CW) excitation with a green laser power of 0.3mW with an 100ms frame time, with a KF concentration of 4nM.

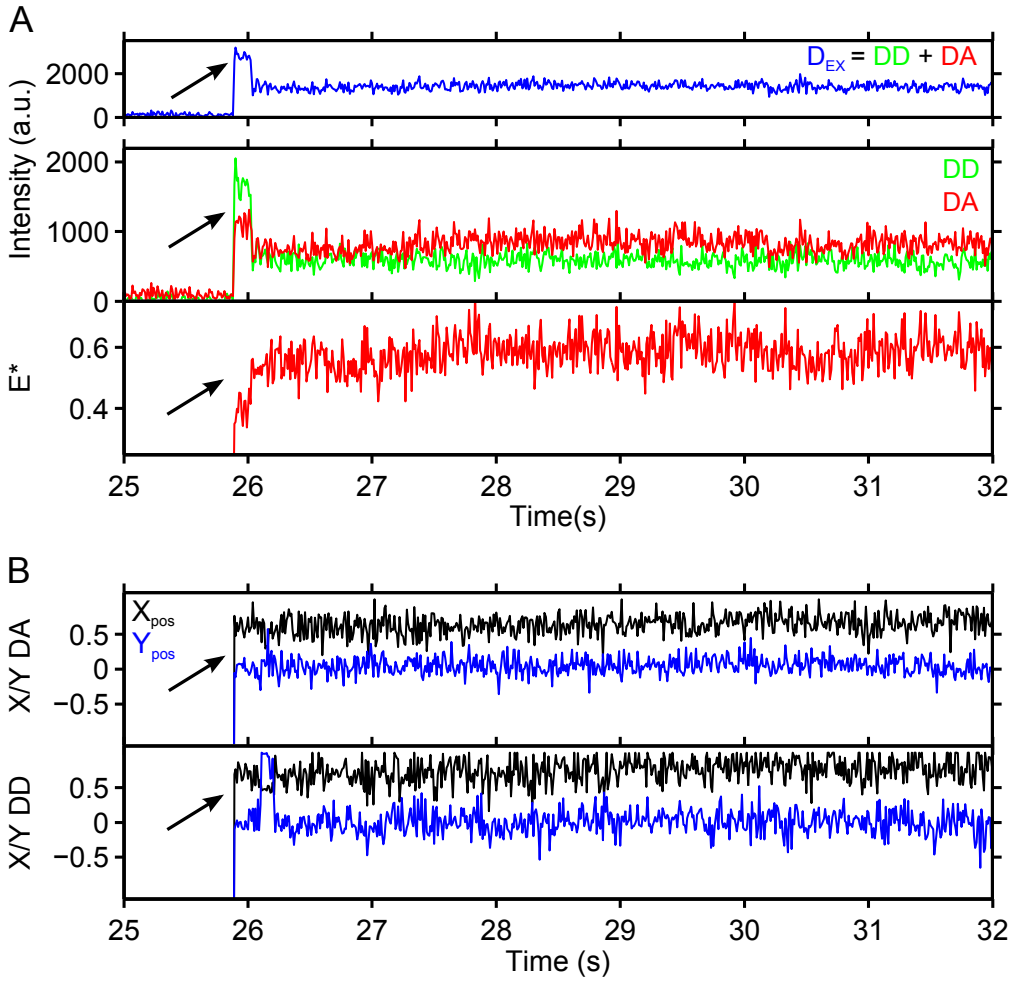


Figure 4.5: Variations in fluorescence intensities at KF-DNA binding events suggesting binding of two KF molecules. (a) Total fluorescence emission under donor-excitation demonstrates the initial frames to have twice the intensity of the following frames (D_{EX} - blue). Donor (DD) and acceptor (DA) emission under donor excitation demonstrate similar differences in intensities. FRET efficiency differs between high and low D_{EX} intensity values (E* - red). (b) X/Y positions of the localised PSFs (black/blue), from which the DD and DA values are extracted (upper/lower panel). Notice the lack of change in localised X/Y positions between the high intensity and low intensity frames. There is however, a change in the DD PSF Y position at $\sim 26.1s$ but this doesn't coincide with any events in the intensity trace. Y position is centred around the localised DNA position (zero), but X position is offset by 0.5 pixels, caused by drift in the microscope. Data was collected using continuous wave (CW) excitation with a green laser power of 0.3mW with an 100ms frame time, with a KF concentration of 4nM.

extract the fluorescence intensities. No such shift in PSF position was seen between the high and low intensity regions (Fig. 4.4b, localised X/Y position of the PSF (blue/black) remain unchanged during changes in D_{Ex}, in both the DD and DA channels. Similarly in 4.5b). This suggested that the fluorescence changes were not caused by nearby surface-bound KF molecules, but it does not eliminate the possibility completely, as surface-bound KF molecules very close to the DNA would not be picked up as differing in position.

Our results are thus consistent with the binding of two KF molecules to the DNA at once. Previous work has suggested that KF can bind to DNA as a dimer[99], although the functional relevance of this phenomenon is unclear. Hypotheses have suggested that a dimer may increase processivity by tethering KF more strongly to DNA, or assist in polymerase-to-exonuclease transfer, or that dimerisation exists as a vestigial function. Double binding has also been observed in other polymerases.[100,101] Our assay supports the notion that KF can bind occasionally as a dimer, as the fluorescence increases are frequently seen to be simultaneous, as well as sequential. Simultaneous increases in fluorescence suggest that both KF molecules have arrived simultaneously, as the likelihood that two molecules bind sequentially within the frame-time is unlikely.

Future work

Our assay does not categorically prove the existence of multiple KF molecules bound to DNA, but the evidence could be improved by investigating the concentration dependence of the binding using a KF concentration titration[i]. A modified assay using FRET between singly-labelled KF molecules or DNA could constrain the locations of the KF molecules more precisely than using PSF localisation, allaying concerns over KF-surface binding or dye photo-physics in our observations above. Such an assay could provide significantly more direct observation than previous work which inferred KF dimers using a complicated combination of gel-electrophoresis, analytical ultra-centrifugation and fluorescence anisotropy.

There exists no crystal structure of KF dimers, suggesting that dimers do not crystallise

[i]TIRF is somewhat limited by the concentration limits of TIRF microscopy ($\sim< 20nM$). This could be circumvented by using the lowered KF-DNA binding K_Ds resultant from the presence of complementary nucleotide.

readily. Structural-FRET[102] allows distances between dyes to be defined by their FRET efficiency, which could be used to solve the structure of KF dimers. Given detailed crystal structures of KF and DNA are already well defined, all that is needed is to position them in space relative to each other, which could be achieved by extracting distances between dye labels.

The materials necessary for such an investigation already exist within our lab (including: DNAs labelled at multiple positions; donor/acceptor fingers-labelled KF; donor/acceptor thumb-labelled KF). A series of relatively straightforward experiments could prove the existence of KF dimerisation, and then follow with a structural model. The function of these dimers might then be elucidated by combining the assays with those described in Chapter 7, which allow observation of the polymerisation rates of individual KF molecules. Despite the relative ease of potential experiments, it is worth noting that the question of the existence of KF dimers is relatively recent, and their functional relevance is currently unknown.

4.2.4 Influence of Atto647N photo-physics on FRET efficiency

In light of the success in observing interconversions between the fingers-open (E*=0.4) and fingers-closed (E*=0.6) conformations (Fig. 4.2c), there was much interest in whether our assay could resolve smaller structural changes in KF. In particular, whether interconversions to the "fingers-partially-closed" conformation could be observed. The fingers-partially-closed conformation was first observed as a small shift in the fingers-open FRET value in diffusion-based experiments, caused by the presence of mismatching nucleotide.[19] It has since been investigated as an important fidelity checkpoint, however the small change in E* makes its observation challenging (open to partially-closed ΔE*=0.03, relative to ΔE*=0.2 for open to closed).[20] Recent single-molecule studies have claimed direct observation of interconversion between the fingers-open and fingers-partially-closed conformations using a KF-DNA FRET pair, although the acceptor dye photophysics was not characterised, nor was the influence of KF-DNA rearrangements or translocations.[77] Here we demonstrate the importance of characterising acceptor dye photophysics, which we achieve using ALEX.

Our assay was able to detect slow interconversions between two similar FRET values

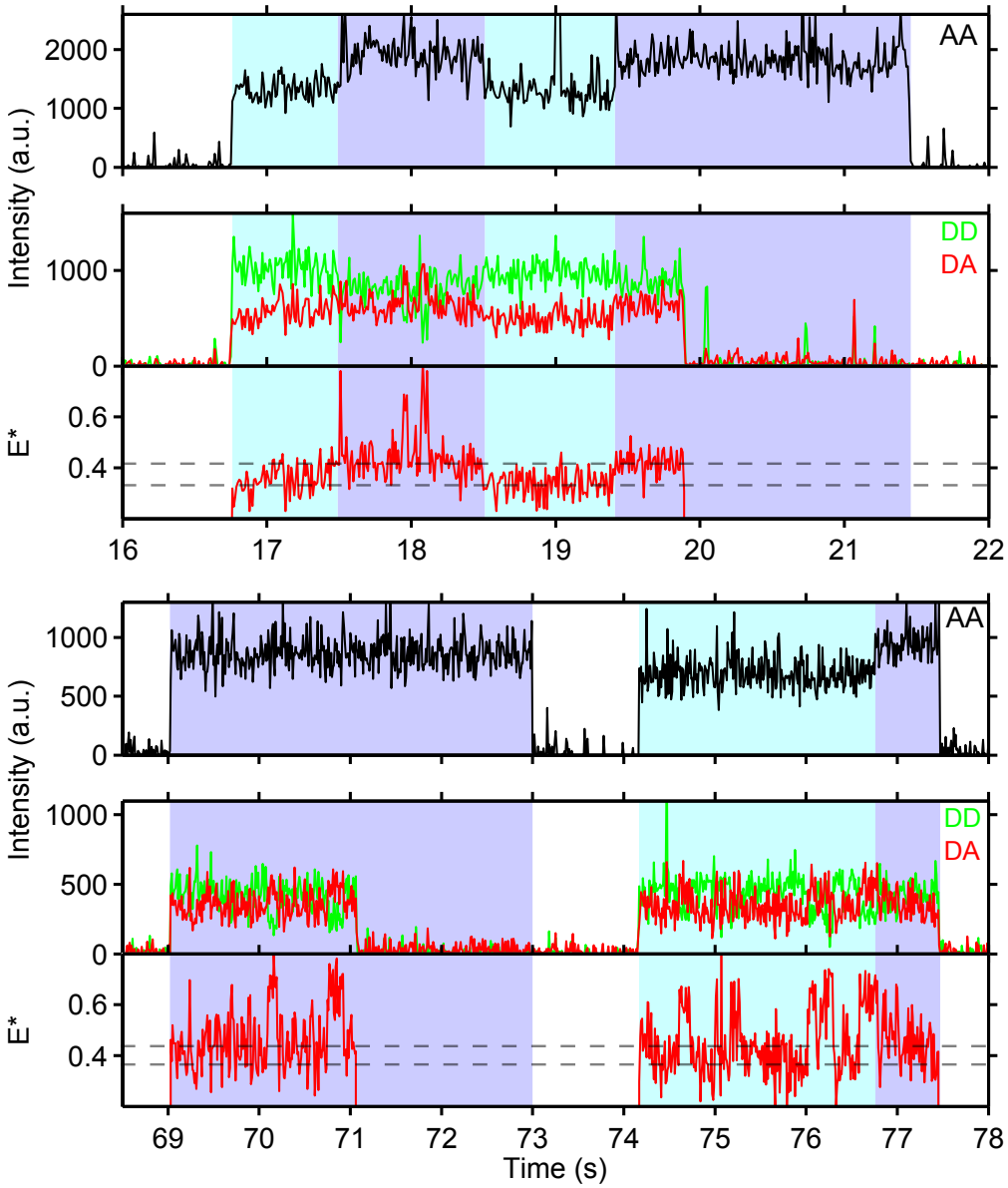


Figure 4.6: Fluorescence intensity traces demonstrating KF-DNA binding events in which small E* changes (dotted lines) correspond with changes in acceptor quantum yield; AA alternating between between low (light blue) and high (purple) intensities (DD-green, DA-red, AA-black). Data for Figures 4.6 and 4.7 were collected using Alternating Laser Excitation (ALEX)[29] at 10ms frame time, with laser powers of 1.5mW and 3mW for red and green lasers respectively, with 4nM KF and 0.125μM dTTP.

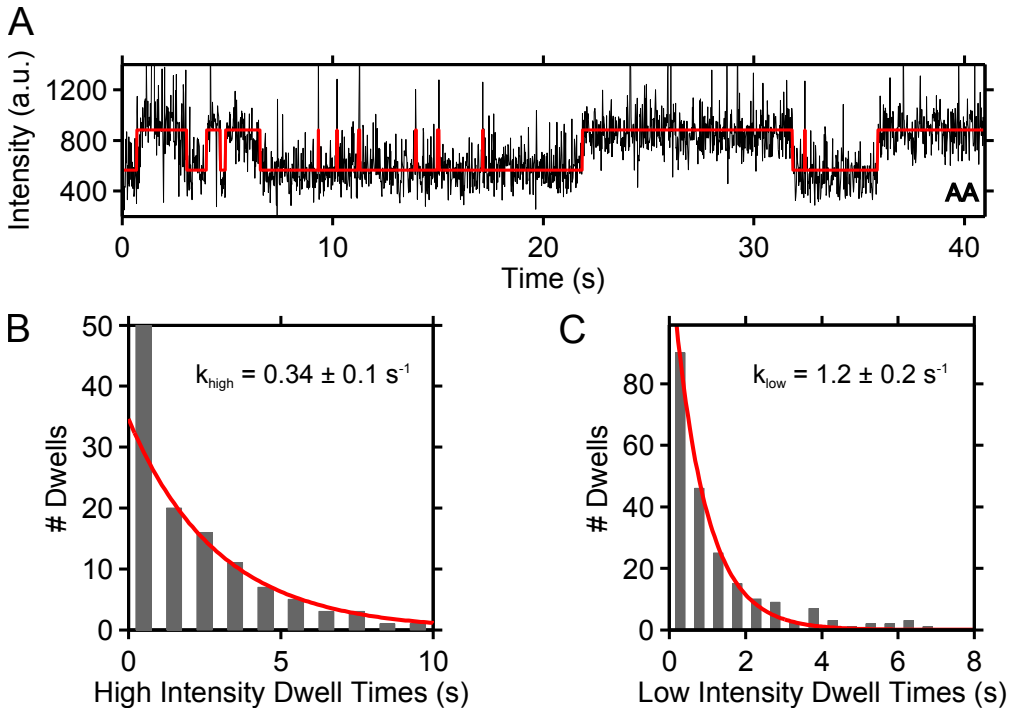


Figure 4.7: Kinetics of Atto647N photo-physics. (a) Acceptor dye fluorescence (AA-black) intensities of concatenated KF-DNA binding events at a single DNA, fitted to a two state Hidden Markov Model (red). Long dwells and short dwells in the high intensity state are seen, the short dwells are assigned to noise caused by changes in nearby fluorescence background. (b) Histogram of high intensity state dwells, fit to a single exponential extracting an exit-rate of $0.34s^{-1}$. The first bin is ignored, as it was assigned to short noise events seen in a. (c) Histogram of low intensity state dwells, fit to a single exponential, extracting an exit-rate of $1.2s^{-1}$.

whilst in the fingers-open conformation (Fig. 4.6a,b, E* shifting between E*=0.39 and E*=0.45). However, these changes in FRET corresponded with changes in acceptor channel fluorescence (AA), demonstrating that changes in acceptor dye quantum yield were causing the E* changes, rather than structural changes in KF (Fig. 4.6a,b, E* changes between light blue and purple regions, which correspond with the AA shifting between low and high intensities). This is consistent with the acceptor dye, Atto647N, having shown intensity state fluctuations in previous work.[97,103]

In order to characterise the time-scale of fluctuations in E* caused by photo-physics of the acceptor, we quantified the acceptor dye intensity variations. Acceptor fluorescence intensities (AA) of KF binding events at 22 DNAs were fit using a Hidden Markov Model (HMM) to extract the dwell times in the low and high acceptor intensity states (see Materials and Methods). (Fig. 4.7a, example of a HMM fit to acceptor fluorescence of concatenated

binding events at a DNA, with excursions between a low and a high intensity states). Dwell times in each intensity state were plotted in histograms to extract the interconversion rates. Dwells in the low intensity state were long ($\sim$ 1s) while the dwells in the high intensity state were longer ($\sim$ 3s) (Fig. 4.7c,d, Fit to low and high intensity dwells). The change in E* due to acceptor fluctuations was sufficiently small (E*$\sim$ 0.08) that we could still distinguish between fingers-open and fingers-closed states.

These results highlight the importance of using Alternating Laser Excitation to characterise the fluorescence intensities of the acceptor dye in single-molecule experiments, to avoid spuriously assigning structural causes to FRET changes which are actually based on changes in dye quantum yield.[104]

4.2.5 Validation using a diffusion based assay

Surface-based experiments can be hampered by artefacts introduced by interactions between the protein and the nearby surface.[105,98] In order to control for such surface effects, we compared the results of a complementary nucleotide titration performed using our surface-based TIRF assay, with a titration measured using a diffusion-based technique.

The single-molecule diffusion assay (described in Section 2.5.1) produces a series of brief observations of E* ($\sim 1ms$), which can be plotted in a histogram to report on the conformation of the KF across the sample. These histograms confirmed the KF-DNA binary complex to be predominantly in the fingers-open conformation, and that increasing dTTP concentration shifted the distribution to predominantly fingers-closed, as dTTP binds the KF-DNA complex, forming the closed ternary complex (Fig. 4.8b, fingers-open peak at E*=0.5 depopulates on increasing dTTP concentration, while the fingers-closed peak at E*=0.7 increases).

A similar titration was performed using our surface-based method, with E* histograms formed from the E* values of frames at KF-DNA binding events. The surface-based titration produced a similar conformational distribution to the diffusion-based titration, with the distribution shifting from predominantly fingers-open, to increasingly fingers-closed with increasing dTTP concentration (Fig. 4.8a, fingers-closed peak at E*=0.6 increases in occu-

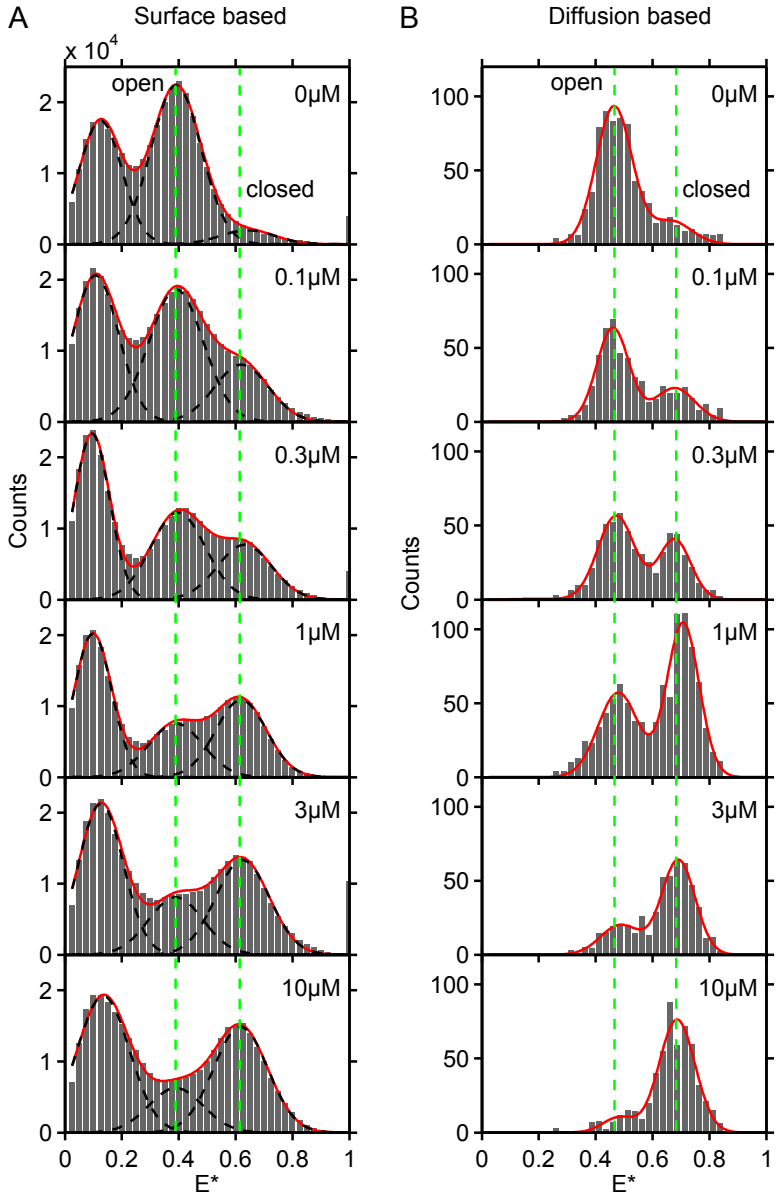


Figure 4.8: Comparison of fingers-conformations measured using surface-based and diffusion-based techniques. (a) E* values at KF-DNA binding events measured on a TIRF microscope across a dTTP titration, performed under continuous wave (CW) excitation with a green laser power of 1.5mW with an 10ms frame time, with a KF concentration of 4nM. CW excitation results in an extra population at E*=0.15 from donor-only molecules. The KF sample shifts from predominantly open (E*=0.4) to predominantly closed (E*=0.6) on increasing dTTP concentration. Occupancies for the donor-only population, fingers-open population and fingers-close population were extracted by fitting 3 Gaussian profiles to the histograms, fit between positions E*=0.08-0.2, E*=0.39-0.45, E*=0.55-0.65. (b) E* values of fluorescence bursts emitted as KF-DNA complexes diffuse through a focused laser beam, measured on a diffusion based microscope performed as described in previous work[20], using the same DNA construct as in Figure 4.1b, but unlabelled. The same buffer was used as for the TIRF experiments, but without glucose oxidase and catalase. Fingers-closed occupancy is seen to increase on increasing dTTP concentration, as the KF molecules close around the complementary nucleotides. Fingers-open and fingers-closed populations were extracted by fitting two unconstrained Gaussian profiles.

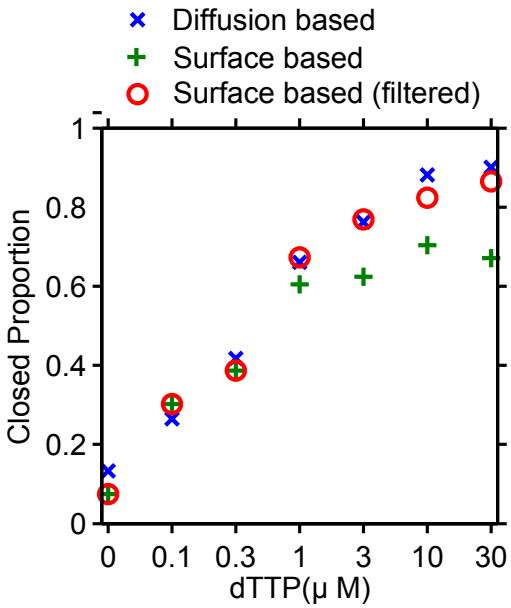


Figure 4.9: Closed state occupancy of histograms in Fig. 4.8 measured using surface-based assay (green crosses) and diffusion-based assay (blue crosses), are seen to diverge from 1-3μM onwards. Closed state occupancy of DNAs subset from the surface-based assay (specified in Fig. 4.10) coincide with the diffusion measurements (red circles).

pancy on increasing dTTP concentration, whilst fingers-open peak at E*=0.4 decreases).

Quantifying the closed state occupancy at each titration point demonstrated that diffusion and surface-based measurements concurred up until 1-3μM, after which point, the surface-based measurements saturate to a slightly lower closed state occupancy than measured with the diffusion method (Fig 4.9 Closed state occupancy measured using diffusion technique (blue crosses) and surface-based (green crosses) diverge at high nucleotide concentrations).

In order to investigate if heterogeneity between DNAs contributes to the discrepancy between diffusion and surface-based measurements, we investigated the fingers-conformation at the scale of individual DNAs, rather than at the experiment ensemble level, as seen above. As we discuss further in Chapter 6, observation of many single-molecules simultaneously produces large volumes of data, making it challenging to represent the entire data-set whilst maintaining the detail provided by individual single-molecules. Here we present a possible visualisation technique to observe behaviours at many single-molecules simultaneously.

In order to observe the fingers-conformations at individual DNAs across the entire experiment, E* histograms of KF binding events at each DNA were sorted according to their mean

E*, normalised, and then plotted in a colour-map (Fig. 4.10, each horizontal line represents an E* histogram at a DNA molecule). These colour-maps demonstrated that up until 1μM dTTP, E* distributions were consistent across most DNAs (Fig. 4.10, colour-maps at 0, 0.1 and 0.3 μM dTTP).

At concentrations 1μM and above, the majority of DNAs demonstrated KF adopting the fingers-closed conformation, however, there was a subset of DNAs at which KF was predominantly open, or had mixed occupancies (Fig. 4.10, at 10μM dTTP, E* histograms for DNAs 1-50 are predominantly high FRET, while DNA 51-68 have low FRET). The majority of DNAs were behaving in a similar fashion, and were selected to form a composite histogram, used to quantify a new closed state occupancy value (Molecules 1:65 at 1μM; 1:50 at 3μM; 1:45 at 10μM). These new closed state occupancies corresponded well with the diffusion based measurements, demonstrating that KF binding events at the majority of DNAs behaved consistently with diffusion based measurements (Fig 4.9, closed state occupancies calculated using DNA subsets (red circles) coincide with diffusion-based measurements (blue crosses)).

The deviation of some DNAs from expected behaviour is unclear. It is entirely possible that this aberrant population is not an artefact, but caused by the differing experimental conditions between TIRF and diffusion-based experiments. For instance, the TIRF measurements used a much higher KF concentration than diffusion-based measurements ($\sim 4nM$ vs $\sim 0.1nM$ KF respectively). In section 4.2.3, we demonstrated it likely that multiple KF molecules can bind DNA at once, which would be more likely at the higher KF concentrations used during the TIRF experiment.

For instance, matching nucleotide is known to increase the stability of the KF-DNA complex, lowering the dissociation rate by a factor of ~ 10.[73] The high time resolution of the TIRF experiment (10ms, necessary to extract the interconversion rates in the next chapter) resulted in short bleaching times. The combination of short bleaching times and increased KF-DNA affinity at high nucleotide concentrations should result in a larger number of invisible KF-DNA binding events, as dyes bleach, and KF remains bound to DNA. These bleached molecules would still be able to bind a second KF molecule. Thus, if our hypothesis about dimerisation is correct, then the observations of this species would be enhanced at

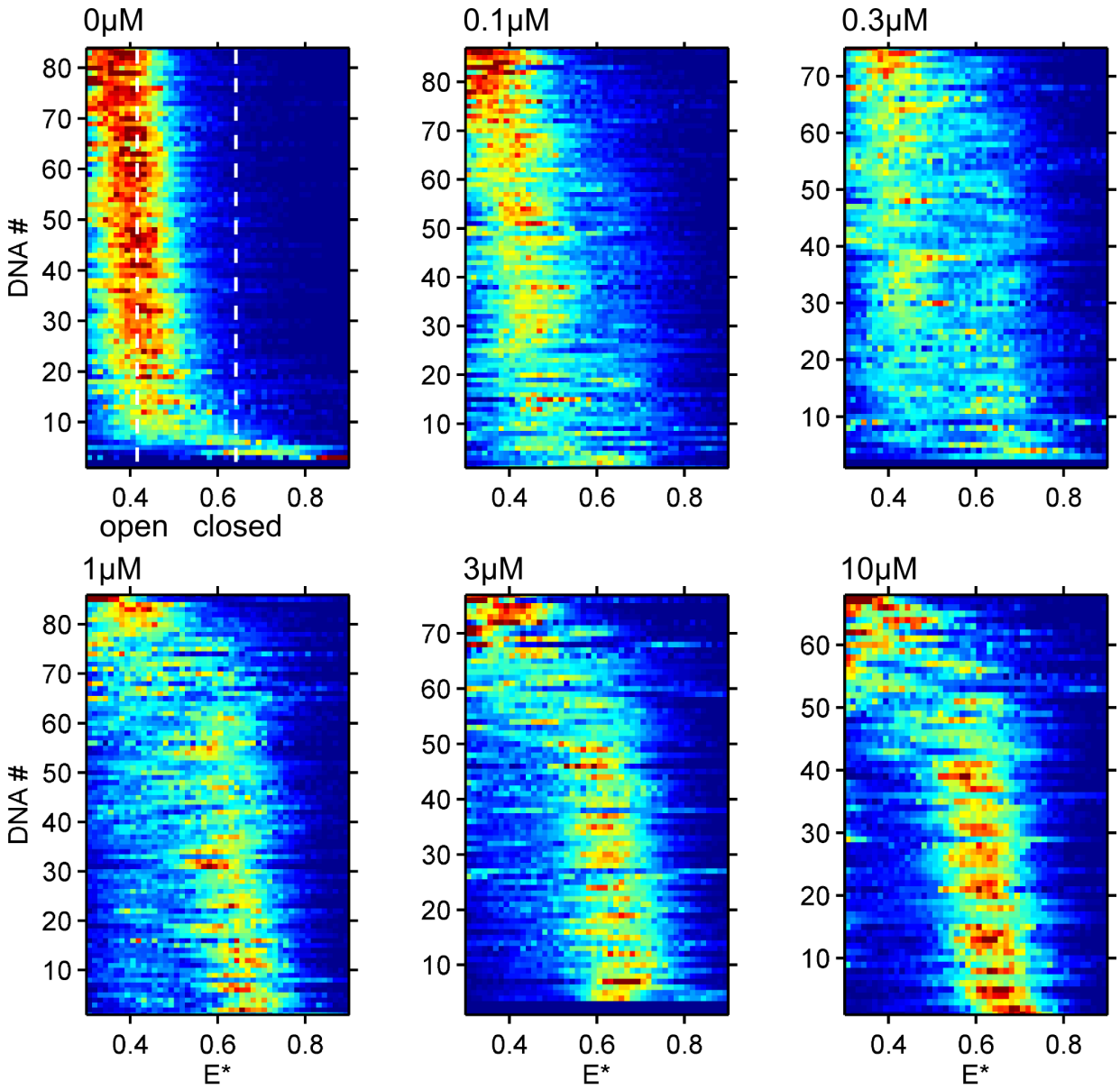


Figure 4.10: Fingers-conformations at individual DNA molecules. (a) E* histograms across all DNAs in each TIRF experiment, normalised, sorted by mean E*, and stacked in a colour-map. (e.g each horizontal line is a E* histogram of KF-DNA binding events at a single DNA). Increasing nucleotide concentration is seen to shift the DNAs from predominantly fingers-open, to predominantly fingers-closed. From $1\mu M$ onwards, KF conformation at a collection of DNAs (upper portion) remain open conformation, whilst most others close.

high dTTP concentrations, as "invisible" KF molecules would be bound to the DNA more frequently, possibly explaining the emergence of the low FRET values.

4.3 Conclusions

In this chapter we have presented an assay to monitor the fingers conformation of single KF molecules for extended periods. We have demonstrated that we achieve transient surface-immobilisation of KF primarily due to DNA binding, although KF-surface binding also contributes. Future work could optimise the immobilisation procedure by lowering the concentrations of neutravidin used, and concomitantly increasing the DNA concentration used to immobilise DNA molecules. We demonstrate that Atto647N photo-physics leads to small changes in E*, whose time-scale we characterise, such that future work can investigate small structural changes such as movement to the fingers-partially-closed state, and differentiate it from photo-physics. Our observations suggest that KF may bind DNA as a dimer occasionally, and we validate our assay by demonstrating that the majority of DNA molecules show behaviour consistent with diffusion based observations.

4.4 Materials and Methods

Sample Preparation

Protein preparation and labelling was performed by researchers in the Grindley lab in Yale University, as described previously.[20,19] Oligonucleotides were synthesised, dye labelled and HPLC purified by IBA Life Sciences.

In order to produce passivated surfaces: Glass coverslips (25x65 mm, thickness number 1, VWR) were burned in a furnace, heated at a rate 90 K/hr, held at 500°C for 1 hour, and cooled at 50 K/hr. Burned glass surfaces were then washed in acetone for 5 minutes, and incubated in a 1:50 Vectabond (Vector Labs) to acetone solution, for 5 minutes. Surfaces were then washed with MilliQ water, and nitrogen dried. Silicone gaskets (Grace Bio-labs) were placed on the dried slides, and each well was incubated with 20μl of 200mg/ml NHS-

PEG (PEG-SVA MW 5000 Da, Laysan Bio) 2.5mg/ml NHS-biotin-PEG (MW 5000 Da, Laysan Bio) in 50mM MOPS pH 7.5. After three hours, the wells were washed and stored in PBS. In order to immobilise DNA, the passivated glass slides were incubated with 0.02mg/ml neutravidin for 1 minute, followed by incubation with 10-100pM biotinilated primer-template DNA for 1 minute.

KF experiments were performed in a solution of 50mM Tris, pH 7.5; 100μg/ml BSA; 1mM DTT; 5% glycerol;1mM Trolox (UV treated[34]); 1.9mg/ml glucose oxidase;0.075mg/ml catalase; 1.5%w/v D-glucose; 10mM $MgCl_2$; which was added to the silicone gasket and sealed using a glass cover slip.

Extracting binding events

Fluorescence time-traces at each DNA were extracted using twoTone software[31] by fitting Gaussian profile to the CCD image at localised DNAs, observed using a custom build TIRF microscope as described in Section 2.5.2. Time-traces were usually extracted at fixed DNA positions, but when the position of the KF-DNA binding event was required (Figures 4.3, 4.4 and 4.5), the binding events were localised by allowing the Gaussian profiles to freely fit their X/Y locations within ± 1 pixels of the initially localised position.

To extract KF binding events from a fluorescence time-trace, a low intensity threshold was used, with single frame events (either association or dissociation) ignored and set to their neighbouring values. A further threshold was set in order to extract the valid KF binding events, with the mean intensity at each KF binding event required to be 1.5 X the low intensity threshold. This allowed the extraction of continuous KF binding events, even if there was a transient dip in fluorescence due to changes in background intensity.

Hidden Markov Modelling

Hidden Markov Modelling was performed using software written by Kristofer Gryte, based on the vbFRET package[106], as documented previously.[97,107] The concatenated KF binding events at 22 DNA molecules were selected on the basis that the concatenation of their binding events was longer than 7 seconds, and fit to a 1 or 2 state HMM model, with the best fit

selected using maximum evidence criteria[106,107], and the dwells in the two states plotted in histograms.

Chapter 5

Binary and ternary complex dynamics in DNA Polymerase I (KF)

In this chapter we extract interconversion rates between the KF fingers-open and KF fingers-closed conformations, in the absence and presence of nucleotides. The binary complex is found to possess intrinsic dynamics, with slow finger-closing and fast fingers-opening rates. KF's fingers are found to close rapidly around a complementary nucleotide, and in the absence of chemistry, nucleotide release is delayed by slow opening of the fingers. We discuss how these fingers kinetics bring into question the existence of the pre-chemistry rate-limiting step suggested in previous work. Additionally, we compare the fingers kinetics to those of T7 DNA polymerase, and apply analysis suggesting that fingers-closing acts to make KF's incorporation fidelity independent of the reaction rate-limiting step.

5.1 Introduction

DNA polymerases maintain genomic integrity by replicating and repairing cellular DNA extremely accurately, rarely incorporating incorrect nucleotides despite the wide range of structurally similar substrates (errors once every ≈10,000-100,000 bases in vitro[38]). In addition to this precision, polymerases possess the flexibility to alter their nucleotide specificity based on context (the templating base), and polymerase malfunctions have been linked to DNA mutations and cancer.[108,109]

Decades of intensive research have elucidated the structure and kinetics of DNA polymerase, but there remain open questions on its fidelity mechanisms. Much attention has

been paid to the rate-limiting step of the polymerisation reaction, which was expected to be a determinant of fidelity.[45] Kinetic studies initially ascribed the rate-limiting step to a pre-chemistry conformational change[44,110,43], with a clear candidate presented shortly after in crystal structures. These structures revealed a large amplitude motion of the polymerase, in which the "fingers" sub-domain closes around an incoming complementary nucleotide before chemistry (Fig. 4.1a). However, bulk fluorescence studies demonstrated this "fingers-closing" conformational change to be rapid and not rate-limiting.[68,42] Ultimately, it remains unclear whether chemistry, or a pre-chemistry conformational change, is rate-limiting for complementary dNTP incorporation by DNA polymerase I.

Furthermore, ensemble experiments on T7 DNA polymerase have demonstrated that whilst fingers-closing around a complementary nucleotide is rapid, the reverse step (pre-chemistry fingers-opening) is slow relative to chemistry, acting to bias the reaction in the productive forward direction after fingers-closing, and removing the rate-limiting step from contribution to fidelity.[93,66,48,52] This forward-biasing after closing is potentially fundamental to the polymerase fidelity mechanism, but it is unclear whether it is a general mechanism between polymerases and whether the rate-limiting step for nucleotide incorporation is less important than expected.

DNA polymerisation is a multi-step process (nucleotide binding, protein conformational change, chemistry etc.) and understanding fidelity requires the quantification of these steps, while experiments on T7 DNA polymerase have highlighted the importance of extracting rates in both forward and backwards directions.[46,48] Single-molecule techniques can resolve such multi-step reactions without the need for synchronisation, and can also observe steps in both forward and backward directions.[111] Resolution of static and dynamic heterogeneity also assists in revealing sub-states unseen in bulk experiments.[19]

Previous single molecule work has inferred conformational changes in T7 DNA polymerase via fluorescence shifts in an environmentally sensitive dye.[84] However, the structural origin of these fluorescence changes were difficult to specify, and the extracted rates didn't correspond to previously measured "fingers-closing" rates.[84,93,52] In order to directly monitor the fingers conformational change in DNA polymerase I we previously employed an

intra-protein FRET strategy in which the mobile fingers and static thumb sub-domains of DNA polymerase I (KF) were labelled with donor and acceptor respectively (Fig. 4.1a).[19,20] Diffusion-based experiments revealed KF to exist in a distribution of states dependent on its ligands in solution. The KF-DNA binary complex was predominantly in the fingers-open state, while the KF-DNA-(A-dTTP) ternary complex (complementary nucleotide) was predominantly in the fingers-closed state. However, quantifying the dynamics between these species was challenging given the short observations times (∼1ms). Surface-immobilisation in conjunction with protein-DNA smFRET has previously allowed extended observations of protein-conformation.[77] However, the rates of structural interconversion were not extracted, and protein-DNA FRET pairs cannot easily deconvolve protein structural changes, DNA rearrangements and translocation events.

Here, we combine our intra-protein FRET strategy with a surface immobilisation technique to achieve the first direct real-time observations of the fingers dynamics of single KF molecules. This is achieved using a TIRF assay to observe dye-labelled KF molecules binding to surface immobilised DNA molecules, as described in the previous chapter (Fig. 4.1c). The high temporal resolution (10ms) of the assay allows us to quantify the structural inter-conversion rates, revealing the fast dynamics of the binary complex for the first time. The kinetics of the pre-chemistry reaction with complementary nucleotide are also extracted, by using dideoxy terminated DNA substrates, suggesting similar fingers kinetics to T7 DNA polymerase, and possibly a general fidelity mechanism between polymerases. Furthermore, we use the reaction rates to quantify the fidelity of the polymerase in relation to its macroscopic error rate. A distinct fluorescence state in the presence of mismatching nucleotide has been seen in KF and T7 DNA polymerase, which promotes faster dissociation of mismatched nucleotides relative to complementary nucleotide.[20] Our data suggest that this "partially closed"[60] structure also promotes a faster dissociation from the DNA, which could contribute a further fidelity mechanism.

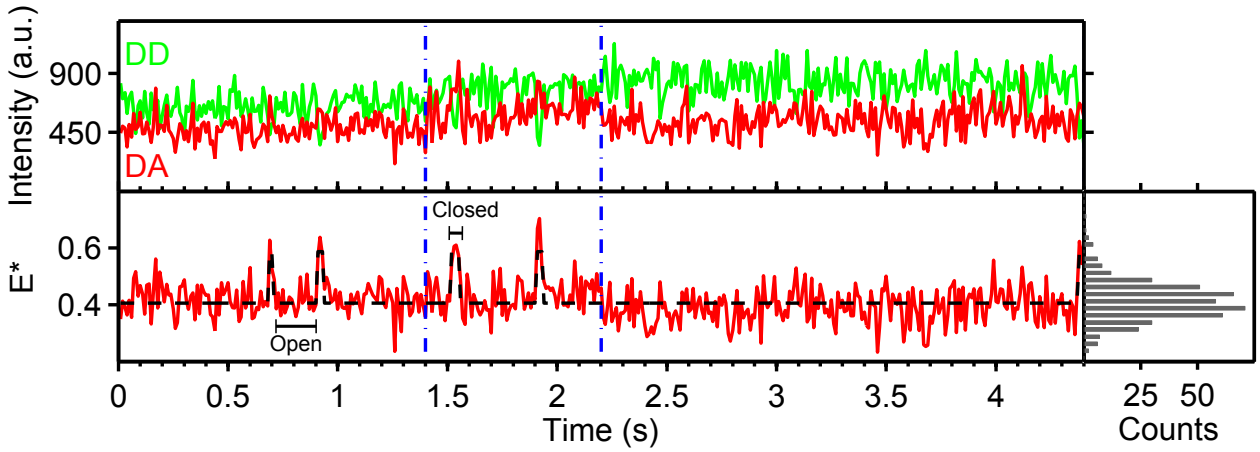


Figure 5.1: Binary complex dynamics of KF. Fluorescence intensity trace of three concatenated Pol I (KF) binding events (upper panel), revealing intrinsic dynamics in the binary complex via alternation of E* between the open (E*≈ 0.4) and closed (E*≈ 0.6) states (lower panel). The upper panel shows donor (DD) and acceptor (AA) channel emission under continuous wave excitation at 10ms frame time. The blue dotted lines denote separate binding events. Interconversions correspond with DD-DA anti-correlation. Closing is seen to be relatively slow compared with subsequent rapid opening, and a histogram of the E* values is seen on the right hand panel. The black dotted line is the Hidden Markov Model (HMM) fit.

5.2 Results and discussion

5.2.1 Binary complex dynamics

By examining KF-DNA binding events in the absence of nucleotides, we studied the binary complex and revealed it to be dynamic. KF-DNA binding events showed the polymerase to exist in the open state for relatively long periods (Fig 5.1: Open state at E*=0.4 between $0 - 0.7s$), with short excursions to the closed state (Fig 5.1: Closed State; E*=0.6 at $0.7s$ and $0.9s$). Binding events were found to be 89% fingers-open by a histogram of E* values across all KF-DNA binding events (frame-based) (Fig. 5.3a: Left hand panel).

Binding events were split into two populations which were assigned to KF-DNA binding and KF-surface binding, by classifying binding events according to their closed state occupancy. The majority of binding events ($\sim$ 90%) were predominantly fingers-open, and assigned to KF-DNA binding (Fig. 5.2a, grey bars). A minor population ($\sim$ 10%) was more broadly distributed, and assigned to KF-surface binding (Fig. 5.2a, cyan bars), as the closed state occupancy distribution corresponded with that measured in the presence of neutravidin

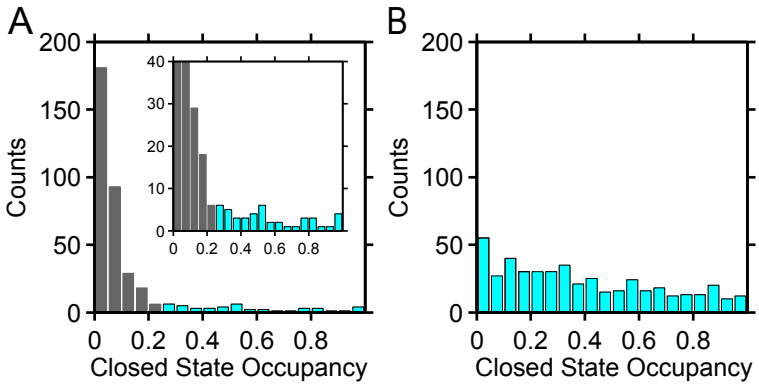


Figure 5.2: Removing the contribution of KF-surface binding events via comparison of the closed state occupancy of KF binding events in (a) the presence of DNA and (b) the absence of DNA. The closed state occupancy of each binding event was classified by first removing bleached acceptor dye frames (E*<0.35), with the closed state occupancy then defined as the fraction of remaining frames with E*>0.55. (a) Closed state occupancy in the presence of DNA shows that the majority of binding events exist in the open conformation (grey). (b) Binding events in the absence of DNA demonstrate a broad closed state occupancy, which are assigned to KF-surface binding events, which account for the cyan population in (a). Inset shows zoom.

but absence of DNA (Fig. 5.2b).

The dynamics of the binary complex were quantified by fitting a Hidden Markov Model (HMM) to the E* time traces (Fig. 5.1, black dotted line), allowing extraction of the dwell times in both the open and closed conformations for the KF-DNA complexes (excluding the KF-surface binding events). The dwell time histograms fit well to a single exponential, extracting a slow fingers-closing rate of $k_1 = 1.1 \pm 0.2s^{-1}$ and fast fingers-opening rate $k_{-1} = 29 \pm 3s^{-1}$ (Fig. 5.3b). These rates defined an open state occupancy of 96%, in good agreement with the occupancy measurement based on the histogram of KF-DNA binding event E* frames (Fig. 5.3b, left hand panel). Comparable rates were extracted when the entire population (KF-DNA and KF-surface binding events), were fit with a HMM, with the open state dwell histograms fit to biexponentials, representing the dynamic contribution of both major and minor populations (Fig. 5.3a).

Discussion

We directly observe the structural interconversion of the binary complex, which had only previously been inferred indirectly. The fingers-closing $k_1 = 1.1 \pm 0.2s^{-1}$ and fingers-opening

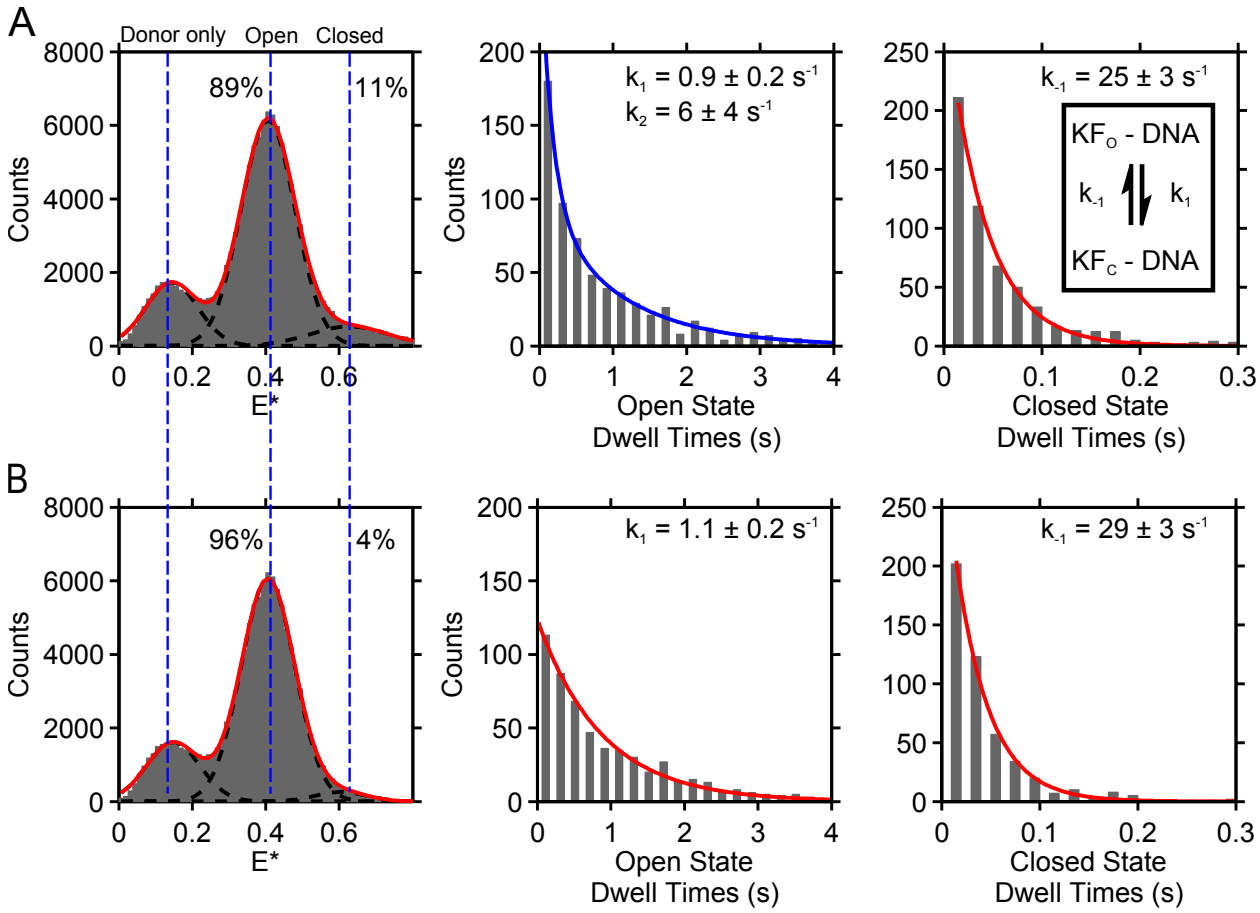


Figure 5.3: Extracting the interconversion rates of the binary complex, for (a) all KF binding events (KF-DNA and KF-surface) and (b) just KF-DNA binding events, selected as specified in Fig. 5.2. Left hand panels quantify the fingers-open, fingers-closed and donor only occupancies via Gaussian distribution fits to E* histograms of binding event frames (black dotted lines, overall fit in red). Middle panels: fingers-open dwell-time histograms extract the fingers-closing rate of the binary complex $k_1 = 1.1 \pm 0.2s^{-1}$ via single exponential fit to (b), however (a) requires a biexponential fit due to contributions from both KF-surface binding events $k_2 = 6 \pm 4s^{-1}$, as well as KF-DNA binding events $k_1 = 0.9 \pm 0.2s^{-1}$. Right hand panels: Fingers-closed dwell-time histograms extract a fingers-opening rate of $k_{-1} = 29 \pm 3s^{-1}$, with a single exponential fit to both (a) and (b). Reaction scheme seen in upper right-hand panel: KF-DNA open state (KF_O) and closed state (KF_C).

$k_{-1} = 29 \pm 3s^{-1}$ rates extracted are significantly more accurate than previous estimates of the fingers-closing ($15s^{-1}$) and opening ($85s^{-1}$) rates, which were extracted indirectly from short observation times ($\sim$1ms) of the KF-DNA complex, using a diffusion based technique. A similar diffusion based assay has also inferred binary complex dynamics in Klentaq1 polymerase.[112] Direct real-time observation of conformational changes have been seen in T7 DNA polymerase, using a similar TIRF technique, however, a dynamic binary complex was not observed, possibly because of the different polymerase; the lower time resolution (200ms relative to our 10ms frame time); or the ambiguous structural nature of the fluorescence change, which was based on an environmentally sensitive dye, providing no clear structural interpretation.[84,52]

The biological significance of the binary complex dynamics/kinetics are unclear, although the relative occupancy of the states is consistent with an enzyme optimizing its open occupancy to maximize substrate binding, whilst minimizing the occupancy of the catalytically active state in the absence of substrate.[48] There has been renewed interest of late in the role that protein motions and fluctuations play in their enzymes function,[113] particularly given observations that cyclophilin A molecules demonstrate intrinsic fluctuations at the reaction turnover rate, despite the absence of substrates.[114] Here we have observed KF-DNA fingers-opening rates ($k_{-1} = 29 \pm 3s^{-1}$) which are very similar to the rate of chemical incorporation of nucleotides ($k_{pol} \approx 40s^{-1}$).[42] However, correlation does not imply causation, and as we demonstrate in Chapter 7, the rate of fingers-opening decreases when using non-terminated "deoxy" DNA substrates, suggesting binary dynamics are not reflecting features of the chemical incorporation, but interestingly suggest that the KF fingers may interact with the terminal 3'-OH of the primer.

Another possibility to explain the decrease in fingers-opening rate when using deoxy substrates over dideoxy substrates, is that fingers-closing is coupled to additional conformational changes at the active site, which change KF's interactions with the primer 3'-OH, leading to changes in the relative stability of the open and closed complexes in the presence or absence of the 3'-OH. Using crystal structures to investigate changes in the 3'-OH interactions is difficult however, as the fingers-closed structures lack the 3'-OH, in order to prevent primer

extension. The fingers-open crystal structures however do contain the 3'-OH, allowing some inferences to be made.

The residue proximal to the 3'-OH which most clearly changes between the open and closed structures is Asp-830, whose Cγ shifts by $\sim$1Å (comparing PDBs 1LV5 and 1L3S). It has been suggested that while the Asp-830 residue forms a hydrogen bond with the 3'-OH in the fingers-open conformation, it loses this interaction on moving to the fingers-closed conformation, as the Asp-830 instead interacts with the Mg^{2+} ion introduced with the incoming nucleotide.[58] However, it is unclear how stabilisation of the open conformation via 3'-OH interaction with Asp-830 could account for our observations of a decrease in opening rates between dideoxy and deoxy primers.

Another possible explanation is the interaction that the 3'-OH is known to form with the second Mg^{2+} ion, which could stabilise the closed binary complex if this Mg^{2+} ion was present in the binary complex closed conformation. However, this second Mg^{2+} ion is thought to arrive later in the reaction pathway, after nucleotide binding and fingers closing, and it is unclear whether its chelation could occur in the binary complex closed conformation, in the absence of phosphate groups of the incoming nucleotide. The possibility of stabilisation of the fingers-closed conformation via Mg^{2+} could be investigated in the future using our assay, by observing the dependence of binary complex fingers dynamics on Mg^{2+} concentration.

5.2.2 Ternary complex dynamics

The ternary complex demonstrated fingers-dynamics dependent on the concentration of complementary nucleotide. As we increased the dTTP concentration, the dwell times in the open state were seen to decrease (Fig. 5.4, compare open states dwells at 0.3μM and 10μM dTTP). The closed state dwells also appeared longer than those of the binary complex (compare closed state dwells in Fig. 5.4 at 0.3μM dTTP , and in Fig. 5.1). The overall equilibrium shifted towards the closed state with increasing nucleotide concentration (Fig. 5.4, right hand panel histograms at 0.3μM , 1μM and 10μM dTTP).

To investigate the fingers dynamics further, we extracted interconversion rates across a range of concentrations of complementary nucleotide (dTTP 0.125μM - 30μM).

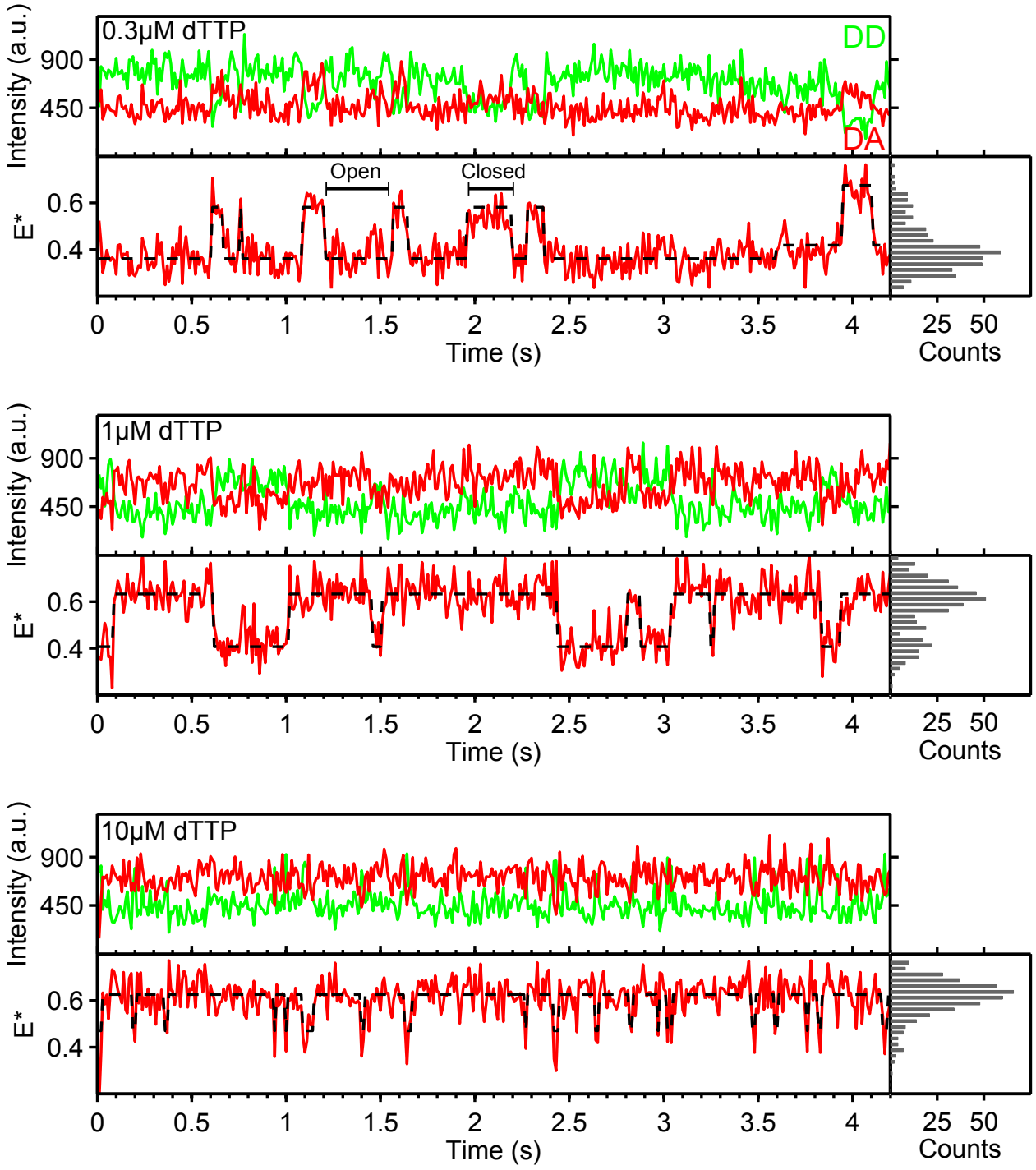


Figure 5.4: Bound Pol I (KF) conformational dynamics in the presence of complementary nucleotide. Fluorescence intensity and E* traces (upper and lower panels) for the KF-DNA complex in the presence of 0.3uM, 1uM and 10uM of complementary nucleotide (dTTP) (taken from a titration range). E* histograms for the binding events are in the right hand panel. The observed closing rate increases with increasing nucleotide concentration, whilst the rate of opening appears slower than binary complex opening. Slow photophysical fluctuations ($\sim 0.3s^{-1}$) in Atto647N lead to small shifts in E*, as seen in the final section of the 0.3uM trace. (See Section 4.2.4)

HMM extracted fingers-open dwell time histograms, which fit well to single exponentials, except at 0μM and 0.1μM dTTP, which fit to a biexponetials, caused by the distinct fingers-closing rate of the KF-surface binding complex, as seen in Figure 5.3 (Fig. 5.5b, biexponential fits at 0μM and 0.1μM (blue), single exponential fits at other nucleotide concentrations (red)). Increasing nucleotide concentration was found to increase the closing rate, from $1.5s^{-1}$ at 0.1μM dTTP up to $\approx 60s^{-1}$ at 10μM dTTP (Fig. 5.6a), interpreted as nucleotide induced closing of KF. Closing rates were too fast to be extracted above 10-30μM due to the limited time resolution (10ms).

To interpret these results, we used a model accounting for binary complex dynamics, nucleotide binding, and subsequent ternary complex dynamics (Fig. 5.6c). This model allowed the polymerase to close (k_1) and open (k_{-1}) according to binary complex kinetics in the absence of bound nucleotide; upon complementary nucleotide binding (with affinity K_2) the polymerase could close (k_3) and open (k_{-3}) according to ternary complex kinetics. Association and dissociation of complementary nucleotide thus modulated the interconversion rates of the polymerase.

The increased closing rate with nucleotide concentration was fit to an equation derived from this model, which was dependent on the binary complex closing rate (k_1); the nucleotide binding affinity (K_2); and the ternary complex closing rate (k_3). (Fig. 5.6c, see derivation in Appendix A):

$$k_{close}^{obs} = \frac{k_1 + k_3 K_2 [dTTP]}{1 + K_2 [dTTP]}$$

The binary complex closing rate was already known, so the data was fit to this equation with fixed $k_1 = 1.1s^{-1}$ (Fig. 5.6a black dotted line). The fit extracted the binding affinity of complementary nucleotide $K_2 = 0.06 \pm 0.04 \mu\text{M}^{-1}$ and the rapid closing of the ternary complex $k_3 = 150 \pm 70 s^{-1}$. Time resolution limits rendered a complete titration curve up to nucleotide saturation impossible; as a result, only fits to a partial curve were possible. However, the extracted value of k_3 were consistent with previous measurements of $k_3 = 140s^{-1}$ and K_2 is similar to that previously measured $K_2 = 0.2\mu\text{M}^{-1}$.[42,20]

The histograms of fingers-closed dwells fit best to biexponentials in the presence of nu-

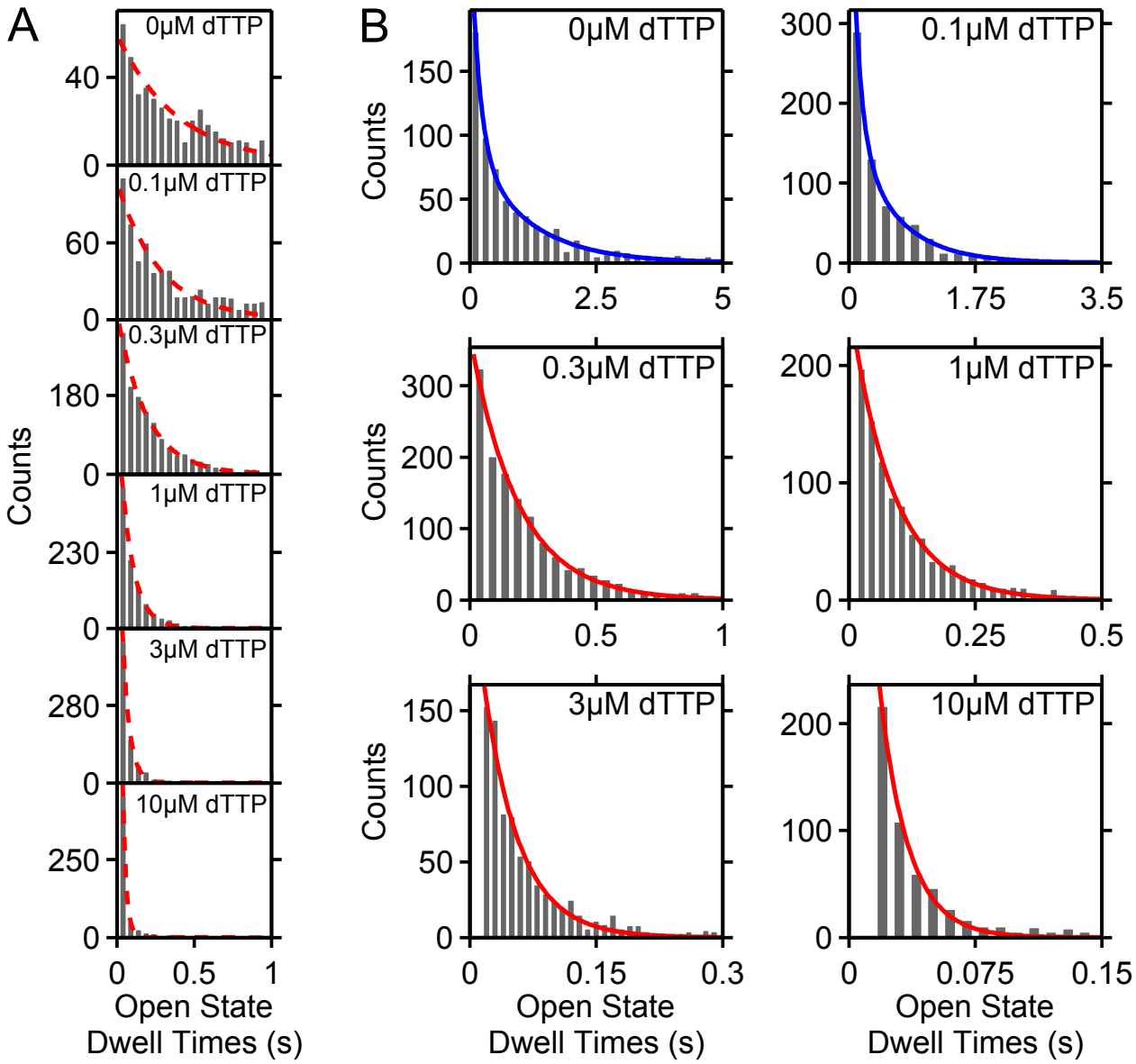


Figure 5.5: Fingers-open dwell times define the fingers-closing rates at different nucleotide concentrations for KF. (a) Histogram of fingers-open dwell times demonstrating the clear decrease dwell-times on increasing nucleotide concentration, by keeping the x-axis scaling constant (b) Fits to fingers-open dwell-time histograms, which are biexponential for dTTP=0, 0.1μM, and single-exponential for other concentrations. Extracted rates are plotted in Fig 5.6

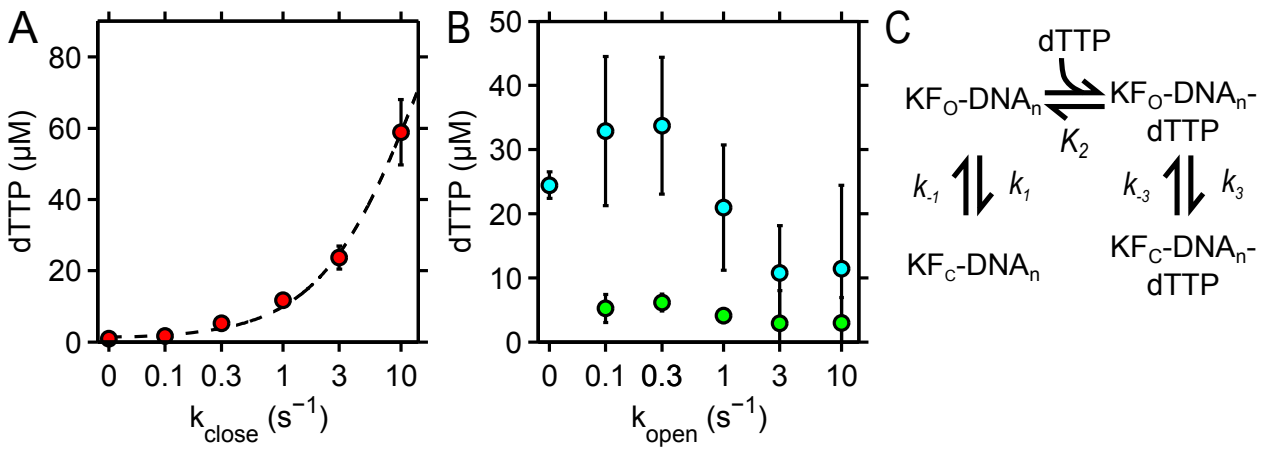


Figure 5.6: The observed closing and opening rates as a function of matching nucleotide concentration. (a) The closing rate increases with nucleotide concentration and is fit to model (c) (black dotted line). Error bars represent the 99% confidence intervals on the fits. (b) The opening rates for binary and ternary complex contributions in cyan and green respectively, with fits becoming less reliable after 1uM, due to missed events caused by fast closing. (c) Model allowing for KF binary and nucleotide induced ternary complex dynamics.

cleotides (Fig. 5.7, compare single-exponential fit at 0μM (red) to the biexponential fit at 0.1μM (blue)), interpreted as the binary complex fast opening rate ($k_{-1} = 29 \pm 3s^{-1}$ at 0.1μM dTTP, consistent with binary complex measurements), and a slower opening rate interpreted as the ternary complex opening ($k_{-3} = 6.1 \pm 1.5s^{-1}$ at 0.1μM dTTP), with complementary nucleotide stabilising the closed state. The fingers-opening rates remained consistent until high nucleotide concentrations at which point the fast fingers-closing resulted in missed open state events, elongating the closed state dwells (Fig. 5.6b closing rates deviate at 3uM and 10uM (cyan). See brief open state dwells and long closed state dwells at 10μM in Fig. 5.4). The rates for binary and ternary complex opening were comparable for nucleotide concentrations below 3uM.

Discussion

The rate-limiting step of the polymerisation reaction has drawn considerable attention over the last decades, due to its expected role in fidelity. Evidence for a pre-chemistry rate-limiting step in KF has been based primarily on two experiments: Firstly, using an unnatural α-thio-nucleotide to slow the rate of chemistry did not change the polymerisation rate relative to natural nucleotide incorporation, suggesting that a step besides chemistry must rate-limit

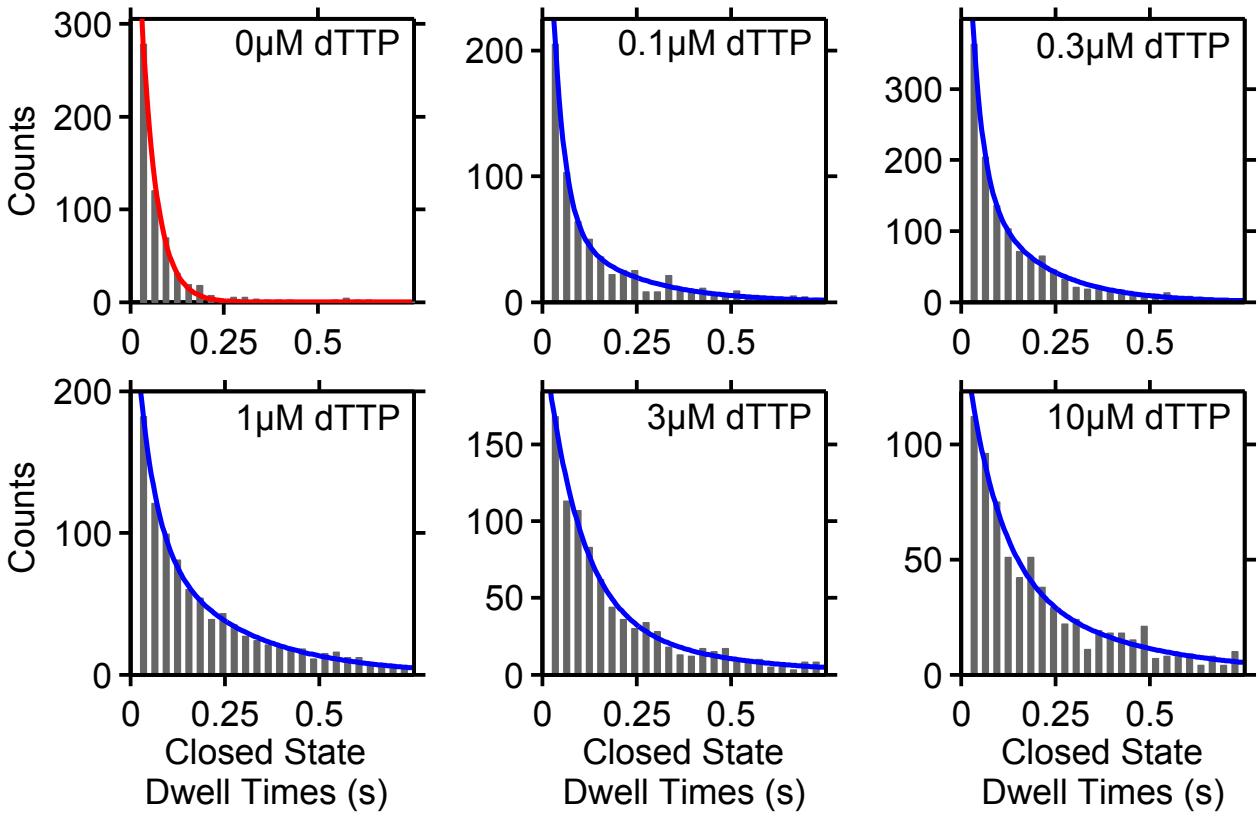


Figure 5.7: Fingers-closed dwells define the fingers-opening rates at different complementary nucleotide concentrations. Single exponential fits are seen in red, demonstrating the binary complex dynamics at 0μM dTTP. Biexponential fits in blue, representing the contributions of the binary (k_{-1}) and ternary (k_{-3}) opening rates in the presence of complementary nucleotide.

polymerisation (experiment referred to as the sulphur-elemental effect, see Section 3.2).[44] Secondly, the DNA product yield of an instantaneously halted polymerisation reaction was compared with a reaction which was halted whilst allowing any pre-formed KF-DNA-dNTP complexes to proceed with their individual reactions - either dissociating or incorporating the nucleotide (see Section 3.2). The difference in these "pulse-quench" and "pulse-chase" yields suggested accumulation of a KF-DNA-dNTP intermediate during the polymerisation reaction, which was explained using the rate-limiting step inferred by the sulphur elemental effect, and by inferring a further slow post-chemistry step, providing slow steps flanking the chemical step, allowing accumulation of nucleotide bound intermediate.[43] Since these experiments the identity of the pre-chemistry rate limiting step has remained unknown, despite investigation of many pre-chemistry conformational changes such as fingers-closing and DNA rearrangements.

However, it has since been noted that the critical requirement for KF-DNA-dNTP intermediate accumulation is that there is slow release of dNTP, which requires a slow step in the reverse direction before chemistry, which would *not* necessarily be manifested in the forward direction rate.[52] Our observations have demonstrated that fingers-closing is rapid in KF, whilst fingers-opening is slow, allowing formation of a long lived KF-DNA-dNTP state. We suggest that this, in combination with the previously posited slow step after chemistry, could explain the difference in pulse-chase and pulse-quench yields. On this basis, and because we believe the sulphur elemental effect to be an unreliable metric (as discussed in Chapter 3), we suggest the existence of a slow pre-chemistry step is no longer required, and should be reconsidered, and suggests that the chemical step of polymerisation may actually be the rate-limiting step.

In Figure 5.8, we consider a reaction scheme in which a polymerase-DNA binary complex (KF_O-DNA_n) can bind a complementary nucleotide (KF_O-DNA_n-dTTP) , undergo a conformational change (KF_C-DNA_n-dTTP), and catalyse the incorporation of nucleotide (KF_O-DNA_{n+1}). We note the fingers kinetics of KF have similarities to a conformational change measured in T7 DNA polymerase. T7 polymerase was found to undergo a rapid conformational change on binding complementary nucleotide ($k_2 = 660s^{-1}$), which demonstrated a

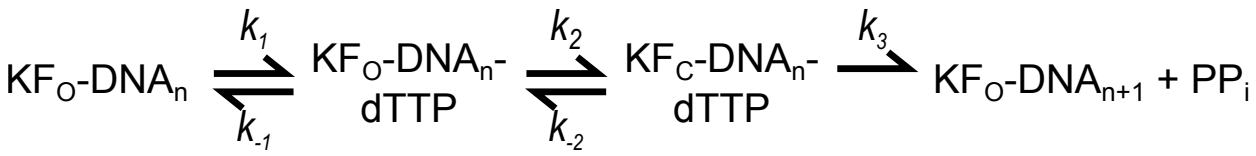


Figure 5.8: Reaction scheme

slow reverse rate ($k_{-2} = 1.6s^{-1}$) relative to the chemistry rate ($k_3 \approx 360s^{-1}$). (Compare these rates with KF fingers closing rate $k_2 = 140s^{-1}$; fingers-opening rate $k_{-2} = 6.1s^{-1}$; and chemistry rate $k_3 \approx 50s^{-1}$)

Tsai and Johnson performed a mathematical analysis demonstrating that the kinetics of this conformational change in T7 DNA polymerase imply that its fidelity is independent of the rate-limiting step of the reaction, which was surprising given that the rate-limiting step was previously thought to determine fidelity. We reproduce this analysis here to demonstrate that the reasoning applies to KF, and that its fidelity is independent of the rate limiting step for KF.

A polymerases discrimination can be quantified by the ratio of its catalytic efficiencies for correct and incorrect nucleotide incorporation (Eq. 5.1). Tsai and Johnson calculated an expression for the catalytic efficiency of the reaction in Figure 5.8:[93]:

$$\frac{k_{cat}}{K_m} = \frac{k_1 k_2 k_3}{k_2 k_3 + k_{-1}(k_{-2} + k_3)}$$

For T7 DNA polymerase, the rate of chemistry was much faster than fingers-opening ($k_3 = 360s^{-1} >> k_{-2} = 1.6s^{-1}$), allowing the catalytic efficiency to be simplified to: $k_{cat}/K_m \approx (k_1 k_2)/(k_2 + k_{-1})$. Notice that this new expression is independent of the chemistry rate, and can be further reduced by assuming nucleotide binding equilibrium ($k_{-1} >> k_2$) leading to:

$$k_{cat}/K_m \approx K_1 k_2$$

To see if this limit applies with KF, we use measured values for fingers-closing k_2 and opening k_{-2} and we use a k_3 value based on literature values. We assign k_3 as approximately equal to the polymerisation rate k_{pol} measured in single turn-over rapid-quench experiments,

as the polymerisation reaction is rate-limited by a single step which occurs after fingers-closing. k_{pol} has been measured many times across the literature, varying from $k_{pol} = 46.5s^{-1}$ for dTTP incorporation; $k_{pol} = 38s^{-1}$ for dGTP incorporation[42]; $k_{pol} = 53s^{-1}$ for dTTP incorporation on a different DNA sequence[68]; $k_{pol} = 50s^{-1}$ for dATP incorporation[53]. This demonstrates that k_{pol} remains constant across the literature for both nucleotide type and surrounding sequence, giving a mean incorporation rate of k_{pol}=$47s^{-1}$, which we assign to $k_3 \approx k_{pol}$

Using these values we noticed that KF has a slow fingers-opening rate relative to the rate of chemistry, similar to T7 DNA polymerase, although the difference in rates is less pronounced:

T7 DNA polymerase : $k_3 = 360s^{-1} >> k_{-2} = 1.6s^{-1}$

KF DNA polymerase : $k_3 = 47s^{-1} >> k_{-2} = 6.1s^{-1}$

The limit calculated by Tsai and Johnson $k_{cat}/K_m \approx K_1 k_2$ still applies to KF, to within a factor of $(1 + k_{-2}/k_3)^{-1} = 0.9$. Thus, we demonstrate that the fingers-closing conformational change acts to remove the rate-limiting step from contributions to fidelity in KF, as seen in T7 DNA polymerase, suggesting that this mechanism could be general between polymerases, and is surprising given decades of research has been performed on elucidating the origin of the rate-limiting step, thought to contribute to fidelity.

Mismatch incorporation is slow[115], and therefore we can assume that equilibrium is reached with nucleotide binding and fingers-closing for mismatch incorporation[93], leading to a catalytic efficiency of $k_{cat}/K_m \approx K_1 K_2 k_3$ for mismatch incorporation. Discrimination is defined by[48]:

$$D = \frac{(k_{cat}/K_m)_{correct}}{(k'_{cat}/K'_m)_{incorrect}} = \frac{K_1 k_2}{K'_1 K'_2 k'_3} \tag{5.1}$$

Using rates extracted in our and previous work, we attempt to quantify the discrimination of KF, in order to relate the microscopic features of the reaction to the macroscopic error rate of the polymerase. Measurements of the intrinsic binding affinity for complementary and mismatched nucleotides $K_1 = 0.06 - 0.2\mu M^{-1}$, and $K'_1 = 0.02\mu M^{-1}$ [20], demonstrate that binding and conformational changes pre-fingers-closing contribute a factor of $\approx 3 -$

10 to fidelity. The inefficiency of fingers closing around mismatching nucleotides ($K_2' = 0.09$) contributes a factor of ≈ 10 to fidelity, while the rapid closing around complementary nucleotide, contributes a factor of $k_3 = 140$, giving a resultant discrimination of 4200 - 14000, compared to the measured error rate of exonuclease deficient KF as 1 in 15,000,[86] although there are no clear estimations for k_3'. Further work should investigate how mutated polymerases change these individual values, and how this effects polymerase fidelity.

The lower pre-fingers-closing affinity for mismatching nucleotide over complementary nucleotide suggests that this partially closed state might actively misalign polymerase residues to increase the dissociation of the nucleotide, as seen in T7 DNA polymerase.[93]

5.2.3 Mismatching nucleotide

Extracting interconversion rates between the fingers-open and fingers-closed conformations in the presence of mismatching nucleotide was not possible given interconversions were rare, and were thus dominated by noise such as interconversions seen in KF-surface binding events.

Previous diffusion based experiments demonstrated that KF shifts to a fingers-partially-closed conformation in the presence of a high concentration of mismatching nucleotide. We reproduced these observations using our TIRF assay, by observing the shift in the fingers-open peak position in E* histograms across a mismatching nucleotide titration (Fig. 5.9b, see shift in fingers-open state peak, extracted and plotted in Fig. 5.10, shifting from E*=0.43 to E*=0.46). The peak was extracted by fitting a single Gaussian to the open state peak between $0.3 < E^* < 0.6$. The E* shift between the fingers-open and fingers-partially-closed conformations was consistent the shift of ΔE*=0.03 observed in previous diffusion based experiments.[20]

In tandem with these observations, we observed that the KF-DNA dissociation rate increased with mismatching nucleotide concentration, consistent with previous observations that mismatching nucleotide destabilises the KF-DNA complex.[73] Dissociation rates were extracted from dwell time histograms, which were biexponential, caused by KF-DNA binding and KF-surface binding components (as described in Chapter 4). KF-surface binding kinetics did not change with mismatching nucleotide concentration, however, the KF-DNA

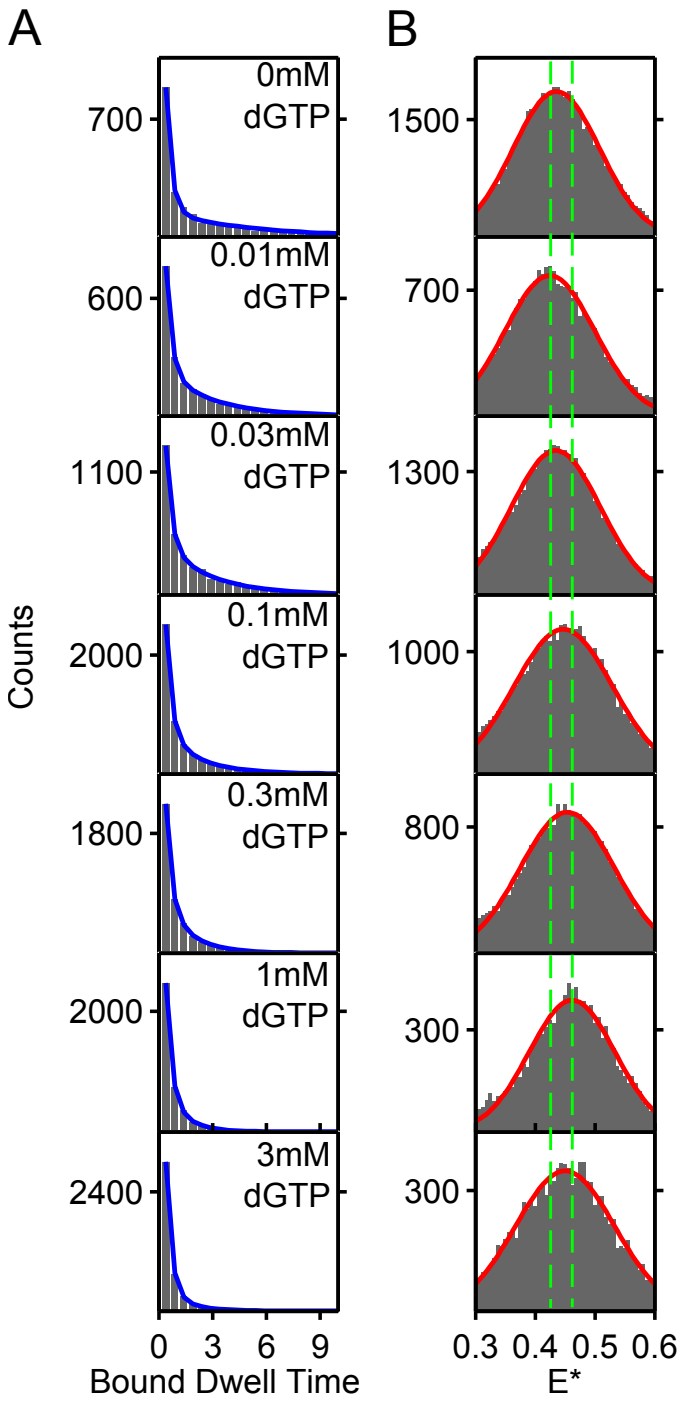


Figure 5.9: Shift open state peak position and KF-DNA dwell time decrease with increasing mismatching nucleotide concentration. (a) Dwell time histogram of all binding events as a function of mismatching nucleotide concentration. Double exponential fit in blue has the fast off rate (KF-surface dissociation) fixed at 3.1 s^{-1}, and the slow off rate (KF-DNA dissociation) freely fit. (b) Histogram of E* frame values for KF-DNA binding events. Peak positions are extracted by fitting a Gaussian profile between E*=0.3 and E*=0.6. Green lines denote the approximate extrema of the peak positions.

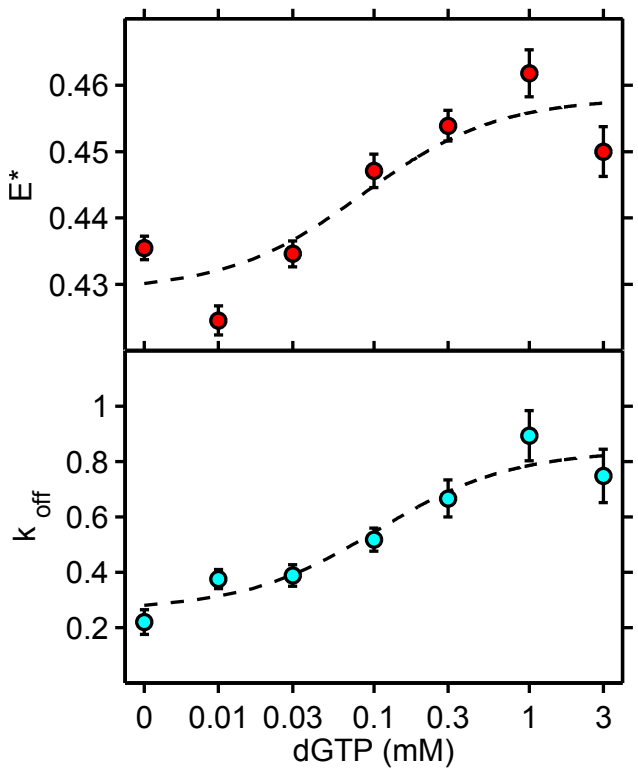


Figure 5.10: Upper panel: Observed open state peak position as extracted from the Gaussian fits in Fig. 5.9b against dGTP concentration. Dotted line is an eye guide. Lower panel: Observed slow dissociation rate (KF-DNA dissociation) as extracted from histograms in Fig. 5.9a against dGTP concentration.

dissociation rate was found to increase from $0.2s^{-1}$ to $0.7s^{-1}$ between 0 and 3mM dGTP (Fig. 5.9a and Fig. 5.10 lower panel). (KF-DNA dissociation rates were extracted with the KF-surface component fixed at $k_{off} = 3.1s^{-1}$)

While both the existence of the partially closed state[19,20,77], and increased KF-DNA dissociation rates in the presence of mismatching nucleotide[73] have has been observed previously, they have not been observed in tandem. The similarity in the form of the open peak position shift, and the increased dissociation rate, suggests that they follow the same underlying dependence (Fig 5.10), suggesting that the partially closed structure could be causing the increased dissociation rate.

Superimposing the crystal structures of the fingers-open (PDB:1L3U) and fingers-partially-closed (PDB:3HPO) KF complexes does not suggest an immediate structural basis for the increased dissociation rate of the partially-closed state. Although there are variations in the structure around the thumb region, which is involved in DNA binding, it appears that this could be caused by the different lengths of DNAs used in the two structures. As mentioned in Chapter 3, a key difference between the fingers-open and partially closed structures is that the first templating base is stacked between the O and the O1 helices in the open con-

formation, and subsequently displaced, and stacked on the primer end in the partially-closed structure. It is possible that removal from this base from contacts with the protein, and displacement of the tyrosine stacked on the primer end, decreases the contacts between the protein and the DNA, thus increasing the dissociation rate, although it is unclear if such a small change could cause the observed magnitude of change in the dissociation rate. Alternatively, the sodium ions introduced with the dNTPs could possibly contribute to the increased dissociation rate of KF, rather than structural changes in KF caused by binding a mismatching nucleotide.

5.3 Materials and Methods

In order to investigate the fingers-dynamics in the presence of absence of nucleotide, DNAs were immobilised on a passivated glass surface, as described in Chapter 4, with the same buffer, except with 4nM KF and a range of dTTP concentrations (dTTP is complementary to the first templating base). Data was acquired for 20 minutes using continuous green excitation at 1.5mW, with a 10ms frame time. Data from 96, 100, 78, 100, 89, 78 and 87 DNA molecules were collected for nucleotide concentrations of 0, 0.1, 0.3, 1, 10, 30μM dTTP respectively.

Fluorescence time-traces at each DNA were extracted using Two Tone software and KF-DNA binding events were extracted using a thresholding algorithm as described in Chapter 4. Only KF-DNA binding events greater than 0.5s in length were used, to minimise contributions from KF-surface binding events. Subsets of binding events were selected as per Fig. 5.2, for investigations of the binary complex.

Hidden Markov Modelling was performed on the concatenated KF-DNA binding events at each DNA separately, fit to 1 to 5 state HMM models, with the best fit selected using maximum evidence criteria.[106,107] The five possible states represented the FRET contributions from: two fingers-open states E*$\approx$ 0.4; two fingers-closed states E*$\approx$ 0.6 (accounting for the photo-physics of the acceptor dye); and one low FRET state in which the acceptor is bleached (E*$\approx$ 0.15). The dwells in the fingers-open and fingers-closed conformation

were extracted from these HMM fits by classifying events according to: $0.3 < E^* \leq 0.53$ as fingers-open; $0.53 < E^* \leq 0.8$ as fingers-closed; and $0 < E^* \leq 0.3$ as donor only KF.

Data for mismatching nucleotide titrations were collected at 100ms frame time using continuous wave green excitation at 0.3mW.

Chapter 6

An assay to monitor KF conformation during DNA polymerisation

Previous chapters investigated the fingers-conformations of DNA polymerase I (KF) during the pre-chemistry DNA polymerisation reaction, aided by inhibition of the polymerisation chemical step. In this chapter, we extend observations to the entire polymerisation reaction, by developing two methods which overcome the difficulties in observing rapid DNA polymerisation, both of which eschew complicated microfluidic solutions. Furthermore, in aid of characterising single-molecule DNA polymerisation, we present a novel visualisation method, which allows concise and intuitive representation of hundreds of single-molecules simultaneously, aiding method development, characterisation and presentation. Initial observations demonstrate that KF remains predominantly fingers-closed whilst polymerising DNA, with more detailed investigations presented in Chapter 7.

6.1 Introduction

Enzymes demonstrate astonishing catalytic efficiencies, catalysing reactions which would rarely occur in their absence, and are hypothesised to enhance reaction rates by binding strongly to the chemical reaction transition state.[1,116] The notion of transition state stabilisation has led to the development of transition state analogues to inhibit enzymes, and the engineering of artificial enzymes based on optimising antibody binding to "transition state" imitator molecules.[117] Despite moderate successes, artificial enzymes have failed to reach efficiencies anywhere near natural enzymes, and enzymology theory cannot yet explain

these deficits, nor has it succeeded in quantifying the origin of rate enhancement in natural enzymes.[118,2,3]

In response to these shortcomings, a number of hypotheses to extend enzymatic theory have been proposed. A contentious, although popular suggestion, is that the "dynamical" features of enzymes may contribute the "missing link" in understanding enzyme catalysis. Indeed, a flurry of recent investigations have demonstrated interesting phenomena linked to enzyme dynamics, such as rate-limiting conformational changes[13]; coupled dynamic networks[119]; intrinsic enzyme dynamics matching the catalytic rate[12]; and intrinsic disorder[120]. There have also been heated debates on whether protein oscillations can enhance quantum tunnelling reactions[121,122].

Conversely, there is criticism that the "dynamics" hypothesis has been overemphasised due to its exotic nature, and that rate enhancement can be explained using standard ideas, such as electrostatics, but whose implementations are challenging due to protein complexity. Additionally, the "dynamics hypothesis" suffers from ill-definitions and lack of clear questions[123,2], with the phrase "protein dynamics" being used to describe a broad variety of enzymes motions, which probably have disparate functions. For instance, the oscillation of atoms around chemical bonds, and concerted large scale conformational changes, both feature under the umbrella of "protein dynamics".[113] Lack of rigorous definitions have been blamed for misunderstandings in the literature.[124,125,126,7]

It is plausible that the great success of x-ray crystallography have somewhat biased enzymology towards static notions, and correlations in conformational dynamics and enzyme turnover seen in some enzymes necessitates the development of further assays to investigate the connection between protein dynamics and catalysis. Here we present such an assay, which allows the observation of conformational changes in the presence of chemistry. Specifically, we observe large-scale motions of the fingers sub-domain of KF during the DNA polymerisation reaction. In the place of dideoxy-terminated DNAs used in previous chapters, which prevented the phosphoryl transfer and hence nucleotide incorporation, here we use regular "deoxy" DNA substrates which allow nucleotide incorporation. This permits investigations of the KF fingers dynamics beyond the pre-chemistry reaction, to the entire

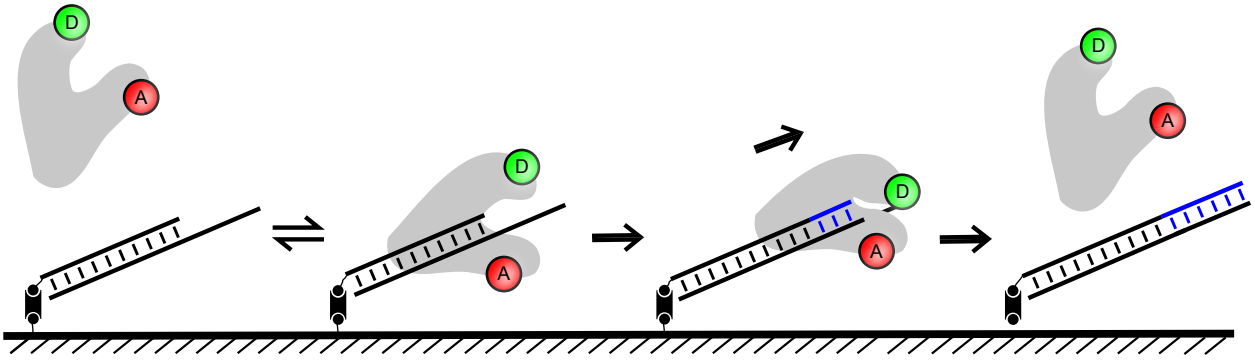


Figure 6.1: Schematic illustrating the single-molecule assay used to observe the fingers-conformation of KF during DNA polymerisation. Non-terminated DNA primer-templates are immobilised onto a glass surface (left), with KF in solution free to bind the DNA (centre-left), incorporate nucleotides into the primer strand (centre-right, blue) and dissociate, leaving an extended DNA or blunt end (right). The DNA sequence used is described in Chapter 7, and consists of a 15bp template overhang.

DNA polymerisation reaction.

Investigating real-time DNA polymerisation presents technical challenges however, as nucleotide incorporations occur rapidly ($k_{pol} \approx 50s^{-1}$)[53]. Single-molecule investigations of DNA polymerisation typically make use of "micro-fluidic chambers" to initiate polymerisation by flowing in reaction buffer, while ensemble studies used sophisticated stopped-flow devices to investigate rapid fluorescence changes. Micro-fluidic chambers, however, complicate an already non-standardised, custom-build single-molecule microscopy system, which can be challenging to use; further slowing the iterative development process of the single-molecule technique. As a result, we devised methods to allow the initiation of the DNA polymerisation reaction on a TIRF microscope, without the need for complications such as microfluidic devices.

In this chapter we present two methods for observing single-molecule DNA polymerisation on a TIRF microscope, visualisation techniques, and initial observations of the KF fingers conformation during DNA polymerisation, which demonstrate important features such as the fingers-closed dwells during polymerisation and the low binding to product DNA after polymerisation. We investigate the fingers-dynamics during polymerisation further in Chapter 7.

6.2 Initiating DNA polymerisation on a TIRF microscope without microfluidics

6.2.1 Challenges

The assay described in previous chapters demonstrated the fingers-conformation of KF molecules as they bound to surface-immobilised DNA molecules, via TIRF microscopy of an intra-protein dye pair on KF. Here we present an extension of this assay, using extendable DNA in the place of terminated DNA, to study the KF fingers-conformation during the polymerisation reaction. DNA polymerisation occurs as KF extends primer DNA by incorporating nucleotides complementary to a single-stranded template overhang, the extended primer then forming a duplex with the template overhang (Fig. 6.1).

Observing real-time DNA polymerisation presented additional challenges for our TIRF assay: DNA polymerisation is known to be rapid (rate of chemistry $\sim 50s^{-1}$ [53]), with a processive polymerisation rate of $\sim 6-15$ bases per second for KF[43,78], requiring continual sample observation after reaction initiation, to avoid missing important aspects of the polymerisation reaction. In addition, while non-extendable DNAs allowed multiple cycles of KF-DNA binding and conformational dynamics, polymerisable DNAs only polymerise once, resulting in much lower statistics for the studying DNA polymerisation compared to the pre-chemistry investigations (Chapters 4 and 5).

A typical method to deal with rapid polymerisation is to use a microfluidic system to initiate polymerisation by flowing in reaction buffer, providing a well-defined start point for the reaction, and continual sample observation by minimally perturbing the focal plane.[127] "Home-made" microfluidic systems involve flow chambers created by sandwiching adhesive tape between glass slides, drilling holes, gluing on flow ports, or melting "Parafilm" as an adhesive. Implementation demonstrated these chambers to be fragile; prohibitively time-consuming to make; made surface passivation difficult; and the adhesive tape caused increased background fluorescence. An alternate method made use of commercial fluidic chambers, although these also required the use of ports, tubes and glue, and the introduc-

tion of materials which are difficult to clear of background fluorescence. These devices also used larger reaction volumes than our silicone wells, and hence significantly more biological material, of the order of several hundred microlitres in the place of $20 - 40\mu l$.

6.2.2 Method

As a result of these difficulties, we developed a method in which the reaction was initiated by manual addition of reaction buffer onto the slide via pipette. The "reaction initiating buffer" contained a solution of nucleotides, and was added to a solution of KF molecules already under observation, thus initiating the reaction mid-experiment. However, initiating the reaction via manual buffer addition caused defocusing of the microscope, losing vital moments of observation of the polymerisation reaction. Additionally, the exact time of buffer addition, hence reaction initiation, was impossible to specify precisely. These issues were accounted for as documented below.

In order to avoid defocusing of the microscope on manual addition of buffer, forces on the coverslide were minimized by avoiding contact between the hard surfaces of the pipette-tip and the slide - when pipetting in reaction buffer, all contacts were made at liquid interfaces. Secondly, an autofocusing system (CRIFF) was used to rapidly return the focal plane to a defined position in the case of defocusing. The autofocusing system relied on the total-internally reflected laser beam to indicate focal plane position, making it important to match refractive indices of the solutions being mixed to avoid perturbations to the reflected beam caused by refractive index changes.

In the following sections we demonstrate the effect of flowing in a high concentration of nucleotides on the overall system and on the individual molecules. Fluorescence from surface-immobilised DNAs were extracted as described previously, except the position of each KF binding event was localised, and binding events further than half a pixel from each localised DNA were removed, to minimize contributions from KF-surface binding. This was possible for these experiments as the microscope drift was low and acquisition times were relatively short (5 minutes).

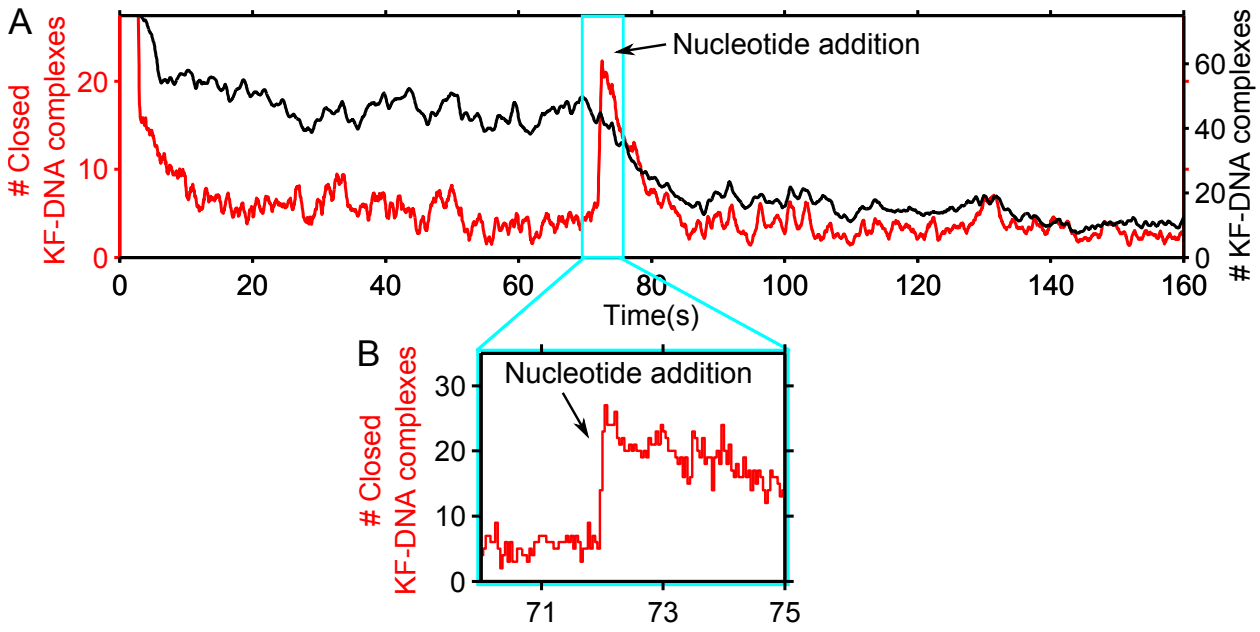


Figure 6.2: (a) Effect of nucleotide addition on a sample containing KF and immobilised DNA, measured by the number of KF-DNA complexes (black trace), and the number of KF-DNA complexes in the fingers-closed conformation over time (red trace). Equilibrium KF-DNA binding occurs up to $\sim 72s$, at which point nucleotide addition occurs and is followed by a rapid increase in the number of fingers-closed complexes, followed by a gradual decrease in number of KF-DNA complexes. The out-of-range large signal at 0-5s is due to the localisation of DNAs by using the red laser, and should be disregarded. (b) Expansion of the cyan region. The number of fingers-closed complexes demonstrates a sharp increase on nucleotide addition.

6.2.3 Results

Prior to reaction initiation, surface-immobilised extendable DNA were localised in a well containing 20μl of buffer solution containing 4nM KF [i]. After $\sim 70s$ seconds, DNA polymerisation was initiated by pipetting on 20μl of "reaction-initiating" buffer to the well buffer. The "reaction-initiating buffer" was equivalent to the "well-buffer" but also contained 200μM of dATP, dTTP, dCTP and dGTP, to rapidly initiate the DNA polymerisation reaction. The refractive indices between "reaction-initiating buffer" and the "well-buffer" were matched by mixing the "well-buffer" with an equal volume of water, to the volume of nucleotide solution mixed in the "reaction-initiating buffer".

[i] As in previous chapters, buffers consisted of: 50mM Tris, pH 7.5; 100μg/ml BSA; 1mM DTT; 5% glycerol;1mM Trolox (UV treated); 1.9mg/ml glucose oxidase;0.075mg/ml catalase; 1.5%w/v D-glucose; 10mM $MgCl_2$

Bulk behaviour

In order to study the effect of the massive increase in nucleotide concentration on the well containing KF and immobilised DNA, two metrics were examined. Firstly, KF-DNA binding and dissociation was quantified by the number of KF-DNA complexes on the surface at each time point (Fig. 6.2, black trace). In addition, the KF fingers-conformation across the sample was quantified via the number of fingers-closed KF-DNA complexes at each time point (Fig. 6.2, red trace, defined by the number of KF-DNA complexes with E*>0.53).

After nucleotide addition, we observed the number of KF-DNA complexes decreasing steadily over $\sim 10s$ (Fig. 6.2, black trace indicates ~ 50 KF-DNA complexes up until ~ 72s, followed by a decrease between $\sim 70 - 90s$ to ~ 15 complexes). We attributed this to the chemical extension of the primer DNA strand, creating a DNA blunt end (or short template overhang, 1-3bp), known to have a significantly reduced polymerase affinity, resulting in a drop in KF-DNA binding[128]. The timescale reflects the time taken to bind and polymerise the 15 base-pair template overhang.

Around the time of nucleotide addition, a sharp increase in the number of KF molecules in closed conformation was observed (Fig. 6.2a, red trace, increase from ~ 5 to ~ 20 closed complexes at $\sim 72s$), and return to a lower value over $\sim 10 - 12s$. We attributed this to the rapid closing of KF on arrival of nucleotides, incorporation of nucleotides into the DNA, with number of closed KF molecules reducing as the DNAs became fully extended. The sharp increase in the number of closed KF molecules (Fig. 6.2b, red trace doubling in one frame at $\sim 72s$) suggests that nucleotide addition and mixing occurs rapidly (within a CCD frame). The simultaneous closing of polymerases allows us to use this event as the starting point of polymerisation, to the precision of the experimental time resolution (40ms), thus circumventing the problem of being unable to specify the reaction-initiation-point manually.

Single-molecule behaviour

The above bulk signal reflected the collection of single-molecule behaviours at individual DNAs. At a single molecule level, individual DNAs would typically demonstrate binding and dissociation of KF molecules prior to nucleotide addition (Fig. 6.3a: Increases in fluorescence

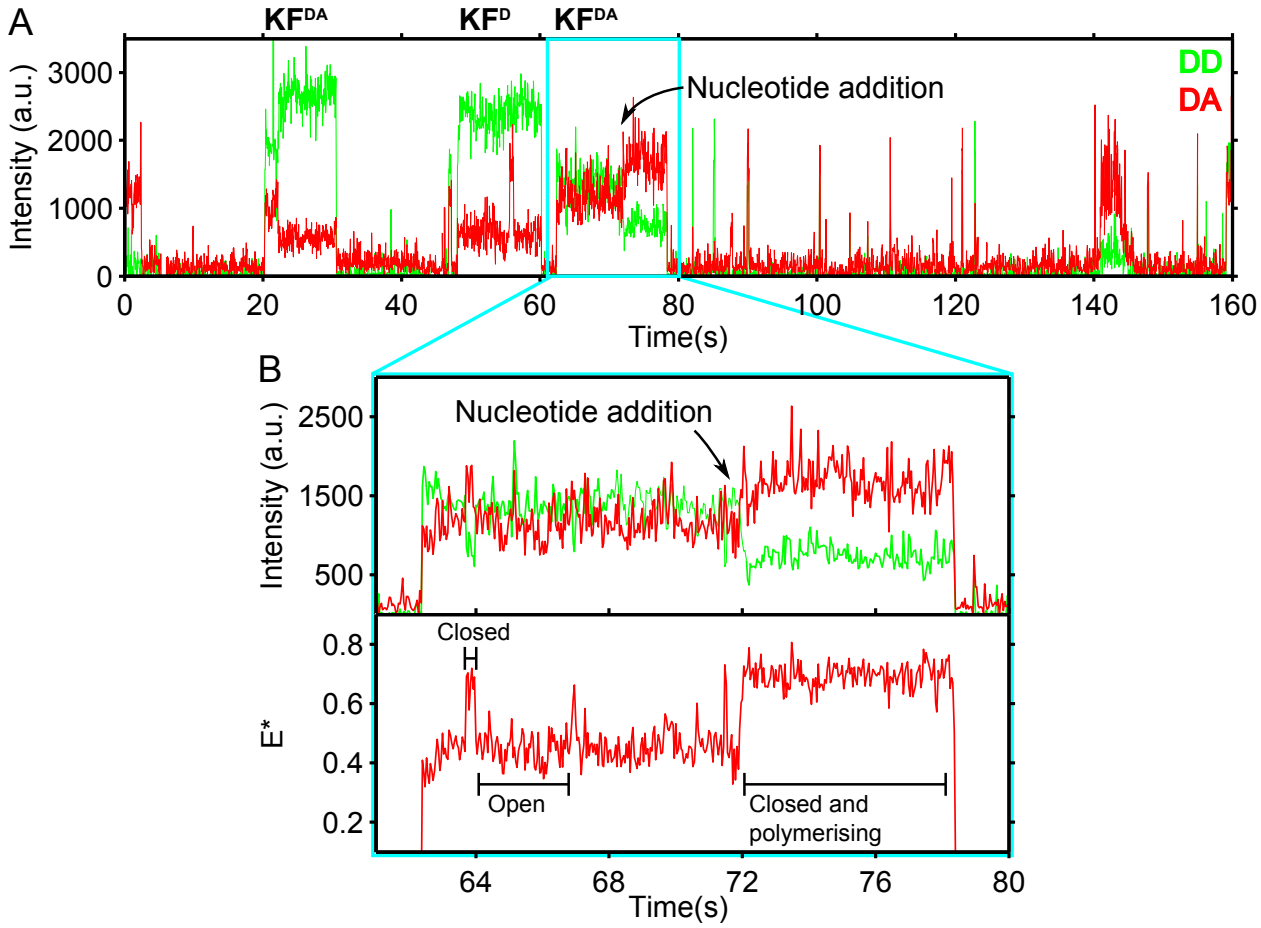


Figure 6.3: Fluorescence time-trace at a single DNA molecule demonstrating KF-DNA binding and DNA polymerisation. (a) Fluorescence increases demonstrate KF-DNA binding prior to nucleotide addition (KF^{DA}, KF^{D}). Decreased fluorescence after nucleotide addition ($>$ 72s) indicates lower KF-DNA binding post-polymerisation. (b) Expanded cyan box, demonstrating the rapid closing of the KF fingers on nucleotide addition at $\sim$ 72s, remaining closed for $\sim$ 6s, followed by dissociation. This is interpreted as KF polymerising the DNA, leaving a short overhang or blunt end. Prior to nucleotide addition the KF-DNA complex undergoes binary complex fingers-dynamics. Donor-excitation, donor-emission (DD - green). Donor-excitation, acceptor-emission (DA - red). FRET efficiency (E* - red)

from $\sim 0 - 72s$). Upon nucleotide addition, KF-DNA complexes interconverted into the fingers-closed state, remaining closed until KF dissociation (Fig. 6.3b: Lower panel, increase in E* from 0.4 (open) to 0.7 (closed) at $\sim 72s$). Thereafter, KF binding to DNA was lower (Fig. 6.3a, lower binding after 78s). The length of the closed state dwell ($\sim$ 6s) was much longer than the fingers-closed dwells seen in the ternary-complex when using dideoxy-terminated DNAs ($\sim 150 - 200ms$), suggesting that KF incorporates multiple nucleotides into the DNA strand. It was notable that the polymerase did not appear to return to the fingers-open conformation at any time during polymerisation, suggesting that any returns to the open state conformation were rapid ($<$ 40ms)

In order to confirm the presence of polymerisation, we also performed a single-molecule translocation assay, which demonstrated that the DNA was being extended via observations of KF translocating along DNA as it sequentially incorporated nucleotides. This work is described in Appendix B.

6.3 Diversity of single-molecule behaviours during DNA polymerisation

6.3.1 Data representation challenges in the single-molecule field

Above, the effect of adding complementary nucleotides to a solution of KF and surface-immobilised DNA molecules was represented by a single fluorescence time-trace and by the behaviour of the DNA "ensemble". This illustrates a key challenge in the single molecule field: How does one faithfully and convincingly convey a data set containing hundreds single-molecules worth of data, when it is impossible to display all the individual molecules? The strength of the single-molecule techniques lies in the ability to measure individual behaviours, but presenting their diversity is difficult because of the quantity of data. Difficulties in data presentation makes reviewing single-molecule work challenging.

Consider a hypothetical scenario using the current system: KF molecules transition between open and closed states, but on occasion, a novel FRET state is seen during open-closed

transitions. An experimenter might sift through hundreds of single-molecules of data; view this behaviour, interpret it as an additional structural state, presenting it in publication via the use of a few "representative" time-traces. The small number of time-traces presentable in a journal (all of which agree with the experimenters interpretation), make it difficult to judge the paper as the majority of the data set is absent. Even if a statistic is given (i.e. X% of the molecules show a given behaviour), the statistic is based on the experimenters interpretation of the entire data set. Ensuring that observed signals are not due to the exceptional ability of humans to spot false patters in noise (apophenia), rigorous statistical tests are required, such as hidden markov modelling, to distinguish genuine signals from noise.

The single-molecule field would benefit greatly from the development of generalisable visualisation techniques, which are concise, intuitive, and preserve individual molecule behaviours, but are minimally reliant on external models. Presenting entire data sets in a concise and intuitive form facilitates the peer-review process by allowing easier assessment of raw data and conclusions drawn. In order to represent and convey the breadth of behaviours during single-molecule DNA polymerisation, we developed a novel visualisation technique which allowed concise representation of hundreds of molecules worth of data simultaneously.

This visualisation technique has proved useful in presenting results to an audience, but also in the day-to-day experiments performed on DNA polymerase. Single-molecule experiments produce a large amount of data, making analysis laborious, requiring sifting through hundreds of molecules of data. Preliminary experiments thus become particularly challenging, as the cycle of experiment and analysis during assay development (optimising factors such as laser intensity, polymerase concentration, imaging buffer and nucleotide concentration) becomes slowed by analysis time. Our visualisation technique assisted in assay development by allowing rapid on-the-fly assessment of an experiment.

6.3.2 A novel visualisation representing hundreds of single-molecules simultaneously

Here we present a visualisation technique which allows the concise and intuitive representation of hundreds of molecules simultaneously, with minimal application of an external model.

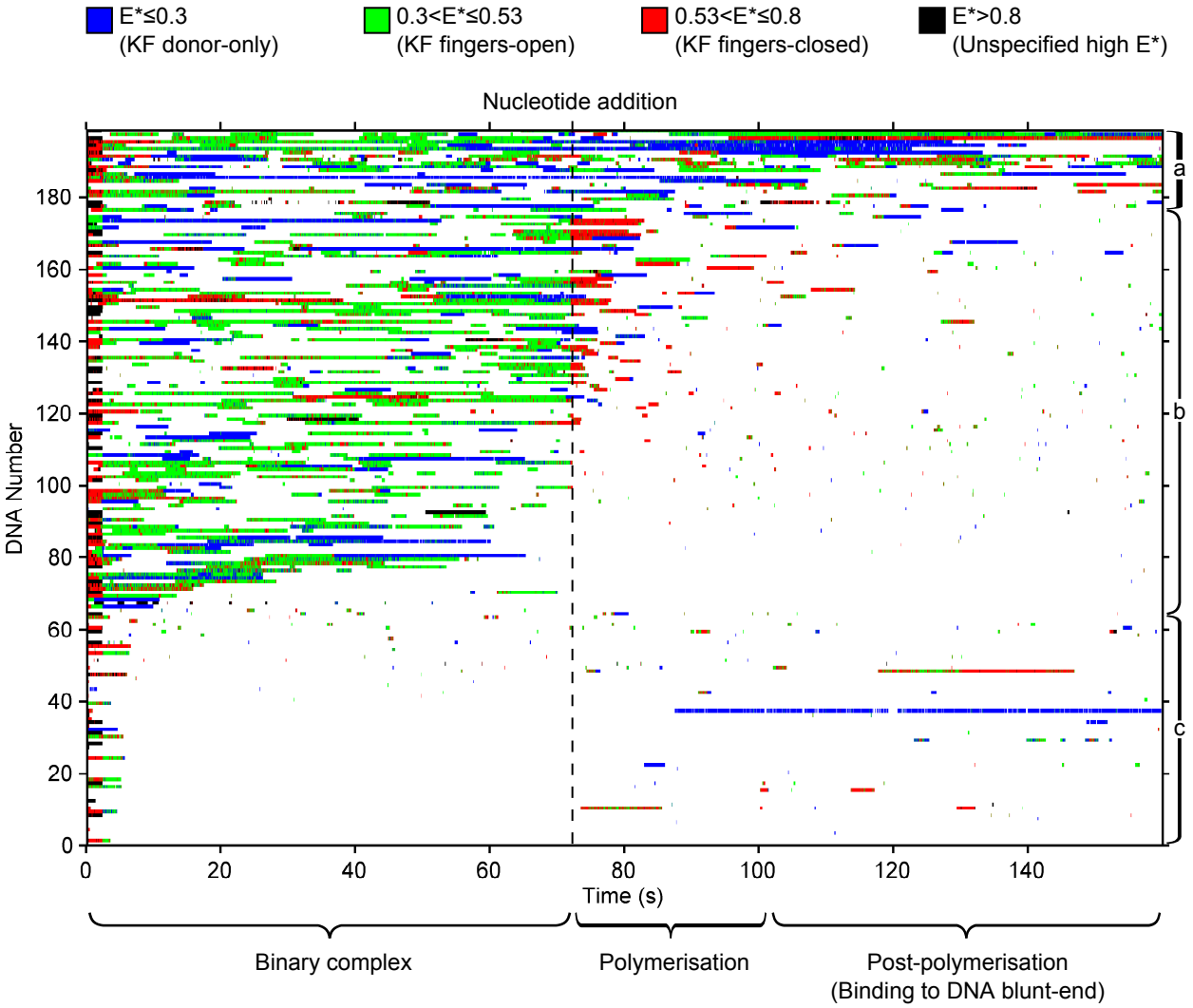


Figure 6.4: Visualisation representing the binding and conformation of KF molecules to an ensemble of $\sim$ 200 DNA molecules. Each horizontal line represents fluorescence time-trace at a DNA, with KF binding events represented in colour, the colour reflecting the E* at each time point. Green: KF fingers-open. Red: KF fingers-closed. Blue: Donor-only KF. Black: Unspecified. Binding events before nucleotide addition (dotted line) are seen to be predominantly open, representing binary complex behaviour. A large increase in fingers-closed KF molecules is seen after nucleotide addition, interpreted as KF molecules polymerising DNA. Decreased KF-DNA binding is seen subsequent to these polymerisations. The period between 0-3 seconds is due to the activated red laser used to localise the DNAs, and should be disregarded. (a) DNA molecules which do not show decreased KF binding after nucleotide addition. (b) DNA molecules which show decreased KF binding after nucleotide addition. (c) DNA molecules which show little KF binding at all.

Each DNA molecule has a fluorescence time-trace cataloguing the binding, dissociation, and the conformational state of KF (e.g. Fig. 6.3). We reduce these fluorescence time-traces into colour-coded 1D time-trace lines, which are white at time-points with no bound KF present, but with KF-DNA binding events represented by colour blocks, with the colour representing the KF conformational state (Fig. 6.4, e.g. one horizontal line). The colour used at each point on the line represents the FRET efficiency of KF, and hence the conformation of the polymerase: Green is used to assign an open polymerase ($0.3 < E^* \leq 0.53$); red a closed polymerase ($0.53 < E^* \leq 0.8$); blue a polymerase without an acceptor dye ($0 < E^* \leq 0.3$) (sometimes due to bleaching); and black is used for a unassigned high FRET values ($0.8 < E^* \leq 1$). This reduces the DNA ensemble into a stack of 1D colour-coded lines, rather than bulky fluorescence time-traces (Fig. 6.4).

This collection of colour-codes lines is then sorted to improve viewability, by stacking the lines according to simple classification rules, resulting in the formation of groups of behaviours. The lines are initially sorted by extent of KF-DNA binding before nucleotide addition. Then, DNAs which show sufficient binding before nucleotide addition are then sorted by extent of binding after nucleotide addition. This results in three groups of DNAs behaviours: DNAs which do not show KF binding (Fig. 6.4c, DNAs $1 - 64$); DNAs which show KF binding, which continues after nucleotide addition (Fig. 6.4a, DNAs $178 - 199$); and DNAs which show KF binding before nucleotide addition and decreased binding thereafter (Fig. 6.4b, DNAs $65 - 177$). This visualisation is used to aid analysis of the DNA polymerisation behaviours across a collection of DNAs, in the next section.

6.3.3 Classifying single-molecules into behavioural groups

Using this visualisation technique, the effect of adding nucleotides to a sample of polymerases and surface-immobilised DNAs becomes clear, whilst also preserving the individual behaviours of the molecules (Fig. 6.4). Before nucleotide addition, the polymerases bind and dissociate to DNA predominantly in the open state, consistent with binary complex behaviour (Fig. 6.4: Green events between $\sim 0 - 72s$ indicate that KF molecules are predominantly fingers-open). Upon nucleotide addition, the number of polymerases in the

fingers-closed state (red) increases dramatically, and remain closed for a few seconds, which we interpreted as polymerisation of DNA (Fig. 6.4: Red binding events between $\sim 72-90s$). There is a stark decrease in the KF-DNA binding after nucleotide addition, interpreted as formation of DNA blunt-ends/short overhangs (Fig. 6.4: White space $\sim 90-160s$)

The visualisation also structures the molecules into behavioural groups:

DNA molecules 1-64 (Fig. 6.4c) show no KF binding: This could be due to DNA primer molecules which are not hybridised to a template; KF molecules which are non-specifically adsorbed to the surface and mistaken for DNA molecules during initial DNA localisation; or DNA molecules which are rendered inactive by some other process.

DNA molecules $\sim$ 65-177 (Fig. 6.4b) show KF-DNA binary complex binding before nucleotide addition, followed by a decrease in binding after nucleotide addition (suggesting DNA polymerisation has occurred). A subset of these molecules $\sim 120-177$ feature a closed (red) KF molecule upon nucleotide addition or shortly after, interpreted as a KF molecule polymerising the DNA. (Some show KF with bleached acceptor dye (blue)). DNA molecules $\sim 65-120$ show decreased binding after nucleotide addition, suggesting the DNA is polymerised, however they lack the presence of a closed polymerase after nucleotide addition. We attribute these "dark polymerisations" to DNA polymerisations by unlabelled KF molecules, or acceptor only polymerases. The protein labelling protocol results in $\sim 25\%$ of KF molecules which are unlabeled[20], while counting acceptor only binding events in ALEX experiments reveals that $\sim 30\%$ of labelled KF molecules are labelled with acceptor only. This gives a total of $\sim 50\%$ of KF molecules which are invisible in non-ALEX experiments, consistent with $\sim 50\%$ of the DNA molecules appearing to have been polymerised by an invisible KF molecule above.

DNA Molecules $\sim$ 178-199 (Fig. 6.4a) do not show the characteristic decrease in KF binding after nucleotide addition. These are categorised as "noise traces", which could be attributed to persistent non-specific adsorption of KF to the surface, possibly caused by defects in the PEG passivation at a specific point, or due to KF-neutravidin binding.

The visualisation technique allows us to categorise our DNA molecules into groups: DNAs which show no polymerase binding (32%); DNAs which demonstrate a polymerisation sig-

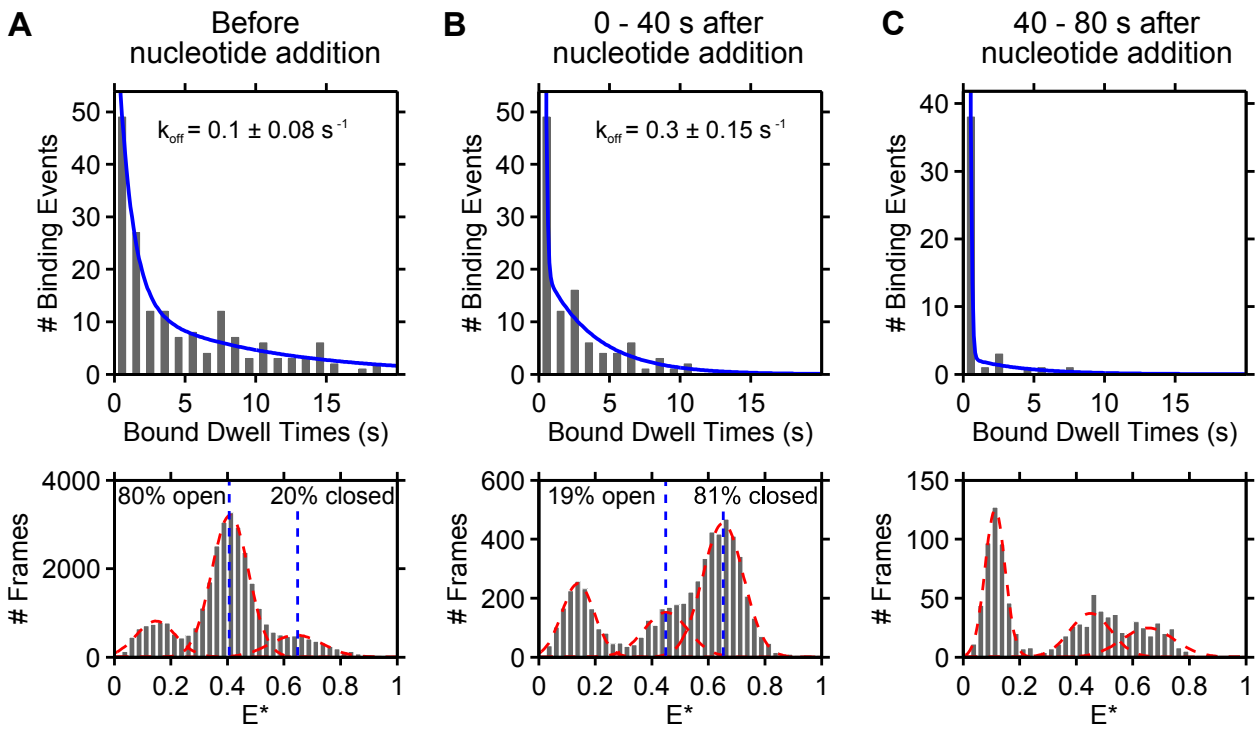


Figure 6.5: KF-DNA dissociation rates and KF fingers-conformation before, during and after DNA polymerisation. Histogram of bound dwell times are fit to double exponentials to extract dissociation rates, with the first exponential representing KF-surface binding as discussed in Chapter 4, with the second representing KF-DNA binding. KF fingers-conformation is represented in histograms of E* values at each frame, fit to 3 Gaussian profiles to extract the population of donor-only KF (E*=0.15), fingers-open KF (E*=0.4), and fingers-closed KF molecules (E*=0.65). (a) Before nucleotide addition KF-DNA complexes show binary complex behaviour, similar to that defined in previous chapters, with slow dissociation rate ($k_{off} = 0.1 \pm 0.08 s^{-1}$), and KF predominantly fingers-open (80%) (b) After nucleotide addition, KF adopts a predominantly fingers-closed conformation (81% closed), and dissociates more readily from DNA ($k_{off} = 0.3 \pm 0.08 s^{-1}$). (c) 40-80s after nucleotide addition, KF binding is significantly decreased.

nature (56%); and DNAs which do not demonstrate a polymerisation signature (12%). In the next section we concentrate our investigation on the polymerised DNAs (DNA molecules $\sim 64 - 177$).

6.3.4 Behaviour of KF molecules before, during and after DNA polymerisation

DNAs molecules $\sim 65 - 177$ were categorised as having been polymerised by KF. The KF fingers-conformation and KF-DNA dissociation rates at these DNAs before, shortly after, and long after nucleotide addition were investigated to further characterise DNA polymerisation

at the single-molecule level.

Before nucleotide addition, the KF fingers-conformation was predominantly open and the KF-DNA dissociation rate was consistent with formation of the binary complex, as described in previous chapters. (Fig. 6.5a, lower panel, FRET efficiency histogram defines KF as 80% open and 20% closed. Upper panel, fit to dwell time histogram with biexponential defines slow KF-DNA dissociation rate ($k_{off} = 0.1 \pm 0.08 s^{-1}$), whilst the second component represents transient KF-surface binding, described in Chapter 4) (This is in comparison with $k_{off} = 0.2 s^{-1}$ and KF being 89% fingers-open on dideoxy terminated DNAs, see Chapter 5).

In the 40 seconds after nucleotide addition, KF shifts to a predominantly fingers-closed conformation as DNA polymerisation occurs (Fig. 6.5b, lower panel shows KF is 19% fingers-open and 81% fingers-closed). Additionally, a faster dissociation rate is seen ($k_{off} = 0.3 \pm 0.08 s^{-1}$), suggesting that DNA polymerisation occurs faster than the time of a typical KF-DNA binding event, with dissociation caused by the decreased KF affinity for fully double-stranded DNA. (Fig. 6.5b, fit to dwell time histogram).

By 40 to 80 seconds after nucleotide addition, binding is significantly decreased, with residual KF-surface binding at short time scales, whose contribution to the histogram was consistent before, during and after nucleotide addition (Fig. 6.5a,b,c, first bin in histogram).

6.3.5 KF fingers conformation during DNA polymerisation

In the previous section, the collection of KF molecules which polymerised DNA molecules 64-170 were seen to be 81% in the fingers-closed state, and 19% in the fingers-open conformation. In order to quantify the KF fingers behaviour during DNA polymerisation in more detail, the fingers behaviour at individual DNA molecules was visualised (DNA molecules 120-177 were selected as they showed polymerisation by a visible KF molecule).

Associated with each DNA was a FRET-efficiency histogram describing the KF fingers-conformation during DNA polymerisation (KF-DNA binding events in the 40s region after nucleotide addition), such as those seen in Fig. 6.5. The collection of FRET efficiency histograms across DNAs 120-177 were sorted by mean FRET value (ignoring donor only frames) and visualised in a stack of histograms, plotted as a colourmap (Fig. 6.6a). Each

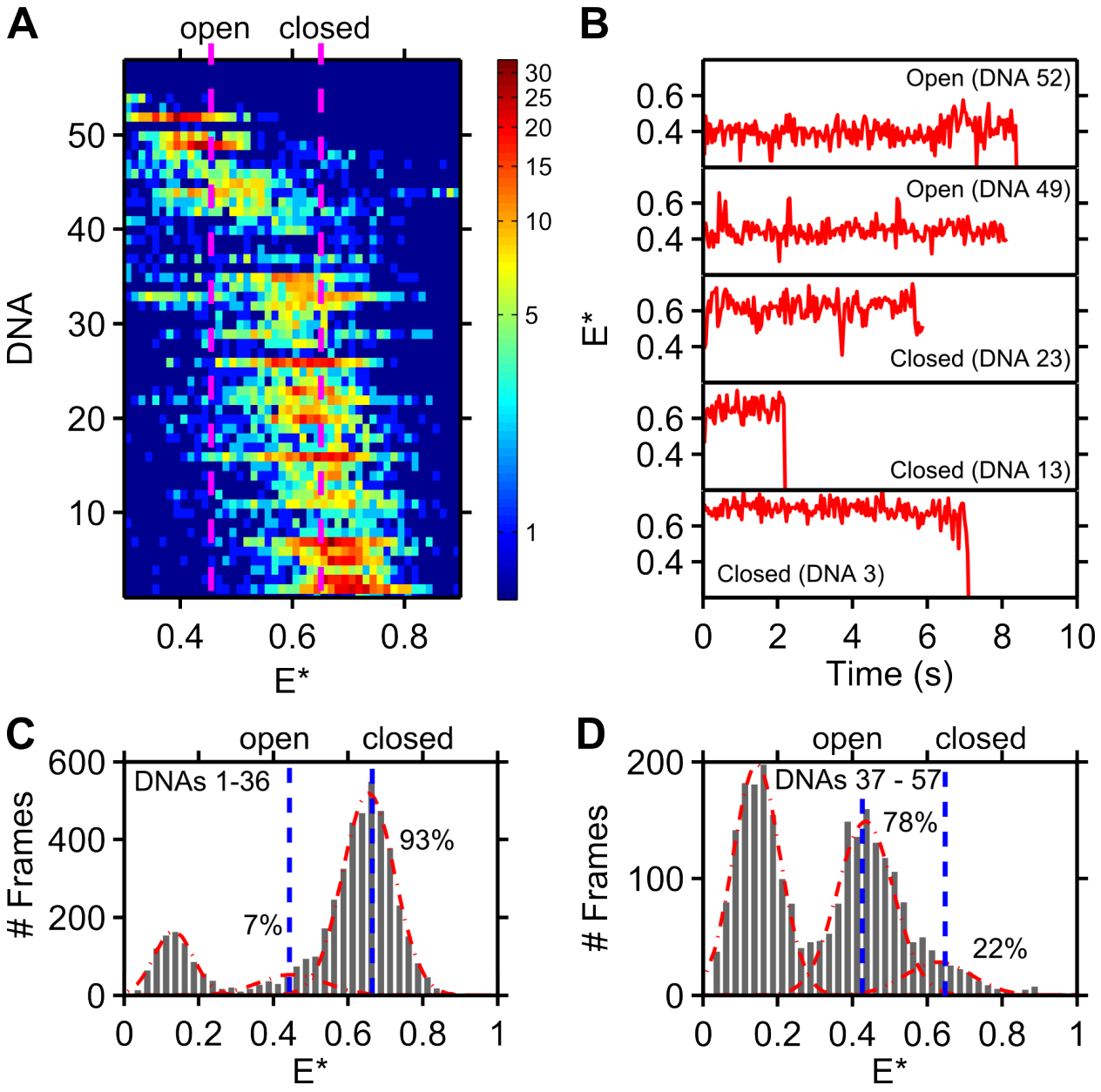


Figure 6.6: (a) Stack of FRET efficiency histograms representing KF fingers-conformation at each polymerised DNA molecule, during DNA polymerisation. Histograms are sorted by mean E*. DNAs are seen to sort into two groups: One group is predominantly fingers-closed (1-36); another is predominantly fingers-open (37-57). The colourbar represents the number of frames spent at each E* value. (b) Sample E* traces of concatenated KF-DNA binding events at DNAs during polymerisation. (c) Summed histograms of FRET efficiencies at DNA molecules 1-36, demonstrating KF to be 7% fingers-open and 93% fingers-closed. Low FRET peak from donor-only molecules. (d) Summed histograms of FRET efficiencies at DNA molecules 37-57, demonstrating KF to be 80% fingers-open and 20% fingers-closed.

horizontal line representing the FRET efficiency histogram of polymerisation at an individual DNA.

The DNAs were seen to cluster into two groups: Those with predominantly high FRET values (Fig. 6.6a, DNAs 1-36), and those with lower FRET values (Fig. 6.6a, DNAs 37-57). The summed histograms of these two groups demonstrated the high FRET group to be 7% fingers-open and 93% fingers-closed (Fig. 6.6c, DNAs 1-36); and the low FRET group to be 78% fingers-open and 22% fingers-closed (Fig. 6.6d, DNAs 37-57).

This demonstrated the overall FRET efficiency during DNA polymerisation is accounted for by two separate groups of KF molecules. One conforms with what is expected of KF molecules polymerising DNA: Binding to DNA, rapidly binding complementary nucleotide and closing, followed by successive polymerisation cycles.

The KF molecules in the predominantly fingers-open conformation do not conform to reasonable expectations of how polymerising KF molecules should appear. However their behaviour is consistent with the sub-population of DNAs whose KF binding events did not respond to the presence of nucleotides, described in Chapter 4, whose lack of correspondence with diffusion-based measurements allows us to discount them.

Thus we demonstrate that, to our time resolution (40ms), KF molecules polymerise DNA predominantly in the fingers-closed conformation, suggesting that any post-chemistry steps after fingers-opening steps are faster than $25s^{-1}$. More importantly though, it demonstrates that these post-chemistry steps contribute negligibly to the overall time taken to polymerise DNA.

6.4 A complementary method for initiating DNA polymerisation without the use of microfluidics.

Above, DNA polymerisation on a TIRF microscope was initiated by manually adding a large concentration of nucleotides to the reaction well. The addition point was then recognised by the sharp switch to the fingers-closed conformation of KF molecules bound to the surface-immobilized DNAs.

This method, however, was limited: it was less useful for investigating DNA polymerisation at low nucleotide concentrations, as low nucleotide concentrations did not induce the rapid simultaneous closing of the KF molecules required to specify the nucleotide addition time-point. Secondly, while this method could observe the polymerisation of a 15 base-pair template overhang, it did not provide sufficient sensitivity to observe rapid events, such as a single-nucleotide incorporation. Thirdly, the extent to which the solution was homogeneously mixed at the start of polymerisation was unclear, making it difficult to assign the precise nucleotide concentration values during polymerisation.

A separate method of initiating DNA polymerisation on a TIRF microscope was developed in order to address these weaknesses, based on using the diffusion-binding time of KF to delay the onset of DNA polymerisation after solution addition. Surface-immobilised DNA molecules were initially localised in a FOV, with no nucleotides or KF in solution. Shortly after DNA localisation, a solution containing nucleotides (at any concentration) and a low concentration of KF molecules ($\sim 0.5 - 1nM$) was added. The low concentration of KF molecules resulted in a time-lag before DNA polymerisation, as KF diffused to bind the DNA. This time-lag allowed the solution to mix fully before polymerisation began, allowing the concentration to be assigned more precisely.

This method of polymerisation initiation was characterised and compared with the previous method to check for consistency between the two methods.

6.4.1 Initiating DNA polymerisation using a low KF concentration

Prior to reaction initiation, surface-immobilised extendable DNA were localised in a well containing 20μl of a buffer solution without nucleotides or KF molecules. Approximately ten seconds after localising DNAs and beginning data acquisition, DNA polymerisation was initiated by pipetting on 20μl of "reaction-initiating" buffer which contained 200μM of dATP, dTTP, dCTP, dGTP and 0.5-1nM of KF. The refractive indices between "reaction-initiating buffer" and the "well-buffer" were matched by using water in the place of nucleotides and KF storage buffer in the place of KF in the "reaction-initiating buffer" (the high glycerol

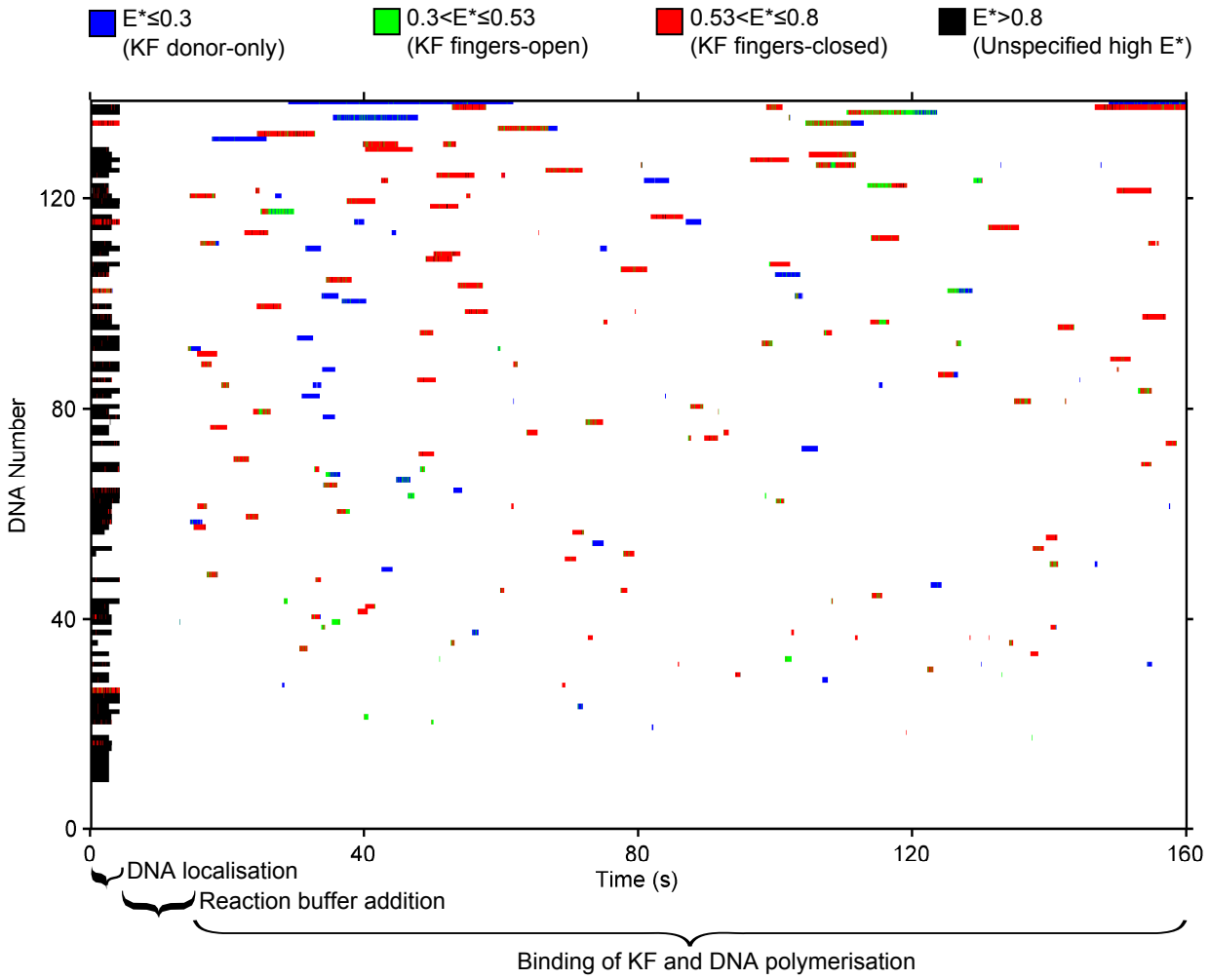


Figure 6.7: Visualisation representing the binding and conformation of KF molecules to an ensemble of DNA molecules. Each horizontal line represents fluorescence time-trace at a DNA, with KF binding events represented in colour, the colour reflecting the E* at each time point. Green: KF fingers-open. Red: KF fingers-closed. Blue: Donor-only KF. Black: Unspecified high FRET. The period between 0-5 seconds is due to the DNAs being localised. KF and nucleotides are added at $\sim$ 10s, and binding events are seen to be distributed after this, representing the stochastic binding of KF as it diffuses and binds to DNA. KF is seen to predominantly be fingers-closed, consistent with it polymerising DNA.

content in the KF storage buffer would otherwise cause a change in refractive index on mixing, resulting in defocusing).

The effect of initiating DNA polymerisation using a limiting concentration of KF is visualised in Fig. 6.7, in which the KF-DNA binding and conformation at each DNA is represented in 1D lines, as specified in Section 6.3.2. In this case, the 1D lines are all sorted according to the amount of KF-DNA binding after solution addition. KF-DNA binding events are predominantly fingers-closed, and there is a clear distribution in the KF-DNA binding times, suggesting that KF binds DNA several seconds after addition of reaction buffer.

The fingers-conformation of KF during DNA polymerisation is investigated as in Section 6.3.5. FRET efficiency histograms characterise the KF fingers conformation at each DNA, and are stacked in a colourmap according to mean FRET values. Here, again, we see two clear groups - DNAs 1-100 which exist predominantly in the closed conformation, and a smaller number of DNAs which exhibit a less clear distribution of FRET values, consistent with previous results. The KF binding events at DNAs 1-100 are 95% fingers-closed (Fig. 6.8).

6.5 Discussion

In this chapter we have presented two complementary methods for initiating DNA polymerisation on a TIRF microscope, which don't require the use of microfluidic systems. The first method relies on manual addition of a large concentration of nucleotides, mid-experiment, to a solution of KF molecules in binding equilibrium with surface immobilised DNAs. Nucleotide addition was seen to result in a rapid closing of KF-DNA complexes, which was used to specify the nucleotide addition time-point to within $\sim 40 - 80ms$, and was followed by a decreased KF-DNA binding. The striking change in behaviour convincingly demonstrated the effect of adding nucleotides, however, the method was limited to investigating relatively slow processes (such as multiple nucleotide additions); has an undefined nucleotide concentration time-profile; and was less effective at low nucleotide concentrations.

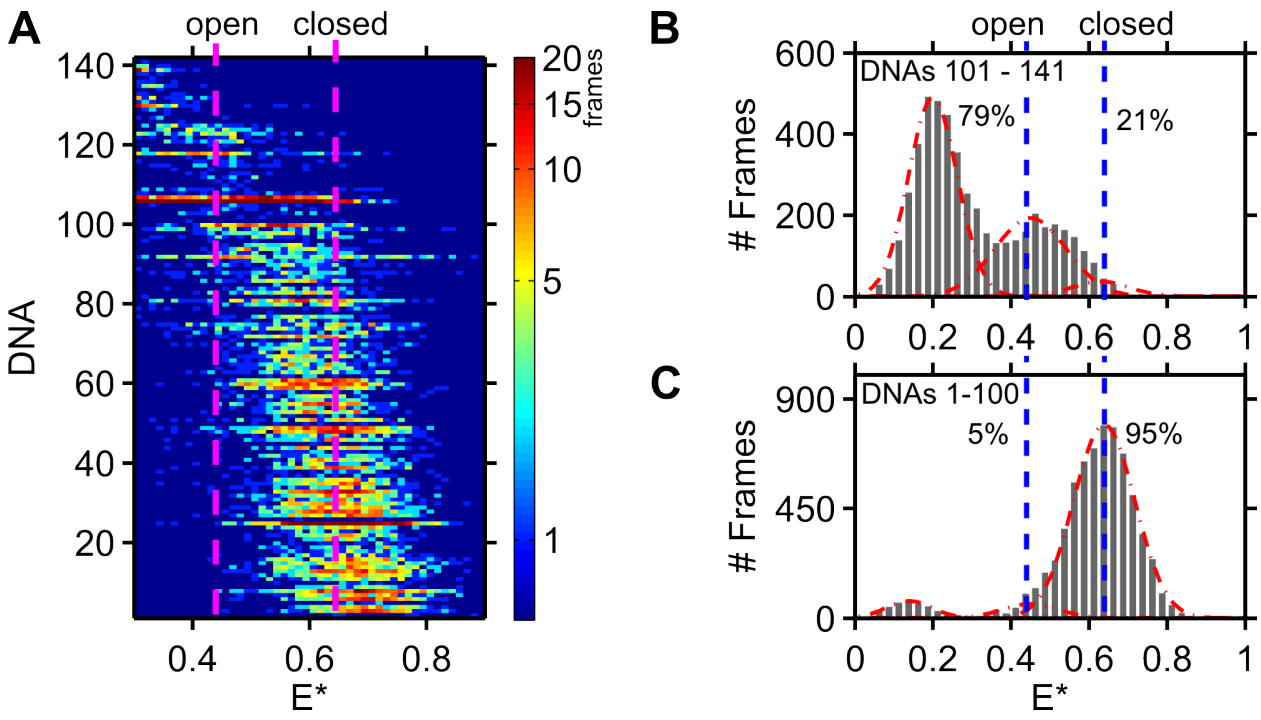


Figure 6.8: (a) Stack of FRET efficiency histograms at each polymerised DNA molecule, representing KF fingers-conformation during DNA polymerisation. Histograms are sorted by mean E*. DNAs are seen to sort into two groups: One group is predominantly fingers-closed (1-100); another is predominantly fingers-open (101-141). The colourbar represents number of frames spent at each E* value. (b) Summed histograms of FRET efficiencies at DNA molecules 101-141, demonstrating KF to be 79% fingers-open and 21% fingers-closed. Low FRET peak from donor-only molecules. (c) Summed histograms of FRET efficiencies at DNA molecules 1-100, demonstrate KF to be 5% fingers-open and 95% fingers-closed.

A second, complementary method, was devised to avoid these limitations. The second method involved reaction initiation via manual buffer addition, but relied on delaying DNA polymerisation by using low KF concentrations. This allowed relaxation and mixing of the system before DNA polymerisation beginning. This method is used in the next chapter in further investigations of the KF fingers conformation during DNA polymerisation.

We have presented a novel and intuitive visualisation technique allowing observation of KF binding and conformation to hundreds of DNAs simultaneously, division into behavioural groups, and classification of behaviours across an entire field-of-view. This visualisation technique is not limited to the current system and other researchers in our group have applied the visualisation technique to the investigation of RNA polymerase, and to investigations involving polymerisation of long DNA strands.

High concentrations of nucleotides were seen to rapidly result in KF fingers-closing, which we interpreted as DNA polymerisation. This was followed by dissociation of KF, and significantly decreased KF-DNA binding thereafter, interpreted as the decrease in KF affinity for the newly polymerised short template overhang/DNA blunt end. However, the assay was unable to categorically state how many nucleotides were incorporated into the strand, or whether the KF dissociates before the DNA it entirely polymerised - we investigate this further in the next chapter.

During DNA polymerisation, KF was found to exist predominantly in the closed state, suggesting that fingers-closing around a complementary nucleotide is much faster than the time resolution (40ms), consistent with previous measurements of the fingers-closing rate ($k_{close} = 140s^{-1}$).[42] It also suggests that any post-chemistry steps after fingers-opening are also rapid ($> 25s^{-1}$).

In the next chapter we investigate single-molecule DNA polymerisation further, by quantifying the times taken to polymerise DNA, observing single-nucleotide incorporations, and quantifying the efficiency at which KF incorporates nucleotides.

Chapter 7

KF fingers conformations during DNA polymerisation

Here we investigate the fingers conformations of KF during the entire DNA polymerisation reaction, without terminating the reaction at the chemical step as seen in previous chapters. We observe processive nucleotide incorporation over different sequence lengths and nucleotide concentrations, and we observe single-nucleotide incorporations. Our results suggest that nucleotide incorporation rates are heterogeneous, which could account for the diversity of polymerisation rates reported in the literature. Additionally, the results suggest that the steps after chemical incorporation that rate-limit processive DNA polymerisation, are caused either by fingers-opening, or by steps between chemistry and the fingers-opening step. We demonstrate that nucleotides are incorporated efficiently once KF is closed around a nucleotide, and that a similarity in intrinsic binary complex dynamics and nucleotide incorporation rates could reflect an underlying mechanism to be investigated in future work.

7.1 Introduction

Investigations of DNA polymerisation mechanisms have generally focused upon the pre-chemistry reaction steps, which define the polymerisation rate (k_{pol}) and the Michaelis constant (K_m), whose values for matching and mismatching incorporation define the fidelity of the polymerase via the discrimination:[129,45]

$$D = \frac{(k_{cat}/K_m)_{match}}{(k_{cat}/K_m)_{mismatch}}$$

Pinpointing and quantifying these steps helps define the kinetic and structural origin of fidelity.

However, there are also important steps in the polymerisation reaction after the chemical incorporation, including: release of pyrophosphate product (PP_i); translocation to the next insertion site; and KF fingers-opening, allowing binding of the next nucleotide. The precise sequence of these post-chemistry steps is still unknown, however kinetic investigations have demonstrated that one of these post-chemistry steps is slow; even slower than the pre-chemistry slow step which defines the single-nucleotide incorporation rate ($k_{slow}^{post-chem} = 15s^{-1}$ vs. $k_{slow}^{pre-chem} = 50s^{-1}$).[43] The structural origin and function of the post-chemistry slow step is unclear, but has been hypothesised to increase fidelity by increasing the time for KF to switch to its 3'-5' exonuclease domain in order to excise any misincorporated nucleotides[51]. Furthermore, kinetic experiments suggest PP_i release occurs after the slow post-chemistry step.[43]

The slow post-chemistry step has yet to be ascribed to any conformational change in KF, despite fluorescence investigations of various structural changes on polymerisable substrates. Bulk fluorescence experiments have detected rates associated with fast DNA rearrangement and fingers-closing pre-chemistry, and fluorescence changes rate-limited by the pre-chemistry slow step.[42,68] No fluorescence changes associated with slow post-chemistry rates have been observed.

Confusion can arise in discussions on the "polymerisation-rate" of KF, dependent on whether single or processive incorporations are being measured. Single-turnover kinetic experiments measure single-nucleotide incorporations, measuring the kinetics up until the chemical step[52], providing the nucleotide incorporation rate ($k_{pol} \approx 50s^{-1}$). However, measurements of multiple nucleotide incorporations are limited by speed at which KF can open and translocate to its next insertions site after chemistry, with these post-chemistry steps being limited by the slow post-chemistry step ($k_{slow}^{post-chem} = 15s^{-1}$).[43] Thus, the nucleotide incorporation rate k_{pol} is limited by pre-chemistry steps, whereas processive nucleotide incor-

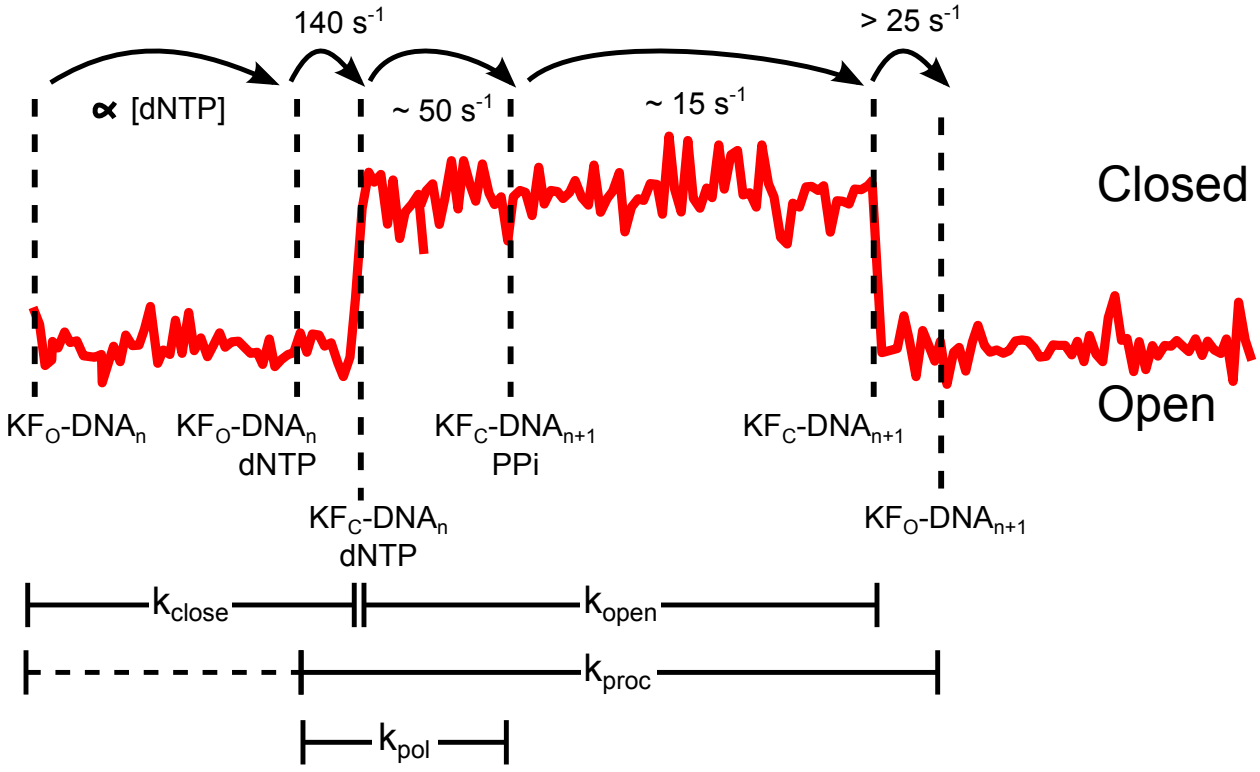


Figure 7.1: Polymerisation reaction scheme, the forward rates between the states, and their relation to dwells in the fingers-open and fingers-closed conformations, illustrated by a schematic E* trace (not actual data). Fingers-open dwells measure the dwell caused by nucleotide binding (dNTP) to the (KF_O-DNA) complex, and subsequent rapid closing forming the KF_C-DNA-dNTP complex. Fingers-closed dwells measure the rate up until nucleotide incorporation, which is dominated by the pre-chemistry slow step ($k_{slow}^{pre-chem}$) inferred from (k_{pol}) in rapid-quench experiments, as well as the post-chemistry steps up until fingers-opening, limited by the post-chemistry slow step ($k_{slow}^{post-chem}$). We demonstrated fingers-opening after chemistry contributes negligibly to the polymerisation-time in the previous chapter. At high nucleotide concentrations, the dotted regions are compressed, leading to the rate of processive nucleotide incorporation (k_{proc}) becoming approximately equal to the fingers-opening rate (k_{open}), given that pre-chemistry fingers-closing is rapid ($140s^{-1}$), and the post-chemistry finger-open dwells are short.

poration rates k_{proc} are limited by the post-chemistry slow step (See Fig. 7.1 for the relation between the reaction scheme and rates).

Measurements of processive polymerisation rates vary significantly in the literature, ranging from 15 base-pairs per second (bp/s) in single turn-over kinetic experiments[43]; 6 bp/s from single-molecule magnetic tweezers experiments[78]; down to 1bp/s in single-molecule FRET experiments[72]. Strand displacement synthesis rates also range from 1.2 b/s to 14b/s.[70,69] The origin of this heterogeneity in observed polymerisation rates is unknown.

Open-questions exist about the post-chemistry DNA polymerisation reaction in part because stochastic multi-step reactions are challenging to investigate using traditional ensemble techniques. Single-molecule techniques are particularly suited to observing the latter stages of multi-step reactions, as desynchronisation of the sample no longer becomes an issue. A notable example is a single-molecule study of KF, which suggested the existence of additional structural states during the DNA polymerisation reaction. These transient intermediates would have been challenging to observe without single-molecule techniques.[72]

Our assay allows observation of KF in the fingers-open and fingers-closed conformations, and in Fig. 7.1 we illustrate how measurement of these dwells relates to steps in the polymerase reaction pathway. Fingers-closing rates (from dwells in the open conformation) contain information on the rate of complementary nucleotide binding, and the fingers-closing rate around this nucleotide. The fingers-opening rate during polymerisation (from dwells in the fingers-closed conformation) contains information about the steps up until chemistry, which are relatively fast ($k_{slow}^{pre-chem} = 50s^{-1}$), and the slower post-chemistry steps ($k_{slow}^{post-chem} =\approx 15s^{-1}$), up until the KF opens again. In the previous chapter, we demonstrated the any steps after fingers opening are rapid, not contributing significantly to the overall polymerisation rate. The long dwells in the fingers-closed conformation during polymerisation suggests that we are observing primarily the effect of the post-chemistry slow step. Given the fingers-closed dwells measure all of the reaction steps besides fingers-closing, which are faster than subsequent steps, we can approximate the fingers-opening rate to the processive polymerisation rate, which measures the rates across the entire polymerisation reaction, not just the steps up until nucleotide addition as measured using single-turnover

kinetics experiments $k_{open} \approx k_{proc}$.

In this chapter we investigate the kinetics of single-molecule DNA polymerisation by applying the methods presented in the previous chapter, in which we observe the fingers-conformations of KF during the polymerisation reaction. Our observations suggest that fingers-opening occurs after, or is responsible for, the slow post-chemistry step. Fingers conformations during synthesis suggest that nucleotides are incorporated highly efficiently, and that heterogeneity in single-nucleotide incorporation rates could account for discrepancies in the polymerisation rates presented in the literature. The dynamics of the KF-DNA binary complex and the rates of single-nucleotide incorporation are seen to be similar, suggesting that binary dynamics could be reflecting an underlying mechanism which limits fingers-opening after nucleotide incorporation, suggesting a further avenue for investigation.

7.2 Results: Processive DNA polymerisation at high nucleotide concentrations

7.2.1 A partially extendable DNA substrate

In the previous chapter, methods for initiating and observing DNA polymerisation on a TIRF microscope were introduced, demonstrating KF to exist predominantly in the fingers-closed conformation while polymerising. Conditions allowed full extension of the 15-bp overhang, although it was unclear whether KF dissociated from the substrate before completing polymerisation. (Shortened template overhangs are known to have decreased affinity for polymerases). Inability to distinguish between full polymerisation, early KF dissociation, or even dye bleaching, made it difficult to assign a precise number of incorporated nucleotides at each DNA, and hence the significance of the temporal length of the fingers-closed dwells during polymerisation was unclear.

In order to avoid complications posed by short template overhangs during the final stages of DNA extension, we designed a DNA template which allowed partial extension of the DNA to known positions (Figure 7.2). The 15 base-pair (bp) template overhang allowed

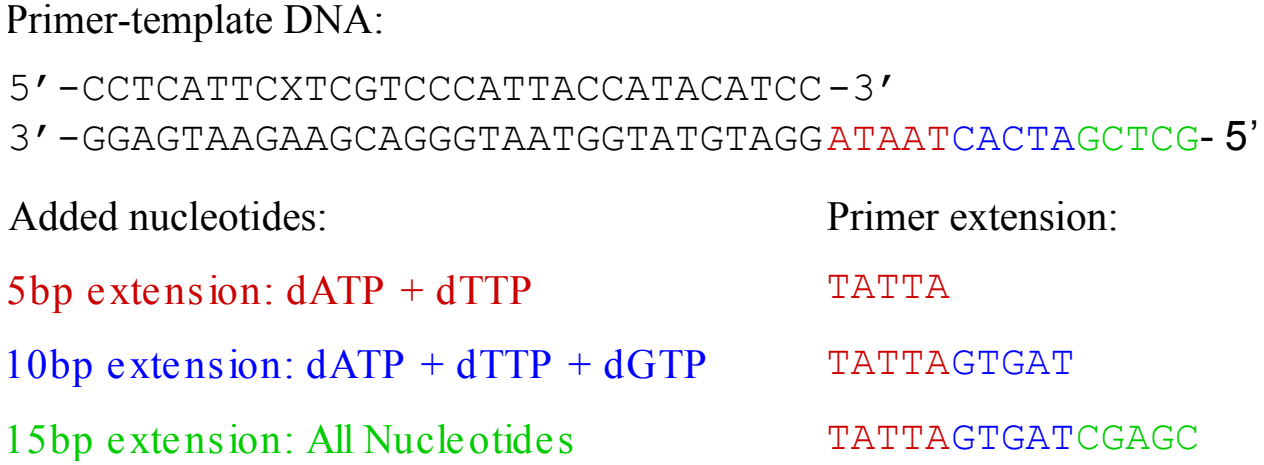


Figure 7.2: DNA primer-template designed to allow staggered DNA extension, dependent on the nucleotide subset used. Template bases 1-5 consist of As and Ts; 1-10 consist of As, Ts, Cs; 1-15 As, Ts, Cs, Gs. Thus the subsets of nucleotides shown permit the illustrated template extensions to 5, 10 and 15 base-pair positions. For instance, the first five nucleobases of the template overhang consisted of only adenine and thymine (red), with a cytosine base at the template-6 position (blue). Thus addition of dATP and dTTP permitted only a five base-pair extension. Similarly, ten base extensions were permitted when nucleotides dATP, dTTP and dGTP were used, which were complementary to the first ten bases (red and blue), but non-complementary to the eleventh (green). Addition of all nucleotides permitted full extension of the strand. 5 and 10 base-pair extensions allowed us to avoid the effects of short template overhangs causing increased dissociation.

extension by 5, 10 and 15-bp, dependent on the nucleotide subset used. For instance, use of the nucleotide dATP and dTTP permitted 5-bp extension, while use of nucleotides dATP, dTTP and dGTP together allowed 10-bp extension, and use of all nucleotides allowed full 15-bp extension.[i]

7.2.2 Partial DNA extension: Single-molecules

Using this DNA sequence, we could investigate the kinetics of DNA polymerisation in greater detail and with more control. The DNA polymerisation reaction was initiated as specified in Chapter 6, Section 6.4, with a reaction-initiating buffer solution added to a well of localised DNAs. The reaction-initiating buffer contained 200μM of each nucleotide, using the subsets: dATP, dTTP for 5bp extension; dATP, dTTP, dGTP for 10bp extension; and dATP, dTTP, dGTP, dCTP for 15bp extension, to produce a final concentration of 100μM. The buffer con-

[i]Henceforth in this chapter we refer to dATP, dTTP, dGTP and dCTP using the nomenclature A,T,G,C respectively.

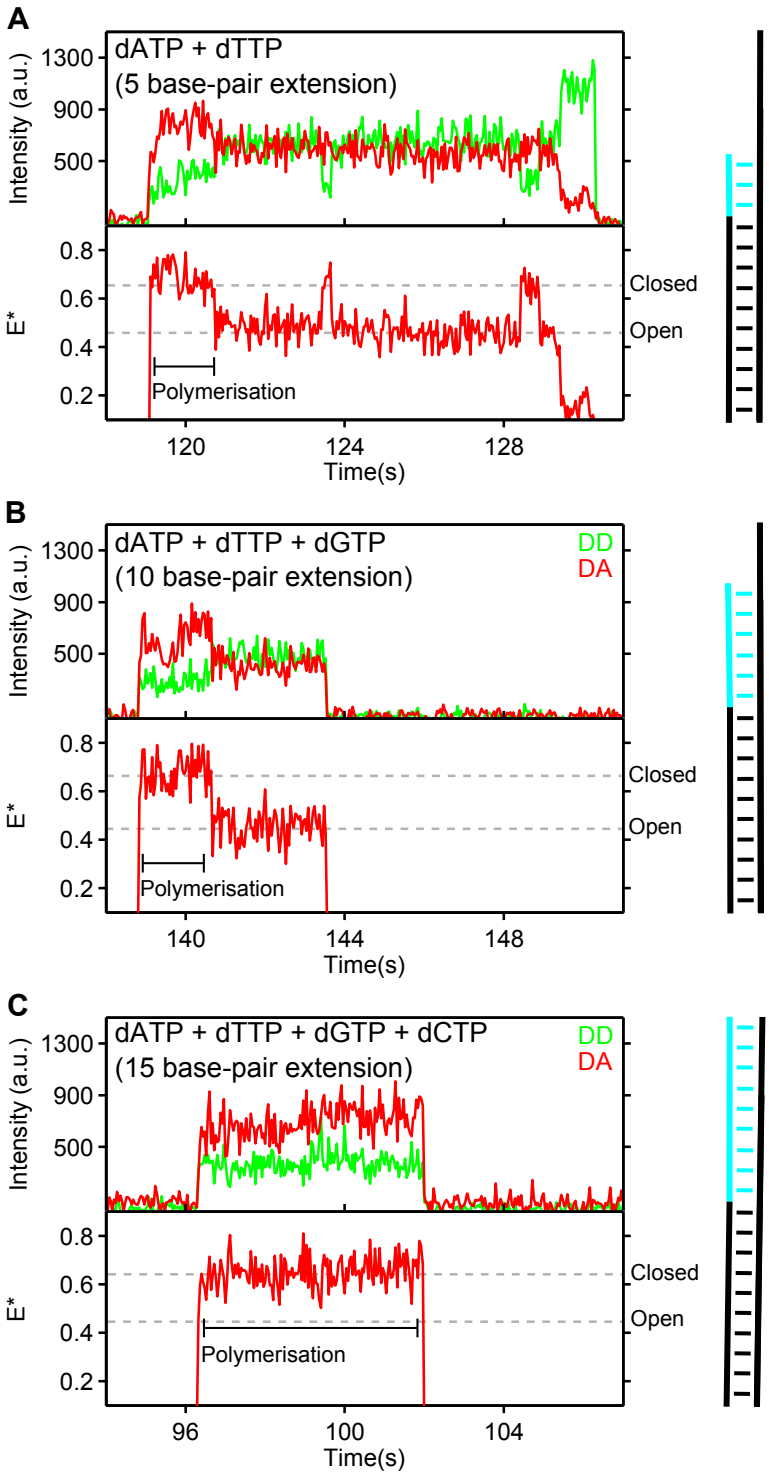


Figure 7.3: Fluorescence time-traces demonstrating KF-DNA binding and DNA polymerisation in the presence of different nucleotide subsets: (a) AT nucleotides, permitting 5-bp extension. KF-DNA binding at $\sim$ 119 seconds shows KF initially closed and the opening at $\sim$ 121 seconds, remaining open before dissociating. Binary complex dynamics seen at $\sim$ 123.5 s and $\sim$ 129 s. (b) 10-bp extension (ATG nucleotides), KF-DNA binding at $\sim$ 138 s shows KF initially closed and then opening at $\sim$ 141 seconds, remaining open before dissociating. (c) 15-bp extension (ATGC nucleotides). KF-DNA binding at $\sim$ 96 s shows KF initially closed and dissociating at $\sim$ 102 seconds. Donor-excitation donor-emission (DD - green); Donor-excitation, acceptor-emission (DA - red); FRET efficiency (E* - red)

tained a low concentration of KF (≈1nM, final conc. 0.5nM) to delay onset of polymerisation via stochastic KF-DNA binding. The high nucleotide concentration began polymerisation immediately upon KF-DNA binding, removing any fingers-open dwells caused by awaiting binding of nucleotide.

Fluorescence traces at DNAs demonstrated that upon KF-DNA binding, KF entered the fingers-closed conformation more rapidly than could be observed using our time-resolution (40ms), and remained closed for a period of time, irrespective of which nucleotide subset was used, which we interpreted as DNA polymerisation (Fig. 7.3, fluorescence time-traces at DNAs demonstrate the initial KF binding event to be in the fingers-closed conformation, irrespective of the nucleotide subset used).

Interestingly, KF showed a distinct behaviour when partial (subsets AT or ATG) or full DNA extension (ATCG nucleotide set) was permitted. During partial extension (AT or ATG subset), KF remained in the closed state for a period of time, then *opened*, and remained in the fingers-open conformation and occasionally undergoing binary complex dynamics. (Fig. 7.3a, during 5bp extension, the first KF-DNA binding event at $\sim$ 119 seconds was initially closed, and opened at $\sim$ 121 seconds. Fig. 7.3b, during 10bp extension the KF binding event at $\sim$ 138 seconds was initially closed and opened at $\sim$ 141 seconds)

During 15-bp extensions, however, KF remained closed for a period of time, and then dissociated, without clearly returning to the fingers-open state. (Fig. 7.3c, first KF binding event at $\sim$ 96 seconds remains in the closed state until dissociating at $\sim$ 102 seconds.)

This is consistent with expectations: 5 and 10-bp extensions produce a primer-template with a sufficient template overhang to remain a polymerisable substrate, and thus KF remains bound to the DNA after polymerisation until dissociating, after which other KF molecules can continue to bind and dissociate the DNA. Conversely, the presence of all nucleotides allows KF to polymerise the DNA until it creates a polymerised substrate to which it can no longer bind efficiently, leading to KF dissociation.

Given KF is known to remain closed whilst polymerising, the fingers-opening point (for 5 and 10-bp extensions) can be assigned as the point at which polymerisation is completed. (This signature is distinct to KF dissociation or dye bleaching.) Thus our designed DNA

allows investigations of the kinetics of processive incorporation, by producing a well defined signal change upon completion of polymerisation at 5 and 10-bp. Henceforth, we refer to the nucleotide subset and its expected extension length interchangeably (AT - 5 base-pair; ATG - 10 base-pair; ATCG - 15 base-pair).

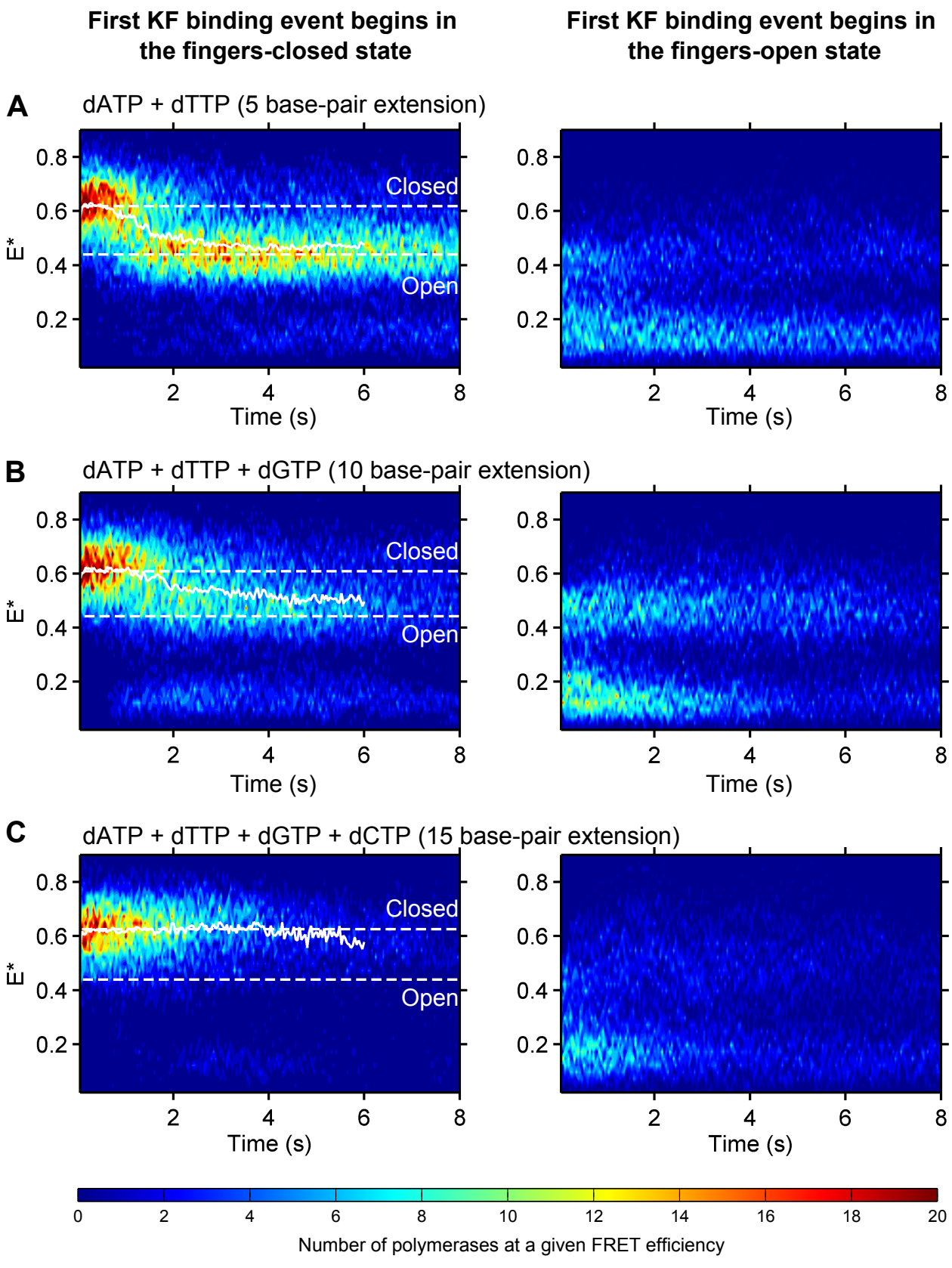
7.2.3 Partial DNA extension: Single-molecule ensemble

To supplement the description provided by the single-molecule traces above, a visualisation technique was used to convey the polymerisation behaviour across the entire DNA ensemble.

Each DNA was associated with a FRET trace describing KF conformation during polymerisation (e.g. FRET efficiency traces in Fig 7.3). Post-synchronising these FRET traces allowed observation as if all polymerisation reactions had begun simultaneously. The collection of FRET traces across all DNAs produced a distribution of FRET efficiencies at each time point, which could be represented in a histogram. Stacking these histograms side-by-side, in a colourmap, demonstrated the time-evolution of the FRET efficiency distribution across the reaction, and hence the KF conformational distribution over time (Fig. 7.4). (Post-synchronisation was achieved by concatenating binding events at each DNA, with the collection of FRET traces across DNAs split into two groups, according to whether first KF binding event begun in the fingers-open or fingers-closed conformation (whether any of the first 4 frames had E*$>$ 0.55, i.e. whether it begun polymerising or not)

These visualisations demonstrated that the ensemble of KF-DNA complexes behave in accordance to the single-molecule traces shown above:

For the ensemble of DNAs undergoing 5-bp extension, KF molecules which begin in the fingers-closed conformation remain closed for $\sim 1s$, and by $\sim 1-2s$ begin returning to the open conformation as they complete polymerisation, and by $\sim 2s$ most KF molecules have returned to the fingers-open conformation (Fig. 7.4a, left hand panel, FRET values distributed around E*=0.65 from $\sim 0-1s$, followed by increasing occupancy around E*=0.45 from $\sim 1s$ onwards). The continued signal after $\sim 1s$ demonstrates that KF continues to bind the DNA. The small population of KF molecules not in the fingers-closed conformation showed no change in behaviour over time (Fig. 7.4a, right hand panel).


First KF binding event begins in the fingers-closed state
First KF binding event begins in the fingers-open state
A
dATP + dTTP (5 base-pair extension)
B
dATP + dTTP + dGTP (10 base-pair extension)
C
dATP + dTTP + dGTP + dCTP (15 base-pair extension)
Closed
Open
E*
Time (s)
Number of polymerases at a given FRET efficiency

For the ensemble of DNAs undergoing 10-bp extensions, a similar behaviour is observed: KF molecules begin in the closed conformation, and return to the fingers-open conformation once polymerisation is completed (Fig. 7.4b, left hand panel, FRET values distributed around E*=0.65 from $\sim 0-1s$, followed by increasing occupancy around E*=0.45 from $\sim 1s$ onwards). Similar to the 5-bp extension, there is a relatively low population of KF molecules which do not begin in the closed conformation, which do not demonstrate DNA polymerisation behaviour (Fig. 7.4b, right hand panel).

It is difficult to extract the polymerisation rates using this representation: the median of E* values show a gradual return to lower E* values, but do not return to the same base-lines, suggesting that the binary complex dynamics are shifted between 5 and 10-bp extensions (Fig. 7.4, solid white lines).

The ensemble of DNAs undergoing 15-bp extension showed a distinct behaviour: The majority began in the fingers-closed conformation, and remained there until dissociating, with no return to the fingers-open conformation. (Fig. 7.4c, left hand panel, FRET values at E*~ 0.65 from $\sim 0-4s$, with a gradual decrease in occupancy from $\sim 2s$ onwards)

7.2.4 Rates of polymerisation by KF

The section above demonstrated a clear distinction between the "full" and partial DNA extension, in which KF continues to bind partially extended DNA, whilst KF dissociated from "fully" extended DNA. While these differences were distinct, it was difficult to specify the time taken to polymerise five, ten and fifteen base-pairs accurately. Naively, it might

---

Figure 7.4 *(preceding page)*: KF polymerising an ensemble of DNA molecules, visualised in a colourmap of the FRET efficiency distributions of post-synchronised FRET time-traces at each DNA. Molecules are split into those that begin in the fingers-closed conformation, and those that don't, representing the entire data set. Contours are used to aid visualisation, in reality the data is quantised in 40ms steps. (a) KF extending DNA by 5 base-pairs with KF in the closed conformation initially, returning to the open conformation by $\sim$ 2s. (b) KF extending DNA by 10 base-pairs with KF in the closed conformation initially, returning to the open conformation by $\sim 3-4$s. (c) KF extending DNA by 15 base-pairs with KF in the closed conformation initially, not returning to the open conformation, instead dissociating by $\sim$ 4s. Results are collated from KF binding events to 279, 306 and 236 DNA molecules for a, b, and c respectively.

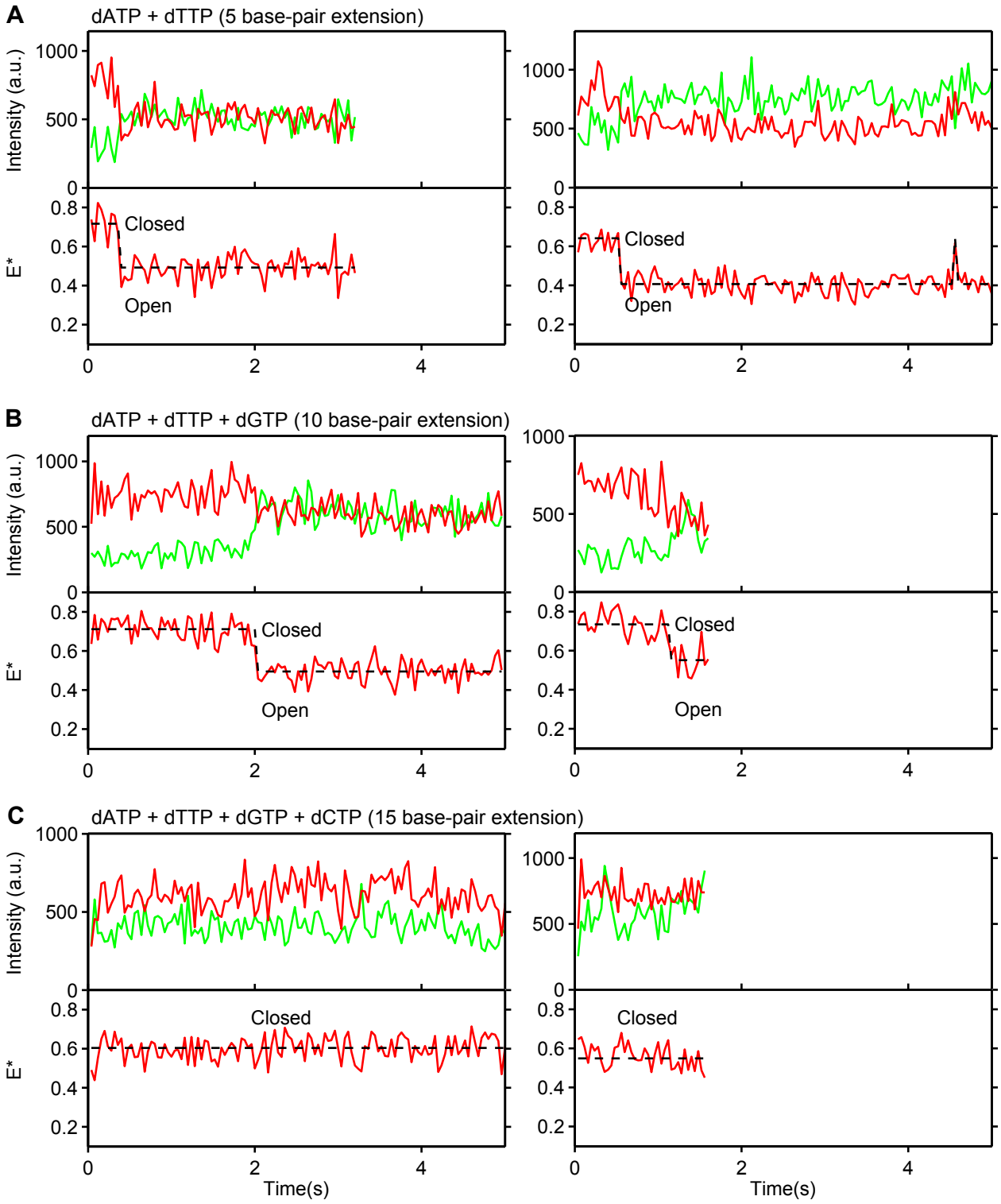


Figure 7.5: Sample HMM fits to KF-DNA binding events during (a) 5-bp extension (b) 10-bp extension and (c) 15-bp extension, extracting the first dwell in the closed conformation, corresponding to the polymerisation-time.

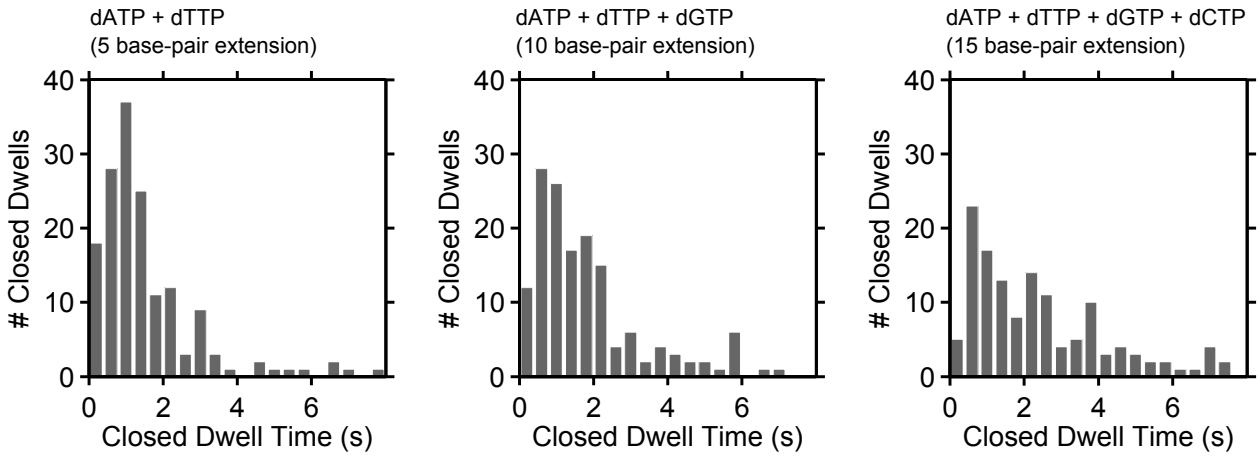


Figure 7.6: Distribution of times taken to polymerise DNAs by 5, 10 and 15 base-pairs, dependent on using nucleotide subsets AT, ATG, and ATGC respectively. Polymerisation-times are the lengths of the first dwell in the KF fingers-closed conformation. Distributions are seen to broaden on increasing incorporation lengths. Results are collated from KF binding events to 279, 306 and 236 DNA molecules for 5, 10, and 15 base-pair extensions respectively.

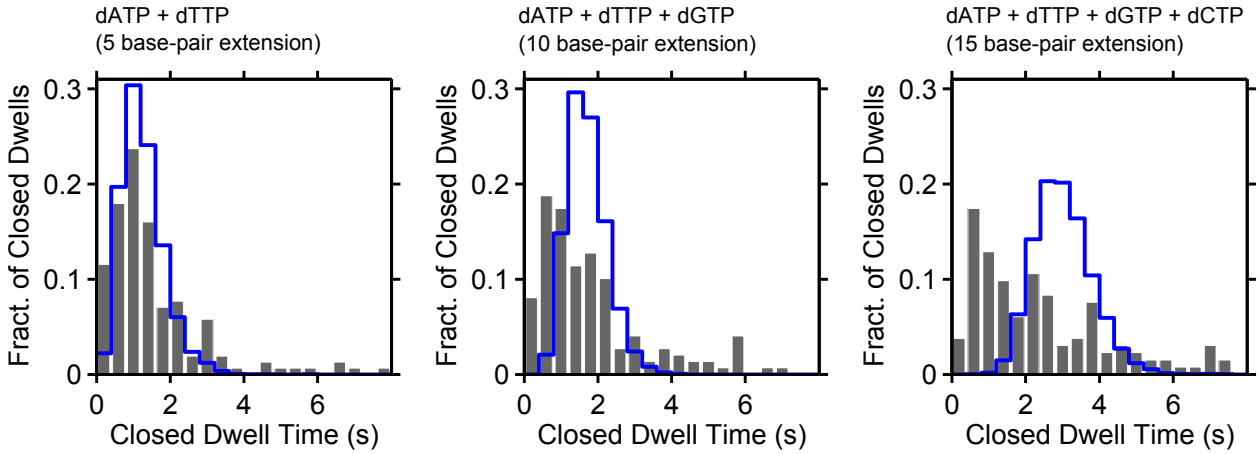


Figure 7.7: Distribution of times taken to polymerise DNAs by 5, 10 and 15 base-pairs as in Fig. 7.6. The histograms are fit to models in which 5, 10 and 15 nucleotides are incorporated at equal rates, with the best fits plotted, extracting incorporation rates of 4, 5.8 and $5s^{-1}$ respectively. While the incorporation fits are reasonably consistent, the form of the fits for 10 and 15 base incorporations are significantly different from the measured distributions, suggesting a model in which all nucleotides are incorporated equally is not the best model.

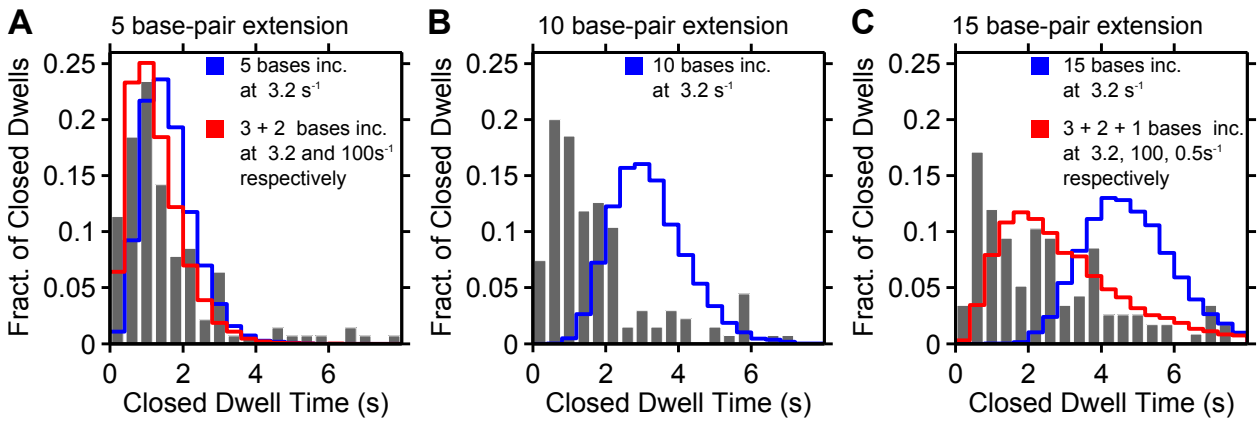


Figure 7.8: Similarly as Fig. 7.6, except that simulated models are superimposed. In blue, a model for polymerisation-times in which (a) 5, (b) 10, and (c) 15 nucleotides are incorporated at a rate $3.2s^{-1}$. This shows reasonable fit to a, but not to b or c, demonstrating the measured polymerisation-times do not conform to a model in which all nucleotides are incorporated at the same rate. In (a), the red line represents a simulation of polymerisation-times in which 3 nucleotides are incorporated at a rate $2.4s^{-1}$ and 2 at a rate $100s^{-1}$, showing a better fit to than the blue model. In (c), the red line represents the same simulation of polymerisation-times except with an additional nucleotide incorporated at a slow rate $0.5s^{-1}$.

be expected that time taken to polymerise five, ten and fifteen base-pair extensions (the "polymerisation time") be one, two and three multiples of the five base-pair polymerisation-time.

In order to investigate the time taken to polymerise 5, 10 and 15 base-pairs, we applied Hidden Markov Modelling as described in Chapter 5 to the first 10s of binding events to each DNA, to extract the first fingers-closed dwell upon KF-DNA binding (see examples fit in Fig. 7.5). This fingers-closed dwell was interpreted as the time taken to polymerise the DNA ("polymerisation time"), given it was shown previously that KF enters the fingers-closed conformation immediately on binding DNA, and remains closed until completing polymerisation, at which point it returns to the open conformation (or dissociates). The "polymerisation time" at each DNA was then plotted in a histogram to quantify the statistical distribution polymerisation-times, under 5, 10 and 15-bp extension (Fig. 7.6).

The histogram of "polymerisation times" for 5, 10 and 15-bp extensions demonstrate distributions with peaks between 1-2 seconds, and tails extending to longer times, which become more pronounced at longer extensions (Fig 7.6). Given the polymerisation-time at each DNA results from the sum of multiple nucleotide incorporations, the resulting polymerisation-time

statistical distribution should follow the combined statistics of these multiple incorporations. The central limit theorem states that the sum of independent, identically distributed random variables will approach a normal distribution as the number of variables increases. This suggests that the "polymerisation times" distributions should approach normal distributions with increasing numbers of incorporated nucleotides, if all nucleotides are incorporated equally.

For the 5-bp extension, the polymerisation-time distribution loosely resembles a normal distribution, however on increasing nucleotide incorporations, the distributions diverge from normal distributions (Fig. 7.6, observe the broadening distribution on increasing extension, moving from the left to right hand panels.).

The polymerisation-time for 5 nucleotides incorporations fit reasonably well to a model in which the 5 nucleotides are incorporated at a rate of $3.2s^{-1}$, based on the incorporation rate of a single nucleotide measured in the next section (Fig 7.8a, compare simulation (blue line) with measurements (grey bars)). However, extending this model to increasing nucleotide incorporations (10 and 15-bp extension at a rate of $3.2s^{-1}$ per nucleotide) provided distributions which diverged from the measurements, suggesting there must be faster incorporations to account for the shorter polymerisation-times measured than those predicted with a uniform incorporation model (Fig 7.8b,c, compare simulation (blue line) with measurements (grey bars)). The divergence of the measured polymerisation-times from normal distributions on increasing extensions, and the simulated model, suggests that the hypothesis of independent, identically distributed variables for nucleotide incorporations, is incorrect.

There are several possibilities which could account for these results. Firstly, each of the four (dATP, dGTP, dTTP, dCTP) nucleotides may have an independent incorporation rate (i.e. are non-identically distributed), or the incorporation rate might be sequence-dependent, such as pausing[69], or switching to the exonuclease mode. Thirdly, the distributions may be accounted for by an artefact in the analysis, for instance caused by KF dissociation or dye photo-bleaching before completion of polymerisation, which might skew the statistical distribution.

In order to control for KF dissociation or photobleaching mid-polymerisation event, these

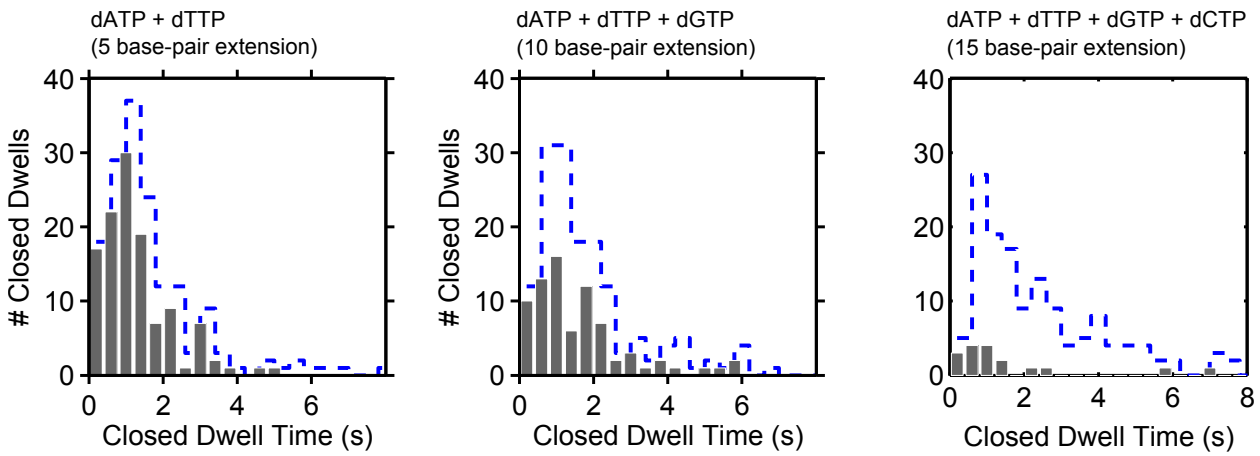


Figure 7.9: Distribution of times taken to polymerise DNAs by 5, 10 and 15 base-pairs, measuring the first fingers-closed dwell (dotted blue line), or only fingers-closed dwells which return to the fingers-open conformation (grey bars), in order to control for KF dissociation or dye-photobleaching. Distributions are comparable apart from at 15 incorporations, as KF does not return to the open conformation after polymerisation.

events were eliminated by only counting molecules which return to the fingers-open state after polymerisation (see Fig. 7.9). The resultant polymerisation-time distributions for 5 and 10-bp extensions are similar to those seen without this control (Compare dotted blue lines and grey bars in Fig. 7.9). The similarity of these histograms suggests that truncation due to photobleaching or early dissociation does not contribute significantly to altering the polymerisation-time distributions. The histogram of polymerisation-times for the 15-bp extension loses most of its signal, consistent with KF molecules dissociating after completing polymerisation and not returning to the open state after full DNA extension.

7.3 Results: Single-nucleotide incorporations

The non-trivial statistical distribution of polymerisation-times (Fig. 7.6) suggested that the kinetics of nucleotide incorporations are more complex than equal incorporation rates for all nucleotides. It was desirable to explain the polymerisation-time distributions by quantifying the kinetics of their constituent single nucleotide incorporations. Measurement of single-nucleotide incorporations could also illuminate the polymerisation mechanisms, and whether there might be complicating factors such as multiple rate limiting steps, or pausing events.

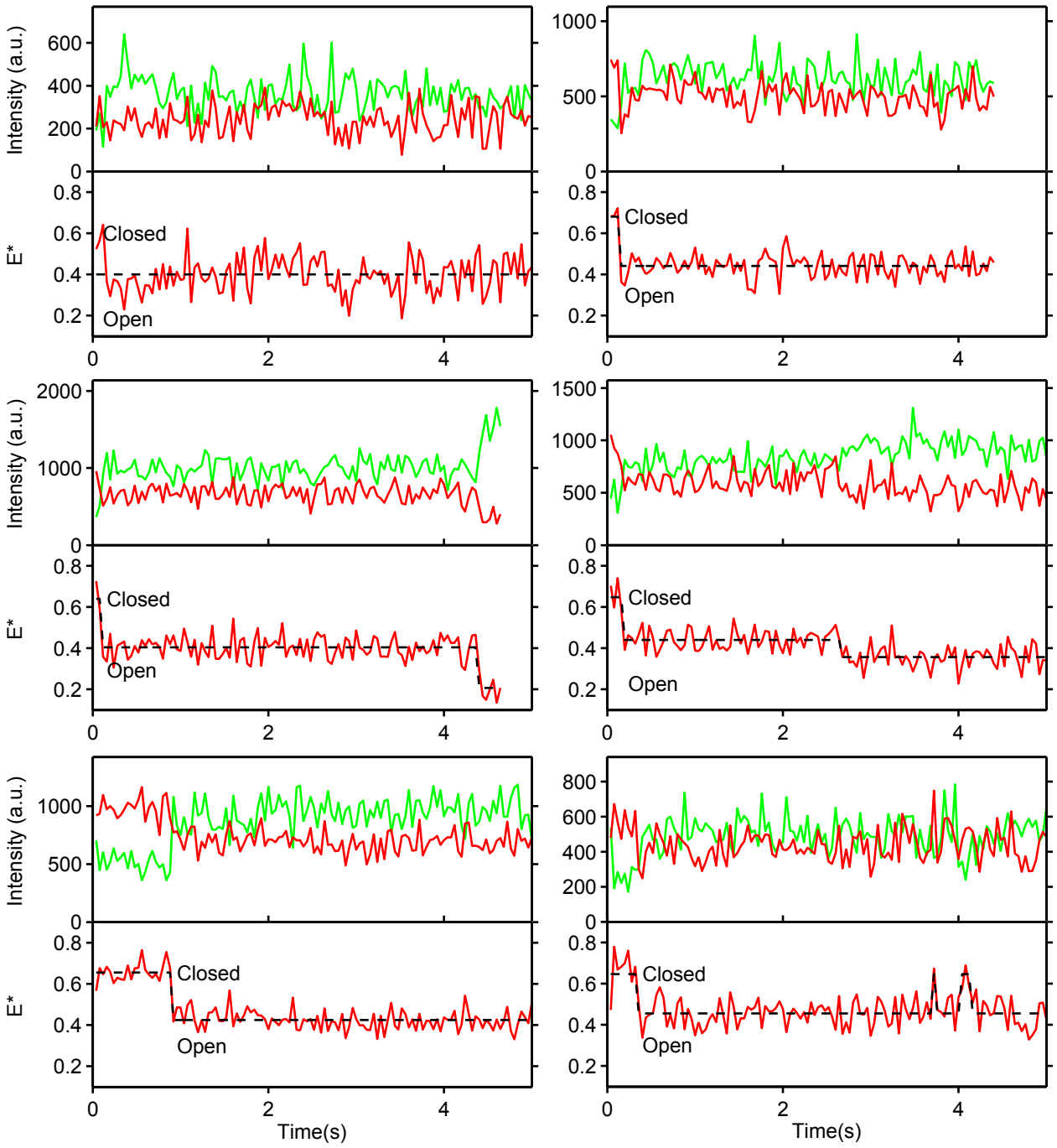


Figure 7.10: Sample HMM fits to KF-DNA binding events for single base-pair extensions, extracting the first dwell in the closed conformation, corresponding with the k_{proc} nucleotide incorporation rate. Note the upper left hand panel has a closed state dwell not fit by the HMM, illustrating the importance of using two techniques to extract the incorporation rate.

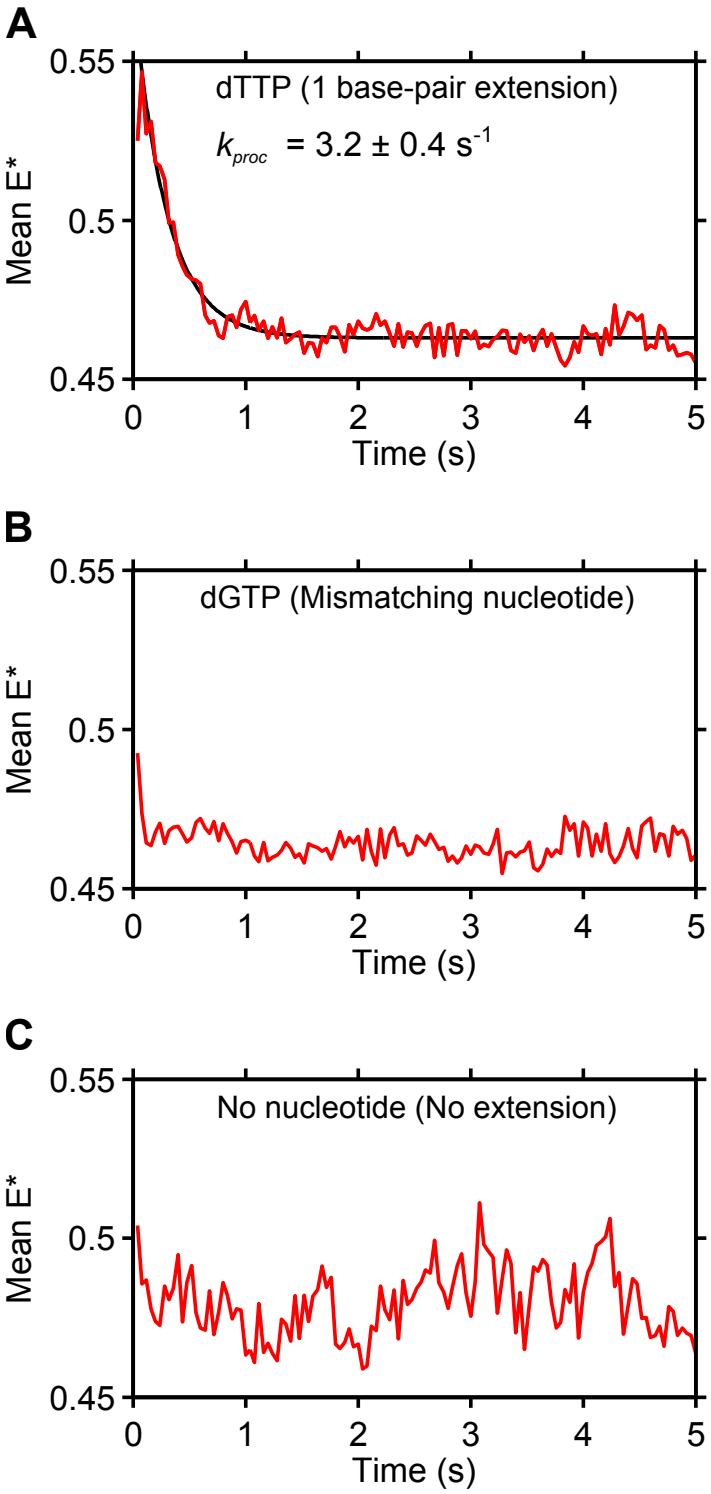


Figure 7.11: Mean E* of post-synchronised KF-DNA binding events across a collection of DNAs in the presence of (a) complementary nucleotide (dTTP) demonstrating an exponential decrease in mean E*, (b) mismatching nucleotide (dGTP), (c) no nucleotides. Results are collated from KF binding events to 453, 568 and 211 DNA molecules for a, b, and c respectively.

Single-nucleotide incorporations occur more rapidly than the multiple incorporations, and therefore are more challenging to measure. Measurements of single-nucleotide incorporations were performed similarly to the previous section, with polymerisation initiated by flowing a reaction-initiating buffer containing an excess of the first complementary nucleotide (500μM dTTP to a final conc. of 250μM dTTP), and a low concentration of KF (1nM to a final conc. of 0.5nM) to limit KF-DNA binding. As controls, polymerisation experiments were also initiated in the absence of nucleotides, and also using an excess of non-complementary nucleotide (500μM dGTP to a final conc. of 250μM dGTP). KF-DNA binding events were then post-synchronised and nucleotide incorporation rates quantified using two independent analysis techniques. (Concatenated KF-DNA binding events are seen in Fig. 7.10)

The bulk behaviour of the post-synchronised polymerisation reaction was first observed using the mean FRET across all KF-DNA complexes over time, which reflected the mean KF conformation across the sample during polymerisation. The collection of KF molecules initially demonstrated high FRET on reaction initiation, accounted for by entry to the fingers-closed conformation and polymerisation (Fig. 7.11a, mean E*=0.55 at t=0). The mean E* decreased as KF molecules returned to the fingers-open conformation after polymerisation. (Fig. 7.11a, mean E*=0.46 by t=1). Fitting a single offset exponential (E*=$A\exp(-k_{open}t)+B$) extracted an KF opening rate of $k_{open} = 3.2 \pm 0.4s^{-1}$ for dTTP, which we ascribe as approximately equal to the rate of "processive" nucleotide incorporation k_{proc} as discussed in the introduction. In contrast, experiments which used mismatching, or no nucleotides, did not show such behaviour (Fig. 7.11b,c mean E* remains at 0.46-0.47).

At the single-molecule level, KF molecules that demonstrated dwells in the fingers-closed conformation upon initial binding to DNA were interpreted as incorporating a single-nucleotide. These fingers-closed dwells were extracted using HMM as described in the previous section (Fig. 7.10), and plotted in histograms (Fig. 7.12). Significantly more dwells in the fingers-closed conformation were seen when using complementary nucleotide than when using mismatching or no nucleotides (Fig. 7.12, compare number of events in left hand panel to middle and right panels.). Fitting the fingers-closed dwells in the presence of dTTP to a single-exponential extracted a incorporation rate of $k_{proc} = 3.4 \pm 0.8s^{-1}$, consistent with

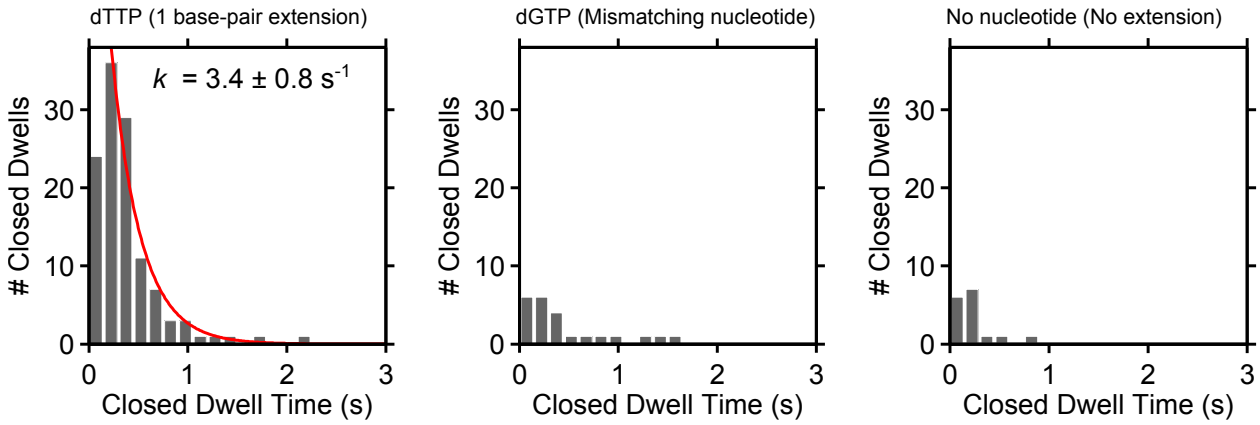


Figure 7.12: Lengths of the first fingers-closed dwell at KF-DNA binding events in the presence of nucleotide complementary to the first template base (left), mismatching nucleotide (middle) and no nucleotide (right). Polymerisation in the presence of dTTP is evidenced by the increased number of events compared with no nucleotides, or mismatching nucleotide. Fit to a single-exponential (left) extracts the rate of dTTP incorporation at $k_{proc} = 3.4 \pm 0.8 s^{-1}$ Results are collated from KF binding events to 453, 568 and 211 DNA molecules for left, middle, and right panels respectively.

the rate extracted in the above paragraph. The single-exponential statistics suggested a single-rate limiting step, suggesting the absence of complicating events such as pausing, or movement to the exonuclease domain, around this specific sequence.

A model of 5, 10 and 15 nucleotide incorporations at the measured rate ($k_{proc} = 3.2 s^{-1}$) produced "polymerisation-time" distributions which fit reasonably well for 5-bp extension measurements, but deviated from the measurements of 10 and 15-bp extensions. (Fig 7.8a, compare simulated distribution (blue trace) with measured distribution (grey)) The large overestimate of polymerisation times for 10-bp extension, and the lack of significant difference between the 5 and 10-bp extension distributions, suggested that the nucleotides between bases 5 and 10 were incorporated significantly more rapidly than $3.2 s^{-1}$.

It is interesting to note that the 5bp polymerisation-time distribution fits better to a model in which there are 3 incorporations at $3.2 s^{-1}$, and 2 rapid incorporations at $100 s^{-1}$ - thus demonstrating that the distributions are consistent with a combination of slow and rapid incorporation events. (Fig 7.8a, red trace).

The measured rate of a single dTTP ($k_{proc} = 3.2 s^{-1}$) incorporation was much slower than previous measurements of the pre-chemistry rate ($k_{pol} = 30 - 50 s^{-1}$), suggesting that nucleotide incorporation is rate limited by steps after the chemical step.[42,53,68,43]

7.4 Results: Processive DNA polymerisation at low nucleotide concentrations

In section 7.2, long dwells in the finger-closed conformation were observed during polymerisation at high nucleotide concentrations. In order to investigate the individual nucleotide incorporations which composed these long dwells, we initiated polymerisation at low nucleotide concentrations. The aim was to cause KF to pause in the fingers-open conformation as it awaited the next complementary nucleotide, splitting the long dwells into their constituent incorporations.

Initial attempts to observe the fingers-closing events associated with polymerisation at low nucleotide concentrations were hampered by the presence of KF-DNA binary complex dynamics, in which KF stochastically enters the fingers-closed conformation in the absence of bound nucleotide. Given binary complex kinetics were very similar to nucleotide incorporation kinetics, it made it difficult to distinguish the two.

We inhibited binary complex dynamics by including mismatching nucleotide in the reaction-initiating buffer, known to reduce the KF fingers-closed occupancy, as KF disfavours the catalytically active state to avoid misincorporation.[19,20]. For instance, initiating DNA polymerisation using the nucleotide subset dATP and dTTP allowed extension of the first 5 nucleotides, but by also including an excess of dCTP to the buffer (which was mismatching up to and including the 6th templating base) dampened any binary complex dynamics, allowing distinction between binary dynamics and nucleotide incorporations. As a result of this strategy, only 5 and 15-bp extensions were possible to investigate, as there was no nucleotide subset which would allow extension to the 10 base-pair position and also include a mismatching base to inhibit binary complex dynamics.

To initiate processive polymerisation at low nucleotide concentrations, the reaction-initiating buffer contained 0.5μM dATP, dTTP and 500μM dCTP for 5bp extension, producing final concentrations of 0.25μM dATP, dTTP and 250μM dCTP. For 15bp extension final concentrations of 0.25μM ATCG were used. The first 5s of binding events to each DNA were extracted and fitted to a HMM, with all dwells counted and measured.

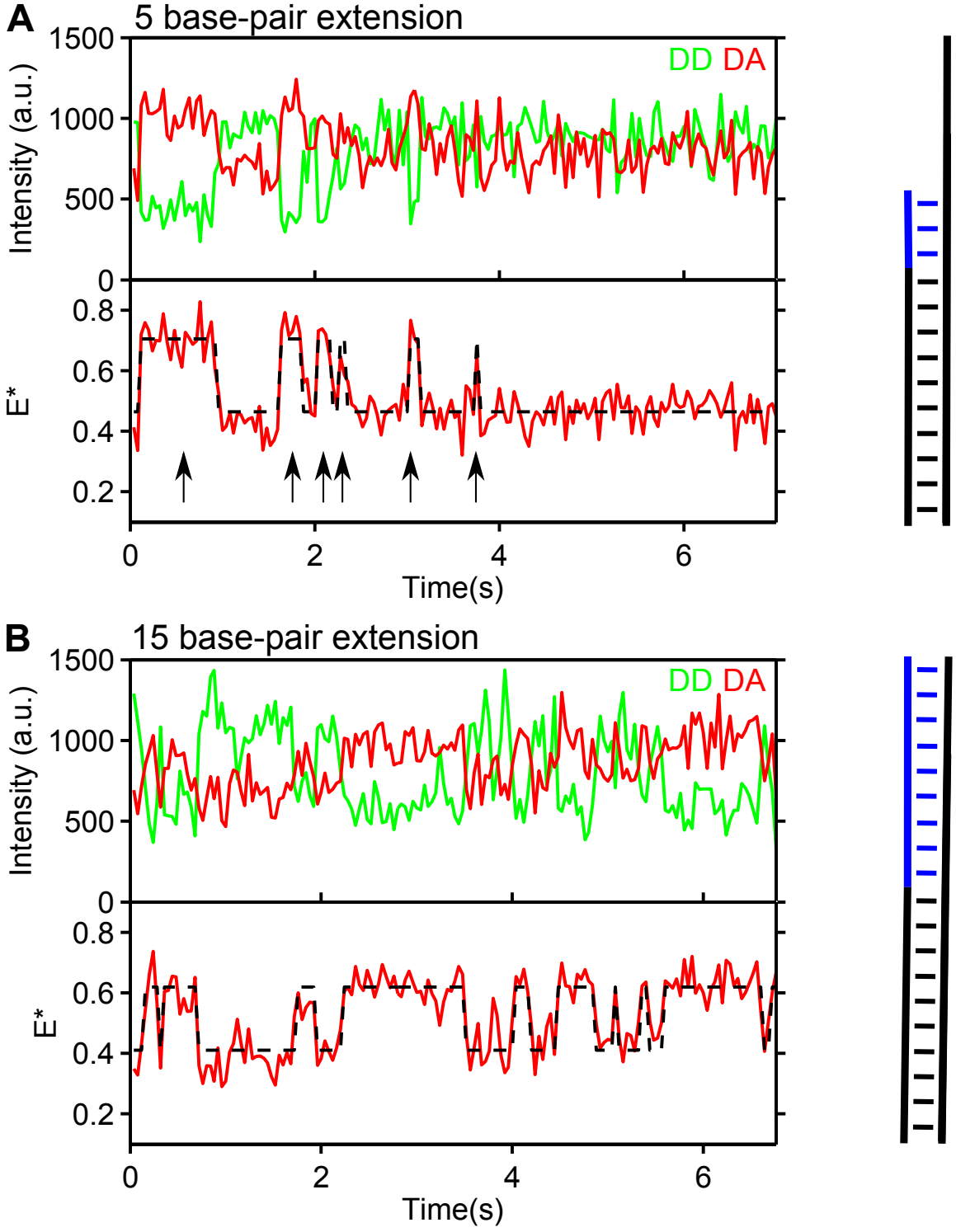


Figure 7.13: DNA polymerisation, at low nucleotide concentrations. (a) Fluorescence trace during 5bp extension, with 6 fingers-closing events seen at the beginning of the E* trace, followed by a long dwell in the fingers-open conformation. Arrows represent potential incorporation events or binary complex dynamics. (b) Fluorescence trace during 15bp extension, with 11 fingers-closing events seen in the E* trace. Donor-excitation, donor-emission (DD - green). Donor-excitation, acceptor-emission (DA - red). FRET efficiency (E* - red). Black traces show HMM fit.

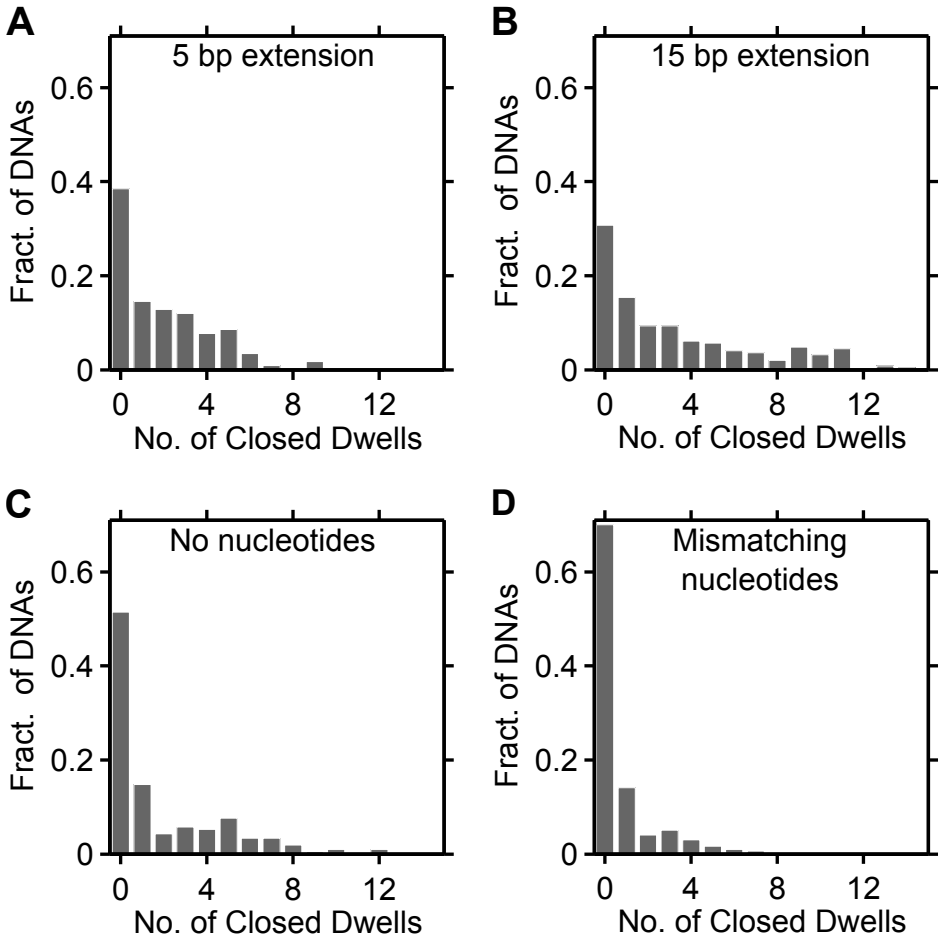


Figure 7.14: Number of fingers-closing events during polymerisation at low nucleotide concentrations, during (a) 5bp extension (AT nucleotides) (b) 15bp extension (ATCG nucleotides). (c) no polymerisation (no nucleotides) (d) no polymerisation in the presence of mismatching nucleotide (250 μM dGTP). Extracted from the first 10 seconds of KF binding at each DNA, across all DNAs. Results are collated from KF binding events to 117, 247 and 211 and 297 DNA molecules for a, b, c and d respectively.

Initiating 5-bp polymerisation resulted in KF-DNA binding events which demonstrated a series of interconversions between fingers-open and fingers-closed conformations, grouped around the start of the KF-DNA binding event (Fig. 7.13a, six open to closed interconversions seen between 0-4 seconds, followed by a long dwell in the open conformation). We interpreted this as the fingers-dynamics as KF extended DNA by 5 base-pairs.

The number of excursions to the fingers-closed conformation during polymerisation was counted using HMM. We compared the number of excursions to the expected number of nucleotide incorporations, in order to investigate the number of fingers-closing events associated with each nucleotide incorporation (Figure 7.14, number of fingers-closing events under different conditions).

During 5-bp extensions, 95% of the traces showed less than or equal to 5 fingers-closing events, suggesting that maximally one fingers-closing event was associated each nucleotide incorporation, allowing us to tentatively assign approximately one fingers-closing event to each nucleotide incorporation (Figure 7.14a, maximum number of fingers-closing events $\sim 5-6$). The distribution also suggested that there were many traces in which the number of fingers-closing events were not being counted fully, perhaps being missed due to their speed, for instance see the anti-correlated FRET changes in Fig. 7.13b which show very brief excursions to the fingers-closed conformation at $\approx 1s$ and $\approx 3.9s$, which are not picked up by the HMM due to their speed.

During 15-bp extensions, $> 99\%$ of the traces showed less than 12 fingers-closing events, consistent with KF dissociating before polymerising to the end of the shortening primer-template. The distribution of lower numbers of fingers-closing events, suggested that some fingers-closing events were being missed.

The missed events could have been caused by rapid nucleotide incorporations, or by dissociation before polymerisation completion. The increase in the number of fingers-closing events counted between 5 and 15 base incorporations suggested that dissociation is not playing a major role for 5-bp extensions.

The number of fingers-closing events were also counted in the absence of nucleotides, and in the presence of mismatching nucleotides (250 μM dGTP). These conditions were distinct to polymerisation, with a larger fraction of DNA traces demonstrating no fingers-closing events (Fig. 7.14d, 70% of DNAs show no fingers-closing events during the first ten seconds of KF-DNA binding in the presence of mismatching nucleotides, relative to $< 40\%$ in the presence of nucleotides, seen in a and b).

The dwells in the fingers-closed conformation during polymerisation were quantified. Fingers-closed dwells during 5-bp extension fit well to a single-exponential, extracting a nucleotide incorporation rate of $k_{proc} = 2.7 \pm 0.3 s^{-1}$, consistent with the single-nucleotide incorporation rate measured above (Figure 7.15a, fingers-closed dwells during 5-bp extension).

The fingers-closed dwells during 15-bp extension, however, fit more appropriately to

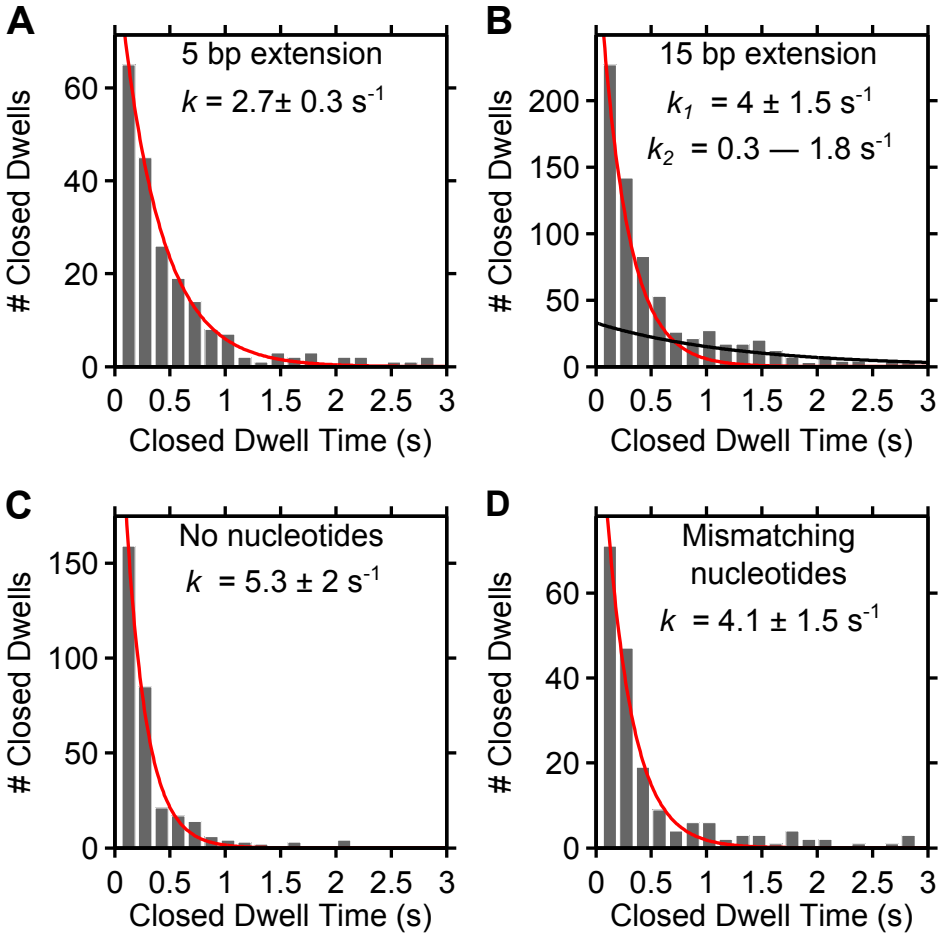


Figure 7.15: Fingers-closed dwells during polymerisation at low nucleotide concentrations, during (a) 5bp extension (AT nucleotides) (b) 15bp extension (ATCG nucleotides). (c) no polymerisation (no nucleotides) (d) no polymerisation in the presence of mismatching nucleotide (dGTP). Finger-closed dwells are extracted from the first 10 seconds of KF-DNA binding, across all DNAs. Fits to single or biexponentials extract the rates of fingers-opening in under each condition. Results are collated from KF binding events to 117, 247 and 211 and 297 DNA molecules for a, b, c and d respectively.

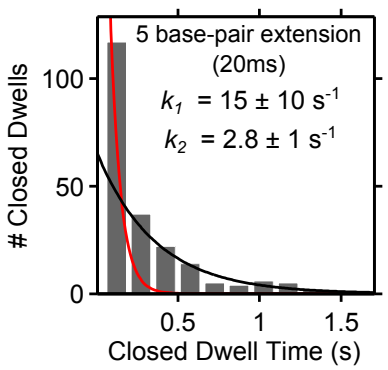


Figure 7.16: Dwells in the fingers-closed conformation during the first 10 seconds of KF-DNA binding, across all DNAs during 5bp extension at 20ms time-resolution. Fits biexponential extract the rates of fingers-opening. Results are collated from KF binding events to 117 DNAs.

a bi-exponential, extracting a nucleotide incorporation rates of $k_{proc} = 4 \pm 1.5s^{-1}$, and $k_{proc} = 0.3 - 1.8s^{-1}$ (Figure 7.15b, fingers-closed dwells during 15-bp extension). These longer dwells could account for the broadly distributed polymerisation-times seen during 15-bp extensions at high nucleotide concentrations in section 7.2.4, and could be caused by slow nucleotide incorporations, or incorporation kinetics which are altered by the shortened template overhang during the final base incorporations. Opening rates from the binary complex fingers-closed conformation are slightly faster ($5.3 \pm 2s^{-1}$), although within error of the rates of nucleotide incorporation ($2.7 \pm 0.3s^{-1}$) (Figure 7.15c, fingers-closed dwells in the absence of nucleotides).

Preliminary investigations at higher time resolution (20ms) demonstrated a similar distribution of number of fingers-closing events during 5bp extensions, but the histogram of fingers-closed dwell times no longer fit well to a single exponential, and suggested that there is a faster process occurring (Figure 7.16, the first bin is higher than expected, whilst the slower incorporation rate is also extracted). Perhaps this represents the faster nucleotide incorporations hypothesised to cause the lack of significant difference in polymerisation times seen between 5 and 10-bp extensions in Fig 7.6, although further controls are needed in order to reliably interpret the 20ms investigations.

7.5 Discussion

Nucleotide incorporation rates may be heterogeneous

In the previous chapter, we demonstrated that any excursions to the open state during DNA polymerisation are rapid ($< 40ms$), and contribute negligibly to the polymerisation rate, with KF effectively appearing closed during processive DNA polymerisation at high nucleotide concentrations. This suggested that any steps after post-chemistry fingers-opening must be relatively rapid.

We thus used the fingers-closed dwells during nucleotide incorporation to infer the time taken to polymerise 5, 10 and 15 base-pair sequences (the "polymerisation times"), and similarly extracted the fingers-closed dwell during a single-nucleotide incorporation, which can be used to approximate the rate of processive nucleotide incorporation k_{proc}. The polymerisation-time distributions were not consistent with a model in which all nucleotides were incorporated at equal rates.

By observing polymerisation at lower nucleotide concentration, we were able to extract the fingers-closed dwells associated with each nucleotide incorporation during processive polymerisation. The first 5 base incorporations demonstrated single-exponential kinetics with an incorporation rate of $k_{proc} = 2.7 \pm 0.3s^{-1}$. Counting the number of fingers-closing events during polymerisation demonstrated that the maximum number of observed fingers-closing events (~ 5) were consistent with 5 base incorporations, however the large number of traces with lower numbers of fingers-closing events, suggested that some incorporations were being missed. For an incorporation rate of $k_{proc} = 2.7 \pm 0.3s^{-1}$, we would expect $\approx 20\%$ of events have dwells shorter than 80ms (40ms frame time), but the distribution of number-of-fingers-closing events suggested that more than 20% of events were being missed, suggesting the presence of faster nucleotide incorporations. Increased numbers of short events during preliminary experiments at higher time resolution (20ms) further hinted at the existence of these rapid incorporations. Future work could increase the time-resolution to 10ms, which could directly demonstrate the existence of these fast incorporation events.

During 15-bp extension, the fingers-closed dwells demonstrated bi-exponential kinetics,

extracting incorporation rates of $k_{proc} = 4 \pm 1.5s^{-1}$, and $k_{proc} = 0.3 - 1.8s^{-1}$, which could account for the broadly distributed polymerisation-times measured for 15-bp extensions (Figure 7.6). The slower incorporations could be due to slow dCTP incorporations, or due to altered incorporation kinetics caused by the short template overhang during incorporation of the final bases.

These combined results suggest that nucleotide incorporation rates are highly variable. The biological relevance of these large differences in incorporation rates, or their structural origin, are unclear, but they are important because generalisations about polymerisation mechanisms are generally made by observing incorporations of single-nucleotides. For these generalisations to be relevant, they have to apply to other nucleotides and sequence contexts, which becomes less clear if there is a large heterogeneity in behaviours between these contexts.

Previous single-molecule work has suggested that nucleotides are incorporated at equal rates[72], however these results were based on observations of a smaller number of nucleotide incorporations. The origin of the incorporation rate heterogeneity can be investigated in future work by measuring the single-nucleotide incorporations rates using different nucleotides and surrounding sequences.

Fingers-opening occurs after the slow post-chemistry step

By observing the fingers-closed dwells during single-nucleotide incorporation, we were able to infer rate of processive nucleotide incorporation $k_{open} \approx k_{proc}$. Incorporation of the first base demonstrated simple exponential kinetics, suggesting a single rate-limiting step. The incorporation rate was much slower ($k_{proc} = 3.2 \pm 0.4s^{-1}$) than previous measurements of the rate up to and including chemical incorporation ($k_{pol} = 30 - 50s^{-1}$)[42,53,68,43]. This suggested that fingers-opening after-chemistry is rate-limited by the post-chemistry slow step, which therefore must occur before fingers-opening, or fingers-opening must be the post-chemistry slow step. These assertions rely on the literature rates, however, in the above section we suggested that incorporation rates may be heterogeneous. We are comfortable with using these literature rates for the steps up until chemistry, as they have been consistent

across multiple investigations and sequence contexts. Our assertions would be more rigorous, however, by making measurements of the incorporation rate in our own sequence context, via the use of rapid-quench apparatus.

Our assertion that fingers-opening occurs after the post-chemistry slow step illuminates a previous fluorescence experiment: A bulk fluorescence experiment used a FRET pair on the KF fingers and DNA respectively, and observed biphasic fluorescence change during polymerisation.[42] The first change was assigned to fingers-closing around complementary nucleotide, but the second change was unassigned, as it was impossible to distinguish between a fluorescence change caused by fingers-opening or KF translocation after chemistry. The second fluorescence change was rate-limited by the pre-chemistry slow step, much faster than our observation of fingers-opening, suggesting their fluorescence change must be due to KF translocation after nucleotide incorporation. This would place fingers-opening after translocation, however, further experiments are needed to verify this hypothesis.

It is interesting to note that processive nucleotide incorporation is rate limited by post-chemistry steps.[43] Is this by an optimised feature of the reaction, or a by-product of optimizing for other features? It has been suggested that post-chemistry dwells allow the polymerase to edit misincorporations, however proofreading from the post-chemistry intermediate has yet to be demonstrated. T4 DNA polymerase operates at higher incorporation rates, and has a more active exonuclease domain, possibly to compensate for the shorter post-chemistry dwells, suggesting a link between exonuclease function and processive polymerisation rates.[46,51,115] To investigate the role of the post-chemistry step, future work could investigate residues which change the kinetics of the post-chemistry step, and see how this effects nucleotide incorporation rates and fidelity.

KF incorporates nucleotides efficiently once closed

When observing polymerisation at lower nucleotide concentrations, we were able count the fingers-closing events and compare with the expected number of nucleotide incorporations. Linking observations of conformational dynamics and catalysis in KF allowed us to demonstrate that the maximum number of fingers-closing events coincided roughly with the ex-

pected number of nucleotide incorporations. This suggested that maximally one fingers-closing event was associated with each nucleotide incorporation, consistent with KF incorporating nucleotides highly efficiently on closing. This is consistent with a quantitative estimate of the nucleotide incorporation efficiency of 85-90%, based on the ratio of forwards rates ($k_{pol} = 30 - 50s^{-1}$) and backwards rates ($k_{open} \sim 5s^{-1}$) after fingers closing.

Similarity of binary dynamics and incorporation rates

Interestingly, the incorporation rate of the first base-pair ($k_{proc} = 3.4 \pm 0.8s^{-1}$) was similar to the rate of fingers-opening of the KF binary complex ($k_{open}^{binary} = 5.3 \pm 2s^{-1}$). Given fingers-opening has to occur after nucleotide incorporation, it is possible that the binary complex fingers-opening may be reflecting structural features limiting fingers-opening rates after nucleotide incorporation, which limits processive polymerisation rates. If this is true, then heterogeneity in incorporation rates should be reflected in heterogeneity in binary complex dynamics between sequences. This requires further investigation: Single-nucleotide incorporation experiments across different sequences could observe any correlation between binary complex dynamics and incorporation rates.

On a related note, the binary complex opening rate on deoxy DNAs was significantly different from the dideoxy DNAs ($k_{open-deoxy}^{binary} \approx 5s^{-1}$ and $k_{open-dideoxy}^{binary} = 25s^{-1}$), suggesting that KF interacts with the terminal 3'-OH of the primer when in the closed conformation. This could suggest that KF in the closed conformation may have an interaction with the newly synthesised base-pair "product" after catalysis, which may be rate limiting the reaction overall. In this case, the incorporation rate would correlate with the binary complex dynamics of the extended DNA, not the pre-extended DNA.

7.6 Materials and Methods

Surface-immobilised extendable DNAs were localised in a FOV, in a well containing 20μl of a buffer solution without nucleotides or KF molecules.[ii] After $\sim$ 10 seconds, 20μl of "reaction-

[ii] As in previous chapters, buffers consisted of: 50mM Tris, pH 7.5; 100μg/ml BSA; 1mM DTT; 5% glycerol;1mM Trolox (UV treated); 1.9mg/ml glucose oxidase;0.075mg/ml catalase; 1.5%w/v D-glucose;

initiating" buffer was pipetted into the well, containing a combination of nucleotides and KF at $\sim$ 1nM, resulting in a final KF concentration of $\sim$ 0.5nM. The refractive indices between solutions were matched by creating the well-buffer with an equal volume of water in the place of nucleotides, and KF storage buffer in the place of KF solution.

Fluorescence time-traces, and individual binding events were extracted as described in Chapter 4, with binding events shorter than 0.5s ignored. Hidden Markov Modelling was performed using the software described in Chapter 4 and 5, with FRET values $0.3 < E^* \leq 0.53$ assigned as KF fingers-open, $0.53 < E^* \leq 0.8$ assigned as KF fingers-closed; and $0 < E^* \leq 0.3$ as a donor only KF.

Simulations

Polymerisation-times were based on the sum of several fingers-closed dwells during polymerisation. We simulated polymerisation-times by summing a collection of random variables sampled from exponential distributions. 5, 10 and 15 exponentially distributed random variables with mean 0.32s were summed, 10000 times each to produce a distribution of polymerisation times based on nucleotide incorporations with equal rates (blue traces in Fig 7.8). The distribution in Fig 7.8a (red line), is produced from the sum of two exponentially distributed random variables with mean 0.01 and 3 variables with mean 0.41s. The distribution in Fig 7.8c (red line) is identical, but with the addition of an additional exponentially distributed random variable with a mean of 1.8s.

10mM $MgCl_2$.

Chapter 8

Concluding remarks

8.0.1 Pre-chemistry investigations

In this thesis we have presented an assay allowing extended observations of the fingers conformation of DNA polymerase I (KF), with the first direct observations of the fingers-dynamics of the KF-DNA binary complex, accurately quantifying the interconversion rates between the fingers-open and fingers-closed conformations. Furthermore, we have quantified the dynamics of the KF-DNA-dTTP ternary complex in the pre-chemistry reaction, demonstrating that complementary nucleotide induces rapid fingers-closing, as seen previously (the extracted rates are described in the context of the reaction scheme in Fig. 8.1).[42] The use of single-molecule techniques allows facile observation of reverse rates, demonstrating for the first time that a slow fingers-opening rate stabilises the nucleotide in the fingers-closed conformation. We suggest that these fingers-kinetics could cause an accumulation of the KF-DNA-dTTP population during the reaction, which could account for the differences in pulse-chase and pulse-quench yields observed previously, in the place of the slow pre-chemistry rate-limiting step proposed previously. It should be noted that this slow pre-chemistry step has been investigated for two decades, but remains elusive.[43,52]

It is notable that the fingers-opening reverse step is slow relative to the rate of chemistry, as slow "fingers-opening" relative to chemistry has also been observed in T7 DNA polymerase[93]. We apply the mathematical analysis developed by Tsai and Johnson for T7 DNA polymerase, demonstrating that the fingers and chemical kinetics of KF result in a

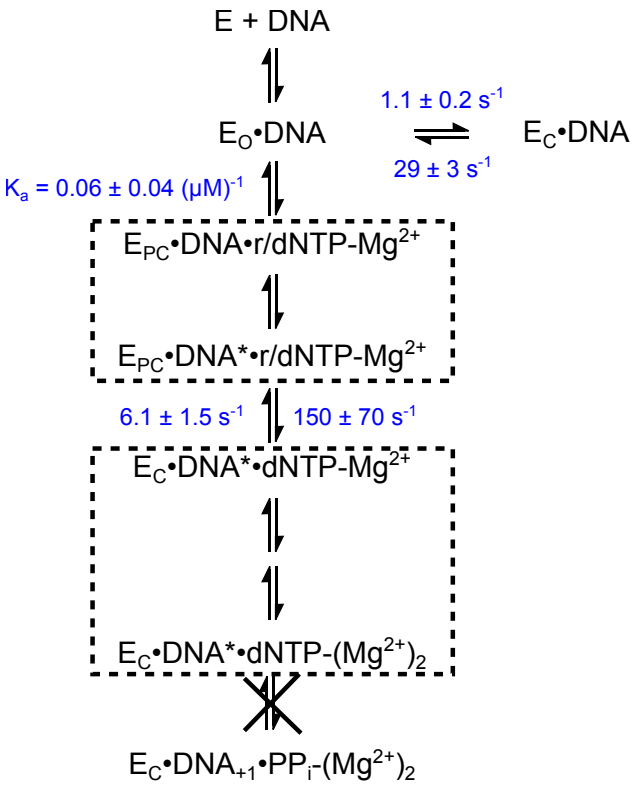


Figure 8.1: Reaction scheme as described in Hohlbein 2013[20] and Chapter 3, including rates and binding affinities (blue) extracted in this thesis during investigations using dideoxy terminated DNA substrates. The schematic is therefore terminated at the phosphoryl transfer step. We observed the rate of fingers closing and opening of the KF-DNA binary complex, and in addition a nucleotide titration allowed us to extract the binding affinity of complementary nucleotide to the open ternary complex, as well as the closing and opening rates of the ternary complex. Upper and lower dotted boxes delineate the collection of substates contained in the open and closed ternary complexes respectively, which we are unable to observe individually using our assay. The exact composition of these composite states is unclear - for instance it is not clear whether the dideoxy terminated complex can bind the second magnesium ion without the 3'-OH to stabilise it.

catalytic efficiency which is independent of the chemical step, as for T7 DNA polymerase. Intuitively, slow "fingers-opening" relative to chemistry biases the reaction in the forwards direction once KF has bound a nucleotide, although the biasing is to lower degree for KF than T7 DNA polymerase, perhaps consistent with the lower fidelity of KF. The removal of the rate-limiting step from contribution to fidelity is surprising, as fidelity has long thought to be dependent on the rate-limiting step of the reaction. The discovery of these kinetics for both KF and T7 DNA polymerase suggest that these kinetics could be general for high-fidelity polymerases.[93,66]

8.0.2 On the role of large scale conformational changes more generally

Considering large-scale conformational changes in enzyme functions more generally, it might be tempting to assign large-scale conformational changes as removing the rate of chemistry from the catalytic efficiency of enzymes. In order to investigate the role of the conformational

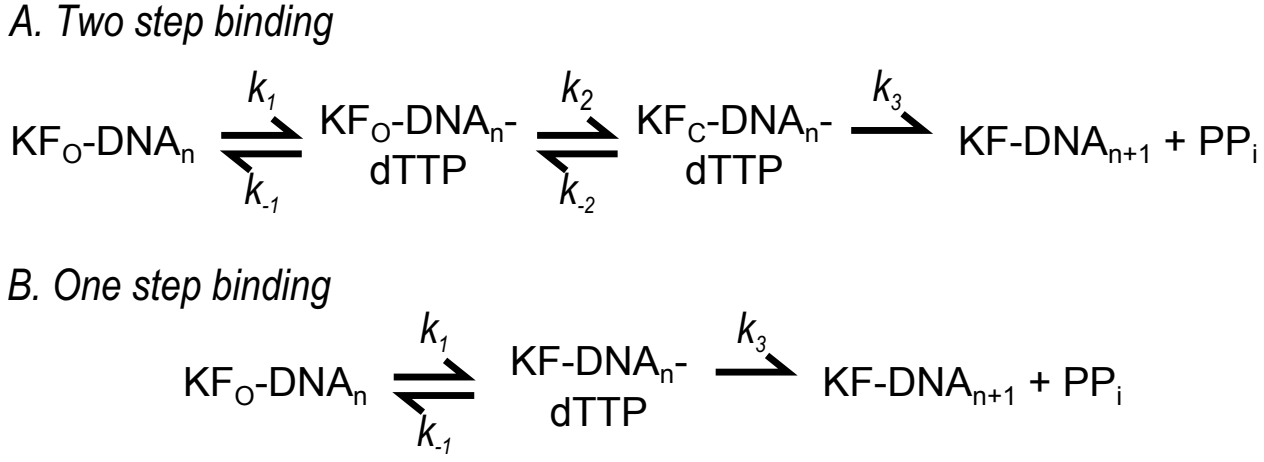


Figure 8.2: Reaction scheme for (a) a two step binding mechanism and (b) a one step binding mechanism.

change more generally, here we briefly compare the reaction mechanism for KF (Fig 8.2a), with a hypothetical reaction mechanism lacking the fingers-closing conformational change (Fig 8.2b), to investigate if a two step binding mechanism can add features lacking in a one step binding mechanism.

Table 8.0.2 describes the k_{cat}, K_M, and k_{cat}/K_M for a one step binding mechanism, and the two step binding mechanism of KF in which fingers-opening (k_{-2}) is much slower than the rate of catalysis (k_3). For a one step binding mechanism, we observe that the catalytic efficiency can also be made independent of the catalytic rate, if the rate of substrate release (k_{-1}), is much slower than chemistry (k_3), leading to a $k_{cat}/K_M = k_1$.

This demonstrates that the conformational change is not strictly necessary to remove the chemistry dependence, if a protein binds very tightly to a substrate in a one-step binding mechanism. However, the conformational change of the two-step mechanism could simply reflect the method by which the polymerase achieves tight substrate binding, by surrounding the substrate from all directions as the protein closes around it. Large-scale conformational changes could simply be a mechanism used by polymerases to bind very tightly to substrates and the transition state of a chemical reaction. However, the result of this is that the protein has to maximise its closing rate (k_2) in order to maximise its catalytic efficiency.

Rate constants

Parameter	Two step binding with $k_3 >> k_{-2}$	One step binding
k_{cat}/k_M	$K_1 k_2$	$\frac{k_1 k_3}{k_{-1}+k_3}$
k_{cat}	$\frac{k_2 k_3}{k_2+k_3}$	k_3
K_M	$\frac{k_3}{K_1(k_2+k_3)}$	$\frac{k_{-1}+k_3}{k_1}$

8.0.3 Investigating the entire DNA polymerisation pathway

We extended our assay to observe the fingers conformations during the entire DNA polymerisation reaction, demonstrating long dwells in the fingers-closed conformation during polymerisation, suggesting that the post-chemistry fingers dynamics may be as important to consider as the pre-chemistry reaction dynamics. These represent the first observations of the fingers conformations of KF during the post-chemistry steps. We observe that the steps after post-chemistry fingers-opening are rapid, and contribute little to the overall reaction rate.

The rate of opening from the finger-closed state during single-nucleotide incorporations was much slower that the rate of nucleotide incorporation (k_{pol}) suggesting that the post-chemistry slow step which limits processive incorporation is caused by fingers-opening, or occurs before fingers-opening (the extracted rates are described in the context of the reaction scheme in Fig. 8.3). We assigned these fingers-opening rates as approximately equal to the rate of processive polymerisation (k_{proc}).

Measurements of the time taken to incorporate 5, 10 and 15 base-pairs did not fit well to a model in which all nucleotides are incorporated an equal rate. Similarly, counting the number of fingers-closing events during polymerisation suggested presence of fast nucleotide incorporations which were being missed, as the maximal number of fingers-closing events were consistent each nucleotide incorporation being associated with one fingers-closing event (i.e. nucleotides are incorporated efficiently), however there were many polymerisations with a lower number of counts. These observations supported the notion that nucleotide

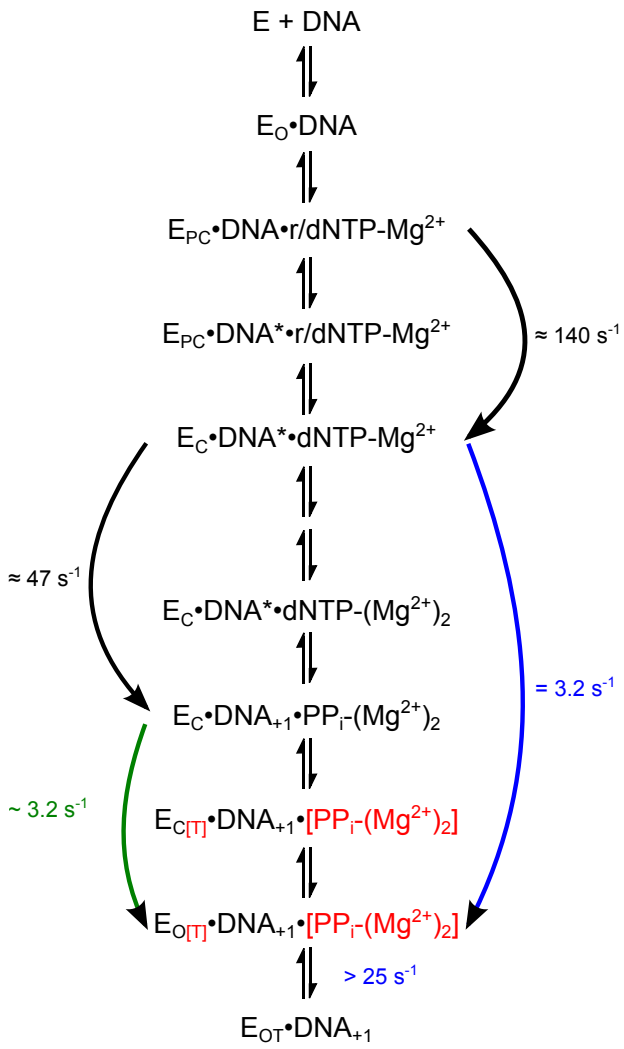


Figure 8.3: Reaction scheme as described in Hohlbein 2013[20] and Chapter 3, extended to include the post-chemistry steps of KF translocation, fingers-opening, and release of pyrophospate and magnesium ions. The order of these steps are unknown, and are therefore highlighted in red and in square brakets to indicate potential presence of absence of these ligands or conformational changes at given points of the reaction. The translocated polymerase is represented by a subscripted T and the DNA with the next incorporated nucleotide is represented by DNA_{+1}. Previously measured rates of fingers-closing[42] and k_{cat} (see Chapter 5) are seen in black, associated with arrows representing the rate of movement between states, with the bracketed substates occuring inbetween. We measure the rate of opening of the fingers during a polymerisation reaction ($3.2s^{-1}$ blue), which can be compared to the previously measured k_{cat} and the steps that it limits during the reaction, to allow us to assign this slow fingers opening to steps occuring after chemistry (green).

incorporation rates are heterogeneous. This heterogeneity could be important, as the kinetics of single-nucleotide incorporations are often used to make generalisations on the enzymes mechanisms, which might not apply if there is significant heterogeneity between nucleotides or sequences.

Additionally, we devised novel visualisation techniques allowing rapid and intuitive characterisation of the behaviours at hundreds of molecules simultaneously, which we hope will be useful to the single-molecule community, or spur further development of visualisation techniques allowing single-molecule experiments to be presented more completely at publication level. Within our laboratory, this visualisation technique has been combined with tethered-fluorophore-motion observations of DNA polymerisation performed together with Justin Pinkney[130], and is currently being used by Diego Duchi in observations of the real-time functioning of RNA polymerase.

Future work

One of the primary sources of noise in our investigations was the binding of KF to the surface nearby DNA molecules. We attributed this primarily to neutravidin binding, and future work could reduce noise by reducing the biotin or neutravidin present on the surface, which is used to immobilise the biotinylated DNA. This would allowing more sensitive measurements, such as the extraction of slow interconversion rates in the presence of mismatching nucleotide, currently challenging because of interference from KF-surface binding events. Kinetic observations in the presence of mismatching nucleotide are desirable as the mechanisms by which mismatches are incorporated are unclear - whether KF incorporates them from the closed conformation, or from a distorted fingers-open conformation. Combining measurements of fingers-closing on extendable substrates, with knowledge of when nucleotide incorporation occurs, would help describe how mismatched incorporations occur mechanistically.

We observed a similarity in the fingers-opening rates of binary complex and fingers-opening rates during nucleotide incorporation. The most imminent future work will investigate if there is any causal connection between these observations. These investigations will be performed using the same experiments as described in Chapters 6 and 7, in which we will observe incorporation rates of different nucleotides with different surrounding sequences, whilst also quantifying the binary complex dynamics to see if the change in incorporation rates between sequences and nucleotides corresponds with changes in binary complex dynamics. Such observations would prove particularly interesting as intrinsic dynamics in other enzymes have been found to rate-limit their turnover, but whether this occurs generally between enzymes is unclear.

We also observed that the fingers-opening rates differed between dideoxy and dioxy DNA substrates, suggesting that the KF fingers may interact with the primer terminus 3'-OH. It is interesting to speculate on this interaction; for instance, could the fingers-interaction with the newly formed base-pair limit fingers-opening after chemistry?

The single-molecule investigations in this thesis have allowed observations of rates which would be challenging or impossible to observe using ensemble techniques, allowing previous

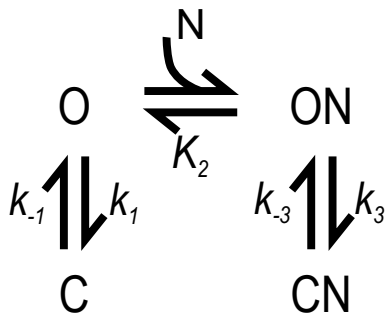
work to be considered in a new light, and have contributed to the understanding of the basis of polymerase fidelity. The single-molecule nature of these investigation have also allowed the observation of rare events suggesting that multiple polymerases may bind DNA at once (potential future investigations in this direction have been discussed in Section 4.2.3). The latter stages of the multi-step polymerisation reaction have been observed without synchronisation, and heterogeneity has been highlighted which could be exploited to observe causal connections between various kinetic parameters in the polymerisation reaction, which could represent a novel advantage of single-molecule techniques.

DNA polymerase is a particularly important enzyme because of its role in mutagenesis and carcinogenesis. Understanding the mechanisms of correct and erroneous replication, as well as DNA repair processes across the wide spectrum of polymerases is important to understanding the various factors contributing to mutagenesis. In this thesis we have contributed knowledge on the reaction mechanism of a specific DNA polymerase, Klenow Fragment, as well as implemented a more general technique for quantifying the fidelity of this polymerase in terms of its reaction substeps. This could be used in the future to understand the susceptibility of this polymerase to misincorporation in terms of its reaction sub-steps, particularly in the presence of carinogenic adducts on the DNA, or other chemical modifications or mutations of the protein.

Appendix A

Deriving an expression for the fingers-closing rate

Notation to describe the reaction scheme in Figure 5.6C, in which KF interconverts between open (O), closed (C) conformations in the binary complex, whose rates change when bound to complementary nucleotide (N):



Total population in open state:

$$[T] = [O] + [ON] \tag{A.1}$$

The rate of fingers closing is given by change in the total population in the fingers-open conformation:

$$\frac{d[T]}{dt} = k_1[O] + k_3[ON] \tag{A.2}$$

We assume nucleotide binding equilibrium:

$$k_2[O] = k_{-2}[ON] \tag{A.3}$$

Using Eq. A.1 and A.3

$$[T] = [O]\left(1 + \frac{k_2}{k_{-2}}\right) \tag{A.4}$$

Using Eq. A.2 and A.3

$$\frac{d[T]}{dt} = \left(k_1 + \frac{k_3 k_2}{k_{-2}}\right)[O] \tag{A.5}$$

Hence, using Eq. A.4 and A.5, we have:

$$\frac{d[T]}{dt} = \left(\frac{k_1 + k_2 K_2[N]}{1 + K_2[N]}\right)[T]$$

Hence, the observed fingers-closing rate is:

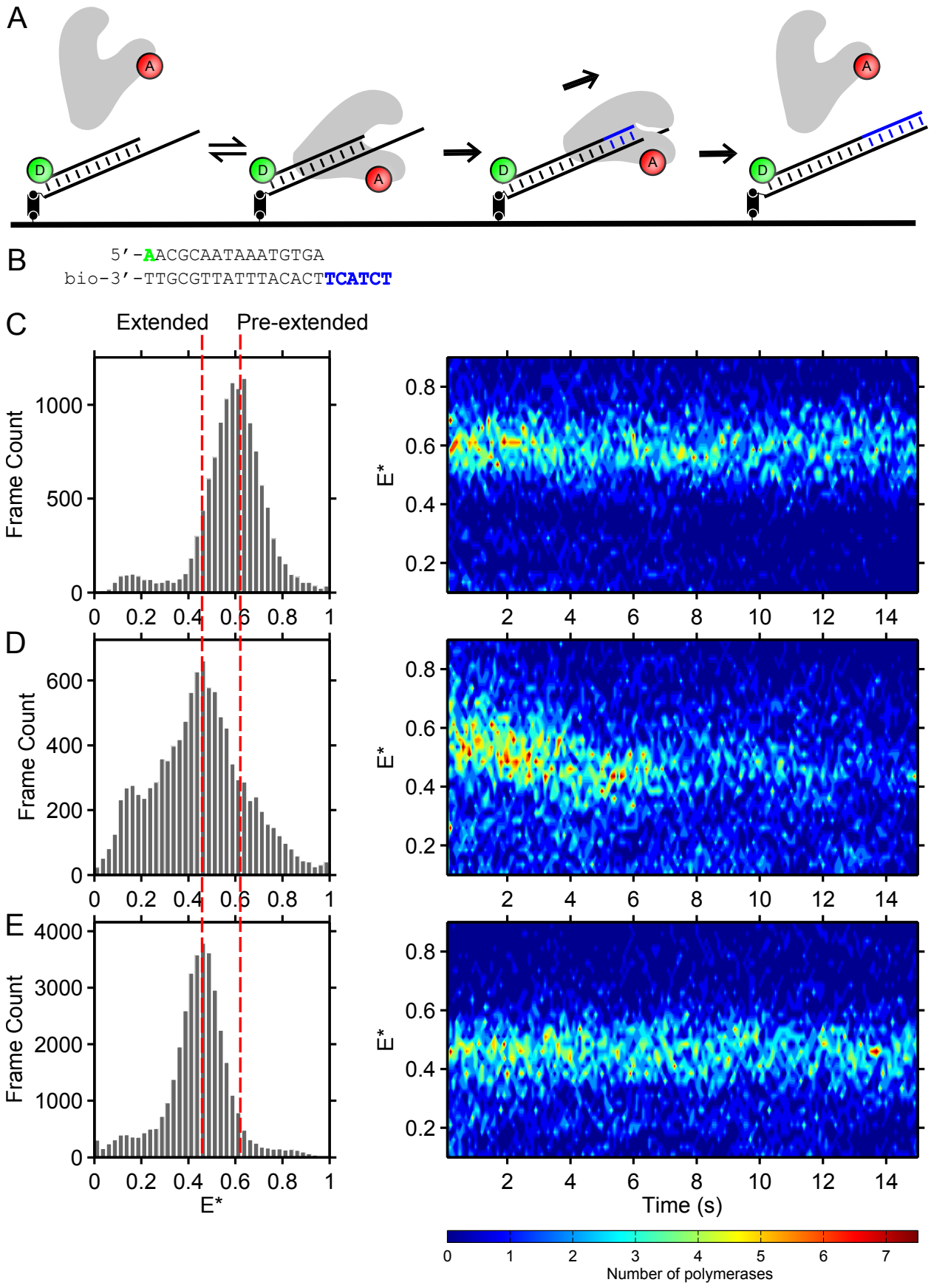
$$k_{close}^{obs} = \frac{k_1 + k_3 K_2[N]}{1 + K_2[N]}$$

Appendix B

Single KF molecules translocating along DNA during polymerisation

In Chapters 6 and 7, we reported experiments observing the conformation of KF's fingers when bound to extendable DNAs in the presence of nucleotides, conditions which allow the full DNA polymerisation reaction to occur. Prior to nucleotide addition, we observed KF binding and dissociating from surface-immobilised DNAs whilst primarily in the fingers-open conformation. After nucleotide addition, a gradual decrease in KF-DNA binding suggested that the DNAs were being polymerised to create DNA blunt ends or short template overhangs, to which KF is known to bind weakly.[128] In addition, the KF fingers were in the closed conformation prior to dissociation, suggesting the binding and processing of nucleotides by the polymerase. In order to further confirm the presence of nucleotide addition reactions and DNA extension, we devised an assay to observe KF translocating along DNA during polymerisation, using methods similar to those described in previous work.[72,73]

During polymerisation, nucleotides are added to the 3' end of the primer, resulting in the KF-DNA bind site moving away from the primer 5' end. To observe this translocation of KF during polymerisation, we used FRET between dyes attached to the DNA and KF. The DNA was labelled with a donor dye (Cy3B) at the -16 primer position, while KF was labelled with an acceptor dye (Atto647N) on the 550 residue of the thumb subdomain (Fig. B.1A,B). We used Alternating Laser Excitation (ALEX) to monitor the fluorescence of both donor


A
D
A
B
5′-AACGCAATAAATGTGA
bio-3′-TTGCGTTATTTACACTTCATCT
C
Extended
Pre-extended
Frame Count
E*
D
E
Time (s)
Number of polymerases

and acceptor dyes at the surface-immobilised DNAs. Using this assay, we observed binding of KF to a DNA via the appearance of acceptor dye fluorescence, as well as the presence of FRET between the donor and acceptor dyes (Fig. B.2 see increase in the acceptor emission under acceptor excitation (AA), as well as an anticorrelated change in donor and acceptor emission under donor excitation (DD/DA)).

In the absence of nucleotides, the FRET efficiency between KF and DNA fluorophores was centered around E*$\approx$ 0.61 (Fig. B.1C, left hand panel). The FRET efficiency distribution remained constant over time, as illustrated in Fig. B.1C (right hand panel), in which the FRET traces of the KF-DNA complexes at each DNA were extracted and post-synchronised, providing a histogram of FRET efficiencies at each time point, whose time evolution is represented in a colormap.

The polymerisation reaction was initiated using the second method described in Chapter 6, in which a solution containing 2nM KF and 0.1μM of all nucleotides was added to a solution of surface-immobilised DNAs, with the polymerisation reaction delayed by the time taken for KF to diffuse and bind the DNA. After initiation of polymerisation, the histogram of E* values demonstrated a new peak at a lower E* (E*$\approx$ 0.46) with a smaller population remaining at the initial high E* (E*$\approx$ 0.61) (Fig. B.1D, left hand panel). We interpreted this as polymerisation of the DNA by KF, in which KF initially binds the DNA in the

---

Figure B.1 *(preceding page)*: (a) Assay to monitor translocation of KF along DNA during polymerisation. KF is labelled with an acceptor on the thumb, and the DNA labelled with a donor. As KF polymerises the DNA, its binding site at the 3' end of the primer becomes increasingly distant from the donor dye. (b) DNA primer-template used in these experiments, labelled with Cy3B via an amine group at the -16 position (green). Template overhang shown in blue. The template is biotinylated for surface-immobilisation. (c) Histogram of E* values at KF-DNA complexes in the absence of nucleotides, and colormap showing the time evolution of the E* values after post-synchronisation of these binding events. Collected at 2nM KF with laser powers of 0.75mW green and 0.3mW red, with a time resolution of 100ms. (d) Histogram of E* values at KF-DNA complexes after initiating polymerisation in the presence of 0.1μM of all nucleotides, and colormap showing post-synchronised binding events with decreasing E* values demonstrating extension of the DNA. Conditions as for c. (e) Histogram of E* values at KF-DNA complexes after allowing polymerisation to complete in the presence of 0.1μM of all nucleotides, and colormap showing post-synchronised binding events with steady low E* values. Data collected at 30nM KF with laser powers of 1mW green and 1mW red, with a time resolution of 80ms.

high E* conformation before extending the DNA, then elongates it, thereby reducing the FRET efficiency as the primer is extended. This interpretation was supported by the post-synchronised FRET traces of the KF-DNA complexes at each DNA. These demonstrated that the FRET efficiency was initially high, representing KF binding to the pre-extended DNA, with a subsequent decrease in FRET over $\approx 6s$ as the DNA was polymerised (Fig. B.1D, right hand panel). The binding of KF to DNA also decreased after this initial E* decrease, as seen by the decrease in the number of KF-DNA complexes after polymerisation (Fig. B.1D, right hand panel).

In order to confirm the end point of the reaction, polymerisation was initiated at high KF concentration, in the presence of 0.1μM of all nucleotides, and allowed to equilibriate for $\approx 3-5$ mins. The high KF concentration then allowed observation of a sufficient number of KF-DNA complexes to create an E* histgram, which demonstrated that after DNA polymerisation, KF bound the DNA in the low E* conformation (Fig. B.1E, left hand panel). Binding in this conformation did not change over time (Fig. B.1E, right hand panel).

The 0.15 change in E* between pre-extended and extended DNAs is consistent with estimates of the distance changes during polymerisation, based on the crystal structures of a related polymerase from *Bacillus stearothermophilus* (PDB 1L3U). Using this crystal structure, the distance between the KF thumb and DNA can be estimated for a primer extended by 0, 4 and 6 bases. This is achieved by interpolation of the bound DNA to create a longer DNA duplex, and then measurement of the distance between the Cβ of the 498 residue (equivalent to the 550 position in *E.Coli* KF) and the C1' of the nucleotide sugar at the -16, -20 and -22 primer positions. Based on $R_0 = 67$Å (determined previously for Cy3B-Atto647N[131]), these distances gave expected FRET changes of 0.13 and 0.19 for a DNA extended by 4 and 6 bases respectively (the template overhang is 6 bases long). The observed 0.15 change lies between these values - consistent with extension of between 4 and 6 bases. This is reasonable as it is not entirely clear whether KF can incorporate nucleotides and create a full blunt end, or whether it can only create short 1-2 template base overhangs under our reaction conditions and time-scales.

To further confirm the extension of DNA, we initiated polymerisation at low KF con-

centration and low nucleotide concentration, and observed the reaction at high temporal resolution (20ms). We were were able to observe single-molecules exhibiting translocation behaviour, in which the first KF-DNA binding event at DNAs demonstrates E* values which decreased gradually from a high E* value, to a lower E* value (Fig. B.2).

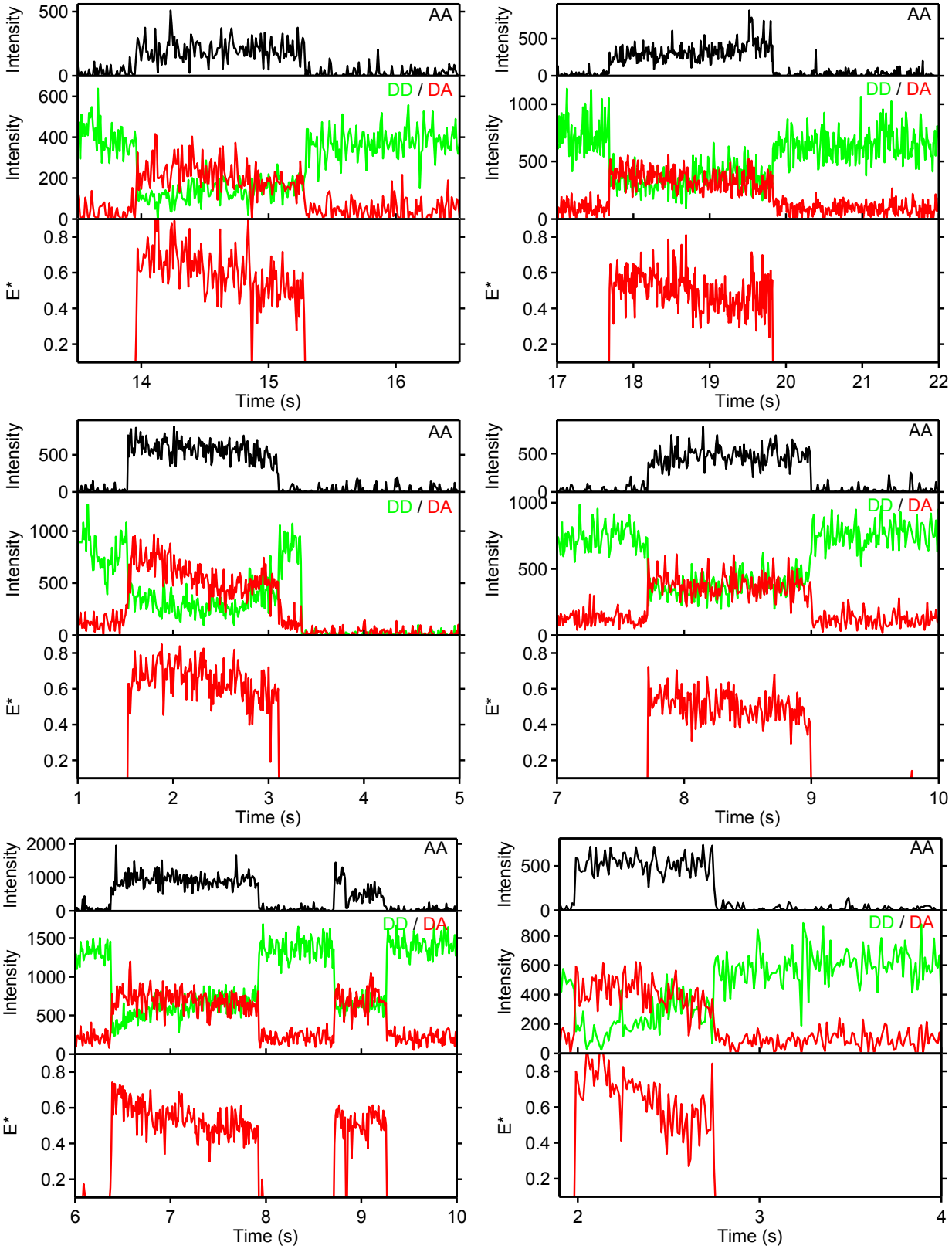


Figure B.2: Fluorescence time traces demonstrating the decrease FRET (E*) seen on the first KF binding event at different DNA molecules, in the presence of 0.1μM of dATP, dCTP, dGTP and dTTP. Donor and acceptor emission under donor exciation shown in green (DD) and red (DA). Acceptor emission under acceptor excitation seen in black (AA). Collected at 2nM KF with laser powers of 2mW green and 1mW red, with a time resolution of 20ms.

Bibliography

[1] R. Wolfenden and M. J. Snider, "The Depth of Chemical Time and the Power of Enzymes as Catalysts," *Accounts of Chemical Research*, vol. 34, pp. 938–945, Dec. 2001.

[2] D. Herschlag and A. Natarajan, "Fundamental Challenges in Mechanistic Enzymology: Progress toward Understanding the Rate Enhancements of Enzymes," *Biochemistry*, vol. 52, pp. 2050–2067, Mar. 2013.

[3] Z. D. Nagel and J. P. Klinman, "A 21st century revisionist's view at a turning point in enzymology," *Nature Chemical Biology*, vol. 5, pp. 543–50, Aug. 2009.

[4] S. Borman, "Much ado about enzyme mechanisms," *Chem. Eng. News*, vol. 82, pp. 35–39, 2004.

[5] J. Villa, M. Strajbl, T. M. Glennon, Y. Y. Sham, Z. T. Chu, and A. Warshel, "How important are entropic contributions to enzyme catalysis?," *Proceedings of the National Academy of Sciences*, vol. 97, pp. 11899–904, Oct. 2000.

[6] M. I. Page and W. P. Jencks, "Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect," *Proceedings of the National Academy of Sciences*, vol. 68, pp. 1678–83, Aug. 1971.

[7] A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. M. Olsson, "Electrostatic basis for enzyme catalysis," *Chemical Reviews*, vol. 106, pp. 3210–35, Aug. 2006.

[8] R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus, "Dynamics of ligand binding to myoglobin," *Biochemistry*, vol. 14, pp. 5355–5373, Dec. 1975.

[9] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, "The energy landscapes and motions of proteins," *Urbana*, vol. 51, no. 61801, p. 61801, 1991.

[10] J. Evenäs, S. Forsén, A. Malmendal, and M. Akke, "Backbone dynamics and energetics of a Calmodulin domain mutant exchanging between closed and open conformations11Edited by P. E. Wright," *Journal of Molecular Biology*, vol. 289, no. 3, pp. 603–617, 1999.

[11] F. Schotte, J. Soman, J. S. Olson, M. Wulff, and P. A. Anfinrud, "Picosecond time-resolved X-ray crystallography: probing protein function in real time," *Journal of Structural Biology*, vol. 147, pp. 235–46, Sept. 2004.

[12] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern, "Intrinsic dynamics of an enzyme underlies catalysis," *Nature*, vol. 438, pp. 117–21, Nov. 2005.

[13] M. Wolf-Watz, V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E. Z. Eisenmesser, and D. Kern, "Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair," *Nature Structural & Molecular Biology*, vol. 11, pp. 945–9, Oct. 2004.

[14] S. Xie, "Single-Molecule Approach to Enzymology," *Single Molecules*, vol. 2, pp. 229–236, Dec. 2001.

[15] R. Yasuda, H. Noji, M. Yoshida, K. Kinosita, and H. Itoh, "Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase," *Nature*, vol. 410, pp. 898–904, Apr. 2001.

[16] X. Zhuang, "A Single-Molecule Study of RNA Catalysis and Folding," *Science*, vol. 288, pp. 2048–2051, June 2000.

[17] X. S. Xie, "Single-molecule Enzymology," *Journal of Biological Chemistry*, vol. 274, pp. 15967–15970, June 1999.

[18] H. P. Lu, "Single-Molecule Enzymatic Dynamics," *Science*, vol. 282, pp. 1877–1882, Dec. 1998.

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

[20] J. Hohlbein, L. Aigrain, T. D. Craggs, O. Bermek, O. Potapova, P. Shoolizadeh, N. D. F. Grindley, C. M. Joyce, and A. N. Kapanidis, "Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion," *Nature Communications*, vol. 4, p. 2131, July 2013.

[21] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Springer, 2007.

[22] M. Dahan, A. A. Deniz, T. Ha, D. S. Chemla, P. G. Schultz, and S. Weiss, "Ratiometric measurement and identification of single diffusing molecules," *Chemical Physics*, vol. 247, no. 1, pp. 85–106, 1999.

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

[24] P. J. Rothwell, S. Berger, O. Kensch, S. Felekyan, M. Antonik, B. M. Wöhrl, T. Restle, R. S. Goody, and C. A. M. Seidel, "Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 1655–60, Feb. 2003.

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

[26] E. Brooks Shera, "Detection of single fluorescent molecules," *Chemical Physics Letters*, vol. 174, pp. 553–557, Nov. 1990.

[27] S. Nie, D. T. Chiu, and R. N. Zare, "Probing individual molecules with confocal fluorescence microscopy," *Science*, vol. 266, pp. 1018–21, Nov. 1994.

[28] J. R. Fries, L. Brand, C. Eggeling, M. Köllner, and C. A. M. Seidel, "Quantitative Identification of Different Single Molecules by Selective Time-Resolved Confocal Fluorescence Spectroscopy," *The Journal of Physical Chemistry A*, vol. 102, pp. 6601–6613, Aug. 1998.

[29] A. N. Kapanidis, T. A. Laurence, N. K. Lee, E. Margeat, X. Kong, and S. Weiss, "Alternating-laser excitation of single molecules," *Accounts of Chemical Research*, vol. 38, pp. 523–33, July 2005.

[30] H. Schneckenburger, "Total internal reflection fluorescence microscopy: technical innovations and novel applications," *Current Opinion in Biotechnology*, vol. 16, pp. 13–8, Feb. 2005.

[31] S. J. Holden, S. Uphoff, J. Hohlbein, D. Yadin, L. Le Reste, O. J. Britton, and A. N. Kapanidis, "Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy," *Biophysical Journal*, vol. 99, no. 9, pp. 3102–3111, 2010.

[32] I. Rasnik, S. A. McKinney, and T. Ha, "Nonblinking and long-lasting single-molecule fluorescence imaging," *Nature Methods*, vol. 3, pp. 891–3, Nov. 2006.

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

[34] T. Cordes, J. Vogelsang, and P. Tinnefeld, "On the mechanism of Trolox as antiblinking and antibleaching reagent," *Journal of the American Chemical Society*, vol. 131, pp. 5018–9, Apr. 2009.

[35] L. A. Loeb and R. J. Monnat, "DNA polymerases and human disease," *Nature reviews. Genetics*, vol. 9, pp. 594–604, Aug. 2008.

[36] U. Hubscher, S. Spadari, G. Villana, and G. Maga, *DNA Polymerases: Discovery, Characterization and Functions in Cellular DNA Transactions*. World Scientific, 2010.

[37] T. Lindahl and R. D. Wood, "Quality control by DNA repair," *Science*, vol. 286, pp. 1897–905, Dec. 1999.

[38] K. A. Eckert and T. A. Kunkel, "DNA polymerase fidelity and the polymerase chain reaction," *Genome Research*, vol. 1, pp. 17–24, Aug. 1991.

[39] T. A. Kunkel and K. Bebenek, "DNA replication fidelity," *Annual Review of Biochemistry*, vol. 69, pp. 497–529, Jan. 2000.

[40] M. J. Bessman, A. Kornberg, I. Lehman, and E. Simms, "Enzymatic synthesis of deoxyribonucleic acid," *Biochim. Biophys. Acta*, vol. 21, pp. 197–198, 1956.

[41] C. M. Joyce and V. Derbyshire, "Purification of Escherichia coli DNA polymerase I and Klenow fragment," *Methods in Enzymology*, vol. 262, pp. 3–13, Jan. 1995.

[42] C. M. Joyce, O. Potapova, A. M. Delucia, X. Huang, V. P. Basu, and N. D. F. Grindley, "Fingers-closing and other rapid conformational changes in DNA polymerase I (Klenow fragment) and their role in nucleotide selectivity," *Biochemistry*, vol. 47, pp. 6103–16, June 2008.

[43] M. E. Dahlberg and S. J. Benkovic, "Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant," *Biochemistry*, vol. 30, pp. 4835–4843, May 1991.

[44] V. Mizrahi, R. N. Henrie, J. F. Marlier, K. A. Johnson, and S. J. Benkovic, "Rate-limiting steps in the DNA polymerase I reaction pathway," *Biochemistry*, vol. 24, pp. 4010–4018, July 1985.

[45] A. Fersht, *Enzyme Structure and Mechanism*. W.H.Freeman & Co Ltd, 1984.

[46] C. M. Joyce and S. J. Benkovic, "DNA polymerase fidelity: kinetics, structure, and checkpoints," *Biochemistry*, vol. 43, pp. 14317–24, Nov. 2004.

[47] A. K. Showalter and M.-D. Tsai, "A Reexamination of the Nucleotide Incorporation Fidelity of DNA Polymerases," *Biochemistry*, vol. 41, pp. 10571–10576, Aug. 2002.

[48] K. A. Johnson, "The kinetic and chemical mechanism of high-fidelity DNA polymerases," *Biochimica et Biophysica Acta*, vol. 1804, pp. 1041–8, May 2010.

[49] J. A. A. Ketelaar, H. R. Gersmann, and K. Koopmans, "The rate of hydrolysis of some p-nitrophenol esters of ortho-phosphoric- and thio-phosphoric acid," *Recueil des Travaux Chimiques des Pays-Bas*, vol. 71, pp. 1253–1258, Sept. 1952.

[50] D. Herschlag, J. A. Piccirilli, and T. R. Cech, "Ribozyme-catalyzed and nonenzymic reactions of phosphate diesters: rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom," *Biochemistry*, vol. 30, pp. 4844–4854, May 1991.

[51] R. D. Kuchta, P. Benkovic, and S. J. Benkovic, "Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity," *Biochemistry*, vol. 27, pp. 6716–6725, Sept. 1988.

[52] C. M. Joyce, "Techniques used to study the DNA polymerase reaction pathway," *Biochimica et Biophysica Acta*, vol. 1804, pp. 1032–40, May 2010.

[53] R. D. Kuchta, V. Mizrahi, P. A. Benkovic, K. A. Johnson, and S. J. Benkovic, "Kinetic mechanism of DNA polymerase I (Klenow)," *Biochemistry*, vol. 26, pp. 8410–8417, Dec. 1987.

[54] D. L. Ollis, P. Brick, R. Hamlin, N. G. Xuong, and T. A. Steitz, "Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP," *Nature*, vol. 313, pp. 762–766, Feb. 1985.

[55] J. R. Kiefer, C. Mao, C. J. Hansen, S. L. Basehore, H. H. Hogrefe, J. C. Braman, and L. S. Beese, "Crystal structure of a thermostable Bacillus DNA polymerase l large fragment at 2.1 Åresolution," *Structure*, vol. 5, pp. 95–108, Jan. 1997.

[56] S. Korolev, M. Nayal, W. M. Barnes, E. Di Cera, and G. Waksman, "Crystal structure of the large fragment of Thermus aquaticus DNA polymerase I at 2.5-A resolution: structural basis for thermostability," *Proceedings of the National Academy of Sciences*, vol. 92, pp. 9264–8, Oct. 1995.

[57] L. Beese, V. Derbyshire, and T. Steitz, "Structure of DNA polymerase I Klenow fragment bound to duplex DNA," *Science*, vol. 260, pp. 352–355, Apr. 1993.

[58] S. J. Johnson, J. S. Taylor, and L. S. Beese, "Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 3895–900, Apr. 2003.

[59] T. Nakamura, Y. Zhao, Y. Yamagata, Y.-j. Hua, and W. Yang, "Watching DNA polymerase η make a phosphodiester bond," *Nature*, vol. 487, pp. 196–201, July 2012.

[60] E. Y. Wu and L. S. Beese, "The structure of a high fidelity DNA polymerase bound to a mismatched nucleotide reveals an "ajar" intermediate conformation in the nucleotide selection mechanism," *The Journal of Biological Chemistry*, vol. 286, pp. 19758–67, June 2011.

[61] C. A. Brautigam and T. A. Steitz, "Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes," *Current Opinion in Structural Biology*, vol. 8, pp. 54–63, Feb. 1998.

[62] C. M. Joyce and T. A. Steitz, "Function and structure relationships in DNA polymerases," *Annual Review of Biochemistry*, vol. 63, pp. 777–822, Jan. 1994.

[63] D. T. Minnick, M. Astatke, C. M. Joyce, and T. A. Kunkel, "A thumb subdomain mutant of the large fragment of Escherichia coli DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity," *The Journal of Biological Chemistry*, vol. 271, pp. 24954–61, Oct. 1996.

[64] S. Doublié, M. R. Sawaya, and T. Ellenberger, "An open and closed case for all polymerases," *Structure*, vol. 7, pp. R31–R35, Feb. 1999.

[65] C. M. Joyce, O. Potapova, A. M. Delucia, X. Huang, V. P. Basu, and N. D. F. Grindley, "Fingers-closing and other rapid conformational changes in DNA polymerase I (Klenow

fragment) and their role in nucleotide selectivity," *Biochemistry*, vol. 47, pp. 6103–16, June 2008.

[66] K. A. Johnson, "Role of induced fit in enzyme specificity: a molecular forward/reverse switch," *The Journal of Biological Chemistry*, vol. 283, pp. 26297–301, Sept. 2008.

[67] E. L. Rachofsky, R. Osman, and J. B. Ross, "Probing structure and dynamics of DNA with 2-aminopurine: effects of local environment on fluorescence," *Biochemistry*, vol. 40, pp. 946–56, Jan. 2001.

[68] V. Purohit, N. D. F. Grindley, and C. M. Joyce, "Use of 2-aminopurine fluorescence to examine conformational changes during nucleotide incorporation by DNA polymerase I (Klenow fragment)," *Biochemistry*, vol. 42, pp. 10200–11, Sept. 2003.

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

[70] K. Singh, A. Srivastava, S. S. Patel, and M. J. Modak, "Participation of the fingers subdomain of Escherichia coli DNA polymerase I in the strand displacement synthesis of DNA," *The Journal of Biological Chemistry*, vol. 282, pp. 10594–604, Apr. 2007.

[71] C. E. Aitken, R. A. Marshall, and J. D. Puglisi, "An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments," *Biophysical Journal*, vol. 94, no. 5, pp. 1826–1835, 2008.

[72] T. D. Christian, L. J. Romano, and D. Rueda, "Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 21109–14, Dec. 2009.

[73] R. P. Markiewicz, K. B. Vrtis, D. Rueda, and L. J. Romano, "Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I," *Nucleic Acids Research*, vol. 40, pp. 7975–7984, June 2012.

[74] K. B. Vrtis, R. P. Markiewicz, L. J. Romano, and D. Rueda, "Carcinogenic adducts induce distinct DNA polymerase binding orientations," *Nucleic Acids Research*, vol. 41, pp. 7843–53, Sept. 2013.

[75] J. P. Torella, S. J. Holden, Y. Santoso, J. Hohlbein, and A. N. Kapanidis, "Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis," *Biophysical Journal*, vol. 100, pp. 1568–77, Mar. 2011.

[76] Y. Santoso, J. P. Torella, and A. N. Kapanidis, "Characterizing single-molecule FRET dynamics with probability distribution analysis," *Chemphyschem*, vol. 11, pp. 2209–19, July 2010.

[77] S. Y. Berezhna, J. P. Gill, R. Lamichhane, and D. P. Millar, "Single-molecule Förster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase I," *Journal of the American Chemical Society*, vol. 134, pp. 11261–8, July 2012.

[78] B. Maier, D. Bensimon, and V. Croquette, "Replication by a single DNA polymerase of a stretched single-stranded DNA," *Proceedings of the National Academy of Sciences*, vol. 97, pp. 12002–7, Oct. 2000.

[79] D. R. Garalde, C. A. Simon, J. M. Dahl, H. Wang, M. Akeson, and K. R. Lieberman, "Distinct complexes of DNA polymerase I (Klenow fragment) for base and sugar discrimination during nucleotide substrate selection," *The Journal of Biological Chemistry*, vol. 286, pp. 14480–92, Apr. 2011.

[80] F. Olasagasti, K. R. Lieberman, S. Benner, G. M. Cherf, J. M. Dahl, D. W. Deamer, and M. Akeson, "Replication of individual DNA molecules under electronic control using a protein nanopore," *Nature Nanotechnology*, vol. 5, pp. 798–806, Nov. 2010.

[81] N. Hurt, H. Wang, M. Akeson, and K. R. Lieberman, "Specific nucleotide binding and rebinding to individual DNA polymerase complexes captured on a nanopore," *Journal of the American Chemical Society*, vol. 131, pp. 3772–8, Mar. 2009.

[82] S. Benner, R. J. A. Chen, N. A. Wilson, R. Abu-Shumays, N. Hurt, K. R. Lieberman, D. W. Deamer, W. B. Dunbar, and M. Akeson, "Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore," *Nature Nanotechnology*, vol. 2, pp. 718–24, Nov. 2007.

[83] T. J. Olsen, Y. Choi, P. C. Sims, O. T. Gul, B. L. Corso, C. Dong, W. A. Brown, P. G. Collins, and G. A. Weiss, "Electronic measurements of single-molecule processing by DNA polymerase I (Klenow fragment)," *Journal of the American Chemical Society*, vol. 135, pp. 7855–60, May 2013.

[84] G. Luo, M. Wang, W. H. Konigsberg, and X. S. Xie, "Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 12610–5, July 2007.

[85] F. Aboul-ela, D. Koh, I. Tinoco, and F. H. Martin, "Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A,C,G,T)," *Nucleic Acids Research*, vol. 13, pp. 4811–24, July 1985.

[86] K. A. Eckert and T. A. Kunkel, "DNA polymerase fidelity and the polymerase chain reaction," *PCR Methods and Applications*, vol. 1, pp. 17–24, Aug. 1991.

[87] J. Petruska, "Comparison between DNA Melting Thermodynamics and DNA Polymerase Fidelity," *Proceedings of the National Academy of Sciences*, vol. 85, pp. 6252–6256, Sept. 1988.

[88] A. C. Olson, J. N. Patro, M. Urban, and R. D. Kuchta, "The Energetic Difference between Synthesis of Correct and Incorrect Base Pairs Accounts for Highly Accurate DNA Replication," *Journal of the American Chemical Society*, vol. 135, pp. 1205–1208, Jan. 2013.

[89] T. A. Kunkel, "DNA replication fidelity," *The Journal of Biological Chemistry*, vol. 279, pp. 16895–8, Apr. 2004.

[90] E. T. Kool, "Replication of non-hydrogen bonded bases by DNA polymerases: a mechanism for steric matching," *Biopolymers*, vol. 48, pp. 3–17, Jan. 1998.

[91] E. T. Kool, "Replacing the Nucleobases in DNA with Designer Molecules," *Accounts of Chemical Research*, vol. 35, pp. 936–943, Nov. 2002.

[92] W. A. Beard, D. D. Shock, B. J. Vande Berg, and S. H. Wilson, "Efficiency of correct nucleotide insertion governs DNA polymerase fidelity," *The Journal of Biological Chemistry*, vol. 277, pp. 47393–8, Dec. 2002.

[93] Y.-C. Tsai and K. A. Johnson, "A new paradigm for DNA polymerase specificity," *Biochemistry*, vol. 45, pp. 9675–87, Aug. 2006.

[94] W. A. Beard and S. H. Wilson, "Structural Insights into the Origins of DNA Polymerase Fidelity," *Structure*, vol. 11, pp. 489–496, May 2003.

[95] Y. Santoso and A. N. Kapanidis, "Probing biomolecular structures and dynamics of single molecules using in-gel alternating-laser excitation," *Analytical Chemistry*, vol. 81, pp. 9561–70, Dec. 2009.

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

[97] L. Le Reste, J. Hohlbein, K. Gryte, and A. N. Kapanidis, "Characterization of dark quencher chromophores as nonfluorescent acceptors for single-molecule FRET," *Biophysical Journal*, vol. 102, pp. 2658–68, June 2012.

[98] H. S. Chung, J. M. Louis, and W. A. Eaton, "Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 11837–44, July 2009.

[99] M. F. Bailey, E. J. C. Van der Schans, and D. P. Millar, "Dimerization of the Klenow fragment of Escherichia coli DNA polymerase I is linked to its mode of DNA binding," *Biochemistry*, vol. 46, pp. 8085–99, July 2007.

[100] K.-H. Tang and M.-D. Tsai, "Structure and function of 2:1 DNA polymerase.DNA complexes," *Journal of Cellular Physiology*, vol. 216, pp. 315–20, Aug. 2008.

[101] K.-H. Tang, M. Niebuhr, A. Aulabaugh, and M.-D. Tsai, "Solution structures of 2 : 1 and 1 : 1 DNA polymerase-DNA complexes probed by ultracentrifugation and small-angle X-ray scattering," *Nucleic Acids Research*, vol. 36, pp. 849–60, Feb. 2008.

[102] T. D. Craggs and A. N. Kapanidis, "Six steps closer to FRET-driven structural biology," *Nature Methods*, vol. 9, pp. 1157–8, Dec. 2012.

[103] J. Vogelsang, R. Kasper, C. Steinhauer, B. Person, M. Heilemann, M. Sauer, and P. Tinnefeld, "A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes," *Angewandte Chemie (International ed. in English)*, vol. 47, no. 29, pp. 5465–9, 2008.

[104] T. Ha and P. Tinnefeld, "Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging," *Annual Review of Physical Chemistry*, vol. 63, pp. 595–617, Jan. 2012.

[105] J. Groll, E. V. Amirgoulova, T. Ameringer, C. D. Heyes, C. Röcker, G. U. Nienhaus, and M. Möller, "Biofunctionalized, ultrathin coatings of cross-linked star-shaped poly(ethylene oxide) allow reversible folding of immobilized proteins," *Journal of the American Chemical Society*, vol. 126, pp. 4234–9, Apr. 2004.

[106] J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, and C. H. Wiggins, "Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data," *Biophysical Journal*, vol. 97, no. 12, pp. 3196–3205, 2009.

[107] S. Uphoff, K. Gryte, G. Evans, and A. N. Kapanidis, "Improved temporal resolution and linked hidden Markov modeling for switchable single-molecule FRET," *Chemphyschem*, vol. 12, pp. 571–9, Feb. 2011.

[108] T. Lang, M. Maitra, D. Starcevic, S.-X. Li, and J. B. Sweasy, "A DNA polymerase beta mutant from colon cancer cells induces mutations," *Proceedings of the National Academy of Sciences*, vol. 101, pp. 6074–9, Apr. 2004.

[109] S. S. Lange, K.-i. Takata, and R. D. Wood, "DNA polymerases and cancer," *Nature Reviews. Cancer*, vol. 11, pp. 96–110, Feb. 2011.

[110] I. Wong, S. S. Patel, and K. A. Johnson, "An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics," *Biochemistry*, vol. 30, pp. 526–537, Jan. 1991.

[111] A. N. Kapanidis and T. Strick, "Biology, one molecule at a time," *Trends in Biochemical Sciences*, vol. 34, pp. 234–43, May 2009.

[112] P. J. Rothwell, W. J. Allen, E. Sisamakis, S. Kalinin, S. Felekyan, J. Widengren, G. Waksman, and C. A. M. Seidel, "dNTP-dependent conformational transitions in the fingers subdomain of Klentaq1 DNA polymerase: insights into the role of the "nucleotide-binding" state," *The Journal of Biological Chemistry*, vol. 288, pp. 13575–91, May 2013.

[113] K. Henzler-Wildman and D. Kern, "Dynamic personalities of proteins," *Nature*, vol. 450, pp. 964–72, Dec. 2007.

[114] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern, "Intrinsic dynamics of an enzyme underlies catalysis," *Nature*, vol. 438, pp. 117–21, Nov. 2005.

[115] B. T. Eger and S. J. Benkovic, "The minimal kinetic mechanism for misincorporation by DNA polymerase I (Klenow fragment)," *Biochemistry*, vol. 31, pp. 9227–9236, Sept. 1992.

[116] L. Pauling, "Chemical achievement and hope for the future," *American Scientist*, vol. 36, pp. 51–8, Jan. 1948.

[117] D. Hilvert, "Critical analysis of antibody catalysis," *Annual Review of Biochemistry*, vol. 69, pp. 751–93, Jan. 2000.

[118] D. D. Boehr, H. J. Dyson, and P. E. Wright, "An NMR perspective on enzyme dynamics," *Chemical Reviews*, vol. 106, pp. 3055–79, Aug. 2006.

[119] K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, and D. Kern, "A hierarchy of timescales in protein dynamics is linked to enzyme catalysis," *Nature*, vol. 450, pp. 913–6, Dec. 2007.

[120] J. H. Fong and A. R. Panchenko, "Intrinsic disorder and protein multibinding in domain, terminal, and linker regions," *Molecular BioSystems*, vol. 6, pp. 1821–8, Oct. 2010.

[121] S. Hay and N. S. Scrutton, "Good vibrations in enzyme-catalysed reactions," *Nature Chemistry*, vol. 4, pp. 161–8, Mar. 2012.

[122] D. R. Glowacki, J. N. Harvey, and A. J. Mulholland, "Protein dynamics and enzyme catalysis: the ghost in the machine?," *Biochemical Society Transactions*, vol. 40, pp. 515–21, June 2012.

[123] S. C. L. Kamerlin and A. Warshel, "At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis?," *Proteins*, vol. 78, pp. 1339–75, May 2010.

[124] A. V. Pisliakov, J. Cao, S. C. L. Kamerlin, and A. Warshel, "Enzyme millisecond conformational dynamics do not catalyze the chemical step," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 17359–64, Oct. 2009.

[125] M. Karplus, "Role of conformation transitions in adenylate kinase," *Proceedings of the National Academy of Sciences*, vol. 107, p. E71; author reply E72, Apr. 2010.

[126] S. C. L. Kamerlin and A. Warshel, "Reply to Karplus: Conformational dynamics have no role in the chemical step," *Proceedings of the National Academy of Sciences*, vol. 107, pp. E72–E72, Apr. 2010.

[127] R. Zhao and D. Rueda, "RNA folding dynamics by single-molecule fluorescence resonance energy transfer," *Methods (San Diego, Calif.)*, vol. 49, pp. 112–7, Oct. 2009.

[128] R. M. Turner, N. D. F. Grindley, and C. M. Joyce, "Interaction of DNA polymerase I (Klenow fragment) with the single-stranded template beyond the site of synthesis," *Biochemistry*, vol. 42, pp. 2373–85, Mar. 2003.

[129] K. A. Johnson, "Conformational coupling in DNA polymerase fidelity," *Annual Review of Biochemistry*, vol. 62, pp. 685–713, Jan. 1993.

[130] J. N. M. Pinkney, P. Zawadzki, J. Mazuryk, L. K. Arciszewska, D. J. Sherratt, and A. N. Kapanidis, "Capturing reaction paths and intermediates in Cre-loxP recombination using single-molecule fluorescence," *Proceedings of the National Academy of Sciences*, vol. 109, pp. 20871–6, Dec. 2012.

[131] V. Mekler, E. Kortkhonjia, J. Mukhopadhyay, J. Knight, A. Revyakin, A. N. Kapanidis, W. Niu, Y. W. Ebright, R. Levy, and R. H. Ebright, "Structural Organization of Bacterial RNA Polymerase Holoenzyme and the RNA Polymerase-Promoter Open Complex," *Cell*, vol. 108, pp. 599–614, Mar. 2002.