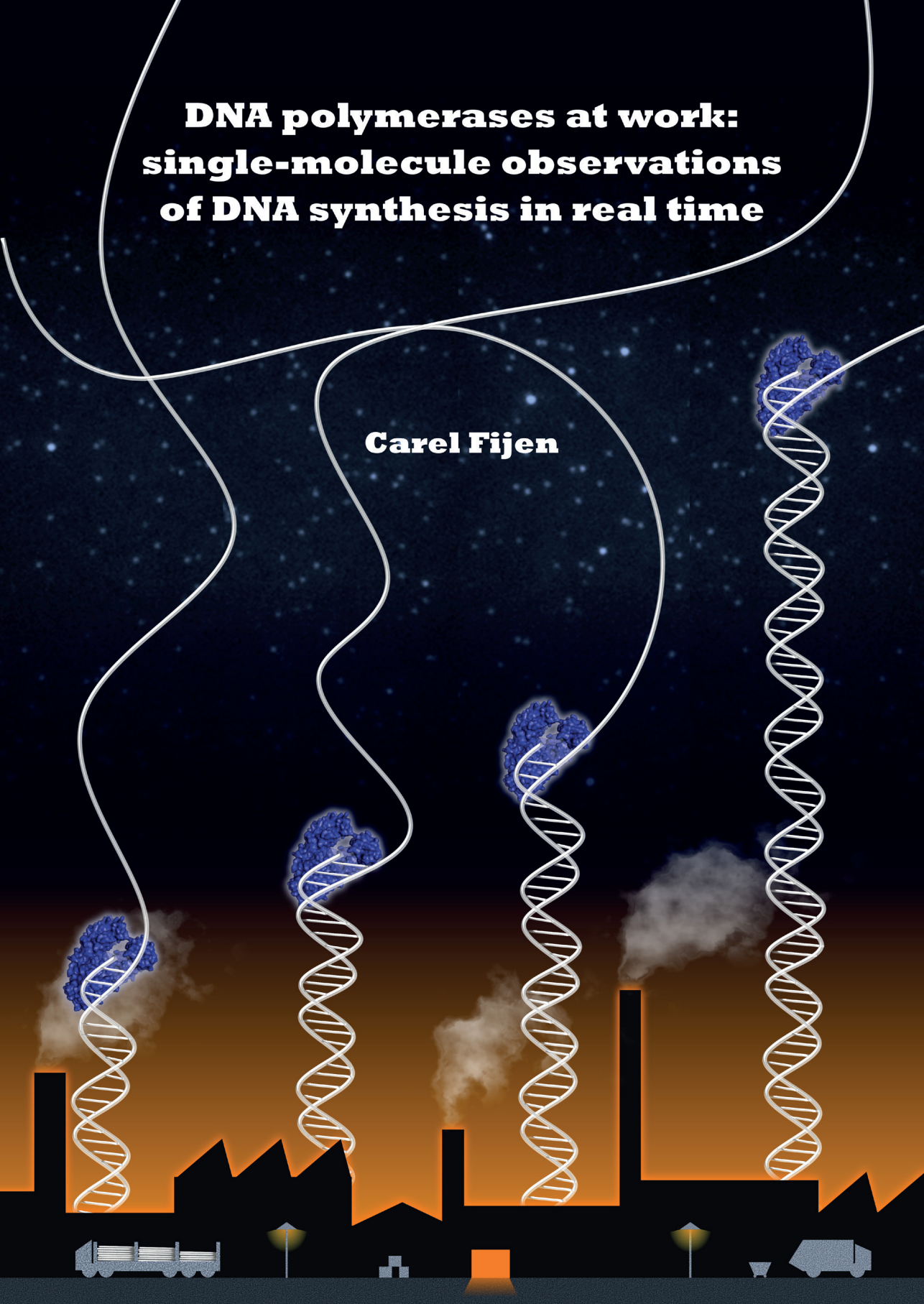


**DNA polymerases at work:  
single-molecule observations  
of DNA synthesis in real time**

**Carel Fijen**





# **DNA polymerases at work: single-molecule observations of DNA synthesis in real time**

**Carel Fijen**

## **Thesis committee**

### **Promotor**

Prof. Dr H. van Amerongen  
Professor of Biophysics  
Wageningen University & Research

### **Co-promotor**

Dr J.C. Hohlbein  
Assistant professor, Laboratory of Biophysics  
Wageningen University & Research

### **Other members**

Prof. Dr J. van der Gucht, Wageningen University & Research  
Prof. Dr M.C. Leake, University of York, UK  
Prof. Dr M. Tijsterman, Leiden University Medical Center  
Dr C.P.M. van Mierlo, Wageningen University & Research

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).



# **DNA polymerases at work: single-molecule observations of DNA synthesis in real time**

**Carel Fijen**

## **Thesis**

submitted in fulfilment of the requirements  
for the degree of doctor  
at Wageningen University  
by the authority of the Rector Magnificus,  
Prof. Dr A.P.J. Mol,  
in the presence of the  
Thesis Committee appointed by the Academic Board  
to be defended in public  
on Wednesday 21 March 2018  
at 1.30 p.m. in the Aula.

Carel Fijen

DNA polymerases at work: single-molecule observations of DNA synthesis in real time,  
174 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2018)  
With references, with summary in English

ISBN: 978-94-6343-743-1

DOI: 10.18174/431813

# Contents

1. General Introduction	<b>p. 7</b>
2. Studying DNA-protein interactions with single-molecule Förster resonance energy transfer	<b>p. 25</b>
3. A single-molecule FRET sensor for monitoring DNA synthesis in real time	<b>p. 59</b>
4. DNA polymerase $\beta$ fingers movement revealed by smFRET suggests a partially closed conformation as a fidelity checkpoint	<b>p. 97</b>
5. Simple nanofluidic devices for high-throughput, non-equilibrium studies at the single-molecule level	<b>p. 119</b>
6. General Discussion	<b>p. 149</b>
7. Summary	<b>p. 163</b>
8. About the author	<b>p. 167</b>
<i>Acknowledgments</i>	
<i>Curriculum Vitae</i>	
<i>Overview of completed training activities</i>	



# General Introduction



## **Background and aim**

Across the three domains of life, genetic information is preserved in DNA. Replication of this genetic code is essential for reproduction, while DNA repair is necessary to maintain genome integrity. In both of these processes, DNA polymerases have an important role: these enzymes synthesize new DNA and often do that with remarkable accuracy. Because of their importance, DNA polymerases have been studied extensively. Moreover, they are widely used as tools in molecular biology, for example in the polymerase chain reaction (PCR), which is used to amplify DNA samples by several orders of magnitude. However, a few fundamental processes are difficult to examine with standard biochemical techniques. Polymerase processivity and pausing as well as fidelity-related conformational changes within a DNA polymerase are hard to probe with an ensemble assay because of their stochastic nature. Therefore, we set out to explore DNA polymerases on a molecule-by-molecule basis.

In the past 20 years, single-molecule biophysics has provided numerous tools to study complex biological reactions and interactions. The main advantage of single-molecule studies is the reduction of averaging effects inherent to ensemble studies. Modern imaging techniques with single-molecule sensitivity can visualize millisecond time scale dynamics because they reduce temporal averaging. A wide array of single-molecule techniques is available, ranging from force-based techniques like magnetic tweezers or atomic force microscopy<sup>1</sup> to fluorescence-based techniques like confocal microscopy<sup>2</sup> and total internal reflection fluorescence (TIRF) microscopy<sup>3-5</sup>. A disadvantage of many of these techniques is low throughput: collecting information from many individual molecules is therefore often a time-consuming process. In diffusion-based confocal microscopy, for example, a typical throughput amounts to only ~100 data points per minute, requiring an hour-long experiment to collect sufficient data.

In this thesis, single-molecule interactions are studied using a combination of TIRF microscopy and Förster Resonance Energy Transfer (FRET). FRET describes a distance-dependent form of non-radiative energy transfer between two fluorophores, which allows to probe distances in the nanometre range<sup>6</sup>. This range is ideal for studying conformational changes of DNA polymerases or DNA itself. TIRF microscopy is a special case of wide-field microscopy that allows the detection of many fluorescent molecules in parallel with a generally high signal-to-noise ratio. This is possible because of a so-called evanescent field generated at the coverglass to water interface, in which only molecules close (~100 nm) to the coverslip surface are efficiently excited<sup>7</sup>. This effect can be exploited by immobilizing a species of interest on the surface, which then allows for detection of ~100 molecules in parallel, with

observation times of minutes (typically until fluorescence disappears, e.g. because of bleaching). A typical acquisition time of 0.1 s then already yields 100 FRET data points. While this parallel setup increases throughput considerably, surface-induced effects may interfere with the measurement and the molecule of interest needs a tag for attachment. Also, the temporal resolution is lower than in confocal microscopy. Therefore, the challenge remains to combine the throughput and long observation times of TIRF microscopy with using non-immobilized molecules at high temporal resolution.

This thesis describes the development of a fluorescent DNA sensor to monitor polymerization in real time without having to use fluorescently labelled polymerases. Using this sensor, we have learned more about (variations in) polymerase speed and processivity. For one specific polymerase, human DNA Polymerase Beta, we later studied a fidelity-related conformational change and were able to establish a “fidelity checkpoint”. Finally, in order to increase throughput in TIRF microscopy, we combine our DNA polymerase work with the introduction of novel nanofluidic devices that can be used to trigger reactions at the single-molecule level.

## **DNA and DNA polymerases**

In all organisms and certain groups of viruses, DNA is the carrier of the genetic code. DNA is a polymer consisting of a sugar and phosphate backbone to which different bases are coupled in a sequential manner. The exact sequence of these bases forms the genetic code. The four canonical bases in DNA are adenine (A), thymine (T), cytosine (C) and guanine (G). Bases A and T are complementary and form two hydrogen bridges; C and G can form three. Because of this sequence-dependent complementarity, two strands of DNA can pair together and form a stable secondary structure. Double-stranded DNA (dsDNA) adopts a right-handed helical structure<sup>8</sup> and is therefore far more rigid than the individual single strands (ssDNA).

DNA polymerases are tasked with synthesizing new DNA by complementing a templating single stranded DNA molecule. Although there is a large variety in their structures and functionalities, most DNA polymerases adopt a general mechanism of action: they catalyse the addition of nucleotides to a growing DNA strand (the primer), based on the sequence dictated by a complementary strand (the template). The structure of some well-studied polymerases resembles a human hand, and the nomenclature for the different subdomains of the protein keeps up this metaphor: DNA binds to the “palm” subdomain, and the “fingers” subdomain closes to transfer a nucleotide to the active site. Because cells contain at least four different deoxyribonucleotides and further (in larger quantities) ribonucleotides, being able to distinguish the “correct” complementary nucleotide from the majority of incorrect ones, is critical to the function of polymerases. This check is thought to occur during



closing of the fingers, but before incorporation, when nucleotides can still be rejected. Interestingly, a crystal structure of bacterial DNA Polymerase I indicated the existence of a partially closed, “ajar” conformation of the fingers, which is linked to fidelity<sup>9</sup>. Additionally, several polymerases contain a so-called “proof-reading (or exonuclease) domain”, which allows the removal of wrong (i.e. non-complementary or damaged) bases that were accidentally incorporated.

The DNA polymerases studied in this thesis are the Klenow Fragment of *E. coli* DNA Polymerase I (KF), human DNA Polymerase Beta (POLB), and the alpha subunit of *E. coli* Polymerase III (POLIII $\alpha$ ). Despite being a cleavage product<sup>10</sup>, KF maintains 5'→3' polymerase activity and 3'→5' exonuclease activity<sup>11</sup>. In this thesis, we used a mutant version of KF in which the aspartic acid at position 424 has been replaced by alanine (denoted D424A); this version does not possess proof-reading activity. The natural role of full-length polymerase I is synthesis of short stretches of DNA: it connects Okazaki fragments during DNA replication and repairs damaged DNA. KF has been studied at the single-molecule level before, providing us with helpful mechanistic references. A recent interesting finding from single-molecule experiments is that KF exhibits the partially closed fingers conformation when supplied with non-complementary nucleotides<sup>12,13</sup>.

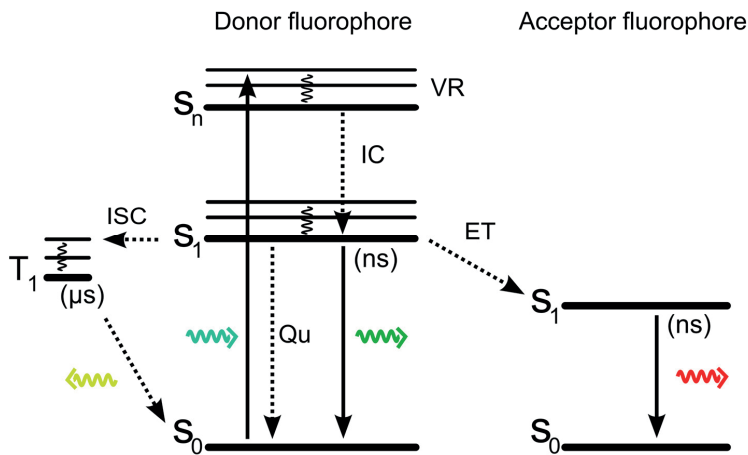
POLB is involved in the eukaryotic Base Excision Repair (BER) pathway: it is a small enzyme (39 kDa) that fills gaps in the DNA that originate from removal of damaged bases<sup>14</sup>. Additional to 5'→3' polymerase activity, it contains a lyase domain, which binds to 5' ends and is therefore involved in recognition of gaps in the DNA. Several crystal structures of PolB bound to DNA are available. In the binary and ternary complexes, the DNA appears to be strongly bent, exposing the template base in the gap<sup>15</sup>. Like KF, POLB possesses a fingers domain which closes for nucleotide incorporation. Recent studies indicate that fingers movement is important for fidelity<sup>16</sup>, as also suggested by a crystal structure of POLB with an incorrect modified nucleotide<sup>17</sup>. Inspired by the previous work on KF, we set out to provide new insights on the existence of a partially closed fingers conformation in POLB.

POLIII is a large complex consisting of multiple subunits, with a function in DNA replication, rather than in DNA repair<sup>18</sup>. The full complex achieves a remarkable speed (1000 nt/s)<sup>19–21</sup> and fidelity ( $\sim 1 : 1.000.000$ )<sup>22</sup>. The  $\alpha$  subunit possesses polymerase activity, but without the rest of the complex, it has lower processivity<sup>23</sup> and is much slower<sup>24</sup>. We used the  $\alpha$  subunit to compare its polymerization properties with those of KF and POLB in the assay that we developed to test for polymerase processivity and speed.

A more elaborate description of DNA polymerases (and other DNA processing enzymes) can be found in chapter 2, in which previous single-molecule work on these enzymes is discussed.

## Fluorescence

When a fluorescent molecule in its ground state ( $S_0$ ) absorbs light, it is excited to a higher energy level ( $S_n$ ). The lifetime of this excited state is generally in the 1-10 ns range. De-excitation to the ground state of the molecule usually occurs in several steps: first, the molecule quickly loses energy by relaxing down to the lowest vibrational level of the excited state, which is a non-radiative process. The energy gap between this level and the ground state, however, is often quite large. Further relaxation may then involve the release of a photon, which is called “fluorescence”. Because the molecule already lost part of its energy before emitting a photon, the wavelength of the emitted light is always red-shifted. Relaxation can also occur through (a series of) non-radiative internal conversions, transition to a triplet state or energy transfer to a neighbouring molecule (for example through FRET, as explained later). All these processes can be summarized in a Jablonski diagram, which schematically shows the different energy levels of a molecule and the possible transitions<sup>25</sup> (Fig. 1.1).



**Figure 1.1.** Jablonski diagram. A fluorophore enters an excited state ( $S_n$ ) upon excitation with light of a specific wavelength. The molecule relaxes down to the lowest excited state level  $S_1$  through a combination of vibrational relaxation (VR) and internal conversion (IC). These processes are non-radiative. De-excitation from  $S_1$  can occur through quenching (Qu), energy transfer to a neighbouring fluorophore (ET), emission of a photon (fluorescence) or inter-

system crossing (ISC) to a triplet state ( $T_n$ ). De-excitation from a triplet state can occur through internal conversions or the emission of a photon (phosphorescence).

## TIRF microscopy

Light travelling through a medium with high refractive index (medium 1) will change direction when it reaches a medium with lower refractive index (medium 2) according to Snell's Law:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} \quad (\text{eq. 1.1})$$

in which  $\theta_1$  is the angle of incident light,  $\theta_2$  is the angle of refracted light (both measured from the normal) and  $n_1$  and  $n_2$  are the refractive indexes of the two mediums. Snell's Law prohibits certain angles of incidence: the largest possible angle for which  $\theta_2$  can still be calculated depends on the ratio  $n_2/n_1$  (and therefore on the materials) and is called the "critical angle". Light approaching the interface at an angle larger than the critical angle will not propagate into medium 2, but will be reflected. Under these conditions, a thin ( $\sim 100$  nm) evanescent field will form at the interface inside medium 2. The intensity of this evanescent field is highest near the surface and decreases exponentially farther into the solution<sup>7</sup>. TIRF microscopy makes use of this evanescent field to specifically excite molecules close to the surface. Fluorescent molecules deeper in the solution are not excited, which reduces background noise when compared to wide-field microscopy.

The TIRF setup used for the experiments in this thesis is a custom-built setup (for a schematic drawing, see Fig. 1.2A). It contains a fibre-coupled laser engine (LightHUB series, Omicron, Germany) equipped with four lasers (405 nm, 473 nm, 561 nm and 638 nm). We primarily used the green (561 nm) and red (638 nm) lasers to excite donor and acceptor fluorophores. Lasers are triggered with a home-written LabVIEW program. Divergent light from the fibre output is collimated by a first lens ( $f = 30, 50$  or  $100$  mm, depending on the application) and focussed by a second lens ( $f = 200$  mm) into the back focal plane of the objective (100x oil immersion,  $NA = 1.49$ , Nikon, Japan). This second lens is used to control the TIRF angle by shifting the beam in relation to the back focal plane. A polychroic mirror (Chroma ZT405/473/561/640rpc) and a bandpass filter (Chroma ZET405/473/561/640m) prevent excitation laser light from entering the emission path. A tube lens ( $f = 200$  mm) focuses the emission light in the aperture, which reduces the field of view to a rectangle. After a second collimating lens ( $f = 200$  mm), the emission is split into three different spectral regions ("channels") by dichroic mirrors. Another lens then focuses all three channels on different regions of the emCCD camera sensor (Andor iXon 897 Ultra, UK). The position of the sample is controlled with a motorized  $x, y$  scanning stage combined with a  $z$ -piezo (ASI, USA).

## Förster Resonance Energy Transfer

Förster resonance energy transfer (FRET) describes the non-radiative energy transfer between an excited donor fluorophore and an acceptor chromophore<sup>6</sup> (in this thesis, we use small, organic fluorophores). The typical distance range in which this phenomenon can occur is 1-10 nm. The basis for energy transfer lies in the resonance of the dipole moments of the molecules involved: the relative orientation of the dipoles and the distance between them are important factors influencing the efficiency of energy transfer. The relative orientation of the dipoles is quantified by orientation factor  $\kappa^2$ , which can vary between 0 and 4: the minimum value of 0 can be reached with a number of configurations, amongst others when the dipoles are perpendicular; the maximum value of 4 is reached when the dipoles are parallel and in line with each other<sup>26</sup>. In many cases, however, it is assumed that free rotation of the dyes causes  $\kappa^2$  to average out to 2/3. Moreover, for FRET to occur, a partial overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor is required ( $J$ ). The distance at which energy transfer is 50% ( $R_0$ ) is furthermore dependent on the quantum yield of the donor dye  $\Phi_D$  and the refractive index of the medium  $n$ <sup>6,27,28</sup>:

$$R_0^6 = 8.8 \cdot 10^{-28} \frac{\Phi_D \kappa^2}{n^4} J \quad (\text{eq. 1.2})$$

Efficiency of energy transfer is then solely dependent on the distance  $R$  between donor and acceptor: when the fluorophores are farther apart, less energy can be transferred. This FRET efficiency  $E$  can be calculated as follows:

$$E = \frac{R_0^6}{R_0^6 + R^6} \quad (\text{eq. 1.3})$$

There are two common ways to determine  $E$  experimentally. The first method uses the fluorescence lifetime of the donor fluorophore in the absence of acceptor ( $\tau_D$ ) and compares it to the lifetime in the presence of an acceptor ( $\tau_T$ ). The higher the FRET efficiency, the more the lifetime of the donor will be shortened:

$$E = 1 - \frac{\tau_T}{\tau_D} \quad (\text{eq. 1.4})$$

The second method uses emission intensities collected from the donor and the acceptor of single molecules. We will denote these emissions as  $DD$  (donor emission upon donor excitation) and  $DA$  (acceptor emission upon donor excitation). The FRET efficiency of each molecule at each time point during the experiment is then calculated as follows:

$$E^* = \frac{f_{Dex}^{Aem}}{(f_{Dex}^{Dem} + f_{Dex}^{Aem})} = \frac{DA}{DD + DA} \quad (\text{eq. 1.5})$$

When calculated in this manner, FRET efficiency is denoted  $E^*$  (with an asterisk). This raw FRET efficiency needs to be corrected for several factors to obtain accurate  $E^{29,30}$ :

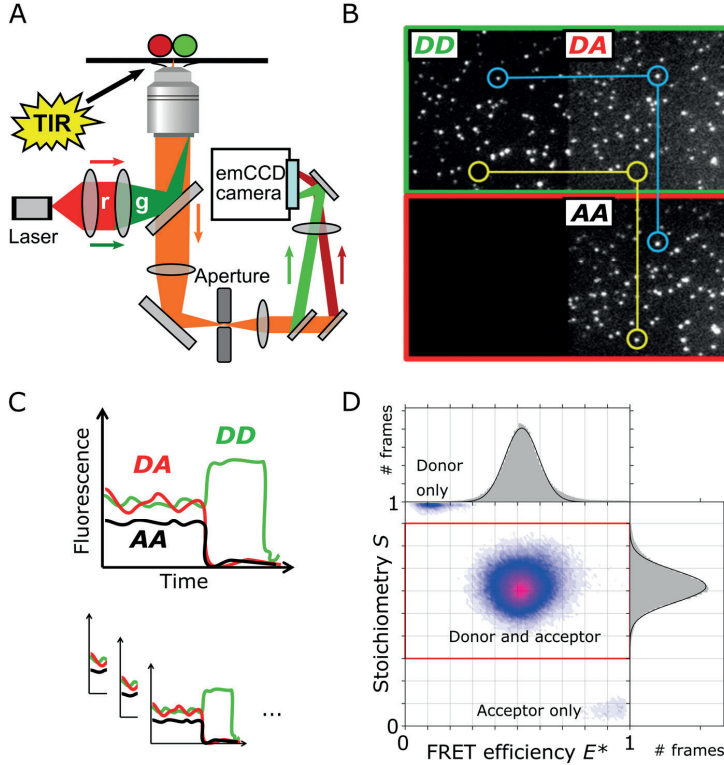
- 1) **Leakage.** An experimental setup demands a sharp cut-off to separate the emissions, but fluorophores emit light in a wide spectral range. Therefore, donor emission can leak into the acceptor detection channel, and is then wrongly identified as acceptor emission.
- 2) **Direct excitation.** The acceptor can be excited at the same wavelength with which the donor is excited. The emission detected from the acceptor is then a sum of FRET and emission caused by direct excitation.
- 3) **Gamma ( $\gamma$ ).** The quantum yields of donor and acceptor are often not the same; neither is the efficiency with which the camera sensor can detect the two spectral regions.  $\gamma$  adjusts the donor emission such that it matches the acceptor emission in terms of quantum yield and detection efficiency.

Access to these three factors is gained by using Alternating Laser Excitation (ALEX)<sup>30,31</sup>. In ALEX, donor and acceptor are excited in an alternating fashion, such that every camera frame in which the donor is excited, is followed by a frame in which the acceptor is directly excited. This introduces AA (acceptor emission upon acceptor excitation, Fig. 1.2B+C). AA will prove useful later to correct  $E^*$  for the effects mentioned above. Using AA, we can calculate stoichiometry ratio  $S$ :

$$S^{\text{raw}} = \frac{(f_{Dexc}^{Aem} + f_{Dexc}^{Dem})}{(f_{Dexc}^{Aem} + f_{Dexc}^{Dem} + f_{Aexc}^{Aem})} = \frac{DA + DD}{DA + DD + AA} \quad (\text{eq. 1.6})$$

Although ALEX halves the effective temporal resolution, it provides useful information on the state of the fluorophores. Using  $S$ , fluorescent species carrying only a donor or acceptor (because of bleaching, or imperfect labelling) can easily be separated from species carrying both. In case a molecule only carries a donor, AA will equal  $\sim 0$  and  $S$  will be  $\sim 1$ . On the other hand, if only an acceptor is present,  $S$  will be  $\sim 0$ . Species with an active donor and acceptor will have a stoichiometry around 0.5. The different populations of molecules can easily be visualized when all single-molecule data is compiled in a two-dimensional histogram. In this so-called  $E^*/S$  histogram, FRET efficiency is plotted along the  $x$ -axis and stoichiometry is plotted along the  $y$ -axis (Fig. 1.2D). Since we are mostly interested in the FRET efficiency of species with both a donor and acceptor, we choose that population to be projected in the one-dimensional histograms on the  $x$  and  $y$  axes (the species within the red box).

A closer look at this  $E^*/S$  histogram reveals that the donor-only population has a FRET efficiency  $> 0$ , and the acceptor-only population has a stoichiometry  $> 0$ . This is the result of contributions from leakage and direct excitation. In this thesis, the conversion from raw FRET efficiencies  $E^*$  to fully corrected FRET efficiency  $E$  (and from there to distances) has been applied several times and is therefore discussed in more detail below.



**Figure 1.2.** Schematic overview of the TIRF microscope used for the work in this thesis (A). A green and red laser create an evanescent field in the area of total internal reflection (TIR). This field is used to specifically excite donor and acceptor fluorophores close to the surface of a glass coverslip. Emission (orange) is collected by the objective and split into green and red spectral regions (channels). These channels are projected side-by-side on the sensor of an emCCD camera. The camera read-out facilitates extraction of DD, DA and AA emission intensities from single molecules (B). The green and red box indicate the field of view upon green and red excitation, respectively. Immobilized particles can easily be identified and linked throughout subsequent frames, yielding emission time traces (C) from which FRET efficiency  $E^*$  and stoichiometry  $S$  can be calculated. These can be binned and plotted in a 2D histogram (D).

## FRET as a molecular ruler: converting FRET efficiencies to distances

To illustrate how the corrections described above are applied in later chapters, we will go through the procedure step by step. This discussion is based on a review by Hohlbein, Craggs and Cordes<sup>30</sup> and data from this thesis.

We start with a so-called  $E^*/S$  histogram (Fig. 1.3, left). This two-dimensional histogram shows FRET efficiency and stoichiometry grouped in  $100 \times 100$  bins, and is compiled from single-molecule data. The histogram shown here was built with data from 570 immobilized DNA sensors after polymerization by KF (see chapter 3, Fig. 3.2B and Table S3.1). The polymerized sensor contains a FRET donor and acceptor and exhibits a low FRET efficiency ( $E^* = 0.35$ , the population in the middle). Two more populations are visible: a small population in the top right corner (low  $E^*$ , high  $S$ ) and another in the bottom left corner (high  $E^*$ , low  $S$ ). The population with low FRET efficiency and high stoichiometry consists of molecules that have an active donor, but no acceptor. Since acceptors are absent, the FRET efficiency of this population should be 0. In practice,  $E_{D-only} > 0$  because of leakage. The population with high FRET efficiency and low stoichiometry contains molecules with only an acceptor. It should be centered around  $S = 0$ . The offset that is visible in the uncorrected histogram is caused by direct excitation of the acceptor, resulting in  $S_{A-only} > 0$ . The number of photons in the  $DA$  channel thus originates from three different sources: FRET, leakage and direct excitation. We need a method to quantify the latter two in order to calculate the actual number of photons due to FRET. To do this, we will make use of the donor-only and acceptor-only populations. Leakage is linearly proportional to the number of photons in the  $DD$  channel with leakage coefficient  $l$ . We can therefore state:

$$Leakage = l \cdot DD \quad (\text{eq. 1.7})$$

Replacing  $DA$  in equation 1.5 with this description of leakage then gives  $E = l / (l + 1)$ . Importantly,  $E$  is here the FRET efficiency of the donor-only population ( $E_{D-only}$ ). We can now use  $E_{D-only}$  to calculate  $l$ :

$$l = E_{D-only} / (1 - E_{D-only}) \quad (\text{eq. 1.8})$$

In our example,  $E_{D-only}$  is found to be 0.12 (Fig. 1.3, left). This yields a leakage coefficient of  $l=0.14$ .

A similar approach can be used to quantify direct excitation. Direct excitation is linearly proportional to the number of photons in the  $AA$  channel with direct excitation coefficient  $d$ :



$$\text{Direct excitation} = d \cdot AA \quad (\text{eq. 1.9})$$

The stoichiometry of the acceptor-only population can then be calculated using  $S_{A\text{-only}} = d / (d + 1)$ , which can be rewritten as:

$$d = S_{A\text{-only}} / (1 - S_{A\text{-only}}) \quad (\text{eq. 1.10})$$

From our example data, we obtain  $S_{A\text{-only}} = 0.08$ .

Every data point in our  $E^*/S$  histogram can now be corrected for contributions of leakage and direct excitation, since the number of photons due to actual FRET can be calculated as follows:

$$FRET = DA - (l \cdot DD) - (d \cdot AA) \quad (\text{eq. 1.11})$$

Replacing  $DA$  in equation 1.5 with this new description of FRET results in:

$$E_{PR} = FRET / (DD + FRET) \quad (\text{eq. 1.12})$$

After correcting  $S$  in a similar fashion, we can compile a new, cross-talk corrected histogram (Fig. 1.3, middle). For clarity, we call the cross-talk corrected FRET efficiency  $E_{PR}$ . Note that the donor-only population is now centered around  $E_{PR} = 0$ , and the acceptor-only population around  $S = 0$ . The position of the main population in the middle of the histogram (species bearing both donor and acceptor) has also shifted.

The last correction is for  $\gamma$ , which accounts for the differences between the quantum yields  $\Phi$  and detection efficiencies  $\eta$  of the donor and acceptor. It is defined as:

$$\gamma = \frac{(\Phi \cdot \eta)_{\text{acceptor}}}{(\Phi \cdot \eta)_{\text{donor}}} \quad (\text{eq. 1.13})$$

If donor and acceptor have equal quantum yield and are equally well detected,  $\gamma = 1$ , and the stoichiometry of populations at low and high FRET efficiency will be the same. In practice, however, this is rarely the case, and species with different FRET efficiencies also have different stoichiometries. Therefore, it follows intuitively that it must be possible to calculate  $\gamma$  from a collection of populations with different FRET efficiencies. In practice,  $\gamma$  is calculated from a plot in which  $1/S$  of multiple FRET populations is plotted against  $E_{PR}$ . Fitting a line through the peak positions yields intercept  $\Omega$  and slope  $\Sigma$ , after which  $\gamma$  can be calculated:

$$\gamma = (\Omega - 1) / (\Omega + \Sigma - 1) \quad (\text{eq. 1.14})$$

In our example, we have 3 more comparable FRET populations. A fit through the 4 peak positions yields  $\gamma = 0.72$  (Fig. 1.3, middle). Please note that  $\gamma$  relies in part on the quantum efficiencies of the dyes, and can vary with pH and temperature. We therefore recalculate  $\gamma$  for every set of measurements. Next, we calculate the fully corrected FRET efficiency using the cross-talk corrected FRET (equation 1.11) and  $\gamma$ :

$$E = FRET / (\gamma \cdot DD + FRET) \quad (\text{eq. 1.15})$$

We correct  $S$  in a similar way. Interestingly, the corrected stoichiometry is independent of  $E$ , and only depends on a factor called  $\beta$  that describes the combined effect of the excitation power ratio (red / green) and the absorption cross sections of acceptor and donor (for a more detailed explanation, please refer to Lee et al.<sup>29</sup>). Importantly, correction for  $\beta$  brings  $S$  to 0.5 (since the excitations are now normalized), without changing  $E$ . We thus obtain the final  $E/S$  histogram. Inter-dye distance  $R$  is now simply calculated using  $E$  and the Förster radius  $R_0$ :

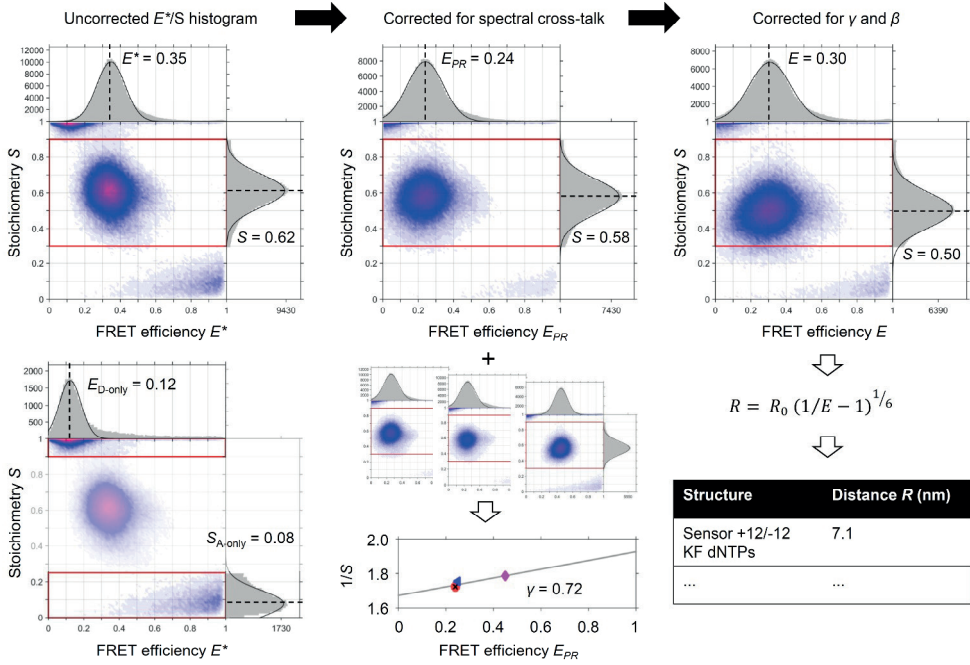
$$R = R_0(1/E - 1)^{1/6} \quad (\text{eq. 1.16})$$

Recently, we participated in a worldwide benchmark study that aimed to compare the experimental results among different labs, and standardize the correction procedure for accurate FRET. To this aim, every participating lab received the same labelled DNA constructs, and was instructed how to measure the FRET efficiency and apply the corrections. Our results were shown to be very similar to TIRF-based results from other labs<sup>32</sup>, showing the robustness of this correction procedure.

## Nanofluidics

Nanofluidics studies the flow behaviour of liquids inside nanometre-sized structures. Integrating nanofluidics in single-molecule studies offers exciting opportunities to study non-immobilized fluorescent molecules. Until recently, nano- and microfluidic devices for studying single molecules were primarily designed for implementation with confocal microscopy. With these devices, it became possible to study mixing at the single-molecule level, although not in a highly parallel fashion and with a dead time after mixing in the order of milliseconds<sup>33,34</sup>. A nanofluidic platform for TIRF microscopy was first introduced in 2014 by Tyagi *et al.*<sup>35</sup>. We developed our own nanofluidic mixing devices in an effort to increase throughput of single-molecule fluorescence assays.

Flow in nanofluidic devices is governed by the hydraulic resistance  $R_{hyd}$ , which in turn is dependent on the viscosity of the liquid and the dimensions of the (nano)channel according to<sup>36,37</sup>:



**Figure 1.3.** Procedure of accurate FRET. First, uncorrected histograms give access to  $E_{D\text{-only}}$  and  $S_{A\text{-only}}$ , which are needed to calculate the contributions of leakage and direct excitation. In the next step, multiple cross-talk corrected FRET populations (in this case found in a total of 4 histograms) give access to  $\gamma$ . Finally, after correcting the FRET efficiency for  $\gamma$ , inter-dye distances can be calculated using  $E$  and  $R_0$ .

$$R_{\text{hyd}} \approx \frac{12\eta L}{1 - 0.630\left(\frac{h}{w}\right)} \cdot \frac{1}{h^3 w} \text{ for } h < w. \quad (\text{eq. 1.17})$$

in which  $\eta$  is the dynamic viscosity,  $L$  is the length,  $h$  is the height and  $w$  is the width of the channel. The smaller the dimensions, the higher the resistance is. To decrease the pressure that is necessary to establish flow in a nanochannel, a channel with larger dimensions can be included in a parallel position. This principle is known as “parallel flow control”<sup>38,39</sup>. The flow rate (volume per time) in each channel scales with the inverse of the hydraulic resistance: the higher the resistance of a channel, the lower the flow through that channel. It thus follows that very low nanochannel flow rates can be established with parallel flow control. We used this insight in the development of our own nanofluidic devices.

## This thesis

The structure of this thesis is as follows:

**Chapter 2** contains an extensive overview of single-molecule studies on DNA polymerases and other DNA-binding proteins. We further introduce FRET-based structural modelling and discuss potential applications.

In **chapter 3**, a DNA FRET sensor to probe polymerase binding, speed and pausing at the single-molecule level is presented. The advantage of this assay is that polymerases do not have to be labelled with a fluorophore. We illustrate the possibilities of this sensor by testing three different polymerases with different processivities. Furthermore, we show that we can pause polymerization at a specific position on the template.

In **chapter 4**, the fingers-closing conformational change of human DNA Polymerase Beta is studied. By labelling the DNA substrate and the polymerase fingers domain with a FRET dye pair, we were able to visualize fingers movement in single polymerase-DNA complexes in real time. We show that the fingers domain of the polymerase closes only partially if the polymerase is supplied with the wrong nucleotide, providing evidence for a “fidelity checkpoint” in this polymerase.

**Chapter 5** presents a new approach to single-molecule detection and screening with the introduction of two nanofluidic devices. The first device is designed for high-throughput equilibrium studies of non-immobilized molecules, while the second device allows mixing of fluorescent molecules without any dead time in detection. We use these channels to study DNA hairpin conformational dynamics and DNA polymerization.

**Chapter 6** discusses a broader implication of the work in this thesis and provides a perspective for future single-molecule research.

**Chapter 7** is a summary of this thesis.

## References

1. Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5**, 491–505 (2008).
2. Deniz, A. A. *et al.* Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci.* **96**, 3670–3675 (1999).
3. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507–516 (2008).

4. Holden, S. J. *et al.* Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophys. J.* **99**, 3102–3111 (2010).
5. Reck-Peterson, S. L., Derr, N. D. & Stuurman, N. Imaging Single Molecules Using Total Internal Reflection Fluorescence Microscopy (TIRFM). *Cold Spring Harb. Protoc.* **2010**, pdb.top73 (2010).
6. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55–75 (1948).
7. Axelrod, D., Burghardt, T. P. & Thompson, N. L. Total Internal Reflection Fluorescence. *Annu. Rev. Biophys. Bioeng.* **13**, 247–268 (1984).
8. Watson, J. D. & Crick, F. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
9. Wu, E. Y. & Beese, L. S. The Structure of a High Fidelity DNA Polymerase Bound to a Mismatched Nucleotide Reveals an ‘Ajar’ Intermediate Conformation in the Nucleotide Selection Mechanism. *J. Biol. Chem.* **286**, 19758–19767 (2011).
10. Klenow, H. & Henningsen, I. Selective Elimination of the Exonuclease Activity of the Deoxyribonucleic Acid Polymerase from *Escherichia coli* B by Limited Proteolysis\*. *Proc. Natl. Acad. Sci. U. S. A.* **65**, 168–175 (1970).
11. Kornberg, A. & Baker, T. *DNA Replication*. (W.H. Freeman, 1992).
12. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nat. Commun.* **4**, (2013).
13. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res.* gkv547 (2015). doi:10.1093/nar/gkv547
14. Beard, W. A. & Wilson, S. H. Structure and Mechanism of DNA Polymerase  $\beta$ . *Biochemistry* **53**, 2768–2780 (2014).
15. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J. & Pelletier, H. Crystal Structures of Human DNA Polymerase  $\beta$  Complexed with Gapped and Nicked DNA: Evidence for an Induced Fit Mechanism. *Biochemistry* **36**, 11205–11215 (1997).
16. Mahmoud, M. M. *et al.* Defective Nucleotide Release by DNA Polymerase  $\beta$  Mutator Variant E288K Is the Basis of Its Low Fidelity. *Biochemistry* (2017). doi:10.1021/acs.biochem.7b00869
17. Batra, V. K., Beard, W. A., Shock, D. D., Pedersen, L. C. & Wilson, S. H. Structures of DNA Polymerase  $\beta$  with Active-Site Mismatches Suggest a Transient Abasic Site Intermediate during Misincorporation. *Mol. Cell* **30**, 315–324 (2008).
18. Kelman, Z. & O'Donnell, M. DNA Polymerase III Holoenzyme: Structure and Function of a Chromosomal Replicating Machine. *Annu. Rev. Biochem.* **64**, 171–200 (1995).

19. Tanner, N. A. *et al.* Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat. Struct. Mol. Biol.* **15**, nsmb.1381 (2008).
20. Tanner, N. A. *et al.* Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Res.* **37**, e27–e27 (2009).
21. Pham, T. M. *et al.* A single-molecule approach to DNA replication in *Escherichia coli* cells demonstrated that DNA polymerase III is a major determinant of fork speed. *Mol. Microbiol.* **90**, 584–596 (2013).
22. Bloom, L. B. *et al.* Fidelity of *Escherichia coli* DNA Polymerase III Holoenzyme THE EFFECTS OF  $\beta$ ,  $\gamma$  COMPLEX PROCESSIVITY PROTEINS AND  $\epsilon$  PROOFREADING EXONUCLEASE ON NUCLEOTIDE MISINCORPORATION EFFICIENCIES. *J. Biol. Chem.* **272**, 27919–27930 (1997).
23. Fay, P. J., Johanson, K. O., McHenry, C. S. & Bambara, R. A. Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of *Escherichia coli*. *J. Biol. Chem.* **256**, 976–983 (1981).
24. Maki, H. & Kornberg, A. The polymerase subunit of DNA polymerase III of *Escherichia coli*. II. Purification of the alpha subunit, devoid of nuclease activities. *J. Biol. Chem.* **260**, 12987–12992 (1985).
25. Jabłoński, A. Über den Mechanismus der Photolumineszenz von Farbstoffphosphoren. *Z. Für Phys.* **94**, 38–46 (1935).
26. van der Meer, B. W. Kappa-squared: from nuisance to new sense. *Rev. Mol. Biotechnol.* **82**, 181–196 (2002).
27. Stryer, L. Fluorescence Energy Transfer as a Spectroscopic Ruler. *Annu. Rev. Biochem.* **47**, 819–846 (1978).
28. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*.
29. Lee, N. K. *et al.* Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys. J.* **88**, 2939–2953 (2005).
30. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
31. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).
32. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements - a worldwide benchmark study. *ArXiv171003807 Phys. Q-Bio* (2017).
33. Wunderlich, B. *et al.* Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat. Protoc.* **8**, 1459–1474 (2013).
34. Dingfelder, F. *et al.* Rapid Microfluidic Double-Jump Mixing Device for Single-Molecule Spectroscopy. *J. Am. Chem. Soc.* **139**, 6062–6065 (2017).

35. Tyagi, S. *et al.* Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nat. Methods* **11**, 297–300 (2014).
36. Mathwig, K. & Lemay, S. G. Mass transport in electrochemical nanogap sensors. *Electrochimica Acta* **112**, 943–949 (2013).
37. Bruus, H. Chapter 1 Governing Equations in Microfluidics. in *Microscale Acoustofluidics* 1–28 (Royal Society of Chemistry, 2014).
38. Liang, H., Nam, W. J. & Fonash, S. J. A Novel Parallel Flow Control (PFC) System for Syringe-Driven Nanofluidics. in 281–283 (2008).
39. Mathwig, K., Mampallil, D., Kang, S. & Lemay, S. G. Electrical Cross-Correlation Spectroscopy: Measuring Picoliter-per-Minute Flows in Nanochannels. *Phys. Rev. Lett.* **109**, 118302 (2012).





# Studying DNA–protein interactions with single-molecule Förster resonance energy transfer

Shazia Farooq\*, Carel Fijen\*, Johannes Hohlbein, Protoplasma 2014, 251, pp 317-332.  
DOI:10.1007/s00709-013-0596-6 (\*equal contributions)



## Abstract

Single-molecule Förster resonance energy transfer (smFRET) has emerged as a powerful tool for elucidating biological structure and mechanisms on the molecular level. Here, we focus on applications of smFRET to study interactions between DNA and enzymes such as DNA and RNA polymerases. SmFRET, used as a nanoscopic ruler, allows for the detection and precise characterisation of dynamic and rarely occurring events, which are otherwise averaged out in ensemble-based experiments. In this review, we will highlight some recent developments that provide new means of studying complex biological systems either by combining smFRET with force-based techniques or by using data obtained from smFRET experiments as constraints for computer-aided modelling.

## Abbreviations

<i>AFM</i>	Atomic force microscopy
<i>ALEX</i>	Alternating-laser excitation
<i>ATP</i>	Adenosine tri-phosphate
<i>BVA</i>	Burst variance analysis
<i>Bp</i>	Base pair
<i>CAP</i>	Catabolite activator protein
<i>DNA</i>	Deoxyribonucleic acid
<i>dsDNA</i>	Double-stranded deoxyribonucleic acid
<i>dNTP</i>	Deoxyribonucleoside tri-phosphate
<i>FRET</i>	Förster resonance energy transfer
<i>FPS</i>	FRET-restrained positioning and screening
<i>MFD</i>	Multiparameter fluorescence detection
<i>NMR</i>	Nuclear magnetic resonance
<i>NPS</i>	Nano-positioning system
<i>PDA</i>	Probability distribution analysis
<i>PIFE</i>	Protein-induced fluorescence enhancement
<i>Pol</i>	Polymerase
<i>quFRET</i>	Quenchable Förster resonance energy transfer
<i>RNA</i>	Ribonucleic acid
<i>RNAP</i>	RNA polymerase
<i>rNTP</i>	Ribonucleoside tri-phosphate
<i>smFRET</i>	Single-molecule Förster resonance energy transfer
<i>SSB</i>	Single-stranded DNA binding protein
<i>ssDNA</i>	Single-stranded deoxyribonucleic acid
<i>TIRF</i>	Total internal reflection fluorescence
<i>TBP</i>	TATA box binding protein
<i>TF</i>	Transcription factor

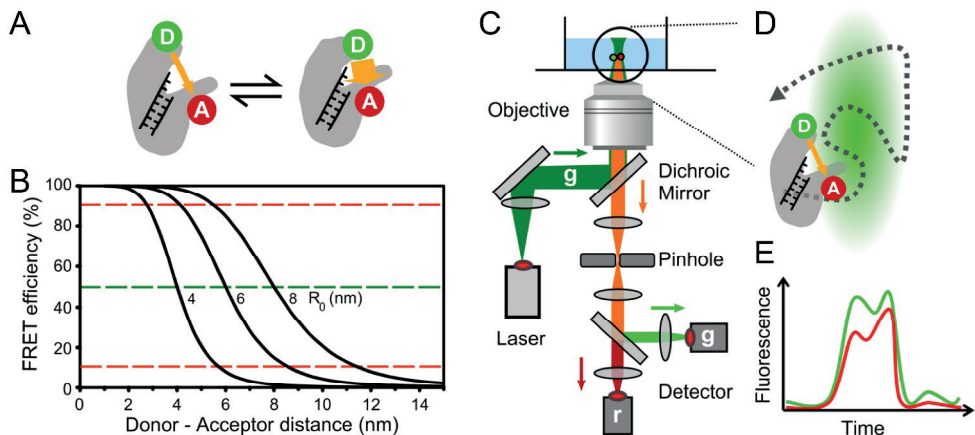
## **Introduction and theoretical background**

In order to understand the structure and function of biomolecular systems despite their often breath-taking complexity, scientists have been developing an ever-growing arsenal of sophisticated instrumentation and analytical methods. Nuclear magnetic resonance (NMR) spectroscopy<sup>1,2</sup> and X-ray crystallography<sup>3</sup>, for example, provide structural information with atomic resolution, but both methods ultimately fall short of resolving dynamic interactions within and especially between biomolecular complexes under physiologically relevant conditions. A major limitation of conventional biochemical analysis originates from ensemble- and time-averaging effects. In other words, the analysis reports on averaged properties of a population rather than the properties of individual species forming this population. With the advent of single-molecule techniques, researchers gained new exciting possibilities to study time-dependent sample distributions, conformational dynamics (Fig. 2.1a), reaction pathways, intermediate states, and asynchronous reactions<sup>4</sup>.

In this review, we will focus on an important member of the class of fluorescence based methods namely the single-molecule Förster resonance energy transfer (smFRET). This methodology allows detecting (relative) changes of distances between two fluorophores in the 2 to 10 nm range thus operating in a range comparable to the size of biomolecules such as proteins, lipids and nucleic acids. We will further limit our review to smFRET-based applications to study structure, dynamics and functions of DNA and DNA/protein interactions. We will also briefly discuss the development of techniques combining smFRET with force-based techniques such as optical and magnetic tweezers. For more general reviews about single-molecule techniques and smFRET, the interested reader is referred to<sup>5–11</sup>.

## **Single-molecule Förster resonance energy transfer**

FRET describes the distance-dependent and non-radiative energy transfer from a donor fluorophore to an acceptor chromophore via a dipole–dipole interaction and was first reported by Theodor Förster more than 60 years ago<sup>12</sup>. Three basic conditions need to be fulfilled for FRET to occur: (1) the spectra for donor emission and acceptor absorption must overlap, (2) donor and acceptor must be in close proximity (<10 nm) and (3) the relative orientation of the donor and the acceptor transition dipole moments must allow transfer of energy<sup>13</sup>.



**Figure 2.1.** Implementation of single-molecule FRET in confocal microscopy. **a** One potential goal of smFRET-based experiments is the study of conformational changes occurring in enzymes. Here, a DNA polymerase bound to DNA is fluorescently labelled with a donor and an acceptor fluorophore. Depending on the conformational equilibrium, two different FRET states are expected (open and closed), whereas ensemble-based experiments would struggle resolving this dynamic heterogeneity. **b** The FRET efficiency  $E$  plotted as a function of Förster radius  $R_0$ . Most common pairs of fluorophores (e.g., Cy3 and Cy5) have a Förster radius around 6 nm allowing handlers to detect distances between 4 and 9 nm. **c** Schematic overview of a confocal setup suitable for detection of smFRET. The laser light is collimated, reflected by a dichroic mirror and focussed with an objective of high numerical aperture to a diffraction limited excitation spot in the sample volume. Fluorescence, originating from excited dyes attached to diffusing proteins or DNA, is collected by the same objective and spatially filtered with a pinhole. Further on, the emitted fluorescence is spectrally split into a green (donor) and a red (acceptor) detection channel. **d** During the transit of a donor-labelled molecule through the focus, some energy can be transferred to the acceptor via FRET. **e** Every burst is characterised by two photon numbers: The number of photons in the donor channel ( $f_{Dex}^{Dem}$ ) and the number of photons in the acceptor channel ( $f_{Dex}^{Aem}$ ).

The FRET transfer efficiency  $E$  can be expressed using two rate constants, where  $k_D$  is the fluorescence emission rate constant of the donor in absence of the acceptor and  $k_T$  is the rate of energy transfer between the donor and the acceptor. These rates can be determined experimentally from the fluorescence lifetime of the donor in absence of the acceptor ( $\tau_D$ ) and in presence of the acceptor ( $\tau_T$ ). The transfer efficiency  $E$ , as shown by Förster, is inversely proportional to the sixth power of the distance  $R$  between the two transition dipoles according to

$$E = \frac{k_T}{k_D + k_T} = 1 - \frac{\tau_T}{\tau_D} = \frac{R_0^6}{R_0^6 + R^6} \quad (\text{eq. 2.1})$$

where  $R_0$  is known as the Förster radius and represents the distance between the transition dipoles corresponding to an energy transfer of 50 % between donor and acceptor (Fig. 2.1b).  $R_0$  is related to the properties of the fluorophores and the relative orientation of their dipole moments and is calculated using

$$R_0^6 = 8.8 \cdot 10^{-28} \frac{\Phi_D \kappa^2}{n^4} \int_0^\infty f_D(\lambda) A(\lambda) \lambda^4 d\lambda \quad (\text{eq. 2.2})$$

where  $\Phi_D$  is the donor quantum yield in absence of a nearby acceptor,  $n$  is the refractive index of the donor–acceptor intervening medium (for a discussion, see<sup>14</sup>), and  $\kappa^2$  is the orientation factor describing the mutual orientation of the two transition dipoles moments. The orientation factor is often set to  $\kappa^2 = 2/3$ , which is justified as long as at least one of the fluorophores has unrestricted rotational freedom<sup>15</sup>. The spectral overlap integral is calculated using the molecular extinction coefficient of the acceptor ( $\epsilon_A$ ) and the wavelength-dependent emission spectrum of the donor ( $f_D$ ).

Ensemble-based FRET techniques have been used to study structural features and dynamics of biological systems<sup>16–18</sup>. The outcome and interpretation of ensemble FRET data, however, is highly affected by the potential presence of dynamic or static heterogeneity in the sample<sup>19</sup>. Observation of FRET at the single molecule level<sup>20,21</sup> has overcome many of the shortcomings of ensemble FRET measurements and allows resolving this heterogeneity. smFRET is now widely applied to study (in vitro) molecular interactions and dynamics<sup>9,22,23</sup>.

In order to detect fluorescence emitted from single fluorophores against any background noise, a number of experimental requirements needs to be fulfilled, as the detectable photon budget from a single fluorophore is limited. Thus, we require a small excitation and detection volume to reduce the background from a scattering or weakly fluorescent medium and to distinguish a molecule of interest from other members of the same species. Often, a low concentration of fluorophores can be achieved by simply diluting the sample. However, as soon as we want to detect dynamic interactions between different fluorescently labelled species, diluting the sample severely limits the number of biomolecular interactions which can be studied, as many interactions require high sample concentrations considering the dissociation constant of their interaction<sup>24,25</sup>. Single-molecule FRET measurements are mostly taken using either confocal microscopy or total internal reflection fluorescence (TIRF) microscopy, the latter being a special case of wide-field microscopy.

## **Diffusion-based confocal microscopy**

Diffusion-based confocal microscopy (Fig. 2.1c) requires dilute solutions containing typically a picomolar concentration of fluorescently labelled species<sup>21</sup>. The molecules

diffuse through a femtolitre-sized excitation volume formed by a focused laser beam and a microscope objective with high numerical aperture. If the donor fluorophore is excited, it can transfer some of its energy to a nearby acceptor and causes the latter to fluoresce. In a confocal detection scheme, the objective used for excitation is also collecting the emitted fluorescence. After the emitted light is spectrally separated from the wavelength of the laser, the fluorescence is spatially filtered by a pinhole eliminating intensity contributions from outside the focus. For FRET detection, the fluorescence is split by a dichroic mirror into two channels, which cover the spectral range of the donor and acceptor emission, respectively. The FRET efficiency for every burst, representing the passage of one molecule through the focus (Fig. 2.1d), can be calculated in two ways: The first option uses the simultaneous recorded fluorescence lifetime of the donor as indicated in Eq. 2.1, but requires more sophisticated instrumentation using a pulsed laser and detectors with picosecond time-resolution. The second option is more common; it utilises the number of photons detected in the donor channel after donor excitation ( $f_{D_{ex}}^{Dem}$ ) and the number of photons detected in the acceptor channel after donor excitation ( $f_{D_{ex}}^{Aem}$ ). For each burst (Fig. 2.1e), the apparent FRET efficiency  $E^*$  is calculated as

$$E^* = f_{D_{ex}}^{Aem} / (f_{D_{ex}}^{Dem} + f_{D_{ex}}^{Aem}) \quad (\text{eq. 2.3})$$

It should be noted that  $E^*$  is not yet corrected for background, spectral crosstalk of the donor into the acceptor-emission channel and the instrument-dependent detection efficiencies of the dyes. For a step-by-step guide for obtaining an accurate FRET measure, the reader is referred to Hohlbein et al.<sup>26</sup>

## Alternating-laser excitation (ALEX)

As described above, excitation with a single laser allows the calculation of an apparent FRET efficiency  $E^*$ . Using common fluorophores, however, the emission spectrum of the donor is often broad and not fully spectrally covered by the donor detection channel. Instead, part of the donor fluorescence is detected in the acceptor detection channel, with the consequence that even a donor-only sample will show a FRET distribution with a mean  $E^*$  peak slightly above zero. The challenge researchers faced was how to discriminate low-FRET molecules with a fluorescently active acceptor from species in which the acceptor is not present or has been photo bleached before. To tackle this issue, Kapanidis and co-workers developed the ALEX scheme in which short periods of donor excitation alter with short periods of direct acceptor excitation<sup>26,27</sup> to verify the presence and state of the acceptor fluorophore in a fluorescently active form. ALEX provides an additional number for each burst  $f_{A_{ex}}^{Aem}$ , which represents the number of photons in the acceptor channel after direct excitation



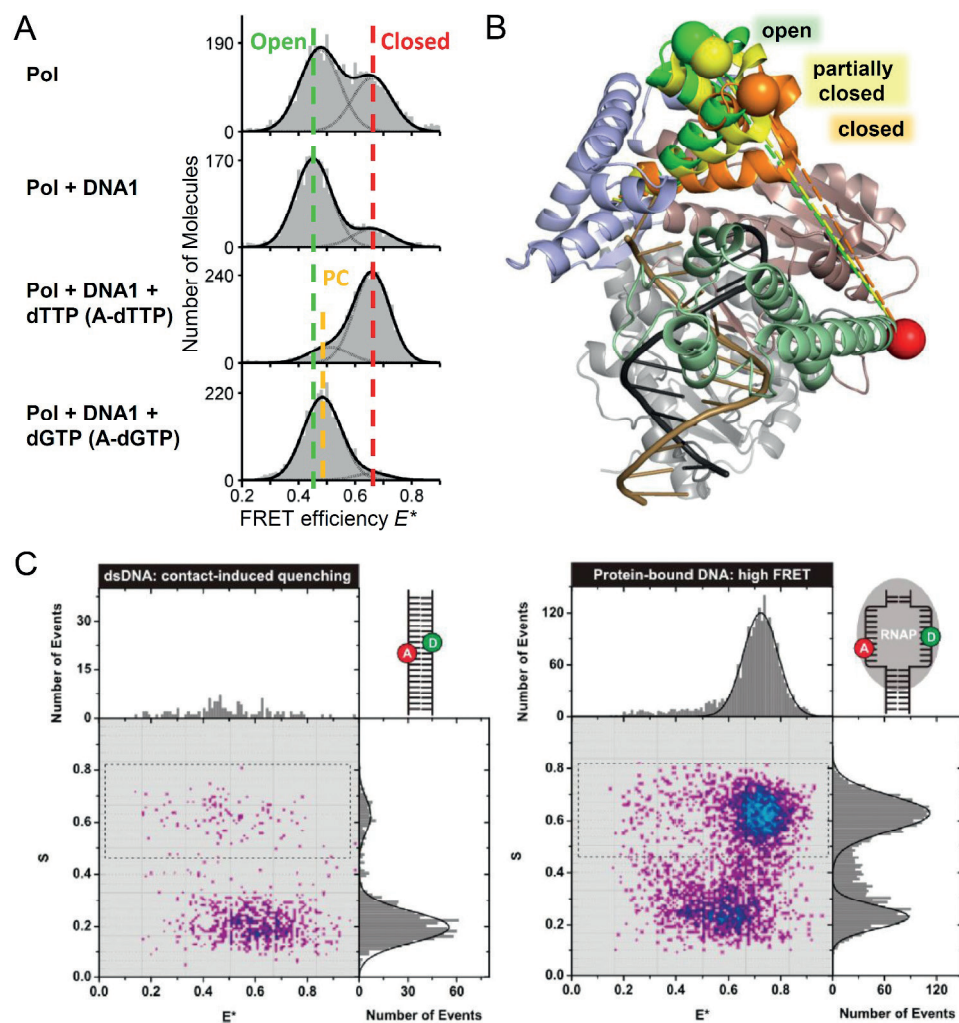
of the acceptor. Using that number, we can calculate the (raw) stoichiometry for each burst according to

$$S^{\text{raw}} = \frac{(f_{D_{\text{exc}}}^{A_{\text{em}}} + f_{D_{\text{exc}}}^{D_{\text{em}}})}{(f_{D_{\text{exc}}}^{A_{\text{em}}} + f_{D_{\text{exc}}}^{D_{\text{em}}} + f_{A_{\text{exc}}}^{A_{\text{em}}})} \quad (\text{eq. 2.4})$$

The stoichiometry represents the ratio of the total number of photons detected after donor excitation divided by the total number of photons detected in each burst. We obtain  $S^{\text{raw}} \sim 1$  for donor-only species (as  $f_{A_{\text{exc}}}^{A_{\text{em}}} \sim 0$ ) and  $S^{\text{raw}} \sim 0$  for acceptor-only species (as  $f_{D_{\text{exc}}}^{D_{\text{em}}} + f_{D_{\text{exc}}}^{A_{\text{em}}} \sim 0$ ). Depending on the relative count rates after donor and acceptor excitation, the (raw) stoichiometry for correctly labelled species bearing both donor and acceptor fluorophores can be tuned such that by plotting the corresponding  $E^*$  values versus the  $S^{\text{raw}}$  values for each burst in a two-dimensional ES histogram we can clearly separate this donor–acceptor species from species bearing only one active fluorophore (see also Fig. 2.2c, right panel). Moreover, we can resolve more complex binding mechanisms if, for example, two acceptor-labelled enzymes bind a donor-labelled DNA<sup>26</sup>.

## Imaging-based TIRF microscopy

A major limitation of solution-based approaches is the short observation time dictated by the transit time of a molecule diffusing through the confocal volume (<3 ms). Therefore, a burst basically represents a snapshot of a molecule, but the history and fate of the particular molecule remains unknown. Immobilising molecules of interest on a surface can overcome the problem of the limited observation time and camera-based schemes such as TIRF microscopy allow for monitoring hundreds of single fluorescent molecules in parallel<sup>32,33</sup>. TIRF microscopy is based on the total internal reflection phenomena, in which an evanescent wave is generated, as light cannot enter from a medium with a high refractive index into a medium with a lower refractive index at an angle greater than a given critical angle. The intensity of the evanescent wave decays exponentially within a few hundred nanometer above the glass surface, hence considerably reducing the background fluorescence from solution. Therefore, this methodology allows for monitoring higher concentrations of analytes in solution than diffusion-based confocal microscopy. One main challenge for imaging-based single-molecule detection is the precise control of the photophysics of fluorophores. Premature photobleaching and photo-induced blinking of fluorophores limit the number of detectable photons, but these difficulties can be largely avoided by using additives for oxygen scavenging and triplet quenching or by using newly developed fluorophores<sup>34–38</sup> making it nowadays possible to detect up to a million photons from a single Cy5 fluorophore<sup>38</sup>.



**Figure 2.2.** Applications of smFRET for studying DNA polymerases (**a**, **b**), adapted from reference<sup>28</sup> and RNA polymerases (**c**). (Reprinted with permission from Cordes et al.<sup>29</sup>. Copyright (2010) American Chemical Society.) **a** Conformational landscapes of DNA polymerase 1 (Klenow fragment)<sup>28</sup>. The unliganded enzymes show a dynamic equilibrium between at least two conformations of the fingers-subdomain. The binary complex of enzyme and DNA (with A as the templating base) shows an equilibrium shifted to the open conformation. Upon addition of correct nucleotide (1 mM dTTP), most molecules are the closed conformation, whereas the ternary complex formed with incorrect nucleotides (1 mM dGTP) adopts a partially closed conformation. **b** Proposed crystal structures of the three conformations. The Pol–DNA binary complex (PDB accession code 1L3U, mobile section in green; Johnson et al.<sup>30</sup>) is superimposed on a ternary complex where the fingers adopt a partially closed conformation (PDB accession code 3HPO, mobile section in yellow; Wu and

Beese<sup>31</sup>) and a ternary complex where the fingers adopt the closed conformation (PDB accession code 1LV5, mobile section in orange; Johnson et al.<sup>30</sup>). The distances between the C $\beta$  positions of residue 550 (red sphere) and 744 (sphere on the fingers) are 5.2, 5.0 and 4.2 nm. The predicted distance change of 0.2 nm upon going from the open to the partially closed conformation is consistent with the observed shift of  $\Delta E^* = 0.04$  seen in **a**. **c** Working principle of *qu*FRET presented based on a two dimensional histogram of the transfer efficiency  $E^*$  versus the stoichiometry as introduced in the main text. Two fluorophores are attached in very close proximity on each strand of a dsDNA so that the fluorescence of both probes is suppressed and only the acceptor-only species is visible (left panel, low stoichiometry). Upon formation of the transcription bubble by a bacterial RNA polymerase, high FRET values are observed as the fluorophores do not quench each other anymore (right panel).

## DNA processing enzymes

A large number of the smFRET studies investigate the interactions between DNA and proteins. Modified DNA, for example labelled with a FRET pair of fluorophores, is commercially available and can be easily immobilised on a modified glass surface. After immobilisation, the DNA acts as a binding target for enzymes freely diffusing in solution. Whereas some interactions can be studied even without labelling the DNA processing enzyme, others use smFRET between two fluorophores attached to the enzyme and the DNA, respectively.

## DNA polymerases

In 1953, Watson and Crick<sup>39</sup> identified the double helix as the main structural element of salt DNA. The authors noted that pairing between both strands might be the basis for a copying mechanism. In fact, only several years later, the group of Arthur Kornberg identified a first enzyme in *Escherichia coli* that synthesized DNA based on a templating DNA strand<sup>40</sup>. The enzyme was simply termed "DNA polymerase" but later classified as DNA polymerase I, when it became evident that five different polymerases coexist in this organism<sup>41,42</sup>. The fidelity by which different polymerases incorporate nucleotides into a (growing) DNA strand is known to vary depending on the cellular role of the specific enzyme<sup>43</sup>. *E. coli* DNA polymerases I and III, which are involved in DNA replication, have high fidelity with a frequency of correct incorporations in the order of 1,000,000:1. To achieve that accuracy, they are utilising their intrinsic 3'-5' exonuclease activity, which allows the removal of wrongly incorporated bases. *E. coli* DNA polymerases IV and V, on the other hand, are involved in translesion DNA synthesis, in which the addition of any nucleotide to resolve the stalling of DNA replication is more important than adding the correct one and have therefore a lower fidelity. Knowledge about polymerase fidelity and its structural basis is important as it concerns the very basic level of information storage and genetic stability in a cell. X-ray crystal structures and NMR studies have been

essential to shed light on nucleotide-polymerase interactions, but static structures are limited in their ability to provide insights in the dynamic processes that occur during nucleotide selection and incorporation.

One of the best studied DNA polymerases is *E. coli* DNA polymerase I (Klenow fragment), which is a cleavage product from bacterial DNA polymerase I<sup>44</sup>. It possesses 5'-3' polymerase activity and 3'-5' exonuclease activity, but it lacks 5'-3' exonuclease activity from full-length Pol I. The structure of the Klenow fragment resembles a human right hand and consists of four subdomains: 3'-5' exonuclease, thumb, palm, and a so-called "fingers" subdomain, which is thought to have a particular important role in nucleotide selection and incorporation (Fig. 2.2a and b). It was shown that the "fingers" close during nucleotide selection, thereby transferring the nucleotide to the active site, where the next step involves incorporation of the nucleotide in the growing DNA strand<sup>30,31,45</sup>. As will be explained below, this "fingers closing" mechanism was found to contribute to the fidelity of the polymerase. A number of smFRET studies have improved our understanding of the conformational changes and dynamics that contribute to polymerase fidelity. Most studies described below used the exonuclease-deficient Klenow fragment (the exonuclease activity can be deactivated with a D424A substitution), as it is easier to handle and has only one internal cysteine, facilitating convenient labelling with organic fluorophores.

In a 2009 study, smFRET-based TIRF microscopy was used to visualise DNA synthesis<sup>46</sup>. A DNA template was labelled with a donor dye and Klenow fragment was labelled with an acceptor moiety at the back of the palm subdomain that is expected to be static towards the DNA during fingers opening and closing. An increase in distance between Klenow fragment and the donor attached to the DNA was observed as soon as the binary complex was provided with nucleotides (dNTPs) complementary to the bases forming the single-stranded DNA. This approach resulted in base pair resolution of the synthesis mechanism, and the observation of several distinct conformational changes related to nucleotide insertion. In a following publication, Markiewicz and co-workers studied the stability of the Klenow fragment-DNA complex in the presence of different nucleotides using single-molecule protein-induced fluorescence enhancement (smPIFE; see<sup>47-49</sup>). It was shown that correct (complementary) dNTPs stabilize the polymerase-DNA complex more than correct rNTPs, while all other incorrect nucleotides destabilize the complex. The authors suggested that a steric clash between the template and an incorrect nucleotide would lead to a higher dissociation constant of the complex as a whole. However, more incorrect nucleotides than correct nucleotides are present under physiological conditions. It was therefore hypothesized that an incorrect nucleotide may be rejected in a step preceding the steric clash to prevent disintegration of the polymerase-DNA complex.

In 2010, it was shown using confocal-based smFRET that the fingers closing mechanism of DNA Pol I does not only occur in a ternary complex of polymerase, (non-extendable) DNA and a correct nucleotide, but also in binary complexes composed of polymerase and DNA and even in the unliganded enzyme<sup>50</sup>. The conformational landscape was probed by labelling the mobile part of the fingers subdomain with a donor and a static position on the thumb subdomain with an acceptor dye. Consequently, a change in FRET efficiency corresponds to a change in conformation: a higher FRET efficiency indicates a "fingers closed" conformation, while a lower efficiency marks an open conformation. Thus, Pol I is able to switch between different conformational states, though it depends on the complex which state is preferred. Binary complexes were mainly found in the open conformation, whereas ternary complexes with correct nucleotides were mostly found in the closed conformation. This study also introduced a new method for analysing smFRET data termed burst variance analysis (BVA). BVA is based on monitoring the standard deviation of FRET calculated from small photon numbers within each burst and allows handlers to distinguish between static and dynamic heterogeneity in a sample<sup>51</sup>. In the experiments on Pol I, BVA revealed the presence of conformational dynamics in the unliganded polymerase in absence of both DNA and nucleotides<sup>50</sup>. This result suggested that fingers closing is not necessarily an indication for successful incorporation of a nucleotide and that other mechanisms are in place to check nucleotides for their compatibility to the base of the templating DNA strand before they are incorporated.

More support for this model came from the fact that the fingers do not close completely in ternary complexes with mispaired dNTPs or when ribonucleotides are added<sup>28,50</sup>. Instead, molecules were observed showing a mean FRET efficiency only slightly higher than the one indicating the open conformation (Fig. 2.2a and b). Seemingly, Pol I can detect a mispair before full fingers closing would occur. This suggests that one or more intermediate states exist, in which the incoming nucleotide is previewed and subsequently rejected or incorporated in the DNA. In 2011, a crystal structure of *Bst* Pol I (a close structural homologue to *E. coli* Pol I) bound to DNA was published<sup>31</sup>. This crystal structure revealed the presence of an intermediate (ajar) conformation of the fingers subdomain when bound to an incorrect nucleotide. Recent smFRET-based studies have identified an intermediate conformation as being present in ternary complexes with incorrect nucleotides<sup>52</sup> and two studies found direct evidence for the population of an intermediate state even in the presence of correct nucleotide substrates<sup>28,53</sup>.

The identification of three instead of two different conformations of the fingers domain raises the question how the polymerase interconverts between these states and how that depends on the type of complex formed. Both, Rothwell and co-workers

and Hohlbein and co-workers analysed the equilibrium conditions and possible transitions and dynamics between the open, closed and intermediate conformation. Variants of probability distribution analysis (PDA) were used for predicting FRET efficiency distributions from a mixture of static or dynamically interconverting FRET species<sup>54–58</sup>. A comparison between experimental data and data generated by PDA allows for the identification and characterization of (dynamic) subspecies. In addition, Hohlbein et al.<sup>28</sup> used wild-type Klenow fragment and mutator derivatives with decreased fidelity to calculate a free energy landscape in which the partially closed state was identified as a major fidelity checkpoint for nucleotide insertion.

The Klenow fragment 3'–5' exonuclease (exo) domain, which catalyses the excision of mismatched nucleotides, has also been subject of single-molecule studies. Using an smFRET approach, it was shown that mismatched primer-template termini bind to the polymerase in a different orientation than matched termini<sup>48</sup>. In a similar study, the static thumb subdomain and the DNA primer strand were labelled with a FRET pair<sup>59</sup>. Using this approach, the binding of Klenow fragment to the template-primer duplex was shown, but also an unexpected switching of the DNA between the pol and exo domains was observed. This finding is in contradiction to the data presented by Markiewicz and co-workers, in which exclusive binding of mismatched DNA in the exo site rather than switching of DNA between two positions was suggested. In both studies, the DNA was labelled several base pairs away from the mismatch, but the labelling position chosen by Lamichhane and co-workers may have been optimized to report on DNA switching. Lamichhane and co-workers found that internal mismatches in the DNA increased the rate of switching between the pol and exo site, which agrees with wrongly inserted nucleotides being excised at the exo domain. Addition of dNTPs also influences pol–exo switching: if the primer terminus is correctly base-paired, a dNTP stabilizes the DNA in the pol site. However, if the primer terminus is mispaired, the same dNTP accelerates switching of the template to the exo domain. Even dNTPs that do not form a correct new base pair (any of the three others) are able to accelerate binding of an already mispaired primer terminus to the exo domain. This means that even incorrect nucleotides have an active role in processes to increase polymerase fidelity. This counterintuitive result is in line with an earlier observation by the same group<sup>52</sup> in which they showed that primer termini are moved to the exo domain when incorrect nucleotides are trapped in the "ajar" conformation.

Taken together, smFRET has been successfully used to study conformational changes and conformational dynamics within DNA polymerases and between DNA polymerases and DNA. For the studied derivatives of DNA polymerase I, the newly characterized partially closed ('ajar') conformation appears to play an important role as a fidelity checkpoint. Whether this intermediate state is present in other DNA



polymerases and whether the conformational dynamics can be linked to fidelity and DNA catalysis will be the subject of upcoming research.

## **RNA polymerases**

RNA polymerases (RNAPs) facilitate the transcription of DNA into RNA, which stands at the basis of protein synthesis. As for DNA polymerases, much of our current knowledge of RNAPs is derived from X-ray crystal structures and ensemble studies. Single molecule studies on RNAPs have mainly focussed on two species: bacterial RNAP and eukaryotic RNA polymerase II (Pol II). Both RNAPs follow the same general sequence of events for transcription initiation<sup>60</sup>. First, the polymerase binds to promoter regions on the DNA to form a so-called "closed complex". Second, local melting of the DNA results in the formation of a so-called "transcription bubble", which forms an "open complex" together with the polymerase. As soon as the open complex starts transcription, it is called an "initial transcribing complex". RNA synthesis is often limited to short RNA products in the first stage of transcription. Once a polymerase produces a longer RNA product (~10 nucleotides for RNAP and ~3 nucleotides for Pol II), it enters a stage of processive RNA synthesis. This event triggers release of initiation factors and the polymerase leaves the promoter region, leading to the formation of a polymerase–DNA "elongation complex" (as reviewed by<sup>61,62</sup>).

In 2006, Kapanidis et al.<sup>63</sup> evaluated three proposed models for initial transcription by the RNAP initial transcribing complex. The models were based on observations showing the formation of RNA products, even when the polymerase did not appear to move along the DNA. The first model stated that RNAP moves along the DNA as a unit, but returns to its initial position after release of RNA<sup>64</sup>. The second model involved stretching of a flexible element in RNAP, resulting in movement relative to the DNA of the leading, but not the trailing edge of the polymerase<sup>65,66</sup>. A third model stated that the polymerase itself does not move. In contrast, this model predicted expansion and contraction of the DNA strand, called "scrunching"<sup>64,67,68</sup>. In order to determine which model is correct, Kapanidis and co-workers monitored FRET changes between different elements of the complex labelled with donor and acceptor dyes by means of smFRET. rNTPs were added to RNAP–DNA complexes to start transcription. A FRET change in agreement with one of the models was only found for a donor–acceptor pair located on the DNA: upon addition of rNTPs an increase in FRET efficiency was observed, indicating a decrease in distance, which ultimately proved that the third "DNA scrunching" model is correct.

A new technique called quenchable FRET (quFRET; Fig. 2.2c) was developed by Cordes et al. to study DNA melting after formation of the open complex<sup>29</sup>. Quenchable FRET relies on contact-induced quenching of two dyes as long as they are in close

proximity ( $<2$  nm). The authors placed a FRET donor (Cy3B) on the first DNA strand and an acceptor (ATTO647N) on the complementary strand. As soon as both strands were annealed, fluorescence detectable from both dyes was greatly diminished. Upon formation of the transcription bubble in the RNAP open complex, dequenching accompanied by a high FRET efficiency was observed. The authors used this principle to derive rate constants for the formation of the open complex and showed that quFRET can be used as a quantitative tool.

RNAP contains a clamp structure that is known to open and close. The dynamics of this opening and closing are thought to be important during RNAP complex formation: the open state allows for accommodation of dsDNA, while the closed state can only accommodate ssDNA<sup>69-72</sup>. A third "collapsed" state does not leave room for any DNA. Chakraborty et al.<sup>73</sup> labelled the clamp with a FRET donor and an immobile part on the enzyme with a FRET acceptor, in order to study the conformational dynamics of the clamp. Free RNAP holo-enzyme showed a FRET efficiency distribution that can be fitted with three Gaussian functions, each of which can be attributed to a different state of the clamp: open, closed and collapsed. The open state appeared to be the predominant state. It remained predominant during formation of the closed complex. Upon formation of the open complex, however, the clamp adopts only the closed state. This is also the case for the initial transcribing complex and the elongation complex. These results suggest that the clamp is triggered to close only after interaction with ssDNA in the open complex.

Eukaryotic Pol II requires the concerted action of several different transcription factors (TFs) to regulate transcription initiation. Studies on Pol II have therefore focussed on the path of the growing RNA strand and the positions of various TFs in the initiation and elongation complexes. TFIIB is such a TF and is associated with the polymerase. It aids in the attachment of promoter DNA and TATA box binding protein (TBP) to the enzyme (the term "TATA box" refers to the name of a recognition element for TFs in the promoter sequence). Andrecka et al.<sup>74</sup> found evidence for eventual complete TFIIB displacement from Pol II, as it showed that the growing RNA chain in the elongation complex interacts with the same dock domain as TFIIB. In their later work, Muschielok et al.<sup>75</sup> used their Nano Positioning System (NPS) to study this interaction again. They found that the TF remains associated to the dock domain longer than they initially expected. This could mean that TFIIB might have a role in guiding the new RNA strand to the Rpb4/7 substructure of the polymerase, which is of physiological importance as this substructure has a role in recruitment of 3' end processing factors<sup>76</sup>.

In 2012, Treutlein et al.<sup>77</sup> published an extensive study in which they constructed a model for a minimal Pol II open promoter complex. This model included a TATA box, a mismatched DNA region, TBP, Pol II and transcription factors TFIIB and TFIIF. The model was constructed using known X-ray crystal structures and smFRET. The NPS



combined data from these two techniques to make accurate predictions about the location of several subunits in the complex. Using this setup, it was found that the B core of TFIIB is displaced in the open complex, but it is likely that it still interacts with DNA and the TATA box binding protein. Furthermore, the authors found that the open complex has TBP and TATA DNA located above the cleft. Downstream DNA was found to switch between a position inside the cleft of the enzyme and a position on top of the cleft. This switch occurs on a timescale of seconds and is therefore considered to be an important kinetic trap. Detection of this kind of flexible behaviour is an excellent example of the strength of smFRET studies: indeed, this switching of DNA is exactly the reason why the Pol II open complex could not be trapped crystallographically before.

## **DNA helicases**

Helicases are motor proteins that separate double stranded nucleic acids such as DNA, RNA or DNA–RNA hybrids by using energy derived from ATP hydrolysis<sup>78</sup>. In DNA replication, for example, DNA helicases unzip dsDNA starting from a position known as replication origin. DNA helicases unwind DNA by breaking the hydrogen bonds that keep the two strands of DNA together, thereby forming the replication fork in which the separated strands serve as template strands for leading and lagging strand DNA synthesis.

Various single-molecule techniques have contributed to a better mechanistic understanding of helicase activity<sup>11,79</sup>. Here we will focus on assays utilising smFRET as a high precision technique to monitor the structural change in DNA upon interaction with DNA helicases.

In 2002, Ha and co-workers studied the mechanism by which *E. coli* Rep helicase initiates DNA unwinding<sup>80</sup>. They immobilised DNA molecules consisting of a short DNA duplex and a single-stranded overhang on a glass surface. TIRF microscopy was applied to detect the fluorophores that were attached on opposite strands at the junction between single- and double-stranded DNA. Rep helicase was found to bind to the single-stranded DNA and then to shuttle towards the junction fuelled by ATP hydrolysis. Upon binding to the junction, fluctuations in the FRET efficiency indicated conformational fluctuations of the DNA, but processive unwinding only occurred after binding of an additional protein. Their analysis suggested that the limited unwinding observed in vitro for Rep is due to the relative instability of the functional complex, caused by DNA rewinding upon complex dissociation and rounds of reinitiation upon reformation of the functional helicase complex.

In 2004, Resnik and co-workers developed an assay for site-specifically labelling of REP helicase<sup>81</sup> that was later used to monitor repetitive shuttling of REP along single-

stranded DNA<sup>82</sup>. Interestingly, after moving in the 3' to 5' direction of the single stranded DNA using ATP hydrolysis, Rep snaps back close to the 3' end, a mechanism likely to be caused by a conformational change of the protein after approaching the DNA junction. It was hypothesised that the shuttling of the DNA helicase along ssDNA might be an effective way of clearing the DNA from unwanted, bound molecules.

The hepatitis C virus NS3 protein is a bifunctional helicase that can unwind both DNA and RNA substrates. It was shown in 2007 that that NS3 unwinds DNA in discrete steps of about 3 bp<sup>83</sup>. The fluorescence assay consisted of a double-stranded DNA, labelled with fluorophores on each strand at the DNA junction where one strand continued with single stranded DNA. Unwinding led to a stepwise decrease in FRET efficiency, and six steps were found for the 18-bp-long double-stranded DNA. As the dwell time histogram showed non-exponential behaviour, the authors fitted a Gamma distribution suggesting that every 3-bp step is composed of three hidden steps of one base pair each. Using additional experiments, the authors suggested a model in which, based on the three domains of NS3, domains 1 and 2 move along the tracking strand (3' to 5') one nucleotide a time, consuming one ATP for each base pair. The third domain stays behind by attaching itself to the DNA until three of such steps have taken place. After the third step, the domain 3 moves forward in a burst motion, unzipping 3 bp as a consequence. NS3 continues unwinding in 3-bp steps until 18 bp. On longer duplexes, the helicases showed repetitive unwinding.

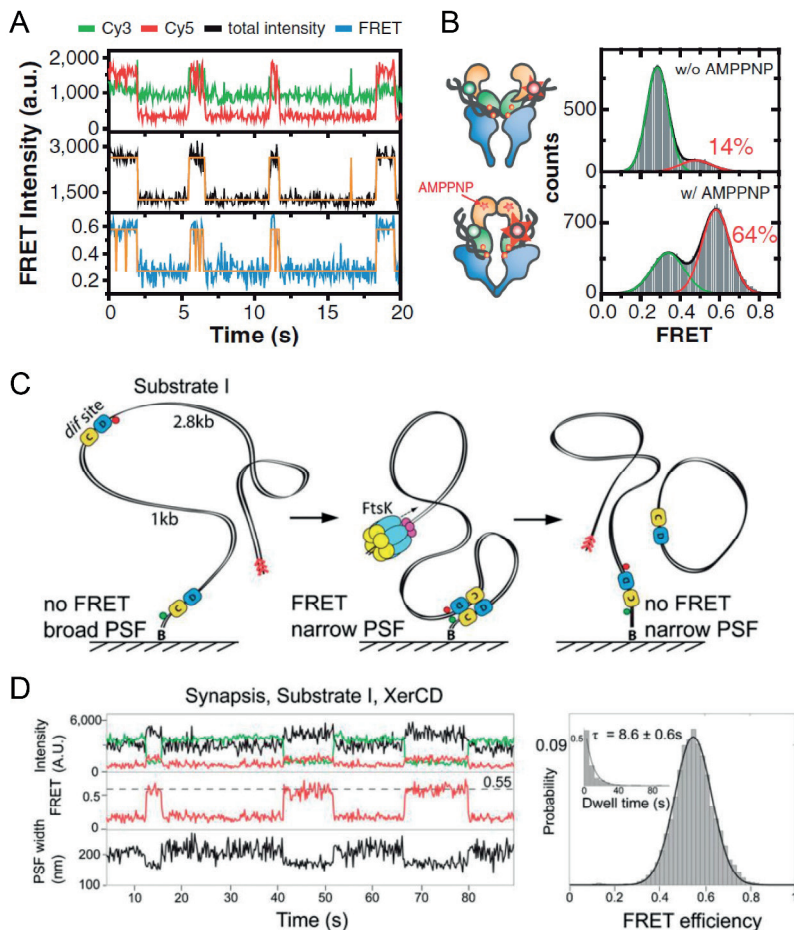
## **DNA topoisomerases, DNA recombinases and transcription factors**

In this part, we will briefly discuss two important classes of DNA processing enzymes that allow the cutting and re-joining of DNA, DNA topoisomerases and DNA recombinases, before continuing with a short discussion about DNA TFs.

DNA topoisomerases regulate DNA supercoiling, which is a consequence of the varying DNA topology ranging from densely packed DNA to accessible DNA required for DNA replication and transcription. DNA gyrase, for example, is a type II topoisomerase found mainly in prokaryotes. The enzyme is capable of introducing negative supercoiling using ATP hydrolysis. Tension that builds up after the unwinding of DNA is revealed by cutting the strands and re-annealing them after the enzyme passed a different DNA segment through. One recent study of this DNA-gate conformation using single molecule FRET claimed that high and low FRET states, corresponding to open and closed conformations of the DNA gate, are equally populated in topoisomerase II<sup>84</sup>. By directly monitoring the conformational state of the DNA gate in DNA gyrase, it was found that the gate is mainly closed and gate

opening is a rare event that occurs only briefly to allow the transfer DNA to pass<sup>85</sup>. Further studies on GyrA, which is a subunit of the heterotetrameric DNA gyrase, elucidated how binding of DNA to the DNA binding region affects the conformational cycle for supercoiling DNA by DNA gyrase<sup>86,87</sup>. A recent report by Lee et al.<sup>88</sup> on human topoisomerase II $\alpha$  revealed that cleavage and opening of DNA is tightly regulated by magnesium ions controlling the bending of gate-DNA. By visualizing the individual steps of the DNA cleaving reaction, the authors shed light on the mechanism, by which the probability of accidental double-strand breaks is minimised (Fig. 2.3a and b).

The second class of enzymes, DNA recombinases, plays an important role in the cellular rearrangement of DNA required, for example, in chromosome segregation. Two recent publications investigated the mechanics of site-specific recombination using Cre-*loxP*<sup>89</sup> and XerCD-*dif* complexes<sup>90</sup>. SmFRET was utilised to monitor the formation of the synaptic complex. A donor and acceptor fluorophore were placed close to both target sites, which were initially well separated (Fig. 2.3c and d). Upon complex formation, both target sites are brought together and FRET can occur between the fluorophores. Interestingly, both publications used up to three independent observables from single fluorescent molecules: (1) FRET, (2) size of the fitted point spread function (PSF), and (3) protein-induced fluorescence enhancement (PIFE<sup>47</sup>) to monitor short- and long-ranged conformational changes of the DNA induced via recombination complexes.



**Figure 2.3.** Applications of smFRET for studying topoisomerases (**a** and **b** reprinted with permission from Lee et al.<sup>88</sup>) and DNA recombination complexes (**c** and **d**, reprinted with permission from Zawadzki et al.<sup>90</sup>). **a** Representative fluorescence intensity and FRET time traces of fluorescently labelled cleavable DNA containing the binding site human topoisomerase IIa. Upon binding of the enzyme and in the presence of AMPPNP (a nonhydrolyzable analogue of ATP), the intensity of the donor increases due to protein induced fluorescence enhancement and an increase in FRET indicates a conformational change of the complex of DNA and enzyme. **b** The FRET histograms with AMPPNP (bottom) and without AMPPNP (top) show that the gate clamping induces a substantial deformation of the DNA gate. **c** Schematic representation of the recombination reaction. The long DNA has the fluorophores (indicated with red and green circles) attached adjacent to each dif site. Recombination can be monitored using FRET between donor and acceptor and the width of the point spread function (PSF) after direct excitation of the acceptor. Successful recombination between the dif sites leads to formation of two DNA molecules. The red arrow heads indicate the preferential binding site for FtsK. **d** Formation of nonproductive synaptic complexes is

*indicated by two observables: (1) an increase in FRET calculated using the intensities of donor (green) and acceptor (red) under donor excitation and (2) an increase in fluorescence intensity detected after direct excitation of the acceptor (black) and an decrease in the fitted PSF width as the acceptor is brought closer to the surface. The histogram (right) show the distributions of FRET efficiencies ( $E^*$ ) and the dwell times (inset) of XerCD–dif synaptic complexes.*

The precise control of gene expression is essential for every living cell and therefore tightly controlled by DNA binding proteins known as transcription factors. TFs act as natural biosensors and switches modulating gene expression of target genes by either promoting or blocking the recruitment of RNAPs. Several human diseases such as diabetes, autoimmune diseases, and cancer have been linked to mutations in TFs such as p53<sup>91</sup>. Therefore, assays for efficient detection of transcriptions factors are highly desirable as they might provide a platform for diagnostics. In 2010, TF-dependent DNA coincidence was detected using ALEX spectroscopy<sup>92</sup>. In the presented assay, two DNA half sites labelled with donor and acceptor, respectively, contain a complementary region of ssDNA that forms the binding site of the catabolite activator protein (CAP) TF. Without the presence of a TF, the binding of the complementary half sites is too weak to form a stable complex as indicated by a single peak in the stoichiometry histogram which reports on the number of different species in the solution bearing an acceptor fluorophore. Upon addition of CAP, the binding of both DNA half sites is stabilised and a second peak can be seen in the histogram. Depending on the concentration of half sites, TF concentrations in the low nanomolar range are detectable. Based on this TF assay, the encapsulation and entrapment of CAP inside a DNA cage was successfully demonstrated<sup>93</sup>. Within the cage, which is based on a DNA tetrahedron, the TF is inactive and cannot bind to cellular DNA. Using acceptor labelled CAP and a donor labelled cage, it was confirmed that CAP is positioned within the cage. Moreover, DNase I was shown to degrade the cage leading to a release of CAP. A different assay was used to detect the binding of CAP via binding-induced bending of a doubly labelled DNA construct containing the CAP binding site<sup>94</sup>.

## **Increasing the information content of smFRET experiments**

In this section, we will highlight some recent developments utilising dual-colour FRET between one donor and one acceptor. For publications introducing three- or four-colour FRET, we refer the reader to reviews elsewhere<sup>9,11,95</sup>.

## **Combining smFRET with force-based techniques**

In the following, we will briefly discuss some recent developments aiming to combine fluorescence-based techniques with force-based techniques such as atomic force microscopy (AFM), optical tweezers and magnetic tweezers. Many designs, often

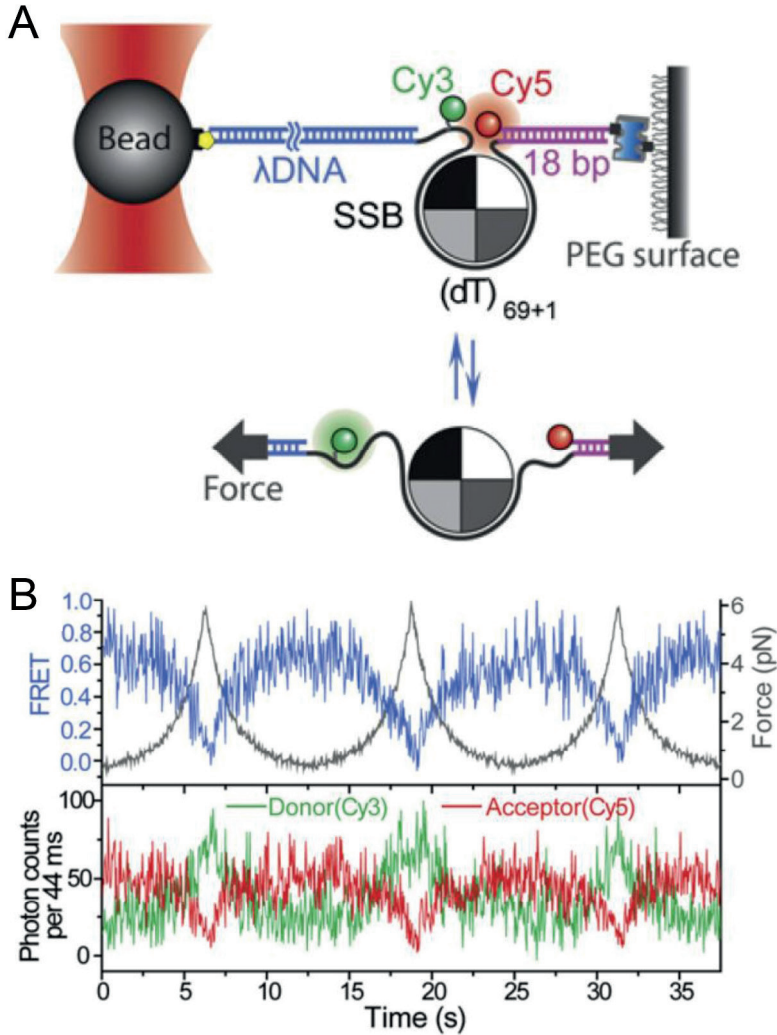
based on DNA spanned between a glass surface and a moveable bead, have been envisioned as early as 1999<sup>22</sup>. The experimental realisation, however, has proven to be very challenging and widespread use has not yet been achieved, despite their promise of providing new means of studying the sub-molecular structure, conformational dynamics and transition states of biological systems. We refer the interested reader to recent reviews, which discuss potential merits in greater detail<sup>4,9,11,95</sup>.

AFM has initially been developed for topographical imaging of molecules<sup>96</sup>. This technique uses a tip that is scanned along the sample surface and the deflection of the tip is measured using a laser and a photo detector. The combination of AFM with single-molecule-based TIRF microscopy was demonstrated by Hugel et al.<sup>97</sup>, who investigated extension of a polymer made of bistable photosensitive azobenzenes, and by Sarkar et al.<sup>98</sup>, who measured the forced unfolding of ubiquitin after calibrating the distance-dependent intensity decay of an evanescent wave using AFM. The promising combination of AFM and TIRF was reviewed in <sup>99</sup>. A combination of (ensemble-based) FRET and AFM was shown by Vickery<sup>100</sup> and by Nakamura et al.<sup>101</sup>, but only very recently was the combination of smFRET and AFM demonstrated<sup>102,103</sup>.

Rapid progress has been achieved for the combination of smFRET with optical tweezers, which utilize an infrared laser to trap and control the position of a bead. By spanning a single DNA molecule between the surface of the bead and the surface of the cover slide (or a pipette tip), strain can be applied simply by moving the bead in respect to the cover slide and the relative position of the bead can be imaged using a camera. The combination of an optical trap and single-molecule fluorescence detection in the visible spectrum was shown in 2003<sup>104</sup>. Brau et al.<sup>105</sup> improved the longevity of the fluorophore, which normally severely suffers from photo damage caused by the infrared laser, by alternating between bead trapping and direct excitation of the fluorophore. Combinations of optical tweezers with smFRET detection were successfully demonstrated in 2007<sup>106,107</sup> by Tarsa and co-workers using smFRET to monitor the opening and closing of a DNA hairpin whilst being under tension applied via the optical trap. Likewise, Hohng et al.<sup>107</sup> mapped the reaction landscape of DNA Holliday junctions, which is a four-stranded DNA structure that switches between open and closed stacking conformations. In a series of beautiful force-fluorescence experiments, Zhou et al.<sup>108</sup> studied the behaviour of the single-stranded DNA binding protein (SSB) interacting with DNA (Fig. 2.4). The authors showed that ssDNA bound to SSB unravels at low forces (<6 pN) and that larger forces lead to the dissociation of SSB. More intriguingly, the authors also found that ssDNA migrates on SSB via reptation rather than that SSB rolls around the DNA.

Another candidate for applying forces are magnetic tweezers, in which a small magnetic bead allows the application of strain and even torque to a surface-





**Figure 2.4.** Combining smFRET with force-based techniques reprinted from Zhou et al.<sup>108</sup>, with permission from Elsevier. **a** The experimental scheme shows a DNA molecule attached between a PEGylated glass surface (right) and a bead (left). The bead is optically trapped and can therefore be used to pull the DNA containing a region of single-stranded DNA to which the single stranded binding protein (SSB) is bound. Unraveling of ssDNA is expected to increase the distance between donor and acceptor. **b** The experimental realisation shows that increasing the pulling force at low levels leads to an unraveling of ssDNA indicated by a decrease in FRET and a clear anticorrelation of the donor and acceptor intensity. The process of unravelling is reversible and larger forces (not shown here) lead to dissociation of SSB.

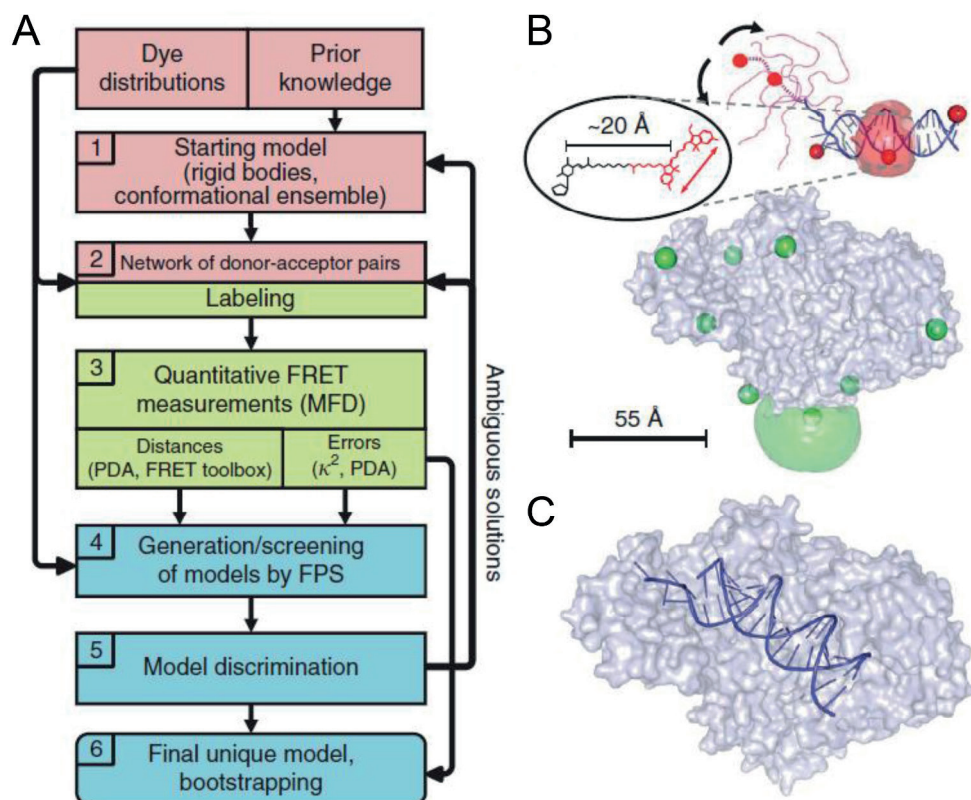
immobilised DNA molecule. In addition, magnetic tweezers do not cause photo bleaching of fluorophores which is advantageous to maximise observation time. In 2005, Shroff et al.<sup>109</sup> showed that changing the distance between a FRET pair can be used to calibrate the force response of a DNA sensor. Lee et al.<sup>110</sup> showed that negative superhelicity of DNA and low tension induced by magnetic tweezers is sufficient to trigger the formation of Z-DNA formation deviating from the canonical B-DNA. By labelling the DNA on both strands of a CG core strand, the appearance of Z-DNA was indicated by an increase in distance between the fluorophores and therefore a decrease of the FRET efficiency. Very recently, Long et al.<sup>111</sup> probed the force-dependent unfolding of G-quadruplex DNA. Again, smFRET was used to probe the conformational change in the nanometer range showing that the transition-state barrier for unfolding is closer to the unfolded state than the folded state of the complex.

## **Quantitative smFRET and computer-aided modelling**

Even though smFRET has been coined as a molecular ruler, converting FRET efficiencies to actual distances requires careful corrections to account for background fluorescence and the spectral properties of the fluorophores<sup>26,112</sup>. An important factor that needs to be considered is that any calculated FRET distance refers to the distance between the emission dipole of the donor and the excitation dipole of the acceptor fluorophore and not to the distance between the points of attachment on the DNA or the enzyme of interest. This notion is particularly important as most fluorophores are attached by means of flexible linkers, thus the position of the fluorophore in respect to the DNA or the enzyme is rarely static. Instead, the fluorophore resembles a certain accessible volume. This will result in a distribution of possible values of the orientation factor  $\kappa^2$  and distances between dipole and residue, causing a potential discrepancy between the distance of interest and the actual distance derived from smFRET.

Several methods have been developed to improve the quality of distance information that can be extracted from smFRET data<sup>113</sup>. In 2008, Muschielok and co-workers introduced a method called NPS (thereby aptly referring to GPS)<sup>75</sup>. It uses a system comprising a so-called antenna dye molecule and several satellite dye molecules to calculate a three-dimensional probability distribution of dye positions. Additionally, information such as the accessible volume derived from X-ray crystallography structures can be taken into account. The calculated dye positions reflect experimental uncertainties, as opposed to previous methods that were only able to show the most likely dye position. In 2011, NPS was extended by taking FRET anisotropy into account. The inclusion of average transition dipole moments of the dyes significantly improved localization accuracy<sup>114</sup>.





**Figure 2.5.** The FRET-restrained positioning and screening (FPS) framework (reprinted with permission from Kalinin et al.<sup>115</sup>, Macmillan Publishers Ltd, copyright 2012). **a** FPS consists of three main parts: (1) experimental design, (2) measurement and analysis, and (3) structural modelling and validation. **b** The structure of HIV-1 reverse transcriptase (PDB accession code 1R0A; Peletskaya et al.<sup>118</sup>) was separated into protein (grey) and dsDNA (blue). The different labelling positions for the acceptor on the DNA (red) and for the donor on the protein (green) are indicated. Clouds around the labelling positions indicate the accessible volume of the fluorophores. Potential solutions of the unresolved ssDNA are shown in magenta. **c** Structural model of the joined complex obtained after rigid-body docking

Another comprehensive framework for combining quantitative smFRET measurements and molecular modelling named FPS (FRET-restrained positioning and screening) was introduced in 2012<sup>115</sup>. FPS consists of overall six steps as shown in Fig. 2.5. In step one, a starting model is created, which includes all known information about a system such as crystal structures or known conformational changes. In step two, a network of donor–acceptor pairs (similar to antennas and satellites as described above) is designed. After sample preparation, single-molecule multiparameter fluorescence detection (MFD) measurements are performed to obtain the required

distances (step three). The (diffusion-based) MFD scheme includes measurements of the fluorescence lifetime and the anisotropy in addition to the fluorescence intensity information<sup>116,117</sup>. The smFRET data is used to generate new models (step four), which are then checked against the initial model in step five. At this point, it might be necessary to revert to step one or two and develop a model that is expected to fit FRET data better. In the final step, a procedure in which noise is added to the calculated set of distances is performed to determine the precision of the model. Using their approach, the authors generated a detailed model for DNA bound to HIV-1 reverse transcriptase, adding a part of the structure that has not been resolved by X-ray crystallography before. More importantly, FRET-based structural modelling allows utilising position data derived from transient conformational states, which are often difficult to trap in X-ray crystallography.

## Conclusions

In the past 20 years, smFRET has emerged as an undisputable powerful tool to elucidate biological processes on the molecular level. Using smFRET-based techniques, researchers are able to study samples under close to physiologically relevant conditions whilst having access to sub-nanometre accuracy to study (dynamic) interactions of and between single-molecules. Especially the study of DNA-protein interactions has greatly benefited, as many interactions can be studied using commercially available, fluorescently labelled DNA. Many previous issues such as premature photo bleaching and the lack of algorithms and software to analyse data have been largely addressed (for an overview of available software packages, see<sup>10</sup>), and we hope that the threshold for applying smFRET can be further reduced especially by developing easier labelling strategies for proteins and providing access and support to software and instrumentation. In the coming years, we expect smFRET to gain further popularity as a tool for structural and molecular biology.

## Acknowledgments

We thank Herbert van Amerongen and Arjen Bader for critically reading the manuscript and providing helpful suggestions. S.F. acknowledges funding from the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO).

## References

1. Wüthrich, K. The way to NMR structures of proteins. *Nat. Struct. Biol.* **8**, 923–925 (2001).
2. Foster, M. P., McElroy, C. A. & Amero, C. D. Solution NMR of Large Molecules and Assemblies. *Biochemistry* **46**, 331–340 (2007).

3. Ilari, A. & Savino, C. Protein Structure Determination by X-Ray Crystallography. in *Bioinformatics* 63–87 (Humana Press, 2008). doi:10.1007/978-1-60327-159-2\_3
4. Kapanidis, A. N. & Strick, T. Biology, one molecule at a time. *Trends Biochem. Sci.* **34**, 234–243 (2009).
5. Moerner, W. E. New directions in single-molecule imaging and analysis. *Proc. Natl. Acad. Sci.* **104**, 12596–12602 (2007).
6. Deniz, A. A., Mukhopadhyay, S. & Lemke, E. A. Single-molecule biophysics: at the interface of biology, physics and chemistry. *J. R. Soc. Interface* **5**, 15–45 (2008).
7. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507–516 (2008).
8. Walter, N. G., Huang, C.-Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: How to use the modern single molecule toolkit. *Nat. Methods* **5**, 475–489 (2008).
9. Hohlbein, J., Gryte, K., Heilemann, M. & Kapanidis, A. N. Surfing on a new wave of single-molecule fluorescence methods. *Phys. Biol.* **7**, 031001 (2010).
10. Preus, S. & Wilhelmsson, L. M. Advances in Quantitative FRET-Based Methods for Studying Nucleic Acids. *ChemBioChem* **13**, 1990–2001 (2012).
11. Kim, H. & Ha, T. Single molecule nanometry for biological physics. *Rep. Prog. Phys. Phys. Soc. G. B.* **76**, 016601 (2013).
12. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55–75 (1948).
13. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (Springer US, 2006).
14. Knox, R. S. & van Amerongen, H. Refractive Index Dependence of the Förster Resonance Excitation Transfer Rate. *J. Phys. Chem. B* **106**, 5289–5293 (2002).
15. Dale, R. E., Eisinger, J. & Blumberg, W. E. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophys. J.* **26**, 161–193 (1979).
16. Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. *Proc. Natl. Acad. Sci. U. S. A.* **58**, 719–726 (1967).
17. Clegg, R. M. Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.* **211**, 353–388 (1992).
18. Jares-Erijman, E. A. & Jovin, T. M. FRET imaging. *Nat. Biotechnol.* **21**, 1387 (2003).
19. Haas, E., Wilchek, M., Katchalski-Katzir, E. & Steinberg, I. Z. Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 1807–1811 (1975).
20. Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6264–6268 (1996).
21. Deniz, A. A. *et al.* Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci.* **96**, 3670–3675 (1999).

22. Weiss, S. Fluorescence Spectroscopy of Single Biomolecules. *Science* **283**, 1676–1683 (1999).
23. McKinney, S. A., Joo, C. & Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **91**, 1941–1951 (2006).
24. Levene, M. J. *et al.* Zero-Mode Waveguides for Single-Molecule Analysis at High Concentrations. *Science* **299**, 682–686 (2003).
25. Holzmeister, P., Acuna, G. P., Grohmann, D. & Tinnefeld, P. Breaking the concentration limit of optical single-molecule detection. *Chem. Soc. Rev.* **43**, 1014–1028 (2014).
26. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
27. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).
28. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nat. Commun.* **4**, (2013).
29. Cordes, T. *et al.* Sensing DNA Opening in Transcription Using Quenchable Förster Resonance Energy Transfer. *Biochemistry* **49**, 9171–9180 (2010).
30. Johnson, S. J., Taylor, J. S. & Beese, L. S. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3895–3900 (2003).
31. Wu, E. Y. & Beese, L. S. The Structure of a High Fidelity DNA Polymerase Bound to a Mismatched Nucleotide Reveals an ‘Ajar’ Intermediate Conformation in the Nucleotide Selection Mechanism. *J. Biol. Chem.* **286**, 19758–19767 (2011).
32. Moerner, W. E. & Fromm, D. P. Methods of single-molecule fluorescence spectroscopy and microscopy. *Rev. Sci. Instrum.* **74**, 3597–3619 (2003).
33. Holden, S. J. *et al.* Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophys. J.* **99**, 3102–3111 (2010).
34. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).
35. Vogelsang, J. *et al.* A Reducing and Oxidizing System Minimizes Photobleaching and Blinking of Fluorescent Dyes. *Angew. Chem. Int. Ed.* **47**, 5465–5469 (2008).
36. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *J. Am. Chem. Soc.* **131**, 5018–5019 (2009).
37. Ha, T. & Tinnefeld, P. Photophysics of Fluorescence Probes for Single Molecule Biophysics and Super-Resolution Imaging. *Annu. Rev. Phys. Chem.* **63**, 595–617 (2012).
38. Zheng, Q. *et al.* Ultra-Stable Organic Fluorophores for Single-Molecule Research. *Chem. Soc. Rev.* **43**, 1044–1056 (2014).

39. Watson, J. D. & Crick, F. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
40. Lehman, I. R., Bessman, M. J., Simms, E. S. & Kornberg, A. Enzymatic Synthesis of Deoxyribonucleic Acid I. PREPARATION OF SUBSTRATES AND PARTIAL PURIFICATION OF AN ENZYME FROM ESCHERICHIA COLI. *J. Biol. Chem.* **233**, 163–170 (1958).
41. Hübscher, U., Spadari, S., Villani, G. & Maga, G. in *DNA Polymerases* 1–58 (World Scientific, 2010).
42. Hastings, P. J. *et al.* Competition of Escherichia coli DNA Polymerases I, II and III with DNA Pol IV in Stressed Cells. *PLoS ONE* **5**, (2010).
43. Hübscher, U., Spadari, S., Villani, G. & Maga, G. in *DNA Polymerases* 59–83 (World Scientific, 2010).
44. Klenow, H. & Henningsen, I. Selective Elimination of the Exonuclease Activity of the Deoxyribonucleic Acid Polymerase from Escherichia coli B by Limited Proteolysis\*. *Proc. Natl. Acad. Sci. U. S. A.* **65**, 168–175 (1970).
45. Joyce, C. M. *et al.* Fingers-Closing and Other Rapid Conformational Changes in DNA Polymerase I (Klenow Fragment) and Their Role in Nucleotide Selectivity. *Biochemistry* **47**, 6103–6116 (2008).
46. Christian, T. D., Romano, L. J. & Rueda, D. Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc. Natl. Acad. Sci.* **106**, 21109–21114 (2009).
47. Hwang, H., Kim, H. & Myong, S. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7414–7418 (2011).
48. Markiewicz, R. P., Vrtis, K. B., Rueda, D. & Romano, L. J. Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucleic Acids Res.* **40**, 7975–7984 (2012).
49. Hwang, H. & Myong, S. Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chem. Soc. Rev.* **43**, 1221–1229 (2014).
50. Santoso, Y. *et al.* Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 715–720 (2010).
51. Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying Molecular Dynamics in Single-Molecule FRET Experiments with Burst Variance Analysis. *Biophys. J.* **100**, 1568–1577 (2011).
52. Berezhna, S. Y., Gill, J. P., Lamichhane, R. & Millar, D. P. Single-Molecule Förster Resonance Energy Transfer Reveals an Innate Fidelity Checkpoint in DNA Polymerase I. *J. Am. Chem. Soc.* **134**, 11261–11268 (2012).
53. Rothwell, P. J. *et al.* dNTP-dependent conformational transitions in the fingers subdomain of Klenotaq1 DNA polymerase: insights into the role of the ‘nucleotide-binding’ state. *J. Biol. Chem.* **288**, 13575–13591 (2013).



54. Antonik, M., Felekyan, S., Gaiduk, A. & Seidel, C. A. M. Separating Structural Heterogeneities from Stochastic Variations in Fluorescence Resonance Energy Transfer Distributions via Photon Distribution Analysis. *J. Phys. Chem. B* **110**, 6970–6978 (2006).
55. Nir, E. *et al.* Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. *J. Phys. Chem. B* **110**, 22103–22124 (2006).
56. Kalinin, S., Felekyan, S., Antonik, M. & Seidel, C. A. M. Probability Distribution Analysis of Single-Molecule Fluorescence Anisotropy and Resonance Energy Transfer. *J. Phys. Chem. B* **111**, 10253–10262 (2007).
57. Kalinin, S., Felekyan, S., Valeri, A. & Seidel, C. A. M. Characterizing Multiple Molecular States in Single-Molecule Multiparameter Fluorescence Detection by Probability Distribution Analysis. *J. Phys. Chem. B* **112**, 8361–8374 (2008).
58. Santoso, Y., Torella, J. P. & Kapanidis, A. N. Characterizing Single-Molecule FRET Dynamics with Probability Distribution Analysis. *ChemPhysChem* **11**, 2209–2219 (2010).
59. Lamichhane, R., Berezhna, S. Y., Gill, J. P., Van der Schans, E. & Millar, D. P. Dynamics of Site Switching in DNA Polymerase. *J. Am. Chem. Soc.* **135**, 4735–4742 (2013).
60. Zhang, Y. *et al.* Structural Basis of Transcription Initiation. *Science* **338**, 1076–1080 (2012).
61. Hahn, S. Structure and mechanism of the RNA Polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**, 394–403 (2004).
62. Saecker, R. M., Record, M. T. & deHaseth, P. L. Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis. *J. Mol. Biol.* **412**, 754–771 (2011).
63. Kapanidis, A. N. *et al.* Initial Transcription by RNA Polymerase Proceeds Through a DNA-Scrunching Mechanism. *Science* **314**, 1144–1147 (2006).
64. Carpousis, A. J. & Gralla, J. D. Interaction of RNA polymerase with lacUV5 promoter DNA during mRNA initiation and elongation: Footprinting, methylation, and rifampicin-sensitivity changes accompanying transcription initiation. *J. Mol. Biol.* **183**, 165–177 (1985).
65. Straney, D. C. & Crothers, D. M. A stressed intermediate in the formation of stably initiated RNA chains at the Escherichia coli lac UV5 promoter. *J. Mol. Biol.* **193**, 267–278 (1987).
66. Krummel, B. & Chamberlin, M. J. RNA chain initiation by Escherichia coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry* **28**, 7829–7842 (1989).
67. Hsu, L. M. Promoter clearance and escape in prokaryotes. *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* **1577**, 191–207 (2002).

68. Pal, M., Ponticelli, A. S. & Luse, D. S. The Role of the Transcription Bubble and TFIIB in Promoter Clearance by RNA Polymerase II. *Mol. Cell* **19**, 101–110 (2005).
69. Zhang, G. *et al.* Crystal Structure of *Thermus aquaticus* Core RNA Polymerase at 3.3 Å Resolution. *Cell* **98**, 811–824 (1999).
70. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural Basis of Transcription: RNA Polymerase II at 2.8 Ångstrom Resolution. *Science* **292**, 1863–1876 (2001).
71. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 Å Resolution. *Science* **292**, 1876–1882 (2001).
72. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex. *Science* **296**, 1285–1290 (2002).
73. Chakraborty, A. *et al.* Opening and Closing of the Bacterial RNA Polymerase Clamp. *Science* **337**, 591–595 (2012).
74. Andrecka, J. *et al.* Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl. Acad. Sci.* **105**, 135–140 (2008).
75. Muschielok, A. *et al.* A nano-positioning system for macromolecular structural analysis. *Nat. Methods* **5**, 965–971 (2008).
76. Runner, V. M., Podolny, V. & Buratowski, S. The Rpb4 Subunit of RNA Polymerase II Contributes to Cotranscriptional Recruitment of 3' Processing Factors. *Mol. Cell. Biol.* **28**, 1883–1891 (2008).
77. Treutlein, B. *et al.* Dynamic Architecture of a Minimal RNA Polymerase II Open Promoter Complex. *Mol. Cell* **46**, 136–146 (2012).
78. Lohman, T. M. *Escherichia coli* DNA helicases: mechanisms of DNA unwinding. *Mol. Microbiol.* **6**, 5–14 (1992).
79. Yodh, J. G., Schlierf, M. & Ha, T. Insight into Helicase Mechanism and Function Revealed through Single-Molecule Approaches. *Q. Rev. Biophys.* **43**, 185–217 (2010).
80. Ha, T. *et al.* Initiation and re-initiation of DNA unwinding by the *Escherichia coli* Rep helicase. *Nature* **419**, 638–641 (2002).
81. Rasnik, I., Myong, S., Cheng, W., Lohman, T. M. & Ha, T. DNA-binding Orientation and Domain Conformation of the *E. coli* Rep Helicase Monomer Bound to a Partial Duplex Junction: Single-molecule Studies of Fluorescently Labeled Enzymes. *J. Mol. Biol.* **336**, 395–408 (2004).
82. Myong, S., Rasnik, I., Joo, C., Lohman, T. M. & Ha, T. Repetitive shuttling of a motor protein on DNA. *Nature* **437**, 1321 (2005).
83. Myong, S., Bruno, M. M., Pyle, A. M. & Ha, T. Spring-Loaded Mechanism of DNA Unwinding by Hepatitis C Virus NS3 Helicase. *Science* **317**, 513–516 (2007).

84. Smiley, R. D., Collins, T. R. L., Hammes, G. G. & Hsieh, T.-S. Single-molecule measurements of the opening and closing of the DNA gate by eukaryotic topoisomerase II. *Proc. Natl. Acad. Sci.* **104**, 4840–4845 (2007).
85. Gubaev, A., Hilbert, M. & Klostermeier, D. The DNA-gate of *Bacillus subtilis* gyrase is predominantly in the closed conformation during the DNA supercoiling reaction. *Proc. Natl. Acad. Sci.* **106**, 13278–13283 (2009).
86. Lanz, M. A. & Klostermeier, D. Guiding strand passage: DNA-induced movement of the gyrase C-terminal domains defines an early step in the supercoiling cycle. *Nucleic Acids Res.* **39**, 9681–9694 (2011).
87. Lanz, M. A. & Klostermeier, D. The GyrA-box determines the geometry of DNA bound to gyrase and couples DNA binding to the nucleotide cycle. *Nucleic Acids Res.* **40**, 10893–10903 (2012).
88. Lee, S. *et al.* DNA cleavage and opening reactions of human topoisomerase II $\alpha$  are regulated via Mg<sup>2+</sup>-mediated dynamic bending of gate-DNA. *Proc. Natl. Acad. Sci.* **109**, 2925–2930 (2012).
89. Pinkney, J. N. M. *et al.* Capturing reaction paths and intermediates in Cre-loxP recombination using single-molecule fluorescence. *Proc. Natl. Acad. Sci.* **109**, 20871–20876 (2012).
90. Zawadzki, P. *et al.* Conformational transitions during FtsK translocase activation of individual XerCD-dif recombination complexes. *Proc. Natl. Acad. Sci.* **110**, 17302–17307 (2013).
91. Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature* **408**, 307–310 (2000).
92. Lymperopoulos, K. *et al.* Single-Molecule DNA Biosensors for Protein and Ligand Detection. *Angew. Chem. Int. Ed.* **49**, 1316–1320 (2010).
93. Crawford, R. *et al.* Non-covalent Single Transcription Factor Encapsulation Inside a DNA Cage. *Angew. Chem. Int. Ed.* **52**, 2284–2288 (2013).
94. Crawford, R., Kelly, D. J. & Kapanidis, A. N. A Protein Biosensor That Relies on Bending of Single DNA Molecules. *ChemPhysChem* **13**, 918–922 (2012).
95. Hohng, S., Lee, S., Lee, J. & Hyun Jo, M. Maximizing information content of single-molecule FRET experiments: multi-color FRET and FRET combined with force or torque. *Chem. Soc. Rev.* **43**, 1007–1013 (2014).
96. Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56**, 930–933 (1986).
97. Hugel, T. *et al.* Single-Molecule Optomechanical Cycle. *Science* **296**, 1103–1106 (2002).
98. Sarkar, A., Robertson, R. B. & Fernandez, J. M. Simultaneous atomic force microscope and fluorescence measurements of protein unfolding using a calibrated evanescent wave. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12882–12886 (2004).



99. Shaw, J. E., Oreopoulos, J., Wong, D., Hsu, J. C. Y. & Yip, C. M. Coupling evanescent-wave fluorescence imaging and spectroscopy with scanning probe microscopy: challenges and insights from TIRF–AFM. *Surf. Interface Anal.* **38**, 1459–1471 (2006).
100. Vickery, S. A. & Dunn, R. C. Combining AFM and FRET for high resolution fluorescence microscopy. *J. Microsc.* **202**, 408–412 (2001).
101. Nakamura, C. *et al.* Enzymatic nanolithography of FRET peptide layer using V8 protease-immobilized AFM probe. *Biosens. Bioelectron.* **22**, 2308–2314 (2007).
102. He, Y., Lu, M., Cao, J. & Lu, H. P. Manipulating Protein Conformations by Single-Molecule AFM-FRET Nanoscopy. *ACS Nano* **6**, 1221–1229 (2012).
103. Lu, H. P. Sizing up single-molecule enzymatic conformational dynamics. *Chem. Soc. Rev.* **43**, 1118–1143 (2014).
104. Lang, M. J., Fordyce, P. M. & Block, S. M. Combined optical trapping and single-molecule fluorescence. *J. Biol.* **2**, 6 (2003).
105. Brau, R. R., Tarsa, P. B., Ferrer, J. M., Lee, P. & Lang, M. J. Interlaced Optical Force-Fluorescence Measurements for Single Molecule Biophysics. *Biophys. J.* **91**, 1069–1077 (2006).
106. Tarsa, P. B. *et al.* Detecting Force-Induced Molecular Transitions with Fluorescence Resonant Energy Transfer. *Angew. Chem. Int. Ed.* **46**, 1999–2001 (2007).
107. Hohng, S. *et al.* Fluorescence-Force Spectroscopy Maps Two-Dimensional Reaction Landscape of the Holliday Junction. *Science* **318**, 279–283 (2007).
108. Zhou, R. *et al.* SSB Functions as a Sliding Platform that Migrates on DNA via Reptation. *Cell* **146**, 222–232 (2011).
109. Shroff, H. *et al.* Biocompatible Force Sensor with Optical Readout and Dimensions of 6 nm<sup>3</sup>. *Nano Lett.* **5**, 1509–1514 (2005).
110. Lee, M., Kim, S. H. & Hong, S.-C. Minute negative superhelicity is sufficient to induce the B-Z transition in the presence of low tension. *Proc. Natl. Acad. Sci.* **107**, 4985–4990 (2010).
111. Long, X., Parks, J. W., Bagshaw, C. R. & Stone, M. D. Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy. *Nucleic Acids Res.* **41**, 2746–2755 (2013).
112. Lee, N. K. *et al.* Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys. J.* **88**, 2939–2953 (2005).
113. Craggs, T. D. & Kapanidis, A. N. Six steps closer to FRET-driven structural biology. *Nat. Methods* **9**, 1157 (2012).
114. Muschielok, A. & Michaelis, J. Application of the Nano-Positioning System to the Analysis of Fluorescence Resonance Energy Transfer Networks. *J. Phys. Chem. B* **115**, 11927–11937 (2011).

- 115.Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).
- 116.Rothwell, P. J. *et al.* Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. *Proc. Natl. Acad. Sci.* **100**, 1655–1660 (2003).
- 117.Widengren, J. *et al.* Single-Molecule Detection and Identification of Multiple Species by Multiparameter Fluorescence Detection. *Anal. Chem.* **78**, 2039–2050 (2006).
- 118.Peletskaya, E. N., Kogon, A. A., Tuske, S., Arnold, E. & Hughes, S. H. Nonnucleoside Inhibitor Binding Affects the Interactions of the Fingers Subdomain of Human Immunodeficiency Virus Type 1 Reverse Transcriptase with DNA. *J. Virol.* **78**, 3387–3397 (2004).





# A single-molecule FRET sensor for monitoring DNA synthesis in real time

Carel Fijen<sup>†</sup>, Alejandro Montón Silva<sup>††</sup>, Alejandro Hochkoeppler<sup>†</sup>, Johannes Hohlbein<sup>†</sup>, *Physical Chemistry Chemical Physics*, 2017, 19, pp 4222-4230. DOI: 10.1039/C6CP05919H



## Abstract

We developed a versatile DNA assay and framework for monitoring polymerization of DNA in real time and at the single-molecule level. The assay consists of an acceptor labelled DNA primer annealed to a DNA template that is labelled on its single stranded, downstream overhang with a donor fluorophore. Upon extension of the primer using a DNA polymerase, the overhang of the template alters its conformation from a random coil to the canonical structure of double stranded DNA. This conformational change increases the distance between the donor and the acceptor fluorophore and can be detected as a decrease in the Förster resonance energy transfer (FRET) efficiency between both fluorophores. Remarkably, the DNA assay does not require any modification of the DNA polymerase and albeit the simple and robust spectroscopic readout facilitates measurements even with conventional fluorimeters or stopped-flow equipment, single-molecule FRET provides additional access to parameters such as the processivity of DNA synthesis and, for one of the three DNA polymerases tested, the detection of binding and dissociation of the DNA polymerase to DNA. We furthermore demonstrate that primer extensions by a single base can be resolved.

## Introduction

DNA polymerases play a pivotal role in DNA replication and DNA repair<sup>1-3</sup>. The complexity of the cellular environment and resultant requirements on adaptability and specificity led to a profound differentiation of DNA polymerases. In humans, 18 different DNA polymerases are currently known that all catalyse the stepwise addition of deoxyribonucleotides to a growing DNA chain<sup>4</sup> but differ greatly in their specific properties. In addition to their fundamental role in living organisms, DNA polymerases have been utilised in many biotechnological applications ranging from polymerase chain reaction (PCR) for amplification of DNA fragments<sup>5</sup> to DNA sequencing<sup>6,7</sup>. Different techniques have been developed to characterize DNA polymerases in terms of their fidelity, processivity and rate of DNA synthesis. At the ensemble level, electrophoretic mobility shift assays (EMSA)<sup>8</sup> and stopped-flow techniques are widely used for quantifying the kinetics of DNA polymerization. Recent fluorescence-based ensemble assays to monitor polymerase activity include monitoring the incorporation of dye-labelled nucleotides into DNA<sup>9</sup> and surface-induced fluorescence quenching of a labelled DNA template on a chip<sup>10</sup>.

In the last two decades, techniques with single-molecule sensitivity and resolution allowed to overcome temporal and spatial averaging inherent to ensemble based characterisations<sup>11,12</sup>. For DNA polymerases, different experimental designs have been applied ranging from optical or magnetic traps<sup>13</sup> (reviewed by Heller *et al.*<sup>14</sup>) to

conductivity measurements on protein nanopores<sup>15,16</sup> and applications utilising Förster resonance energy transfer (FRET), a process in which the distance-dependent energy transfer from a donor fluorophore to an acceptor chromophore allows to resolve changes of distances in the nanometre range<sup>17,18</sup>. Single-molecule FRET experiments on DNA polymerases have been designed to measure the rate of DNA polymerization by strand-displacing DNA polymerases<sup>19</sup>, to identify sliding characteristics and binding orientations of HIV reverse transcriptase<sup>20,21</sup>, to identify translocation of DNA polymerases with single base pair resolution<sup>22</sup> and to determine conformational dynamics and the free-energy landscapes of pre-chemistry nucleotide selection in *E.coli* DNA polymerase I<sup>23,24</sup>. Most of the experimental approaches, however, required the positioning of one or two fluorescent labels on the polymerase, which is a time consuming and complicated procedure.

Moreover, DNA polymerase-based DNA sequencing was demonstrated at the single-molecule level using a sequential approach involving the successive addition of dye-labelled nucleotides<sup>25</sup> and non-sequentially by utilising so-called zero-mode waveguides as optical waveguides that provide a confined excitation and detection volume which is small enough to monitor polymerization reactions in presence of all four nucleotides and in real time<sup>6</sup>. As each of the four dNTPs is conjugated to a different fluorescent dye, the temporal immobilisation of the nucleotide during synthesis is sufficient for base calling. Alternative approaches for DNA sequencing utilise a polymerase-template complex attached to the tip of an atomic force microscope to sample different nucleotide pools<sup>26</sup> or measuring nucleotide-dependent binding kinetics of singly labelled DNA polymerases<sup>27</sup>.

Despite the progress, a simple and versatile assay for characterizing DNA polymerization at the single-molecule level, in real time and without requiring fluorescently labelled DNA polymerases is still missing. Here, we propose a DNA assay that allows probing both the speed of DNA polymerization and processivity, and for one species, even the binding and dissociation constants of the DNA polymerases to DNA. Whereas a comparable assay has been successfully used to screen for HIV-1 reverse transcriptase on the ensemble level<sup>28</sup>, further optimization of the assay and its application at the single molecule level allowed us to study dynamic and static heterogeneity with up to single base pair resolution. We compared different polymerases, namely the  $\alpha$  subunit of *E. coli* DNA polymerase III (here after POLIII $\alpha$ ), *E. coli* DNA polymerase I Klenow fragment (KF) and human DNA polymerase beta (POLB). Albeit we performed the measurements at the single-molecule level, the design is suitable for any instrument featuring a fluorescent read out mode.

## Materials and Methods

### Design principles of the DNA sensor

Our DNA assay consists of a DNA primer labelled with an acceptor fluorophore annealed to a DNA template labelled with a donor fluorophore (Fig. 3.1a), inspired by DNA constructs featuring recessed DNA that have been used before to assess SSB protein binding<sup>29</sup>, flexibility of single-stranded DNA<sup>30</sup> or dynamics of the HIV transcriptase<sup>20</sup>. As the single stranded overhang of the DNA template is randomly coiled, we expect a high FRET efficiency  $E^*$  defined as

$$E^* = f_{Dex}^{Aem} / (f_{Dex}^{Dem} + f_{Dex}^{Aem})$$

where  $f_{Dex}^{Dem}$  represents the donor emission intensity after donor excitation and  $f_{Dex}^{Aem}$  the acceptor emission intensity after donor excitation. Upon addition of DNA polymerase, interactions between the DNA polymerase and the single stranded template might change the distance between donor and acceptor, thereby modulating the FRET efficiency (Fig. 3.1b). The addition of dNTPs starts the DNA polymerization reaction that decreases the FRET efficiency as the distance between the fluorophores increases (Fig. 3.1c). After complete polymerization, the extended DNA molecule is rigid, yielding the lowest possible FRET signature depending on the chosen labelling positions (Fig. 3.1d and e). All experiments were performed at room temperature (21 °C).

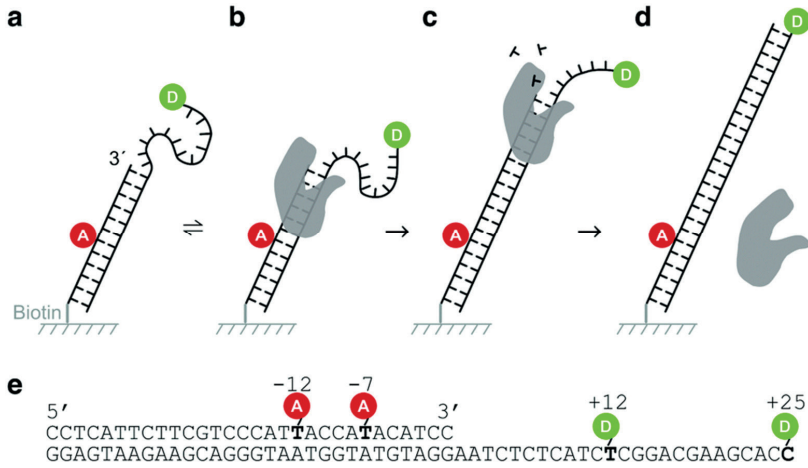
### DNA polymerases

We used three different DNA polymerases: Klenow fragment from *E. coli* DNA Polymerase I (KF), human Polymerase Beta (POLB) and the  $\alpha$  subunit from *E. coli* DNA Polymerase III (POLIII $\alpha$ ). Both KF and POLB are involved in DNA repair and have been studied extensively<sup>1,31-34</sup>. DNA polymerase III, on the other hand, is a large enzyme complex responsible for DNA replication<sup>35-37</sup>. Its  $\alpha$  subunit (POLIII $\alpha$ ) possesses polymerase activity, but no exonuclease or proof-reading activity.

### DNA sequences and modifications

Oligonucleotides were ordered from IBA, Germany. 5-Amino-C6 modifications, internal (on dT) or terminal (on dC), were used to label positions with NHS-reactive dyes. A 30-mer primer sequence (biotin-5'-CCTCATTCTTCGTCGCCATTACCATACATCC-3') was labelled with ATTO647N at either position -12 or -7, counted from the 3' end. This primer was annealed to a 55-mer template (3'-GGAGTAAGAAGCAGGGTAATGGTATGTAGGAATCTCTCATC-TCGGACGAAGCACC-5'), labelled with Cy3B at either position +12 or +25 of the overhang. We used two labelling configurations: +25/-7 (donor/acceptor) and +12/-12 (donor/acceptor) (Fig. 3.1e).





**Fig. 3.1** Design of the DNA sensor. (a) The DNA sensor consists of acceptor labelled DNA primer and a donor labelled DNA template immobilized on a cover slip using a biotin/neutravidin/biotin linkage. (b) Addition of polymerases establishes a dynamic equilibrium in which polymerases bind and dissociate from the DNA sensor. Interactions between the single-stranded overhang and the polymerase may cause a change in the distance between the donor and the acceptor. (c) Upon addition of nucleotides, the polymerase starts the polymerization reaction extending the DNA primer. During the process, the overhang is taken up in a rigid DNA helix, increasing the distance between donor and acceptor. (d) The final DNA shows a lower transfer efficiency between the donor and the acceptor than the native DNA sensor. (e) Sequence of the DNA sensor with indicated labelling positions. We used the combinations of +12/-12 and +25/-7 for our experiments.

### Total-internal reflection fluorescence (TIRF) microscopy

For imaging, we used the alternating-laser excitation scheme (ALEX)<sup>38–40</sup> in combination with TIRF microscopy on a home-built microscope described in more detail elsewhere<sup>41</sup>. Briefly, our microscope is equipped with a laser engine (Omicron, Germany), a 100x oil immersion objective with N.A. = 1.49 (Nikon, Japan), and an Ixon Ultra 897 emCCD camera for image acquisition (Andor, UK). The sensors were immobilized on PEGylated coverslips through a neutravidin-biotin interaction. All samples were measured in either gasket wells or flow channels formed by sticky-slides VI<sup>0.4</sup> (Ibidi, Germany). Images of immobilized DNA sensors were taken at a rate of 20 frames per second, using excitation powers of 0.75 mW (642 nm) and 1.5 mW (561 nm). Imaging was performed in a buffer containing 50 mM Tris HCl (pH7.5), 100 µg mL<sup>-1</sup> BSA, 10 mM MgCl<sub>2</sub>, 5% glycerol, 1 mM DTT, 1 mM Trolox, 1% glucose oxidase/catalase and 1% glucose. Trolox is a triplet state quencher, while glucose oxidase and catalase serve as an oxygen scavenger system; both prevent premature bleaching of the fluorophores<sup>42,43</sup>. After addition of the imaging buffer, DNA sensors were imaged for 50 seconds per movie (DNA binding studies) or for 500 seconds per

movie (real-time polymerization studies). To increase the total number of observed molecules in binding studies, three consecutive movies were recorded for each sample condition.

#### **Analysis of completed DNA synthesis: end-point assays**

Polymerases (40 nM), dNTPs (100  $\mu$ M each) and sensors (10 nM) were mixed together in an aliquot using a buffer resembling our imaging buffer, but without the oxygen scavenger system. A mixture without any dNTPs was used as a negative control. To ensure the completion of all polymerization reactions, the mixtures were left to react at RT for 90 minutes, after which we stopped the reaction by adding EDTA to a final concentration of 0.1 M. For data acquisition, the DNA products were immobilized and imaged as described above.

In a similar experiment, which was only performed for KF, polymerization reactions were stopped at specific positions along the template by adding only a subset of the dNTPs needed for full polymerization.

#### **Monitoring DNA binding and DNA polymerization**

For DNA binding studies, we mixed the respective polymerases into the imaging buffer (final concentrations: KF 500 pM, POLIII $\alpha$  2 nM and POLB 10 nM). The concentration of KF was chosen such that about 50% of all molecules are bound to DNA as verified using titrations of the DNA sensors with polymerase (Fig. S3.1, Supporting Information). The concentrations of POLIII $\alpha$  and POLB for the DNA binding studies were chosen such that most time traces of the DNA polymerization experiments showed DNA synthesis implying that binding at these concentrations occurs.

For real-time polymerization experiments, polymerases and dNTPs were directly provided in the imaging buffer. Since polymerization events can happen shortly after adding DNA polymerases and dNTPs, we started data acquisition of the immobilised molecules in standard imaging buffer before quickly adding polymerases and dNTPs during acquisition in a similar volume of imaging buffer. Final concentrations in the sample were identical to the ones chosen for the binding experiments. dNTPs were added to a concentration of 100  $\mu$ M each, except for a set of experiments in which dNTPs were added to successively create artificial pausing sites.

At the ensemble level, the polymerization of the +25/-7 sensor by KF was recorded in a 96 wells plate, using a SpectraMax M2 plate reader. For experimental details, please refer to the Supporting Information.

### **Time trace analysis of single-molecule FRET data**

Analysis of the acquired movies was performed in MATLAB (MathWorks, UK) using custom-built software packages<sup>44</sup> yielding time traces of both donor and acceptor intensity, as well as FRET efficiency  $E^*$  and stoichiometry  $S$  (Supporting Information) for each DNA. For calculation of exact distances within the polymerized DNA construct,  $E^*$  was corrected for background, cross-talk between the channels, detection efficiencies of the dyes and quantum yield (Supporting Information). Theoretical distances were calculated based on a static model of the polymerized DNA construct<sup>45</sup> (3D-DART). We used the FPS software developed by the Seidel lab<sup>46</sup> to model the dye clouds representing possible positions of the fluorophores attached to modified bases around the DNA duplex and calculated the predicted inter-dye distance  $\langle RDA \rangle_E$ .

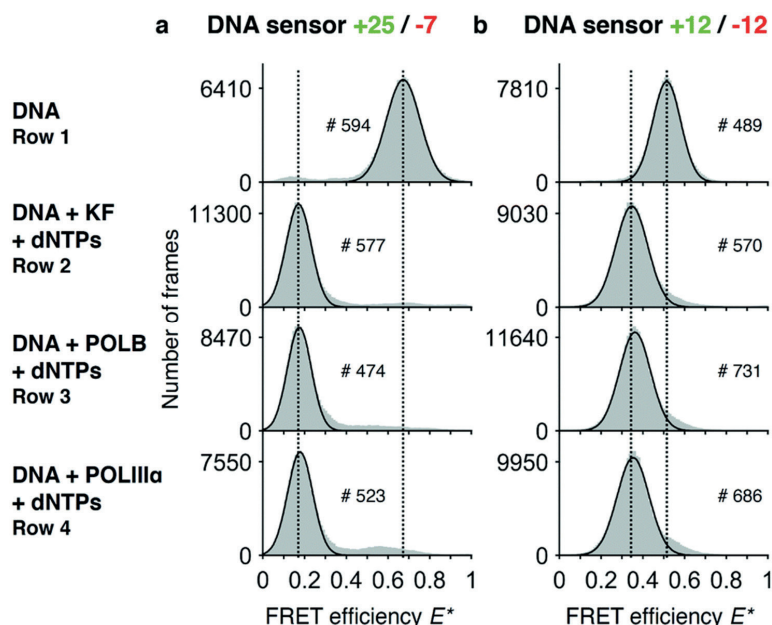
### **Hidden Markov modelling (HMM)**

Single-molecule time traces of binding and polymerization experiments were analysed with the ebFRET software package<sup>47</sup>. Using a Bayesian approach, ebFRET obtains the best fit *via* an iterative process after introducing a prior distribution of parameters such as mean  $E^*$  values, dwell times and a general noise level. We used ebFRET to determine the number of steps involved in the polymerization of individual sensors as follows. HMM states lasting shorter than 1 s were removed from further analysis. To calculate the duration  $\Delta t_i = t_2 - t_1$  of individual polymerization events  $i$ , we assigned the first intersection of the high FRET HMM state with the decreasing FRET efficiency as  $t_1$ . The intersection of the decreasing FRET efficiency with the next lower FRET HMM state was taken as  $t_2$ . We repeated that procedure for cases in which more than one polymerization step was visible and summed up the polymerization times for each individual time trace. For our +12/-12 polymerization sensor in combination with measurements using KF, we used a slightly modified approach to account for the presence of a short increase in  $E^*$  just before each polymerization event. In these cases,  $t_1$  was defined as the last position of the high FRET peak before the FRET decrease due to DNA synthesis. For further information on HMM the reader is referred to the Supporting Information.

## **Results**

### **DNA sensors indicate completion of DNA synthesis**

To test whether our DNA sensors can be polymerized by KF, POLB and POLIII $\alpha$ , we mixed the sensors with the respective polymerases and dNTPs in an aliquot and imaged the DNA products after an incubation of 90 minutes at room temperature using our TIRF microscope ("end-point assays"). The native +25/-7 sensor showed a single peak at  $E^* = 0.65$ , while the +12/-12 sensor shows a peak at  $E^* = 0.52$ , both representing the respective  $E^*$  value before DNA polymerization (Fig. 3.2a and b, row



**Fig. 3.2** Assay to test completion of DNA synthesis. Histograms of FRET efficiencies  $E^*$  for the +25/-7 (a) and the +12/-12 (b) DNA sensor. Each histogram (100 bins shown in grey) was derived from individual time traces obtained from fitting an indicated number of molecules (#). The histograms were fitted to a single-Gaussian model (black line) to indicate the mean peak position of the dominant FRET species. Vertical dashed lines were added for visual guidance. First row: FRET efficiencies of the native DNA sensor. Second to fourth row: FRET efficiencies after incubating the DNA sensor for 90 minutes in presence of dNTPs and DNA polymerases (second row: DNA polymerase I, Klenow fragment; third row: DNA polymerase beta; fourth row: DNA polymerase III alpha subunit). The histograms indicate successful polymerization of DNA as the FRET efficiencies are shifted towards lower values. Peak positions of the fits and the corresponding errors can be found in Supporting Information Table S3.1.

1). Sensors that were incubated in presence of polymerases and dNTPs showed a peak at lower FRET efficiency than the native sensors ( $E^* = 0.17$  for +25/-7, and  $E^* = 0.35$  for +12/-12) (Fig. 3.2a and b, rows 2-4). The +25/-7 sensor showed the largest relative change in FRET efficiency of  $\Delta E^* = 0.48$  (+12/-12:  $\Delta E^* = 0.17$ ).

As expected, we did not see any significant differences in the  $E^*$  histograms of the various DNA polymerases after DNA synthesis and we note that >90% of all DNA molecules show a shift towards the lower FRET conformation. We calculated the distances between the fluorophores using accurate FRET ( $\langle RDA \rangle_{E^*,+12/-12} = 7.1$  nm,  $\langle RDA \rangle_{E^*,+25/-7} = 10.7$  nm) and compared these values with the distances computationally derived from the structural model of a double stranded DNA helix

( $\langle \text{RDA} \rangle_{E, \text{mod}, +12/-12} = 7.3 \text{ nm}$ ,  $\langle \text{RDA} \rangle_{E, \text{mod}, +25/-7} = 10.7 \text{ nm}$ ). Moreover, we found similar distances when measuring a separately ordered dsDNA construct with the same sequence as our fully polymerized DNA primer (Supporting Information Table S3.4). These values suggest a successful and complete polymerization of the DNA sensors.

### **KF induces conformational changes upon binding to the DNA template strand**

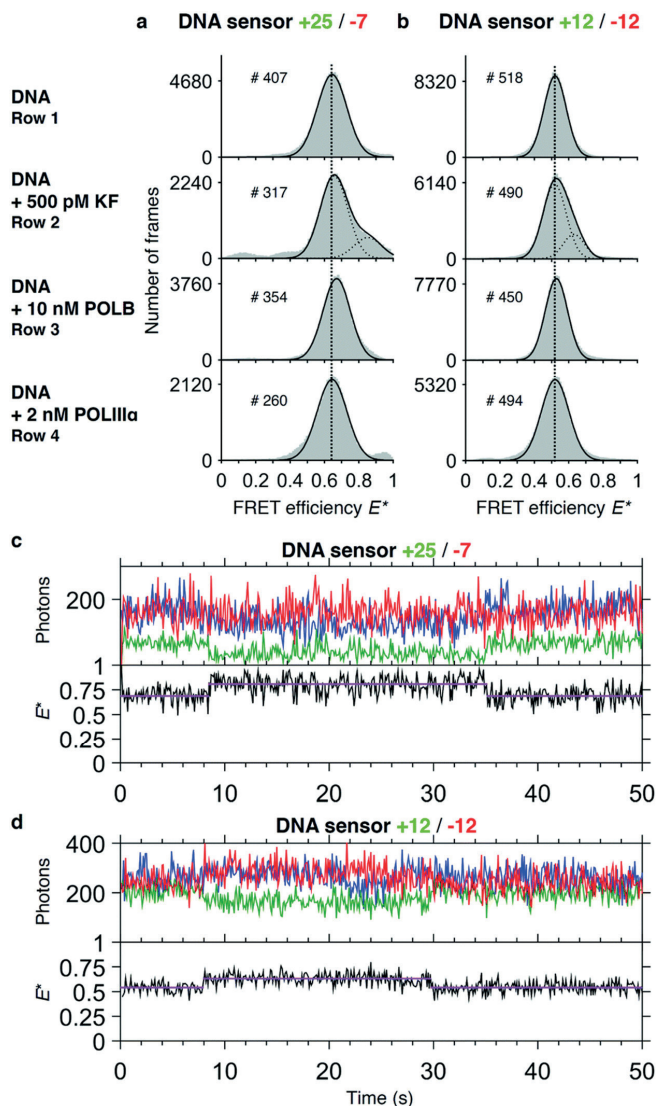
We asked whether the sensors show a change of FRET efficiency upon binding of DNA polymerases in the absence of nucleotides. Of the three DNA polymerases tested at the concentrations mentioned above, only binding of KF showed a new species in the FRET histogram with an increased  $E^*$  (+25/-7:  $E_{\text{DNA}}^* = 0.65$  and  $E_{\text{DNA}+\text{KF}}^* = 0.85$ ; +12/-12:  $E_{\text{DNA}}^* = 0.52$  and  $E_{\text{DNA}+\text{KF}}^* = 0.65$ ) (Fig. 3.3a and b).

The increase in  $E^*$  of POLB of the major species using the +25/-7 sensor is too small to be clearly attributed to a conformational change upon binding. For both POLB and POLIII $\alpha$  we further tested higher concentrations of the polymerase without seeing any response on our DNA sensor (data not shown).

For KF, the population of sensors in the higher FRET state is concentration dependent: during titrations, the plateau value for the percentage of sensors bound to KF was found to be ~50% for the +25/-7 sensor and ~60% for the +12/-12 sensor (Fig. S3.1, Supporting Information), well below the expected maximum of 100% for the high-FRET state. We therefore hypothesise that different binding modes must exist which cannot be detected by a change in FRET efficiency. The titrations revealed apparent dissociation constants of 0.41 nM (sensor +25/-7) and 0.12 nM (sensor +12/-12) indicating strong binding of KF to the DNA (Fig. S3.1, Supporting Information).

For KF, analysis of the corresponding single-molecule time traces revealed dynamic switching between high and low FRET efficiencies (Fig. 3.3c and d), implying that binding and dissociation occurs at a timescale of seconds. As can be seen in the trace of the +25/-7 sensor, which was designed to maximize  $\Delta E$  upon polymerization, the acceptor fluorescence after direct acceptor excitation appears to be slightly quenched upon binding of KF. During the binding event, the FRET signal (red line) remained constant, but the donor fluorescence signal decreased leading to an effective increase in FRET efficiency. As a consequence of this convolution of acceptor quenching and a change in the distances between the fluorophores, the correct interpretation of the FRET efficiency is complicated<sup>48</sup>. We attribute the acceptor quenching to the proximity of the -7 labelling position to the polymerase binding site, which causes an interaction between the dye and the protein indicated by a slightly higher dissociation constant of KF for this sensor ( $K_d = 0.41 \text{ nM}$  versus  $K_d = 0.12 \text{ nM}$  for the +12/-12 sensor, Fig. S1, Supporting Information). Our attempts to perform global dwell time analysis on the binding events remained inconclusive, however, likely due to a combination of

relatively long dwell times ( $>1$  s) with small changes in  $\Delta E^*$  ( $<0.2$ ) and base line FRET values changing slightly between individual traces.



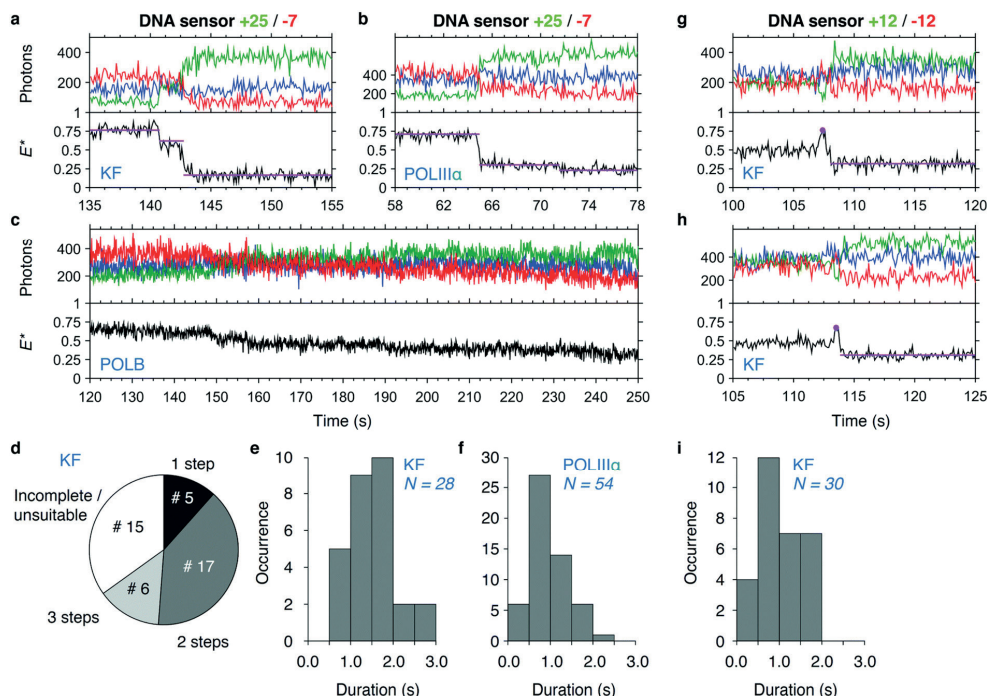
**Fig. 3.3** Binding studies. Histograms of FRET efficiencies  $E^*$  for the +25/-7 (a) and the +12/-12 (b) DNA sensor. Each histogram (100 bins shown in grey) was derived from individual time traces obtained from fitting an indicated number of molecules (#). The histograms were fitted to a single-Gaussian model (black line) to indicate the mean peak position of the dominant FRET species. Vertical dashed lines were added for visual guidance. First row: FRET efficiencies of the native DNA sensor. Second to fourth row: FRET efficiencies after addition of DNA polymerases to the sensor (second row: DNA polymerase I, Klenow fragment; third row: DNA polymerase beta; fourth row DNA polymerase III alpha subunit).



Addition of KF gives rise to a population of sensors that exhibit a higher  $E^*$  than the native sensor. Peak positions of the fits and the corresponding errors can be found in Supporting Information Table S1. **(c and d)** Representative time traces of the DNA sensors +25/-7 **(c)** and +12/-12 **(d)** in presence of 500 pM KF. Upper panel: Acceptor fluorescence upon direct acceptor excitation (AA; blue trace). Donor (DD, green trace) and acceptor (DA, red trace) fluorescence upon donor excitation. Lower panel: FRET efficiency  $E^*$  (black trace) with fitted HMM model (magenta trace).

### Monitoring DNA synthesis in real time

As we have verified using the end-point experiments, the transfer efficiency  $E^*$  decreases to stable values after successful DNA synthesis. For monitoring the synthesis in real time, we first focused on the +25/-7 sensor as it showed the largest change in  $E^*$  upon polymerization. Polymerization of the sensor by KF showed a stepwise decrease of FRET efficiency (Fig. 3.4a): at around 140 s, we saw a short increase in FRET efficiency from the base level of  $E^* \sim 0.75$  to  $E^* \sim 0.80$ , which was too small to be picked up by HMM, before dropping to  $E^* \sim 0.60$ , a level which was held for around 1.5 s. After that period, the FRET efficiency dropped to  $E^* \sim 0.20$  representing fully polymerized DNA. The example time trace for POLIII $\alpha$  shows fast synthesis requiring two synthesis steps (Fig. 3.4b). Polymerization by POLB, on the contrary, is found to be slow: time traces show a gradual decrease over a long time span (>100 s, Fig. 3.4c) in line with the slow DNA synthesis shown for POLB on recessed DNA substrates<sup>49</sup>. In some cases, we observed an increase in the fluorescence of the acceptor after direct excitation upon polymerization. We attribute this to the close proximity of the dye to the incorporation site, where its local environment may change significantly during primer elongation. Our data showed that polymerization is in some cases not completed during the time of data acquisition. For KF and the +25/-7 sensor, 65% of all analysed time traces, representing 28 out of 43 DNA molecules in total, showed complete polymerization during data acquisition; 5 molecules completed DNA synthesis in one step and 23 molecules required two or three steps indicated by a considerable 'lag time' between regions with decreasing FRET efficiency (Fig. 3.4d). This lag time may be caused by either pausing or dissociation with subsequent re-association of the polymerase. Of the remaining 35% of time traces (15 molecules), 10 traces showed incomplete synthesis indicated by a final  $E^* > 0.3$  and 5 traces could not be analysed due to fluorophore blinking events or other complications (more example time traces are presented in Fig. S3.2, Supporting Information). Plotting the durations of all 28 completed polymerization events excluding the lag time revealed an average duration of  $1.6 \pm 0.5$  s (mean  $\pm$  standard deviation, Fig. 3.4e). The faster polymerization by POLIII $\alpha$  is reflected in a lower average duration of  $1.0 \pm 0.4$  s for the same sensor ( $N = 54$ , Fig. 3.4f). A similar analysis for POLB was not possible due to the low processivity, preventing us to fit the data with HMM.



**Fig. 3.4** DNA synthesis monitored in real time. (a–c) Single-molecule time traces showing DNA polymerization in real time of sensor +25/–7 by KF (a), POLIIIa (b) and POLB (c). Upper panel: Acceptor fluorescence upon direct acceptor excitation (AA; blue trace). Donor (DD, green trace) and acceptor (DA, red trace) fluorescence upon donor excitation. Lower panel: FRET efficiency  $E^*$  (black trace) with fitted HMM model (magenta trace). (d) Pie chart of the number of detectable steps in KF polymerization traces (sensor +25/–7). A total of 43 polymerization events were analysed. (e and f) Histograms of polymerization durations in KF (e) and POLIIIa (f) polymerization traces (sensor +25/–7). (g and h) Single-molecule time traces showing polymerization of sensor +12/–12 by KF. Peak positions and the HMM fit to the polymerized state (used for calculation of durations) are indicated. (i) Histogram of polymerization duration in KF polymerization traces (sensor +12/–12). More polymerization traces of KF, POLIIIa and POLB can be found in Fig. S3.2–S3.5 (Supporting Information).

Using the +12/–12 sensor and KF, we observed that polymerization is often preceded by a clear increase in  $E^*$  (Fig. 3.4g and h). We attribute this increase to reorganization of the template just before polymerization, similar to the increase in  $E^*$  we observed in the binding experiments (Fig. 3.3). It is unclear, however, to what extent the polymerization reaction has already started before we begin to observe the signature decrease in  $E^*$ . We determined the duration of the polymerization to be  $1.1 \pm 0.5$  s ( $N = 30$ , Fig. 3.4h).

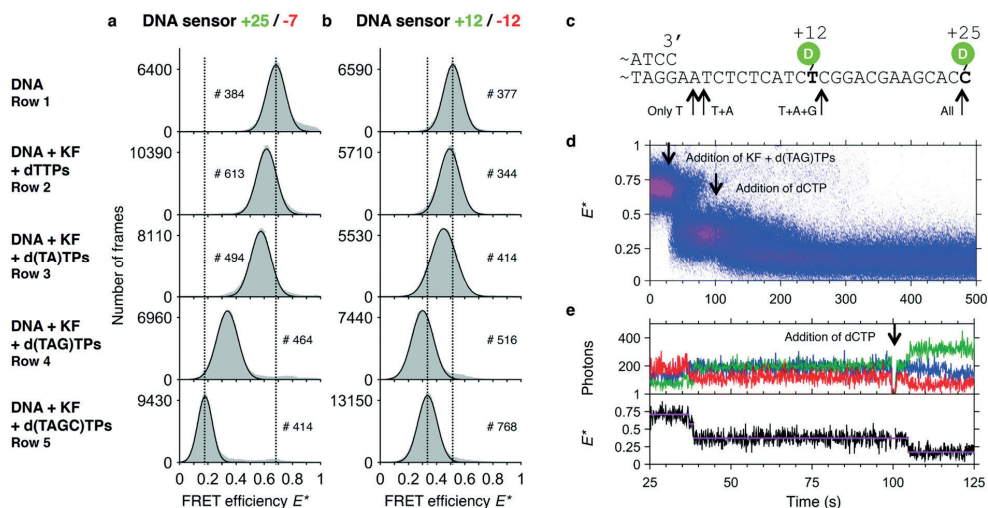


Additionally, we tested the performance of the +25/−7 sensor at the ensemble level for KF using a microwell plate reader (Fig. S3.6, Supporting Information). We further examined the influence of the fluorescent labels on polymerization reaction rates using gel-shift assays by comparing labelled and unlabelled DNA constructs (Supporting Information Fig. S3.7–S3.9). From the data, we conclude that the rates of DNA synthesis are comparable, with a ~2 fold decrease seen for the labelled constructs measured at 6 °C. Furthermore, we note that the general evaluation of processive polymerization remains challenging<sup>50,51</sup>.

### **Polymerization can be paused at specific points by tuning dNTP concentrations**

As shown above, the limited processivity of KF often prevents the completion of DNA synthesis in a single step. We asked whether we could pause DNA synthesis at specific positions and whether we could resolve the addition of single bases to the primer template. We therefore designed the sequence of our template such that omitting certain dNTPs in the reaction mixture will prevent the polymerase from continuing DNA synthesis at well-defined positions. While dTTP is necessary to synthesise the first two bases, addition of dATP will lead to an extension of the primer by a single base. Additional dCTP allows for continuation of DNA synthesis halfway through the template. To determine the FRET efficiencies of such artificially created pausing sites, we first ran a series of polymerization reactions in a cuvette in which we left out different subsets of dNTPs. As expected, the  $E^*$  histograms show partial polymerizations indicated by a decrease of the mean FRET value upon addition of a subset of dNTPs (Fig. 3.5a and b). For convenience, we added the template sequence and indicated the sites where synthesis can be halted (Fig. 3.5c). Remarkably, our data show for the +25/−7 sensor that extensions by two bases (addition of dTTP) and even by a single base (addition of dTTP and dATP) lead to a significant shift in the peak position (Fig. 3.5a, for a detailed analysis of the peak position and associated fitting errors please see the Supporting Information). An interesting deviation from the general trend is seen for the +12/−12 sensor, where the synthesis after addition of 3 out of 4 dNTPs leads to a peak position lower than after full synthesis (Fig. 3.5b). We attribute this to an altered mean position of the donor fluorophore when polymerization continues to form a double helix around the dye.

In a real-time polymerization experiment in which dCTP is initially omitted from the reaction mixture, we can show that polymerizations only finish after addition of dCTP at  $t = 100$  s (Fig. 3.5d and e).



**Fig. 3.5** Assay to test pausing of DNA synthesis. Histograms of FRET efficiencies  $E^*$  for the +25/-7 (a) and the +12/-12 (b) DNA sensor. Each histogram (100 bins shown in grey) was derived from individual time traces obtained from fitting an indicated number of molecules (#). The histograms were fitted to a single-Gaussian model (black line) to indicate the mean peak position of the dominant FRET species. Vertical dashed lines were added for visual guidance. First row: FRET efficiencies of the native DNA sensor. Second to fourth row: FRET efficiencies after incubating the DNA sensor for 90 minutes in presence of Klenow Fragment and the indicated subset of dNTPs. The histograms indicate that polymerization is only completed if all 4 dNTPs are provided, and that polymerization is otherwise halted at different positions along the template. Peak positions of the fits and the corresponding errors can be found in Supporting Information Table S1. (c) Part of the DNA sensor sequence with the stop positions on the template indicated by arrows. (d) Plot of 74 accumulated polymerization traces of the +25/-7 sensor, obtained from a single experiment. Polymerizations are only finished ( $E^* < 0.25$ ) after dCTP is added to the reaction mixture. (e) Example trace of a polymerization reaction from (d).

## Discussion

We developed a fluorescently labelled DNA sensor reporting on DNA synthesis by a change in the FRET efficiency between two fluorophores placed on the DNA template and the DNA primer, respectively. Using single-molecule TIRF microscopy, we characterized three different polymerases that vary in speed and cellular function. We showed for all three polymerases (KF, POLIII $\alpha$  and POLB) that the FRET efficiency of the sensor decreases upon polymerization of DNA. Examination of the inter-dye distances in the final products indicated that the sensors are fully polymerized. Real-time single-molecule experiments showed expected differences in polymerization speed and processivity. While polymerization by POLIII $\alpha$  was fast, KF and POLB

were noticeably slower. This finding can be explained by the role of these two polymerases in DNA repair, which does not require fast DNA synthesis as it is the case for DNA replication. Rates obtained from single-molecule studies for primer extension by KF vary between 1 and 30 nt s<sup>-1</sup>, but are known to depend strongly on template sequence, the base that is being incorporated and whether the synthesis is in the initial or steady state<sup>13,19,22,52</sup>. Our assay finds broad distributions of polymerization durations, which is indicative of the large variation in speed even between polymerase molecules of the same species. We found the mean duration for polymerization of a 25 nt long template to be  $1.6 \pm 0.5$  s for KF, which is in the same order of magnitude as the other single-molecule studies mentioned above. We found that most polymerization events required more than one step for completion caused by either pausing or dissociation of the DNA polymerase. Ensemble studies on a homopolymeric template showed that full-length POL I incorporates 40–50 nucleotides per binding time<sup>53</sup>.

Our data also showed a conformational change in the single-stranded template overhang upon binding of KF, as indicated by a population of molecules at higher  $E^*$ . Apparently, KF can cause a structural rearrangement that results in a decreased donor–acceptor distance. In fact, co-crystal structures of a structural homologue of KF suggest that the single stranded overhang of the templating DNA might be aligned with the help of residues of the fingers-subdomain<sup>54</sup>. A potential explanation for this template folding could lie in the accessibility of the substrate allowing the incoming dNTP to be more easily incorporated.

Our assay provides an easy and robust way of testing the effect of polymerase inhibitors or compare polymerase mutants with each other. Recently, Sharma *et al.*<sup>28</sup> proved with a comparable sensor design that dequenching of donor fluorescence is a reliable method to test the effect of reverse transcriptase inhibitors. We note that the sequence of the template strand may be modified to include sequence specific pausing sites such as Pyr-G-C motifs<sup>55</sup>, triplet repeats<sup>56,57</sup>, or secondary structures introduced for example as hairpins<sup>6</sup> or G-quadruplexes<sup>58,59</sup>. The extent to which these modifications itself may influence the change in FRET efficiency needs to be carefully assessed, but especially smFRET based on the alternating-laser excitation scheme provides an excellent opportunity to distinguish between FRET or intensity changes caused by either photophysical effects or conformational changes. We further note that even though placing a fluorophore close to the binding site can weaken the binding of the DNA polymerase, this spatial proximity could be used to develop assays based on protein induced fluorescence enhancement, which can indicate the presence of bound proteins without necessarily changing the FRET efficiency between the donor and the acceptor fluorophore<sup>48,60,61</sup>. Additionally, using a dark quencher<sup>62</sup> as an acceptor would allow experiments with single (continuous) colour

excitation in which the extent of donor dequenching becomes the measure for the progression of the polymerization reaction.

In our current implementation, the achievable temporal resolution is limited by the rate of data acquisition to 100 ms as two consecutive camera frames of 50 ms duration are computationally merged to provide the three photon streams as discussed in the Supporting Information. Recently published work suggests, however, that smFRET measurements with 1–3 ms time resolution and simultaneous detection of hundreds to thousands of individual DNA molecules are feasible<sup>41,63</sup>. In terms of spatial resolution, both the large FRET change upon full polymerization achievable with the +25/–7 sensor and the single-base pair sensitivity in particular regions of the templating DNA for both sensors are highlighting the robustness and the flexibility of the DNA assay. We note, however, that maximising the FRET change upon polymerisation requires one of the dyes to be placed closer to the 3' end of the DNA primer. In fact, using the +25/–7 sensor revealed a weakening of the DNA polymerase binding to DNA compared to the +12/–12 sensor thereby potentially affecting the processivity of the polymerase for the first few added bases. We suggest to use a more conservative sensor design for applications aimed at synthesising only a few bases.

We expect that additional improvements on the design of the sensor can further increase the range of potential applications. Labelling the templating DNA, for example, with additional fluorophores on the single stranded overhang, thereby creating a multi-colour FRET arrangement, could increase the range of processivities beyond the 25 bases which we can currently monitor.

An additional advantage of the assay is that it does not require the labelling of DNA polymerases with fluorophores and that experiments on the ensemble level can be carried out with standard spectrophotometric lab equipment. With microwell plate imaging instruments reaching levels of sensitivity sufficient to detect fluorescence from samples present in femtomolar quantities, the time dependent imaging of increasing donor (or decreasing acceptor) intensity upon polymerization of DNA allows to evaluate large DNA polymerase libraries utilising an instantaneous readout. As such, this type of assay promises to be a simple and straightforward tool to characterize speed and processivity of DNA polymerases both at the ensemble and single-molecule level.

## **Funding**

J. H. acknowledges support from a Marie Curie Career Integration Grant [630992]. Funding for open access charge: Marie Curie Career Integration Grant [630992].

## Acknowledgments

DNA Polymerase I KF was kindly supplied by Catherine Joyce, Tim Craggs and Achillefs Kapanidis. Human Polymerase Beta was kindly supplied by Jamie Towle-Weicksel and Joann Sweasy. We thank the Laboratory of Toxicology (Hans van den Berg) at Wageningen University and Research for providing access to the plate reader. J. H. acknowledges support from a Marie Curie Career Integration Grant (#630992).

## References

1. Kornberg, A. & Baker, T. *DNA Replication*. (W.H. Freeman, 1992).
2. Hübscher, U., Spadari, S., Villani, G. & Maga, G. in *DNA Polymerases* 1–58 (World Scientific, 2010).
3. Hübscher, U., Spadari, S., Villani, G. & Maga, G. in *DNA Polymerases* 59–83 (World Scientific, 2010).
4. García-Gómez, S. *et al.* PrimPol, an Archaic Primase/Polymerase Operating in Human Cells. *Mol. Cell* **52**, 541–553 (2013).
5. Saiki, R. K. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491 (1988).
6. Eid, J. *et al.* Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* **323**, 133–138 (2009).
7. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135 (2008).
8. Fried, M. & Crothers, D. M. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**, 6505–6525 (1981).
9. Walsh, M. T., Roller, E. E., Ko, K.-S. & Huang, X. Measurement of DNA Polymerase Incorporation Kinetics of Dye-Labeled Nucleotides Using Total Internal Reflection Fluorescence Microscopy. *Biochemistry (Mosc.)* **54**, 4019–4021 (2015).
10. Langer, A. *et al.* Polymerase/DNA interactions and enzymatic activity: multi-parameter analysis with electro-switchable biosurfaces. *Sci. Rep.* **5**, (2015).
11. Farooq, S., Fijen, C. & Hohlbein, J. Studying DNA–protein interactions with single-molecule Förster resonance energy transfer. *Protoplasma* **251**, 317–332 (2014).
12. Kapanidis, A. N. & Strick, T. Biology, one molecule at a time. *Trends Biochem. Sci.* **34**, 234–243 (2009).
13. Maier, B., Bensimon, D. & Croquette, V. Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12002–12007 (2000).

14. Heller, I., Hoekstra, T. P., King, G. A., Peterman, E. J. G. & Wuite, G. J. L. Optical Tweezers Analysis of DNA-Protein Complexes. *Chem. Rev.* **114**, 3087–3119 (2014).
15. Olasagasti, F. *et al.* Replication of Individual DNA Molecules under Electronic Control Using a Protein Nanopore. *Nat. Nanotechnol.* **5**, 798–806 (2010).
16. Cockroft, S. L., Chu, J., Amarin, M. & Ghadiri, M. R. A Single-Molecule Nanopore Device Detects DNA Polymerase Activity with Single-Nucleotide Resolution. *J. Am. Chem. Soc.* **130**, 818–820 (2008).
17. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55–75 (1948).
18. Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. *Proc. Natl. Acad. Sci. U. S. A.* **58**, 719–726 (1967).
19. Schwartz, J. J. & Quake, S. R. Single molecule measurement of the ‘speed limit’ of DNA polymerase. *Proc. Natl. Acad. Sci.* **106**, 20294–20299 (2009).
20. Liu, S., Abbondanzieri, E. A., Rausch, J. W., Grice, S. F. J. L. & Zhuang, X. Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science* **322**, 1092–1097 (2008).
21. Abbondanzieri, E. A. *et al.* Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature* **453**, 184–189 (2008).
22. Christian, T. D., Romano, L. J. & Rueda, D. Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc. Natl. Acad. Sci.* **106**, 21109–21114 (2009).
23. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res.* **43**, 5998–6008 (2015).
24. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nat. Commun.* **4**, (2013).
25. Harris, T. D. *et al.* Single-Molecule DNA Sequencing of a Viral Genome. *Science* **320**, 106–109 (2008).
26. Kim, Y. *et al.* Reading Single DNA with DNA Polymerase Followed by Atomic Force Microscopy. *J. Am. Chem. Soc.* **136**, 13754–13760 (2014).
27. Previte, M. J. R. *et al.* DNA sequencing using polymerase substrate-binding kinetics. *Nat. Commun.* **6**, (2015).
28. Sharma, K. K., Przybilla, F., Restle, T., Godet, J. & Mély, Y. FRET-based assay to screen inhibitors of HIV-1 reverse transcriptase and nucleocapsid protein. *Nucleic Acids Res.* **44**, e74 (2016).
29. Roy, R., Kozlov, A. G., Lohman, T. M. & Ha, T. SSB diffusion on single stranded DNA stimulates RecA filament formation. *Nature* **461**, 1092–1097 (2009).



30. Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M. & Ha, T. Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy. *Biophys. J.* **86**, 2530–2537 (2004).
31. Barnes, D. E. & Lindahl, T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* **38**, 445–476 (2004).
32. Beard, W. A. & Wilson, S. H. Structure and Mechanism of DNA Polymerase  $\beta$ . *Biochemistry (Mosc.)* **53**, 2768–2780 (2014).
33. Towle-Weicksel, J. B. *et al.* Fluorescence Resonance Energy Transfer Studies of DNA Polymerase  $\beta$  THE CRITICAL ROLE OF FINGERS DOMAIN MOVEMENTS AND A NOVEL NON-COVALENT STEP DURING NUCLEOTIDE SELECTION. *J. Biol. Chem.* **289**, 16541–16550 (2014).
34. Klenow, H. & Henningsen, I. Selective Elimination of the Exonuclease Activity of the Deoxyribonucleic Acid Polymerase from Escherichia coli B by Limited Proteolysis\*. *Proc. Natl. Acad. Sci. U. S. A.* **65**, 168–175 (1970).
35. McHenry, C. & Kornberg, A. DNA polymerase III holoenzyme of Escherichia coli. Purification and resolution into subunits. *J. Biol. Chem.* **252**, 6478–6484 (1977).
36. Fay, P. J., Johanson, K. O., McHenry, C. S. & Bambara, R. A. Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of Escherichia coli. *J. Biol. Chem.* **256**, 976–983 (1981).
37. McHenry, C. S. & Crow, W. DNA polymerase III of Escherichia coli. Purification and identification of subunits. *J. Biol. Chem.* **254**, 1748–1753 (1979).
38. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
39. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).
40. Lee, N. K. *et al.* Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys. J.* **88**, 2939–2953 (2005).
41. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17**, 27862–27872 (2015).
42. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *J. Am. Chem. Soc.* **131**, 5018–5019 (2009).
43. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).
44. Holden, S. J. *et al.* Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophys. J.* **99**, 3102–3111 (2010).
45. van Dijk, M. & Bonvin, A. M. J. J. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res.* **37**, W235–W239 (2009).
46. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).

47. van de Meent, J.-W., Bronson, J. E., Wiggins, C. H. & Gonzalez, R. L. Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophys. J.* **106**, 1327–1337 (2014).
48. Ploetz, E. *et al.* Förster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. *Sci. Rep.* **6**, (2016).
49. Chagovetz, A. M., Sweasy, J. B. & Preston, B. D. Increased Activity and Fidelity of DNA Polymerase  $\beta$  on Single-nucleotide Gapped DNA. *J. Biol. Chem.* **272**, 27501–27504 (1997).
50. Datta, K., Wowor, A. J., Richard, A. J. & LiCata, V. J. Temperature Dependence and Thermodynamics of Klenow Polymerase Binding to Primed-Template DNA. *Biophys. J.* **90**, 1739–1751 (2006).
51. Rentergent, J., Driscoll, M. D. & Hay, S. Time Course Analysis of Enzyme-Catalyzed DNA Polymerization. *Biochemistry* **55**, 5622–5634 (2016).
52. Olsen, T. J. *et al.* Electronic Measurements of Single-Molecule Processing by DNA Polymerase I (Klenow Fragment). *J. Am. Chem. Soc.* **135**, 7855–7860 (2013).
53. Bryant, F. R., Johnson, K. A. & Benkovic, S. J. Elementary steps in the DNA polymerase I reaction pathway. *Biochemistry* **22**, 3537–3546 (1983).
54. Johnson, S. J., Taylor, J. S. & Beese, L. S. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3895–3900 (2003).
55. Mytelka, D. S. & Chamberlin, M. J. Analysis and suppression of DNA polymerase pauses associated with a trinucleotide consensus. *Nucleic Acids Res.* **24**, 2774–2781 (1996).
56. Kang, S., Ohshima, K., Shimizu, M., Amirhaeri, S. & Wells, R. D. Pausing of DNA Synthesis in Vitro at Specific Loci in CTG and CGG Triplet Repeats from Human Hereditary Disease Genes. *J. Biol. Chem.* **270**, 27014–27021 (1995).
57. Samadashwily, G. M., Raca, G. & Mirkin, S. M. Trinucleotide repeats affect DNA replication in vivo. *Nat. Genet.* **17**, 298 (1997).
58. Noer, S. L. *et al.* Folding dynamics and conformational heterogeneity of human telomeric G-quadruplex structures in Na<sup>+</sup> solutions by single molecule FRET microscopy. *Nucleic Acids Res.* **44**, 464–471 (2016).
59. Schiavone, D. *et al.* PrimPol Is Required for Replicative Tolerance of G Quadruplexes in Vertebrate Cells. *Mol. Cell* **61**, 161–169 (2016).
60. Hwang, H., Kim, H. & Myong, S. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7414–7418 (2011).
61. Lerner, E., Ploetz, E., Hohlbein, J., Cordes, T. & Weiss, S. A Quantitative Theoretical Framework For Protein-Induced Fluorescence Enhancement–Förster-



- Type Resonance Energy Transfer (PIFE-FRET). *J. Phys. Chem. B* **120**, 6401–6410 (2016).
62. Le Reste, L., Hohlbein, J., Gryte, K. & Kapanidis, A. N. Characterization of Dark Quencher Chromophores as Nonfluorescent Acceptors for Single-Molecule FRET. *Biophys. J.* **102**, 2658–2668 (2012).
63. Juette, M. F. *et al.* Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat. Methods* **13**, 341–344 (2016).

## Supporting information

### Alternating-laser excitation (ALEX)

Alternating-laser excitation (ALEX) is a technique in which both donor and acceptor dyes are excited in an alternating fashion<sup>38,39</sup>. In addition to  $f_{D_{exc}}^{Dem}$  and  $f_{D_{exc}}^{Aem}$  detected in standard smFRET experiments, the ALEX scheme also monitors the acceptor emission after acceptor excitation  $f_{A_{exc}}^{Aem}$  allowing to separate donor-only and acceptor-only species from molecules that have both a donor and acceptor. The value in use to identify these species is called the stoichiometry ratio  $S$ :

$$S^{raw} = (f_{D_{exc}}^{Aem} + f_{D_{exc}}^{Dem}) / (f_{D_{exc}}^{Aem} + f_{D_{exc}}^{Dem} + f_{A_{exc}}^{Aem})$$

A second advantage of using ALEX is the easy application of correction factors to calculate accurate FRET efficiency  $E$  from apparent FRET efficiency  $E^*$ .

Calculating intramolecular distances using accurate FRET

The FRET efficiency  $E$  describes a distance-dependent energy transfer according to

$$E = \frac{1}{1 + (R/R_0)^6}$$

in which  $R$  is the inter dye distance and  $R_0$  is the Förster radius.  $R_0$  depends on the dye pair in use and corresponds to the distance at which  $E$  equals 0.50. To obtain  $E$ , the apparent FRET efficiency  $E^*$  is corrected for direct excitation of the acceptor, leakage of the donor into the acceptor channel and the gamma factor, which is a measure for the detection efficiencies of the dyes and their quantum yields<sup>38,40</sup>. Correction factors for leakage and direct excitation can be deduced from the peaks of donor-only and acceptor-only species, respectively. A third correction needs to be done for the effect of the quantum yields and detection efficiencies of the dyes, which is represented by the gamma factor. The value of gamma is determined by comparing the stoichiometries of populations at different FRET efficiencies: a value of gamma other than 1, will cause these stoichiometries to be different. In practice, gamma can be calculated from the slope when plotting  $1/S$  against  $E^{40}$ . For the distance determination of the polymerized sensors, we used  $R_0 = 6.2$  nm<sup>44</sup>.

### Step by step guide for Hidden Markov Modelling (HMM) analysis using ebFRET

- 1) Hand selected time traces that indicated one or more polymerization events were loaded into ebFRET.
- 2) The selected time traces are filtered for photo bleaching: Parts that indicated a bleached acceptor were excluded from further analysis.

3) Due to the long duration of measurements, some traces were partially effected by focus drift. We therefore decided to exclude these parts, in which no polymerisation reaction was visible, from further analysis.

4) The priors needed for HMM were taken from the  $E^*$  values of the non-polymerised and the fully polymerised DNA sensor, respectively.

5) Before running ebFRET, the ideal number of states (related to the number of steps in the polymerization reaction) is unknown. Fitting was therefore performed successively for two to six states. Even within a single experiment, the noise level can be different between single traces. Therefore, care should be taken not to overestimate the number of present states especially within noisy traces. After running ebFRET for a respective number of states, all traces were manually inspected for cases of over- or underestimation of the number of present states and excluded from further analysis when required.

6) For the +25/-7 sensor, the start time of DNA synthesis  $t_1$  was determined using ebFRET as the time point between the initial state and the second fitted state. In the case of the +12/-12 sensor, we decided to pick the highest  $E^*$  value as the starting time  $t_1$  just before polymerization as ebFRET was not reliably able to pick up these characteristic short peaks in  $E^*$ .

7) In cases with two low final  $E^*$  values present in a time trace, the lowest one was chosen to determine the polymerization time  $t_2$ . The main consideration here is that there is not one specific FRET efficiency value that corresponds to a polymerized construct. Rather, the FRET efficiencies of polymerized molecules appear to be normally distributed around a mean value of  $E^* \approx 0.18$  (for the +25/-7 sensor). We looked at the widths of the Gaussian fits in Figure 3.2, and decided to set the cut-off value at 0.25. This corresponds to the mean value plus approximately one standard deviation.

### **Polymerization at the ensemble level monitored in a plate reader**

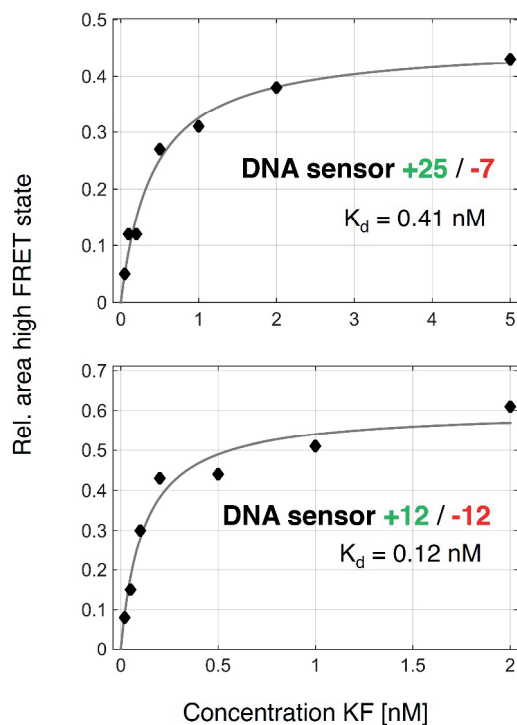
The polymerization of the +25/-7 sensor by KF was recorded in a 96 wells plate, using a SpectraMax M2 plate reader. dNTPs (100  $\mu$ M each) and sensors (10 pmoles, to reach a concentration of 100 nM) were mixed together in an aliquot resembling our imaging buffer, but without the oxygen scavenger system. An emission spectrum (excitation at 500 nm) was taken before addition of polymerases (see Fig. S3.6a, dotted line). By exciting at 500 nm instead of 560 nm, the excitation efficiency is decreased by a factor of  $\sim 3$ . After mixing polymerases (20 nM) into the buffer to create a total reaction volume of 100  $\mu$ L, emission intensities at 575 and 665 nm were recorded for 600 s (Fig. S3.6b). Immediately after this recording, another emission spectrum was taken (Fig. S3.6a, solid line).

**Denaturing PAGE gel to assess influence of fluorescent labels on polymerase speed**

We performed a series of bulk primer-extension polymerizations on our +25/-7 sensor, our +12/-12 sensor and on an unlabelled version of the same construct. DNA (final concentration: 25 nM) and KF (final concentration: 100 nM) were mixed together in a buffer similar to our imaging buffer, but without the oxygen scavenger system. A fraction of this solution was taken as measurement  $t = 0$  s. Addition of dNTPs (final concentration: 500 nM each) started the reactions. Reactions were quenched with EDTA after 15 s, 30 s, 60 s, 120 s and 300 s. Because reaction speed of DNA synthesis is too fast at room temperature, we performed the reactions in a cold room at 6 °C. This temperature does not considerably change the  $K_d$  of KF for primer-template DNA,<sup>50</sup> but it does slow down the polymerization reaction<sup>51</sup>.

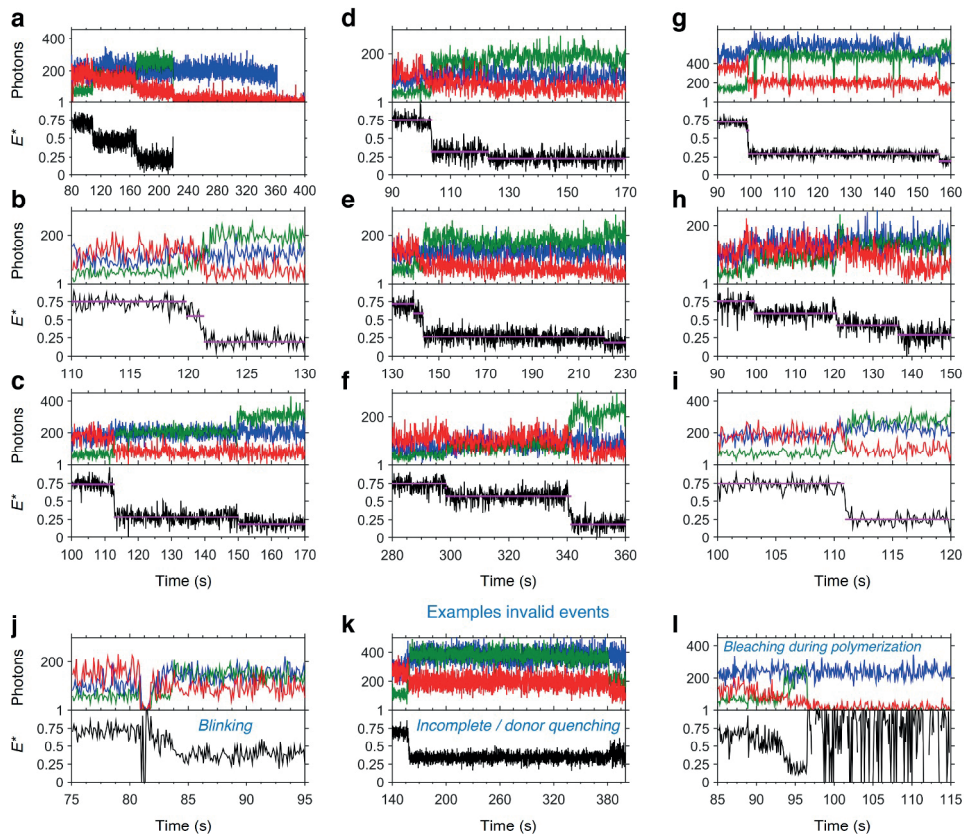
Products were mixed with formamide (to 50% v/v), heated for 5 minutes at 95 °C unless otherwise indicated and loaded onto a denaturing PAGE (20%, 6 M urea). After running the gel, DNA bands were visualized using SYBR Gold stain (Fig. S3.8+9).

Upon inspection of the band pattern, it became clear that the primer is invisible at the concentrations that were used for the assay. Some other constructs were also difficult to detect, probably because of a quenching mechanism between SYBR Gold and Cy3B/ATTO647N. Additionally, unexpected bands running slightly slower than the template were visible. We hypothesized that these bands correspond to non-denatured sensors. To test this hypothesis, we used a few native and fully polymerized sensors that we did not subject to the aforementioned heat treatment. The gel shows that this heat treatment does not always have the desired effect (Fig. S3.8), leading to the conclusion that non-denatured fractions indeed cause the unexpected bands. Moreover, we noted that fully polymerized sensors are harder to denature than native sensors. We used this principle to identify the progress of polymerization at each time point. To this aim, we measured the increasing intensities of the upper bands (corresponding to fully polymerized product) over time and fitted them to Michaelis-Menten kinetics to extract reaction half-time  $t_{1/2}$  (Fig. S3.9). We conclude from our data that the rates of polymerisation are comparable between the labelled and unlabelled constructs, but we note that further investigation and, ideally, the use of stopped flow equipment and radioisotope labelled DNA would be required to evaluate the DNA sensors further.

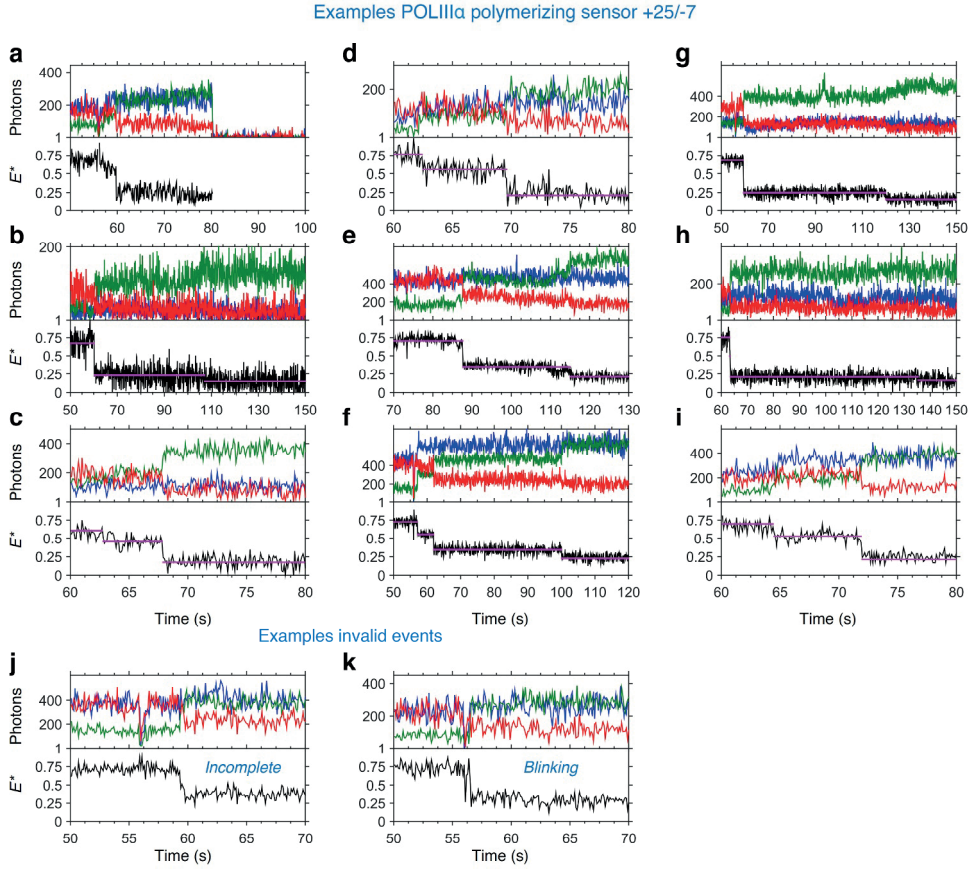


**Figure S3.1** Binding of KF to the sensors reveals a dissociation constant  $K_d$  of 0.41 nM for the +25/-7 sensor, and a  $K_d$  of 0.12 nM for the +12/-12 sensor (data fitted to Michaelis-Menten kinetics:  $\text{rel. area high FRET state} = (\text{maximum occupancy} * [\text{KF}]) / ([\text{KF}] + K_d)$ ).

## Examples KF polymerizing sensor +25/-7

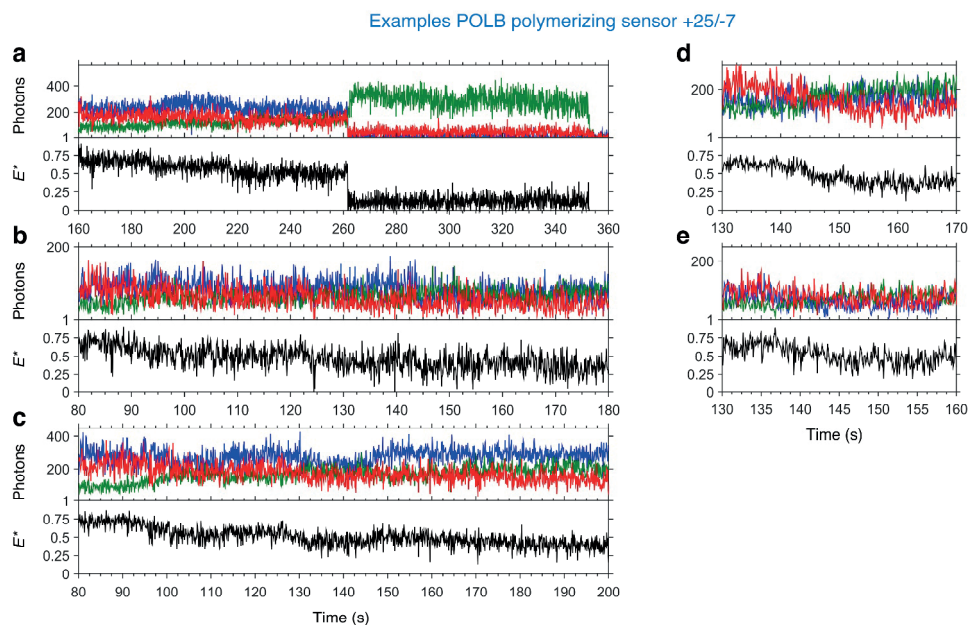


**Figure S3.2 (a-l)** Examples of time traces showing polymerization of the +25/-7 sensor by KF, including a trace showing donor and acceptor bleaching (a). These data were obtained during the same experiment as the trace in Fig. 3.4a. A few examples of invalid events are shown as well. Events are considered invalid if it is impossible to accurately determine the polymerization time because of either dye blinking (j) or incomplete polymerization ( $E^*_{final} > 0.3$ ) (k) or quenching or bleaching during a polymerisation event (k, l). Upper panel: Acceptor fluorescence upon direct acceptor excitation (AA; blue trace). Donor (DD, green trace) and acceptor (DA, red trace) fluorescence upon donor excitation. Lower panel: FRET efficiency  $E^*$  (black trace) with fitted HMM model (magenta trace).



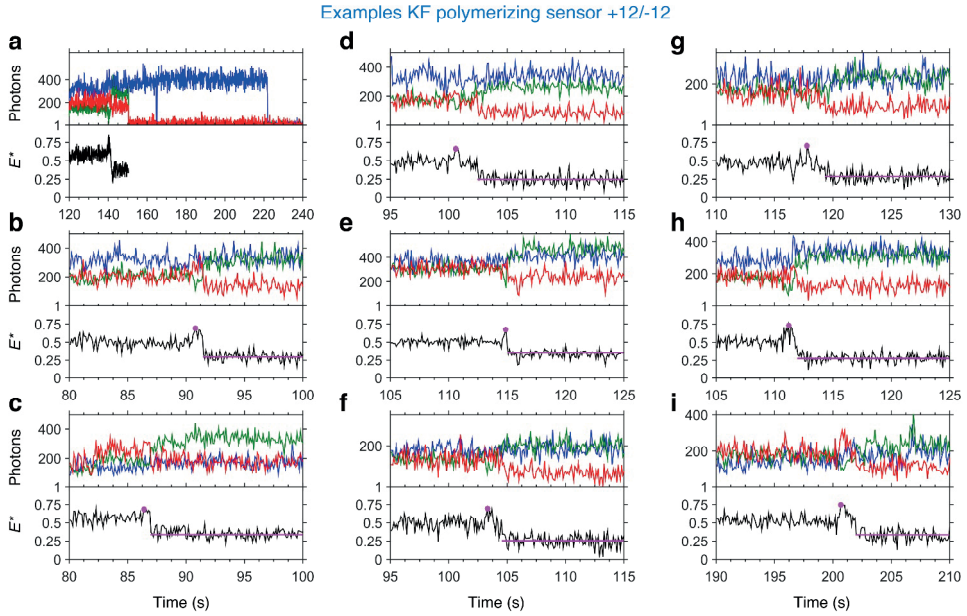
**Figure S3.3 (a-k)** Examples of time traces showing polymerization of the +25/-7 sensor by PolIIIa, including a trace showing donor and acceptor bleaching (a). These data were obtained during the same experiment as the trace in Fig. 3.4b. Examples of invalid events are shown as well. Events are considered invalid if it is impossible to accurately determine the polymerization time because of an incomplete reaction (j) or dye blinking (k). Upper panel: Acceptor fluorescence upon direct acceptor excitation (AA; blue trace). Donor (DD, green trace) and acceptor (DA, red trace) fluorescence upon donor excitation. Lower panel: FRET efficiency  $E^*$  (black trace) with fitted HMM model (magenta trace).



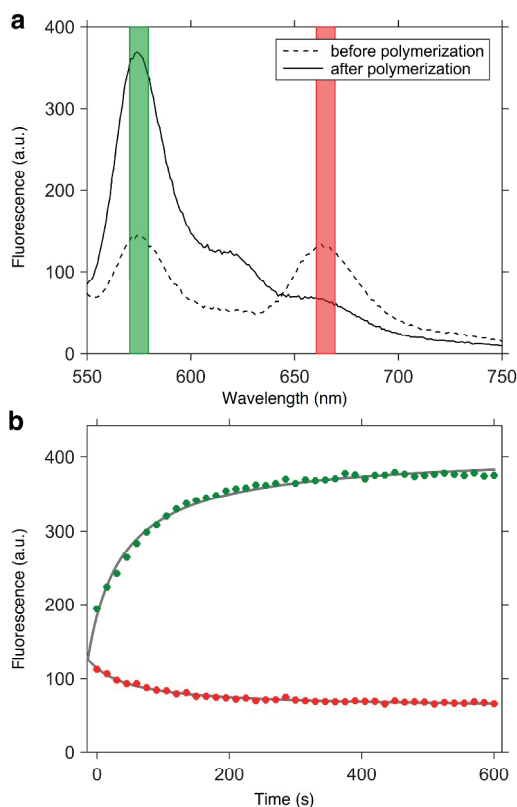


**Figure S3.4 (a-e)** Examples of time traces showing polymerization of the +25/-7 sensor by PolB, including a trace showing donor and acceptor bleaching (a). These data were obtained during the same experiment as the trace in Fig. 3.4c. Upper panel: Acceptor fluorescence upon direct acceptor excitation (AA; blue trace). Donor (DD, green trace) and acceptor (DA, red trace) fluorescence upon donor excitation. Lower panel: FRET efficiency  $E^*$  (black trace).





**Figure S3.5 (a-i)** Examples of time traces showing polymerization of the +12/-12 sensor by KF, including a trace showing donor and acceptor bleaching (a). These data were obtained during the same experiment as the traces in Fig. 3.4g+h. Upper panel: Acceptor fluorescence upon direct acceptor excitation (AA; blue trace). Donor (DD, green trace) and acceptor (DA, red trace) fluorescence upon donor excitation. Lower panel: FRET efficiency  $E^*$  (black trace) with fitted HMM model (magenta trace).

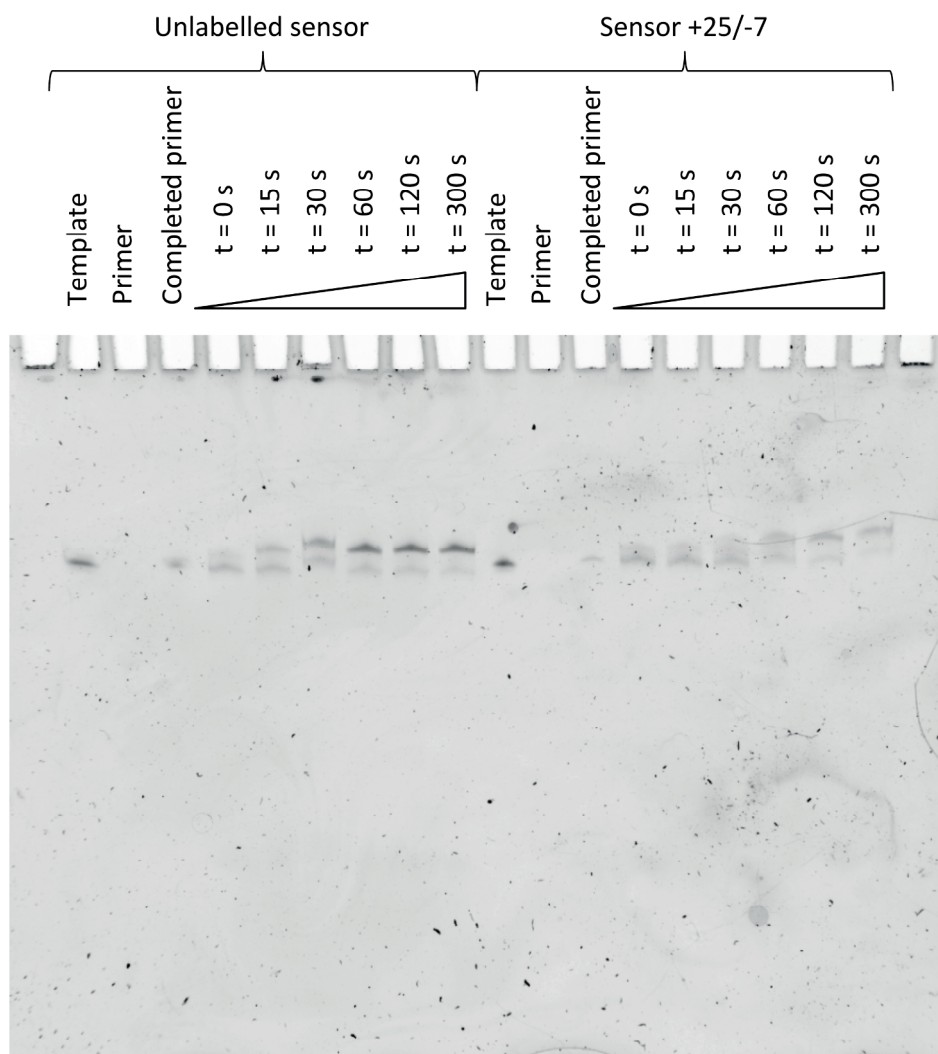


**Figure S3.6** Polymerization of the +25/-7 sensor by KF, measured in a well plate reader. The emission spectrum recorded before addition of KF shows peaks from emission of Cy3B and ATTO647N (a). The other emission spectrum, taken ~ 600 s after initiation of the reaction, shows an increase in Cy3B fluorescence intensity and a decrease in FRET. Green and red lines indicate the detection wavelengths during the kinetic measurements: 575 nm and 660 nm. Fluorescence intensity followed over time after addition of KF (b) Data fitted to Michaelis-Menten kinetics:

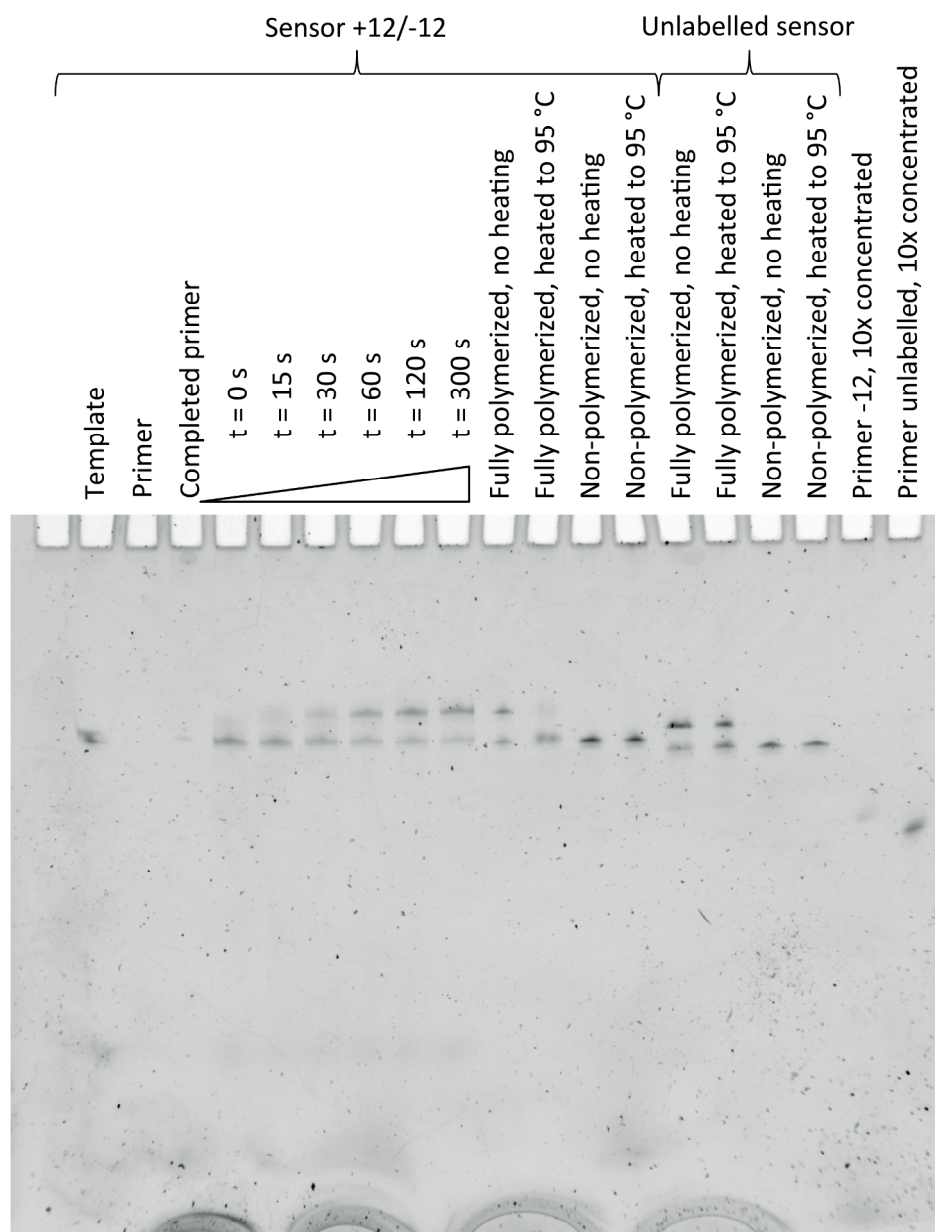
Donor intensity  $I = I_{\max} * (\text{time} + \text{horizontal offset}) / (\text{halftime} + \text{time} + \text{horizontal offset}) + \text{vertical offset}$

Acceptor intensity  $I = -I_{\max} * (\text{time} + \text{horizontal offset}) / (\text{halftime} + \text{time} + \text{horizontal offset}) + \text{vertical offset}$

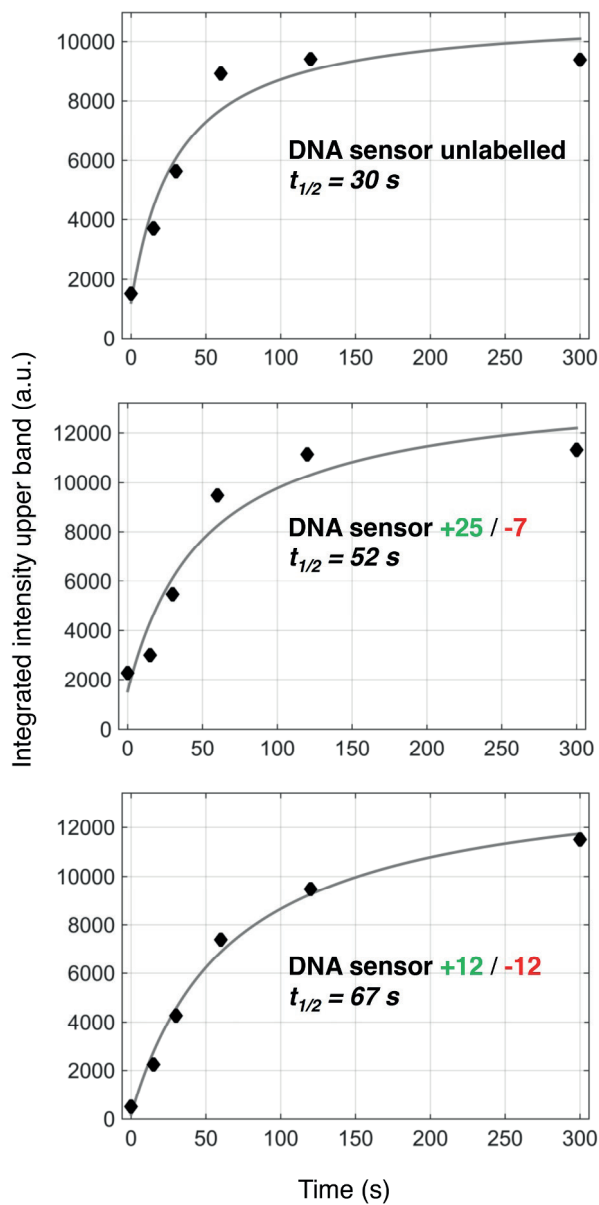
This fitting shows that the reaction was 50% finished at  $t = 60$  s.



**Figure S3.7** Polymerization reaction rates were determined from band intensities in denaturing polyacrylamide gels. Gel images show bands corresponding to denatured and non-denatured polymerization products. The increasing intensities of the bands corresponding to fully polymerized product were measured and taken as an indicator for the progress of the reaction.



**Figure S3.8** Polymerization reaction rates were determined from band intensities in denaturing polyacrylamide gels. Gel images show bands corresponding to denatured and non-denatured polymerization products. The increasing intensities of the bands corresponding to fully polymerized product were measured and taken as an indicator for the progress of the reaction.



**Figure S3.9** Polymerization reaction rates determined from band intensities in denaturing polyacrylamide gels. Intensities determined from the non-denatured upper bands in figures S7 and S8 were plotted against time and fitted to Michaelis-Menten kinetics:

$$\text{Intensity upper band} = (\text{maximum intensity} * \text{time}) / (\text{time} + t_{1/2}) + \text{offset}$$

*Supplementary Table 3.1 Peak positions of E\* histograms that were shown in the main text.*

	Sensor +25/-7	Sensor +12/-12
<b>Assay to test completion of synthesis (Figure 2 a+b)</b>	Peak position $\pm 2$ * standard error of fit (= 95% confidence interval)	Peak position $\pm 2$ * standard error of fit (= 95% confidence interval)
<b>DNA</b>	$0.676 \pm 0.002$	$0.514 \pm 0.001$
<b>DNA + KF + dNTPs</b>	$0.171 \pm 0.002$	$0.346 \pm 0.002$
<b>DNA + PolB + dNTPs</b>	$0.175 \pm 0.002$	$0.361 \pm 0.002$
<b>DNA + PolIII<math>\alpha</math> + dNTPs</b>	$0.178 \pm 0.003$	$0.352 \pm 0.002$
<b>Binding studies (Figure 3 a+b)</b>		
<b>DNA</b>	$0.643 \pm 0.001$	$0.521 \pm 0.001$
<b>DNA + KF</b>	$0.661 \pm 0.004$	$0.521 \pm 0.030$
	$0.884 \pm 0.016$	$0.620 \pm 0.119$
<b>DNA + PolB</b>	$0.671 \pm 0.001$	$0.528 \pm 0.001$
<b>DNA + PolIII<math>\alpha</math></b>	$0.644 \pm 0.002$	$0.521 \pm 0.002$
<b>Assay to test polymerase pausing (Figure 5 a+b)</b>		
<b>DNA</b>	$0.683 \pm 0.002$	$0.515 \pm 0.001$
<b>DNA + KF + dTTPs</b>	$0.617 \pm 0.001$	$0.493 \pm 0.001$
<b>DNA + KF + dTTPs + dATPs</b>	$0.576 \pm 0.002$	$0.442 \pm 0.001$
<b>DNA + KF + dTTPs + dATPs + dGTPs</b>	$0.339 \pm 0.002$	$0.301 \pm 0.001$
<b>DNA + KF + dTTPs + dATPs + dGTPs + dCTPs</b>	$0.180 \pm 0.001$	$0.336 \pm 0.001$

*The indicated range corresponds to the 95% confidence interval of the Gaussian fit.*

**Supplementary Tables 3.2, 3.3 and 3.4 (overleaf)** Raw and corrected values for E and S. Corrections for leakage and direct excitation were applied, as well as a correction for the gamma factor.

Sensor +25/-7	Uncorrected		Corrected for Leakage + Direct Excitation		Corrected for gamma factor (=0.85)		R (nm)
	$E^*$	$S$	$E_{PR}$	$S$	$E$	$S$	
<b>DNA</b>	0.68	0.61	0.65	0.59	0.68	0.50	5.5
<b>DNA + KF + dNTPs</b>	0.17	0.65	0.026	0.62	0.035	0.51	10.8
<b>DNA + POLB + dNTPs</b>	0.18	0.65	0.0053	0.62	0.013	0.51	12.8
<b>DNA + POLIII<math>\alpha</math> + dNTPs</b>	0.18	0.63	0.027	0.61	0.037	0.50	10.7

Sensor +12/-12	Uncorrected		Corrected for Leakage + Direct Excitation		Corrected for gamma factor (=0.72)		R (nm)
	$E^*$	$S$	$E_{PR}$	$S$	$E$	$S$	
<b>DNA</b>	0.51	0.59	0.45	0.56	0.53	0.50	6.1
<b>DNA + KF + dNTPs</b>	0.35	0.62	0.24	0.58	0.30	0.50	7.1
<b>DNA + POLB + dNTPs</b>	0.36	0.60	0.25	0.57	0.32	0.49	7.0
<b>DNA + POLIII<math>\alpha</math> + dNTPs</b>	0.35	0.61	0.24	0.58	0.31	0.50	7.1

Sensor +12/-12	Uncorrected		Corrected for Leakage + Direct Excitation		Corrected for gamma factor (=1.15)		R (nm)
	$E^*$	$S$	$E_{PR}$	$S$	$E$	$S$	
Sensor +25/-7	0.74	0.77	0.72	0.76	0.69	0.50	5.4
Sensor +25/-7, fully extended primer	0.22	0.76	0.04	0.74	0.03	0.49	10.9
Sensor +12/-12	0.56	0.75	0.51	0.74	0.48	0.46	6.3
Sensor +12/-12, fully extended primer	0.39	0.77	0.30	0.74	0.27	0.50	7.3







# DNA polymerase $\beta$ fingers movement revealed by smFRET suggests a partially closed conformation as a fidelity checkpoint

Carel Fijent<sup>†</sup>, Mariam Mahmoud<sup>‡</sup>, Rebecca Kaup<sup>†</sup>, Jamie Towle-Weicksel<sup>‡</sup>, Joann Sweasy<sup>‡</sup>, Johannes Hohlbein<sup>†</sup>

*Manuscript in preparation*



<sup>†</sup>Laboratory of Biophysics, Wageningen University & Research, Wageningen, The Netherlands

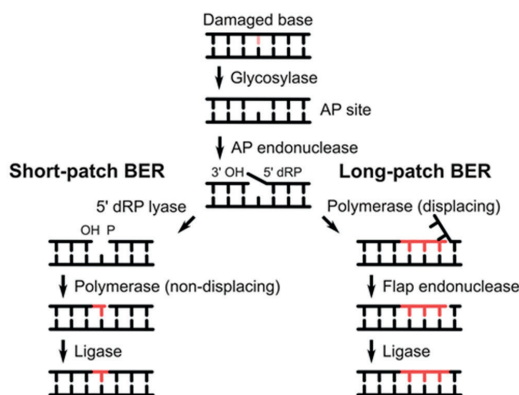
<sup>‡</sup>Department of Therapeutic Radiology, Yale University School of Medicine, New Haven, U.S.A.

## Abstract

The eukaryotic DNA polymerase beta (POLB) plays an important role in cellular DNA repair as it fills up gaps in double-stranded DNA that result from removal of damaged bases. Since defects in DNA repair may lead to cancer and genetic instabilities, POLB and its mechanism of action have been extensively studied, especially substrate binding and fidelity-related conformational changes called “fingers closing”. In this chapter, we applied single-molecule Förster resonance energy transfer (smFRET) to study the conformational dynamics of POLB. Using an acceptor labelled polymerase and a donor labelled DNA substrate, we measured distance changes associated with DNA binding and fingers closing upon addition of nucleotides. Our findings suggest that POLB does not bend its gapped DNA substrate to the extent POLB related crystal structures indicate: instead, the extent of DNA bending seems to be significantly less profound. Furthermore, we visualized dynamic fingers closing in single POLB-DNA complexes upon addition of complementary nucleotides and derived rates of conformational changes. Additionally, we provide evidence that the fingers close only partially when an incorrect nucleotide is bound. This “ajar” conformation found in POLB, a member of the X-family of DNA polymerases, suggests the existence of an additional fidelity checkpoint similar to what has been previously proposed for a member of the A-family, the bacterial *E. coli* DNA polymerase I.

## Introduction

DNA repair is pivotal for maintaining genome integrity<sup>1</sup>. Among the most common damages of DNA are base lesions, in which the chemical structure of a single base has been altered<sup>2,3</sup>. These modifications may disturb proper base pairing and can lead to harmful mutations in the genome. In eukaryotes, the base excision repair (BER) pathway is responsible for replacing these damaged bases<sup>4,5</sup> (Fig. 4.1). Within the BER, damaged base(s) and the corresponding part of the backbone are removed creating a gap of one or multiple nucleotides in the DNA<sup>6</sup>. DNA polymerase Beta (shortened POLB) then binds to the gap and subsequently fills the gap by adding new nucleotides to the 3' end of the primer strand<sup>7,8</sup>. Any remaining part of the former backbone is removed by the lyase activity of POLB. Eventually, a DNA ligase catalyses the formation of a phosphodiester bond to connect the DNA strands<sup>9,10</sup>.



**Figure 4.1.** Schematic overview of the Base Excision Repair (BER) pathway. A damaged base is removed from the DNA by a DNA glycosylase, after which the remaining backbone structure is incised by an AP endonuclease. From there, two different pathways are possible: in short-patch BER, the remaining backbone structure is first removed by a lyase, creating a gap in the DNA. Polymerase beta incorporates the correct nucleotide(s), and a ligase closes the backbone. In long-patch BER, strand displacement DNA synthesis takes place, followed by removal of the old backbone by a flap endonuclease before the DNA is being ligated.

POLB is one of the smallest eukaryotic polymerases and belongs to the X-family of DNA polymerases<sup>11</sup>. It consists of two domains: a polymerase domain and the aforementioned lyase domain<sup>12</sup>. POLB was shown to have an elongated structure in solution<sup>13,14</sup>. Upon binding to gapped DNA, the lyase domain interacts with the 5' phosphate on the downstream strand, while the polymerase domain adopts a structure that has been compared to a hand<sup>12</sup>. Crystal structures have suggested that POLB bends its DNA substrate with an angle of  $\sim 90^\circ$ <sup>15</sup>. Incoming nucleotides then bind to a subdomain known as the “fingers”, forming the ternary complex. A conformational change called “fingers closing” positions the nucleotide closer to the active site to facilitate chemistry. Studies on *E. coli* DNA polymerase I Klenow Fragment (KF), which undergoes a very similar conformational change from an open to a closed conformation, suggested that the fingers do not close entirely when a non-complementary nucleotide is bound. Instead, an intermediate “ajar” conformation was identified, which serves as a “fidelity checkpoint”<sup>16,17</sup>.

In cells, non-complementary nucleotides and ribonucleotides vastly outnumber correct nucleotides; in cancerous cells, nucleotide concentrations can be an order of magnitude higher than in healthy cells<sup>18</sup> further highlighting that effective mechanisms for discriminating correct from non-complementary nucleotides are pivotal for faithful DNA repair. However, for POLB the existence of additional fidelity checkpoints that can be associated to a ajar conformational state is not widely

accepted, although a crystal structure with a mismatched nucleotide suggested that an intermediate conformation might exist<sup>19</sup>.

To study fingers movement of POLB in more detail, Towle-Weicksel *et al.* introduced an assay based on Förster Resonance Energy Transfer (FRET) to monitor fingers-closing using stopped-flow<sup>20</sup>. This approach used a fluorescent label on the DNA substrate and a label on POLB, the latter being attached to the fingers domain after site-directed mutagenesis of valine-303 to cysteine. By fitting their stopped-flow traces to a multi-step kinetic model, the authors were able to extract rates for fingers closing and opening in presence of the complementary nucleotide. Non-complementary nucleotides were not found to induce fingers-closing, leading the authors to hypothesize that discrimination between correct and incorrect nucleotides already takes place before fingers-closing. In later work, the authors showed that a low-fidelity POLB mutant found in cancer cells exhibits altered fingers dynamics<sup>21</sup>.

Here, we developed two single-molecule assays to study the DNA binding behaviour and fingers movement of POLB, for which we used a combination of FRET and Total Internal Reflection Fluorescence (TIRF) microscopy allowing the monitoring of hundreds of molecules in parallel and in real-time. The first assay uses a doubly-labelled gapped DNA substrate to report on binding of unlabelled, wild-type POLB (wtPOLB). A second assay, inspired by single-molecule work on *E. coli* DNA Pol I (KF) by Evans *et al.*<sup>22</sup>, employs a similar design as the stopped-flow experiments discussed above: the fingers domain of POLB is labelled with an acceptor dye, whereas a gapped DNA substrate bears the donor fluorophore. The labelling position on the DNA was chosen such that open and closed conformations of the fingers exhibit different FRET efficiencies ( $E$ ) when POLB is bound to the surface-immobilized DNA substrate. This approach allowed us to visualize fingers movement of individual DNA polymerases, which we used to study the response of the polymerase to both complementary and non-complementary nucleotides added to the buffer. We found that non-complementary nucleotides and complementary ribonucleotides lead to partial fingers closing, while complementary nucleotides induce complete fingers closing. This finding supports the idea that a partially closed “fidelity checkpoint” is a more widely employed mechanism by DNA polymerases than initially anticipated as it can be found in both bacterial and eukaryotic DNA polymerases.

## Materials and methods

### Polymerase purification and labelling

In the following, we use the term “wild-type POLB” to refer to POLB bearing the substitutions C239S, C267S and V303C, introduced to have a single cysteine residue

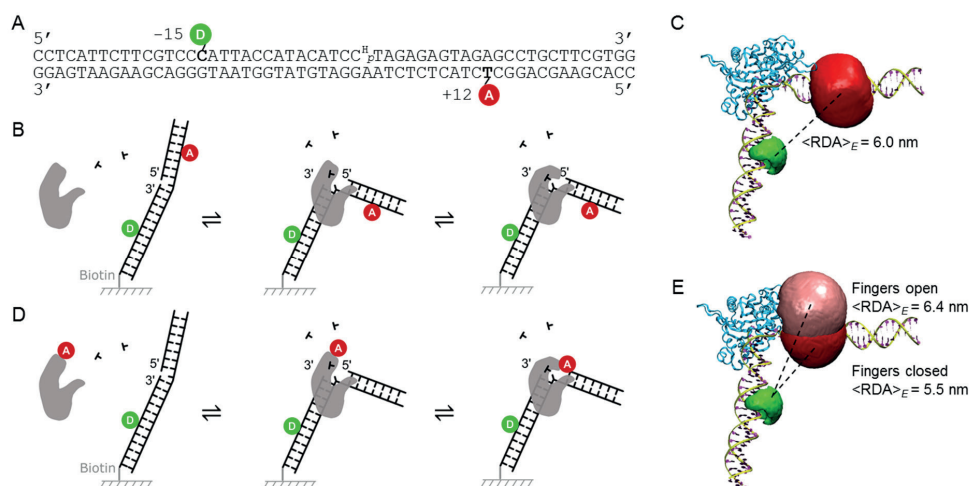
on the fingers subdomain that can react with the fluorophore bearing a maleimide moiety<sup>20</sup>. For the assays in which the fingers conformational change is studied, the V303C was labelled with Alexa Fluor 647 following procedures described before<sup>20</sup>. The labelling efficiency was 60-70% as determined by SDS-polyacrylamide gel electrophoresis (data not shown). For experiments with *E. coli* DNA Polymerase I Klenow Fragment, we used the D424A mutant (hereafter referred to as simply “KF”) that abolishes the 3' to 5' exonuclease activity of KF.

### DNA substrate design

As a first step to construct a gapped DNA construct labelled at adequate positions with fluorophores, we examined crystal structures 1BPX and 1BPY<sup>15</sup> that represent POLB bound to gapped DNA with open and closed fingers, respectively. We extended the DNA from the crystal structures on both sides of the polymerase with a B-DNA helix of our preferred sequence using the 3D-DART server<sup>23</sup>. Next, we used FPS (short for *FRET-restrained Positioning and Screening*) software to model the accessible volumes of the fluorophores at potential labelling positions on the DNA and determine inter-dye distances  $\langle RDA \rangle_E$ <sup>24</sup>. Modelling parameters include the dimensions of the fluorophore and the dimensions of the linker (Supplementary Table 4.1). We selected two labelling positions (-15 and +12, see Fig. 4.2A-C) that are located outside the binding region of POLB. Using Cy3B as a donor fluorophore on the -15 position and Cy5 as an acceptor on +12, these positions are within the distance range for FRET ( $R_{0, Cy3B \rightarrow Cy5} = 6.9$  nm,  $\langle RDA \rangle_{E, model} = 6.0$  nm). Additionally, the Cy3B on the -15 position is also close enough to the fingers subdomain to exhibit FRET with the Alexa Fluor 647 ( $R_{0, Cy3B \rightarrow Alexa\ Fluor\ 647} = 6.9$  nm,  $\langle RDA \rangle_{E, fingers\ open} = 6.4$  nm,  $\langle RDA \rangle_{E, fingers\ closed} = 5.5$  nm, Fig. 4.2D-E). Importantly, these distances translate to a large difference in FRET efficiency ( $E$ ) between the open ( $E = 0.60$ ) and closed ( $E = 0.79$ ) conformations of the fingers.

We annealed the 1 nt gapped DNA complex using a template A from a 30-mer primer sequence (biotin-5'-CCT CAT TCT TCG TCC CAT TAC CAT ACA TCC<sub>H</sub>-3'), a 55-mer template sequence (5'-CCA CGA AGC AGG CTC TAC TCT CTA AGG ATG TAT GGT AAT GGG ACG AAG AAT GAG G-3') and a 24-mer downstream complementary strand (5'-phosphate-TAG AGA GTA GAG CCT GCT TCG TGG-3'), which we ordered from IBA Life Sciences (Germany) and Eurogentec (Belgium). All oligos were HPLC or gel purified prior to use. The primer sequence was internally labelled with donor dye Cy3B through a C6 linker at the previously determined -15 cytosine base; for experiments with wtPOLB, also the template was internally labelled with Cy5 through a C6 linker at the +12 thymine base.





**Figure 4.2.** Experimental design. **A)** Sequence and labelling positions of the gapped DNA construct. The 3' end of the primer is dideoxyterminated and a phosphate group is present at the 5' end of the complementary strand. **B)** Schematic overview of the assay with the bending sensor. A gapped DNA substrate is thought to bend upon binding of POLB, which translates to a change in FRET efficiency. **C)** Extended crystal structure (pdb: 1BPX). Accessible volumes of Cy3B (green) and Cy5 (red) are outside the putative binding region of the polymerase. **D)** Schematic overview of the "fingers closing" assay. POLB, labelled with Alexa Fluor 647 on the fingers domain, binds to a gapped DNA substrate that is labelled with Cy3B. Fingers movement results in a change in FRET efficiency. **E)** Structure 1BPX with the accessible volumes of Alexa Fluor 647, in the open (pink) and closed (red) conformation. (The accessible volume of the closed conformation was modelled using structure 1BPY, which is not shown here.) Cy3B is on the primer. The distance between the donor and acceptor dyes decreases with fingers closing.

### TIRF experiments

Labelled DNA molecules were immobilized on PEGylated glass coverslips using a protocol described before<sup>22</sup>. We used flow channels formed by Ibidi sticky-Slides VI<sup>0.4</sup>. Molecules were imaged on a home-built TIRF microscope, described elsewhere in this thesis. All experiments were performed using Alternating Laser Excitation (ALEX) in which the direct excitation of the donor alternates with the direct excitation of the acceptor fluorophore<sup>25,26</sup>. Experiments on wtPOLB and doubly labelled DNA were performed with laser powers of 1.5 mW ( $\lambda = 561$  nm) and 1.5 mW ( $\lambda = 638$  nm). Excitation time and camera frame time were set to 50 ms. Raw FRET efficiency ( $E^*$ ) was calculated using  $E^* = DA / (DD + DA)$ , in which  $DD$  is donor emission intensity after donor excitation, and  $DA$  is acceptor emission intensity after donor excitation (FRET). Acceptor emission intensity after acceptor excitation  $AA$ , as obtained during ALEX, was used for time trace selection. Experiments with fluorescently labelled

POLB were performed with laser powers of 1.5 mW ( $\lambda = 561$  nm) and 0.75 mW ( $\lambda = 638$  nm). Excitation time and frame time were 25 ms. Surface-immobilized DNA molecules were imaged in a buffer containing either wtPOLB (10, 30, 100 and 300 nM) or labelled POLB-V303C-Alexa Fluor 647 (10 nM). Imaging buffer further contained 50 mM Tris (pH 7.5), 10 mM  $MgCl_2$ , 100 mM NaCl, 100  $\mu$ g/mL BSA, 5% glycerol, 1 mM DTT, 1 mM Trolox, 1% gloxy and 1% glucose. Trolox is a triplet state quencher<sup>27</sup>; gloxy and glucose form an enzymatic oxygen scavenger system to prevent premature fluorophore bleaching<sup>28</sup>. When specified, complementary dTTPs were added to achieve final concentrations of (0.1, 0.5, 1, 2, 5, 10 and 50)  $\mu$ M; concentrations of non-complementary dGTPs and rUTPs were (10, 30, 100, 300, 1000 and 3000)  $\mu$ M. For the experiments with KF, an imaging buffer without NaCl was used.

### Time trace selection and Hidden Markov Modelling

Time traces from individual molecules were collected to measure polymerase binding times and to extract dwell times of open and closed fingers conformations. Because of variations in the signal-to-noise ratio of molecules as well as the presence of bleaching and blinking, an initial selection of molecules was made by hand: only DNA molecules that showed a constant  $DD+DA$  signal with sudden transitions (within 1 frame) from the free to the bound state were selected. To determine the binding times, we first applied a 5 frame moving median filter to all selected traces, before applying additional selection criteria: 1) the sum of  $DD$  and  $AA$  is higher than 50 photons and 2) the FRET efficiency is higher than 0.4. Additionally, settings were such that disappearance of donor signal (bleaching) was interpreted as the end of the trace, and disappearance of acceptor signal (bleaching or polymerase dissociation) was interpreted as the end of a binding event. Filtering traces following these criteria sometimes resulted in longer binding events being cut in multiple shorter events due to remaining noise. To prevent these cases, an exception was added to allow single point excursions to lower intensities or FRET efficiencies. The final algorithm was found to identify most binding events that are also detectable by visual inspection. Extremely short events, however, are often not detected because of the strong median filter. These events may therefore be under-represented in our dwell time histograms. For extraction of fingers conformational changes, binding events from experiments with dTTPs were loaded into ebFRET, a software package for Hidden Markov Modelling<sup>29</sup>. Because the final point of some binding events may have a donor or acceptor that is already decreasing in intensity (just before the cut-off that we set for dissociation or bleaching), we removed these points from the traces by applying a padding of 10 time points. Next, the prior for the minimal centre position (open fingers) was set to  $E^* = 0.4$ , and that for the maximal centre position (closed fingers) to  $E^* = 0.8$ . The convergence threshold was set to  $10^{-5}$ .

## Results

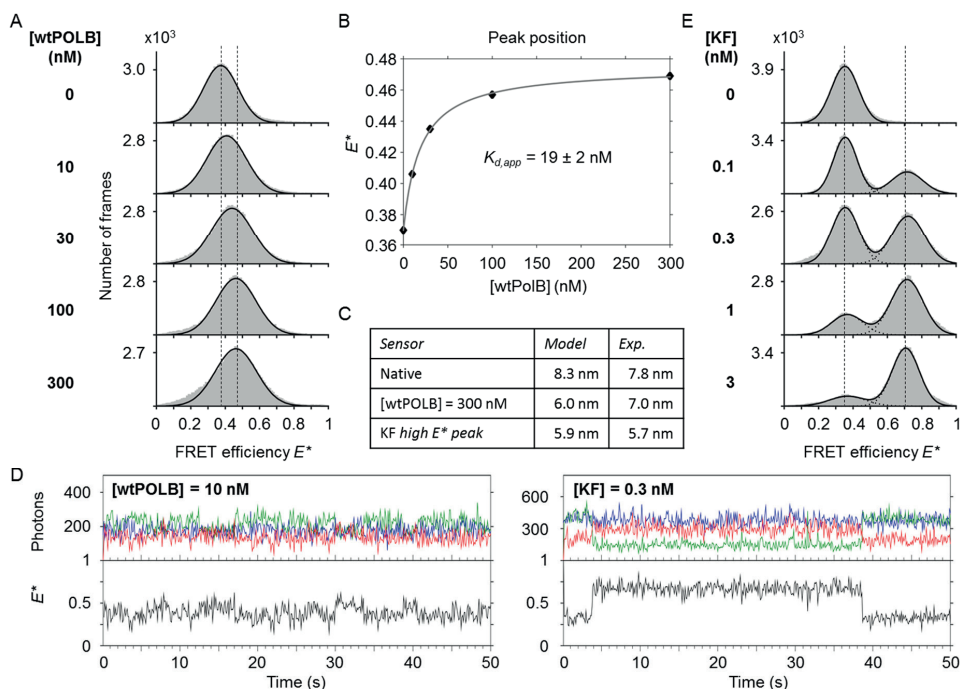
### **POLB bends its substrate, but in a different fashion than crystal structures suggest**

First, we assessed binding of wtPOLB to dsDNA with a 1 nt gap, mimicking the BER pathway intermediate that is the natural and preferred substrate of POLB<sup>12</sup>. Crystal structures 1BPX and 1BPY suggested that the DNA adopts a sharply bent conformation ( $\sim 90^\circ$ ) after binding of POLB. Bending of the DNA is therefore an indicator for polymerase binding. We labelled our DNA substrate with a donor dye on the primer and an acceptor dye on the template, at positions that are outside the binding region of the polymerase (as judged from crystal structures 1BPX and 1BPY), thus creating a “bending sensor”.

The native DNA substrate has a FRET efficiency  $E^* = 0.37$  (Fig. 4.3A) which corresponds to an inter-fluorophore distance of 7.8 nm after correcting for leakage, direct excitation and gamma<sup>26</sup> (see Supplementary Table 4.2 for step-by-step correction). A model of a rigid double-stranded DNA helix in PyMOL predicted an inter-dye distance  $\langle RDA \rangle_E$  of 8.3 nm. These values are slightly different, but we note that our static DNA model cannot account for the flexibility introduced by the gap.

Increasing the concentration of POLB lead to a peak shift, levelling off towards a FRET efficiency of  $\sim 0.47$  (Fig. 4.3B), reflecting the bent conformation of the DNA and, after corrections, a distance of 7.0 nm (Fig. 4.3C). The distance calculated from structures 1BPX and 1BPY is significantly shorter (6.0 nm). We used the peak shift to determine the dissociation constant  $K_d$ , describing the binding affinity of DNA and POLB, and obtained  $K_d = 19 \pm 2$  nM, in good agreement with values published before<sup>21,30</sup>. A trace of a single DNA molecule at [wtPOLB] = 10 nM shows excursions between the low and high FRET states (Fig. 4.3D). Interpretation, however, is complicated by the noise in combination with a small  $\Delta E^*$ .

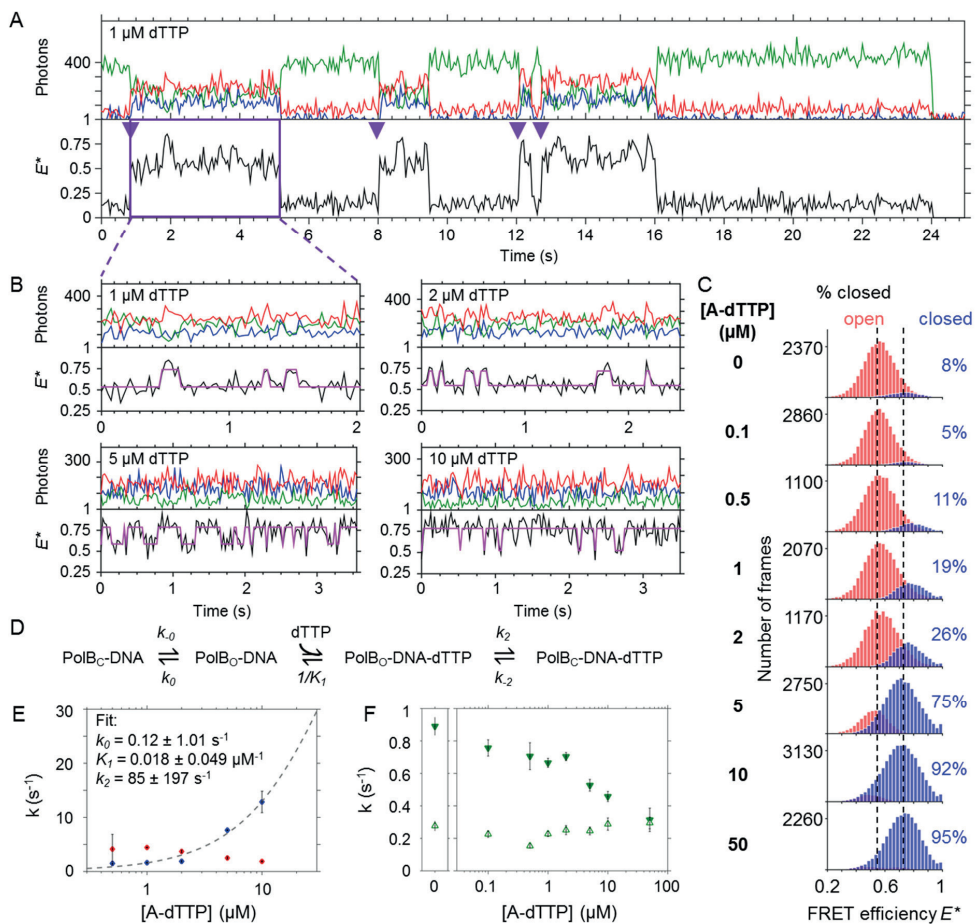
Additionally, we tested the response of our bending sensor to *E. coli* DNA Polymerase I Klenow Fragment (KF). Like POLB, this polymerase bends a gapped DNA substrate to an angle of approximately  $90^\circ$  (as determined by FRET measurements and structural modelling by Craggs *et al.*, data not yet published). Upon increasing the concentration of KF, we did not observe a peak shift; instead, we saw a new peak appearing at  $E^* = 0.71$  (Fig. 4.3E). Accurate FRET calculations revealed a corresponding inter-fluorophore distance of 5.7 nm (Fig. 4.3C, Supplementary Table 4.2) in excellent agreement with a FRET restrained structural model by Craggs *et al.* predicting a distance of 5.9 nm.



**Figure 4.3.** Response of the gapped DNA bending sensor to wtPOLB and KF. **A)** The mean FRET efficiency  $E^*$  increases with increasing concentrations of wtPOLB. **B)** Peak positions from **A)** plotted against wtPOLB concentration, fitted to a Langmuir binding isotherm (grey line) reveal an apparent  $K_d$  of  $19 \pm 2$  nM. **C)** Modelled and experimentally determined inter-dye distances of both the native bending sensor and the bent conformation. **D)** Time traces of a single DNA bending sensors at [wtPOLB] = 10 nM and [KF] = 0.3 nM. **E)** Bending sensors respond differently to increasing concentrations of KF, showing a growing occupancy of a new high-FRET species.

### Hidden Markov Modelling resolves fingers movement in presence of complementary dTTPs

Next, we studied the ability of labelled POLB to report on the conformation of the fingers subdomain. We performed a titration of labelled POLB with increasing [A-dTTP] concentrations. Time traces of single DNA molecules showed binding events of labelled POLB as an increase in AA and the appearance of FRET (Fig. 4.4A). Hidden Markov Modelling (HMM) was used to identify the open (low  $E^*$ ) and closed (high  $E^*$ ) conformations within time traces of individual binding events (Fig. 4.4B). At [dTTP] = 0.5  $\mu$ M, traces predominantly show low FRET efficiency, with only short excursions to the high FRET efficiency that is associated with closed fingers. At higher dTTP concentrations, longer residence times in the closed state are observed. We constructed FRET efficiency histograms and indicated the open and closed populations, as determined by HMM (Fig. 4.4C). Without dTTPs being present, the



**Figure 4.4.** Fingers opening and closing of POLB revealed by smFRET. **A)** Time trace of a single DNA molecule in presence of 10 nM labelled POLB and 1  $\mu\text{M}$  complementary dTTTPs. POLB binding events are indicated with purple triangles. At  $t = 24$  s, the donor bleaches. **B)** Time traces of labelled POLB-DNA complexes, at various concentrations of dTTP. The first trace ( $[\text{dTTP}] = 1 \mu\text{M}$ ) is a binding event taken from A). FRET efficiency  $E^*$  (black trace) is calculated from the DD signal (green trace) and the DA signal (red trace) in the upper panel. The AA signal (blue trace) is shown here as well to indicate that the observed events are not due to acceptor photophysics. An HMM fit (magenta) indicates the open and closed conformation of the fingers. **C)** Corresponding FRET efficiency histograms (32 bins between  $E^* = 0.2$  and 1) of the POLB-DNA complexes. The FRET efficiencies of the open (shown in red) and closed (shown in blue) conformation were plotted after the states have been assigned via HMM. **D)** Schematic model used to describe the dynamics of fingers movement in the binary ( $k_0/k_{-0}$ ) and ternary complex ( $k_2/k_{-2}$ ). **E)** The observed closing rates  $k_{\text{close,obs}}$  (blue) and opening rates  $k_{\text{open}}$  (red) plotted against  $[\text{dTTP}]$ . Data were fit to a function described in the main text (dashed line) and derived from the model depicted in C). Error bars represent the

95% confidence interval obtained from fitting the dwell times. **F)** Complementary dTTPs stabilize the POLB-DNA complex. Rates  $k_{on}$  (green open triangles) and  $k_{off}$  (green solid triangles) plotted against [dTTP].  $k_{off}$  decreases with increasing concentrations of dTTP, while  $k_{on}$  remains constant. Error bars represent the 95% confidence interval obtained from fitting the dwell times.

fingers mostly adopt the open conformation (92%). With increasing dTTP concentration, the closed conformation is increasingly populated. At [dTTP] = 50  $\mu$ M, the fingers are mostly closed (95%).

Using accurate FRET, we determined the distances associated with open and closed fingers (Supplementary Table 4.3). We found an inter-fluorophore distances of 6.5 nm for the open and 5.6 nm for the closed conformation. We note that these distances are in excellent agreement with the distances of 6.4 nm and 5.5 nm as predicted from structural modelling with FPS software (see Methods).

Dwell time histograms of the open and closed conformations were constructed and fitted with exponential decay curves (concentrations of 0, 0.1 and 50  $\mu$ M were left out due to an insufficient number of transitions). The rate of fingers closing is extracted from the dwell times in the open conformation and vice versa. Plotting  $k_{obs,close}$  and  $k_{open}$  against the concentration of dTTPs showed that the closing rate is concentration-dependent; the opening rate varies as well, but to a lesser extent. A simple model that links fingers closing to the affinity of complementary nucleotides was previously described (Fig. 4.4D)<sup>22</sup>. Here, the concentration of dTTPs relates to  $k_{obs,close}$  as follows:

$$k_{obs,close} = \frac{k_0 + [dTTP]k_2K_1}{1 + [dTTP]K_1}$$

in which  $k_0$  is the closing rate of the binary complex,  $K_1$  is the association constant for dTTPs and  $k_2$  is the closing rate with dTTP bound to the fingers. As fingers closing was very rare in our traces at [dTTP] = 0  $\mu$ M, we could not provide the model with a value for  $k_0$ . We therefore fitted our data with no constraints for  $k_0$ ,  $K_1$  and  $k_2$  (Fig. 4.4E) and found a  $k_2$  of  $85 \pm 197$  s<sup>-1</sup> (mean  $\pm$  standard deviation),  $K_1$  of  $0.018 \pm 0.049$   $\mu$ M<sup>-1</sup> (corresponding to a  $K_{d(dTTP)}$  of 56  $\mu$ M) and a  $k_0$  of  $0.12 \pm 1.01$  s<sup>-1</sup>.

We noted that the duration of POLB-DNA binding events increased with increasing [dTTP] by observing an increase of  $k_{off}$ , while  $k_{on}$  is not affected (Fig. 4.4F).

### **Fingers close only partially in presence of non-complementary dNTPs and complementary rNTPs**

To investigate the potential existence of a “partially closed” or “ajar” conformation in POLB, we studied the fingers conformational change in presence of non-complementary dNTPs and complementary rNTPs. Previous work has shown that increasing concentrations of an incorrect nucleotide shift the position of the “fingers



open" peak of KF towards a slightly higher FRET efficiency, likely caused by the polymerase quickly screening and rejecting incorrect nucleotides<sup>17,22</sup>. For POLB, we performed a similar experiment: we tested a range of different concentrations of dGTPs and rUTPs, and plotted the position of the main peak (Fig. 4.5A-C). Indeed, for both incorrect nucleotides a shift in  $E^*$  from  $\sim 0.55$  to  $\sim 0.59$  is observed, suggesting that POLB (like KF) has a fidelity checkpoint at a partially closed fingers conformation. The associated  $\Delta E^*$ , however, is too small for clear detection of this state in individual single-molecule time traces.

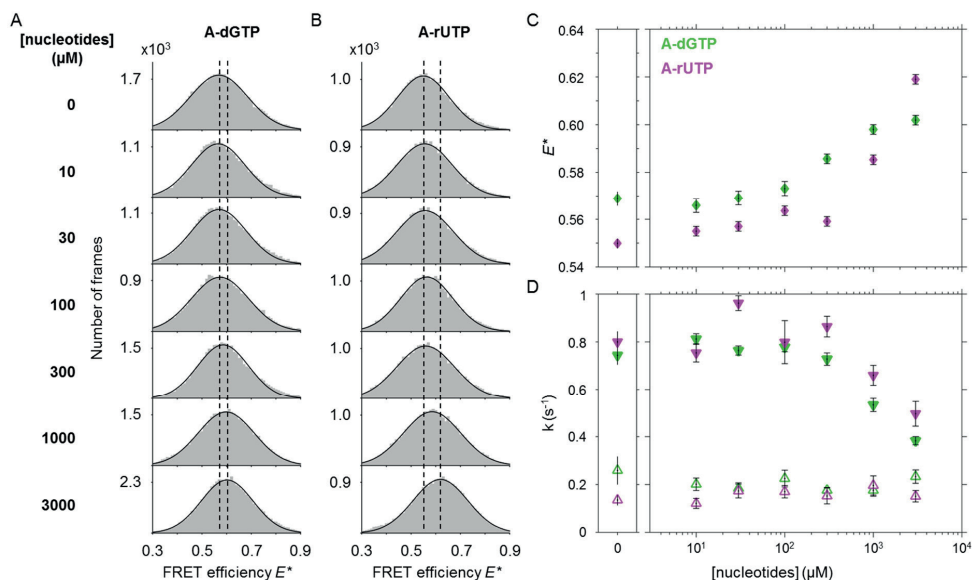
We asked if incorrect nucleotides have an influence on the stability of the ternary complex. Evans *et al.* and Markiewicz *et al.* showed that non-complementary nucleotides increase  $k_{off}$  for KF<sup>22,31</sup>. We identified all polymerase binding events in our time traces with labelled POLB and constructed dwell time histograms. Fitting with an exponential decay function yielded values for  $k_{on}$  and  $k_{off}$  for every nucleotide concentration of the titration series (Fig. 4.5D). While  $k_{on}$  remained constant,  $k_{off}$  decreased at high nucleotide concentrations, indicating that for POLB bound to gapped DNA even incorrect nucleotides stabilize the ternary complex albeit only at higher concentrations than for the correct dNTP.

## Discussion

The use of smFRET allowed us to observe and analyse conformational changes of individual POLB-DNA complexes in real-time, thereby overcoming some of the ensemble averaging inherent to conventional fluorescence-based techniques, such as stopped-flow.

Our experiments with wtPOLB showed substrate binding with a  $K_d$  of  $19 \pm 2$  nM. Other studies have also reported values in the low nanomolar range for gapped DNA constructs: fitting stopped-flow data resulted in a  $K_d$  of 5 nM<sup>21</sup>, and a titration based on single-turnover analysis at different DNA concentrations revealed a  $K_d$  of 22 nM<sup>30</sup>. Surprisingly, strong DNA bending upon polymerase binding, as predicted by various structures resolved with X-ray crystallography, does not occur to the expected extent suggesting that the bend angle in solution is larger than  $90^\circ$ . Compared to conditions required for successful crystallisation, we were able to use a DNA sequence with longer up- and downstream strands (here: upstream region 30 base pairs, downstream region 24 base pairs; pdb: 1BPX is 10 and 5 base pairs, respectively). While an upstream binding region of  $11 \pm 1$  nt was revealed with fluorescently labelled DNA previously<sup>32</sup>, the short downstream region in 1BPX raises the question whether the crystal structures obscure interactions of the polymerase with parts of the DNA that are farther away from the gap.





**Figure 4.5.** Incorrect nucleotides lead to partial fingers closing and stabilize the ternary complex. **A)** FRET efficiency histograms of POLB-DNA complexes at increasing  $[A\text{-dGTP}]$ . Histograms were filtered based on colocalization of donor (DD) and acceptor (AA) fluorophores (see Supplementary Information). Dashed lines are added for visual guidance. **B)** FRET efficiency histograms of POLB-DNA complexes at increasing  $[A\text{-rUTP}]$ . **C)** Position of the main peak plotted against  $[nucleotide]$ . Both dGTPs (green diamonds) and rUTPs (magenta diamonds) cause a shift in the peak of the open conformation. Error bars indicate the 95% confidence intervals on the fits. **D)** Rates  $k_{on}$  (open triangles) and  $k_{off}$  (solid triangles) at increasing concentrations of dGTPs (green) and rUTPs (magenta). Error bars indicate the 95% confidence intervals on the fits.  $k_{off}$  decreases with increasing concentrations of incorrect nucleotides, while  $k_{on}$  remains constant.

We further studied the conformational change associated to “fingers closing” using a fluorescently labelled version of POLB. We observed an increase in the rate of fingers closing with increasing concentrations of the complementary dNTP, as expected for an induced fit mechanism. Previous studies<sup>20,21</sup> found a rapid fingers closing rate of  $98\text{ s}^{-1}$ , close to the maximum rate of  $85\text{ s}^{-1}$  that we determined from our fit. The opening rate is concentration-dependent as well, and decreases with increasing concentrations of dTTPs. This may indicate that  $k_{-2}$  is lower than  $k_{-0}$ , i.e. the closed state is stabilized when a complementary nucleotide is bound. However, we note that the likelihood of missing events increases with dTTP concentration, leading to lengthening of the closed-state dwell times and extraction of a lower apparent opening rate. Generally, because of our time resolution of 50 ms, the dynamics of opening and closing become increasingly difficult to analyze at dTTP concentrations higher than  $10\text{ }\mu\text{M}$ .

The  $K_d$  of the incoming nucleotide ( $1 / K_1 = 56 \mu\text{M}$ ) is higher than determined before using stopped-flow and chemical quench analysis ( $K_d = 2.5 \mu\text{M}$ )<sup>20,21</sup>. The reason for this difference is that we fitted our observations to a model that accounts for fingers opening and closing in presence and absence of dNTPs. An analysis based on the assumption that fingers close only and immediately after nucleotide binding, however, would find a  $K_d$  at the concentration where the open and closed conformations are equally populated (in our data: between 2 and 5  $\mu\text{M}$ ).

We found evidence that the fingers adopt a partially closed conformation when supplied with incorrect nucleotides, reminiscent of KF. Both addition of dGTPs and rUTPs result in partial closing, indicating that at this ‘fidelity checkpoint’, incoming nucleotides are both screened for complementarity and backbone structure. Ensemble studies on KF suggested that complementary rNTPs proceed farther in the reaction pathway than non-complementary dNTPs<sup>32,33</sup>, but the difference in  $E^*$  is too small in our assay to investigate that further. A crystal structure of POLB with a dG-dAMPCPP mismatch in the active site initially suggested that fingers close at least partially<sup>19</sup>. It will be interesting to see in follow-up single-molecule studies whether this checkpoint is more dominantly populated in mutator variants of POLB, as it has been shown for bacterial DNA Pol I<sup>17</sup>.

Furthermore, we found that non-matching nucleotides do not destabilize the polymerase-DNA complex, as was found to be the case for KF<sup>22,31</sup>. Instead, at high concentrations ( $> 1 \text{ mM}$ ) these nucleotides stabilize the POLB-DNA complex by lowering  $k_{\text{off}}$ . This seems counterintuitive, since prolonged binding times of a flawed ternary complex may increase the chance of actual incorporation of the wrong base. Non-matching nucleotides, however, are always more abundant in cells than the correct dNTP. Considering nucleotides are randomly sampled, repair of damaged DNA may take much longer if a ternary complex with a non-complementary nucleotide is destabilized.

The main advantage of using smFRET to study fingers movement is the direct observation of this conformational change in individual molecules, making it straightforward to calculate rates for opening and closing. In addition, smFRET provides access to the associated intermolecular distances, which support our interpretation that the observed fluctuations in FRET efficiency are indeed switching between an open and closed state. An even more direct method to study fingers opening and closing would see both fluorophores on the polymerase, to avoid convolution of fingers movement with any potential movement in the DNA substrate.

In summary, our work offers novel insights in the mechanisms that govern POLB fidelity. With a doubly-labelled POLB variant currently being designed, it is a promising start for future single-molecule work on POLB.

## Acknowledgments

*E. coli* DNA Polymerase I Klenow Fragment was kindly supplied by Catherine Joyce, Tim Craggs and Achillefs Kapanidis.

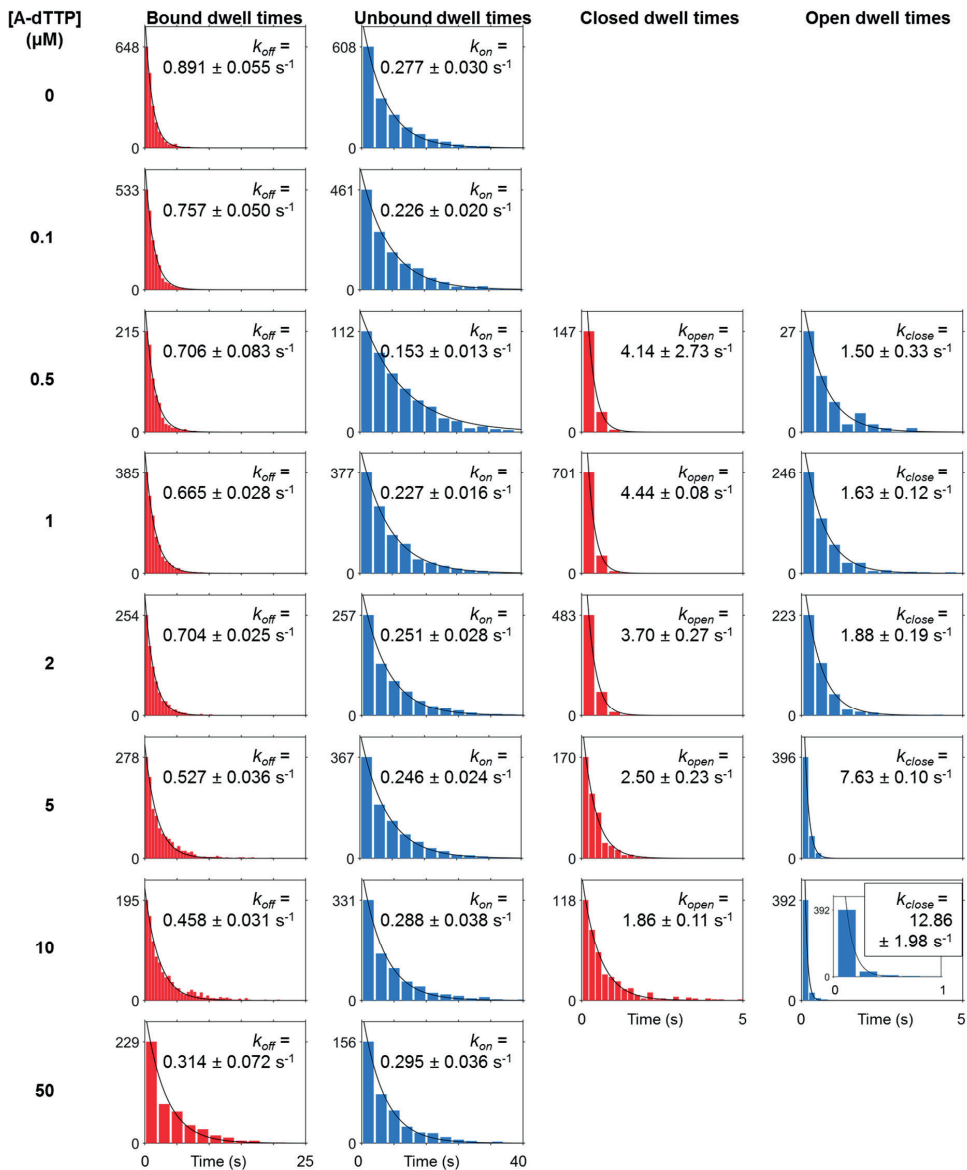
## References

1. Lindahl, T. & Wood, R. D. Quality Control by DNA Repair. *Science* **286**, 1897–1905 (1999).
2. Barnes, D. E. & Lindahl, T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* **38**, 445–476 (2004).
3. Bauer, N. C., Corbett, A. H. & Doetsch, P. W. The current state of eukaryotic DNA base damage and repair. *Nucleic Acids Res.* **43**, 10083–10101 (2015).
4. Dianov, G. & Lindahl, T. Reconstitution of the DNA base excision—repair pathway. *Curr. Biol.* **4**, 1069–1076 (1994).
5. Krokan, H. E. & Bjørås, M. Base Excision Repair. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).
6. Fromme, J. C., Banerjee, A. & Verdine, G. L. DNA glycosylase recognition and catalysis. *Curr. Opin. Struct. Biol.* **14**, 43–49 (2004).
7. Sobol, R. W. *et al.* Requirement of mammalian DNA polymerase- $\beta$  in base-excision repair. *Nature* **379**, 183–186 (1996).
8. Singhal, R. K. & Wilson, S. H. Short gap-filling synthesis by DNA polymerase beta is processive. *J. Biol. Chem.* **268**, 15906–15911 (1993).
9. Prasad, R. *et al.* Specific Interaction of DNA Polymerase  $\beta$  and DNA Ligase I in a Multiprotein Base Excision Repair Complex from Bovine Testis. *J. Biol. Chem.* **271**, 16000–16007 (1996).
10. Cappelli, E. *et al.* Involvement of XRCC1 and DNA Ligase III Gene Products in DNA Base Excision Repair. *J. Biol. Chem.* **272**, 23970–23975 (1997).
11. Yamtich, J. & Sweasy, J. B. DNA Polymerase Family X: Function, Structure, and Cellular Roles. *Biochim. Biophys. Acta* **1804**, 1136–1150 (2010).
12. Beard, W. A. & Wilson, S. H. Structure and Mechanism of DNA Polymerase  $\beta$ . *Biochemistry* **53**, 2768–2780 (2014).
13. Kim, S.-J. *et al.* Characterization of the Tryptophan Fluorescence and Hydrodynamic Properties of Rat DNA Polymerase  $\beta$ . *J. Mol. Biol.* **244**, 224–235 (1994).

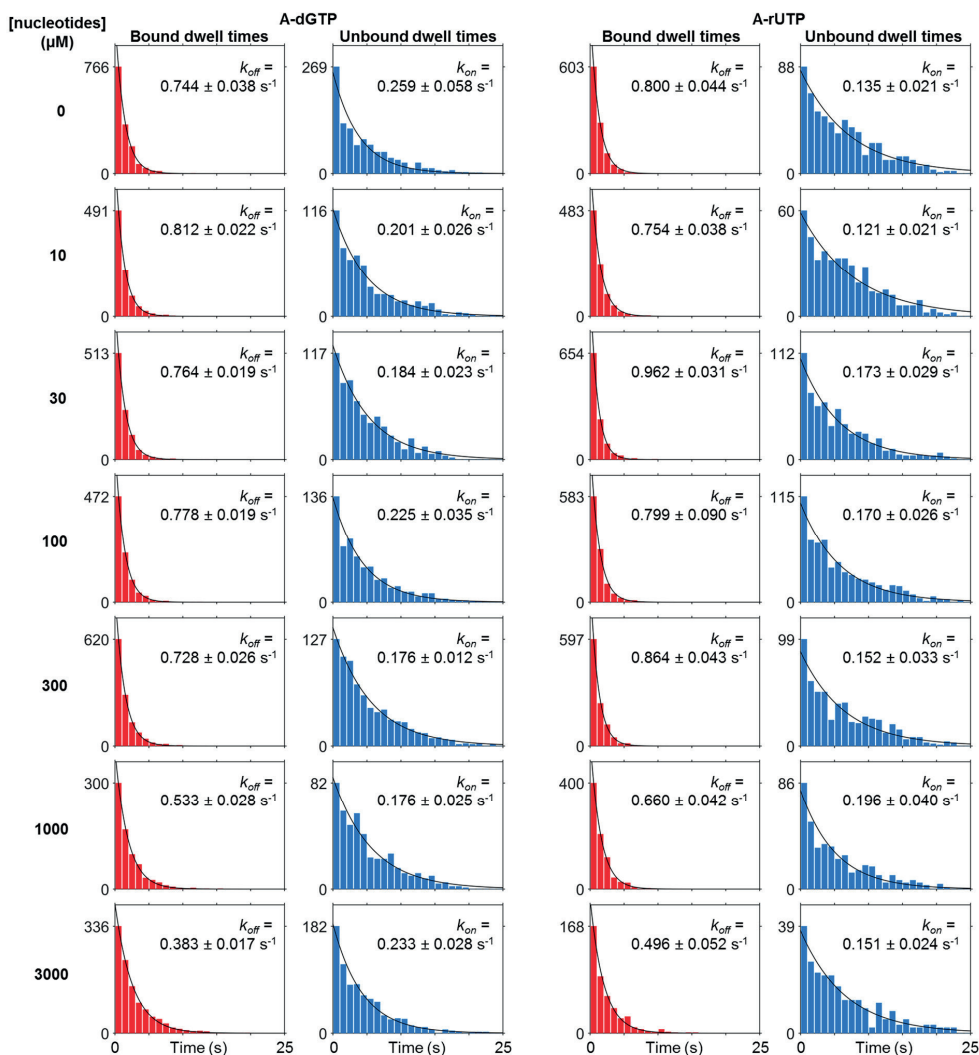
14. Tang, K.-H., Niebuhr, M., Aulabaugh, A. & Tsai, M.-D. Solution structures of 2 : 1 and 1 : 1 DNA polymerase-DNA complexes probed by ultracentrifugation and small-angle X-ray scattering. *Nucleic Acids Res.* **36**, 849–860 (2008).
15. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J. & Pelletier, H. Crystal Structures of Human DNA Polymerase  $\beta$  Complexed with Gapped and Nicked DNA: Evidence for an Induced Fit Mechanism. *Biochemistry* **36**, 11205–11215 (1997).
16. Wu, E. Y. & Beese, L. S. The Structure of a High Fidelity DNA Polymerase Bound to a Mismatched Nucleotide Reveals an ‘Ajar’ Intermediate Conformation in the Nucleotide Selection Mechanism. *J. Biol. Chem.* **286**, 19758–19767 (2011).
17. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nat. Commun.* **4**, (2013).
18. Traut, T. W. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1–22 (1994).
19. Batra, V. K., Beard, W. A., Shock, D. D., Pedersen, L. C. & Wilson, S. H. Structures of DNA Polymerase  $\beta$  with Active-Site Mismatches Suggest a Transient Abasic Site Intermediate during Misincorporation. *Mol. Cell* **30**, 315–324 (2008).
20. Towle-Weicksel, J. B. *et al.* Fluorescence Resonance Energy Transfer Studies of DNA Polymerase  $\beta$  THE CRITICAL ROLE OF FINGERS DOMAIN MOVEMENTS AND A NOVEL NON-COVALENT STEP DURING NUCLEOTIDE SELECTION. *J. Biol. Chem.* **289**, 16541–16550 (2014).
21. Mahmoud, M. M. *et al.* Defective Nucleotide Release by DNA Polymerase  $\beta$  Mutator Variant E288K Is the Basis of Its Low Fidelity. *Biochemistry* **56**, 5550–5559 (2017).
22. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res.* gkv547 (2015). doi:10.1093/nar/gkv547
23. van Dijk, M. & Bonvin, A. M. J. J. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res.* **37**, W235–W239 (2009).
24. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).
25. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).
26. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
27. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *J. Am. Chem. Soc.* **131**, 5018–5019 (2009).

28. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).
29. van de Meent, J.-W., Bronson, J. E., Wiggins, C. H. & Gonzalez, R. L. Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophys. J.* **106**, 1327–1337 (2014).
30. Berg, B. J. V., Beard, W. A. & Wilson, S. H. DNA Structure and Aspartate 276 Influence Nucleotide Binding to Human DNA Polymerase  $\beta$  IMPLICATION FOR THE IDENTITY OF THE RATE-LIMITING CONFORMATIONAL CHANGE. *J. Biol. Chem.* **276**, 3408–3416 (2001).
31. Markiewicz, R. P., Vrtis, K. B., Rueda, D. & Romano, L. J. Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucleic Acids Res.* **40**, 7975–7984 (2012).
32. Joyce, C. M. *et al.* Fingers-Closing and Other Rapid Conformational Changes in DNA Polymerase I (Klenow Fragment) and Their Role in Nucleotide Selectivity. *Biochemistry* **47**, 6103–6116 (2008).
33. Garalde, D. R. *et al.* Distinct Complexes of DNA Polymerase I (Klenow Fragment) for Base and Sugar Discrimination during Nucleotide Substrate Selection. *J. Biol. Chem.* **286**, 14480–14492 (2011).

## Supplementary Information

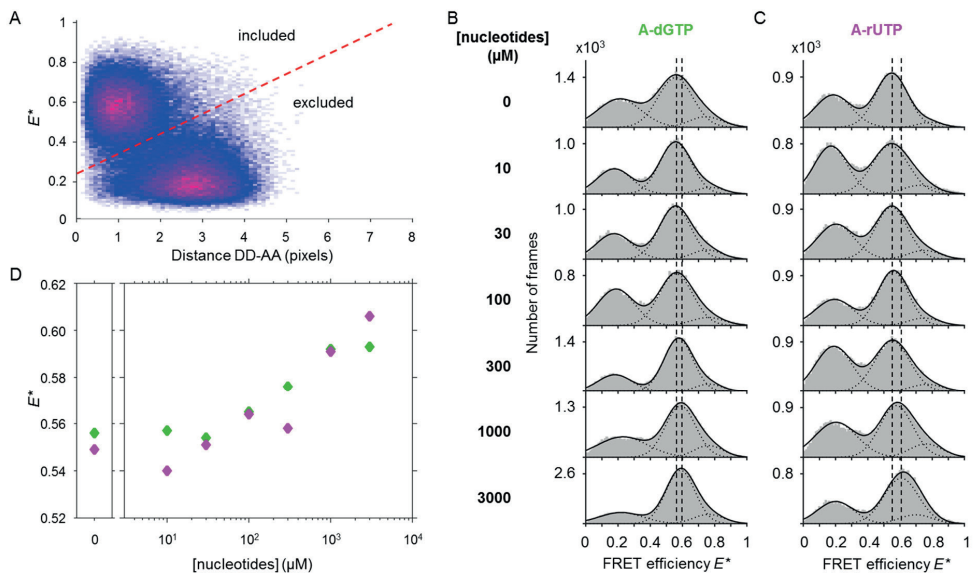


**Supplementary Figure 4.1.** Left: dwell time histograms of POLB binding events in presence of increasing concentrations of correct nucleotide dTTP. Rates  $k_{on}$  and  $k_{off}$  are plotted in Figure 4.4E. Right: dwell time histograms of open and closed fingers. Rates  $k_{close,obs}$  and  $k_{open}$  are plotted in Figure 4.4D. All rates are given with the 95% confidence interval.



**Supplementary Figure 4.2.** Dwell time histograms of POLB binding events in presence of increasing concentrations of incorrect nucleotides. Rates  $k_{on}$  and  $k_{off}$  are plotted in Figure 4.5D. All rates are given with the 95% confidence interval.





**Supplementary Figure 4.3.** **A)** Implementation of filtering FRET efficiency data points based on monitoring the spatial localization of the donor and acceptor in their respective detection channels on the camera. The positions of molecules in the DD and AA detection channels are determined with subpixel accuracy using a free elliptical 2D Gaussian fit. The observed FRET efficiency of each molecule is then plotted against the Euclidean distance between each DD and AA localization. The resulting plot reveals two well-separated populations: a population of actual POLB-DNA complexes (at  $0.55 < E^* < 0.60$ ), and a population of non-specifically adsorbing POLB molecules (at low  $E^*$ ). Note that the first population does not center around Distance = 0 pixels because alignment of the channels is not perfect. Data points above the dashed red line are used to construct the FRET efficiency histograms seen in Figure 4.5A+B. **B-D)** Raw data, not filtered based on distance. These histograms show a similar trend as the filtered histograms shown in Figure 4.5. **B)** FRET efficiency histograms of POLB-DNA complexes at increasing [dGTP]. These unfiltered histograms were best fit with 3 Gaussians with mean values around  $E^* = 0.2$  (non-specific adsorption),  $E^* = 0.57-0.60$  (open / partially closed conformation) and  $E^* = 0.8$  (closed conformation as determined from measurements with dTTP). Dashed lines are added for visual guidance. **C)** FRET efficiency histograms of POLB-DNA complexes at increasing [rUTP]. **D)** Position of the main peak plotted against nucleotide concentration. Both dGTPs (green diamonds) and rUTPs (magenta diamonds) cause a shift in the peak of the open conformation.

**Supplementary Table 4.1.** Parameters used for calculation of accessible volumes in FPS software.

Dye	Type	Length (nm)	Width (nm)	R1 (nm)	R2 (nm)	R3 (nm)
<b>Cy3B</b>	C6 NHS	1.42	0.45	0.82	0.33	0.22
<b>Cy5</b>	C6 NHS	2.30	0.45	1.10	0.30	0.15
<b>Alexa Fluor 647</b>	C2 maleimide	2.10	0.45	1.10	0.47	0.15

**Supplementary Table 4.2.** Calculation of distances from raw FRET efficiencies for the bending sensor.

Sensor	Uncorrected		Corrected leakage and direct excitation		Corrected for $\gamma$ and $\beta$		Distance $\langle RDA \rangle_E$ (nm)
	$E^*$	S	$E_{PR}$	S	E	S	
<b>Native</b>	0.37	0.61	0.21	0.58	0.32	0.49	7.8
<b>[wtPOLB] = 300 nM</b>	0.47	0.59	0.34	0.56	0.47	0.49	7.0
<b>[wtKF] = 3 nM</b>	0.71	0.58	0.66	0.55	0.76	0.49	5.7

**Supplementary Table 4.3.** Calculation of distances from raw FRET efficiencies for complexes of labelled POLB and labelled DNA

[dTTP] ( $\mu$ M)	Uncorrected		Corrected leakage and direct excitation		Corrected for $\gamma$ and $\beta$		Distance $\langle RDA \rangle_E$ (nm)
	$E^*$	S	$E_{PR}$	S	E	S	
<b>0 (fingers open peak)</b>	0.54	0.71	0.49	0.72	0.59	0.44	6.5
<b>50 (fingers closed peak)</b>	0.72	0.71	0.68	0.71	0.78	0.46	5.6



# Simple nanofluidic devices for high-throughput, non-equilibrium studies at the single-molecule level

Carel Fijen<sup>†\*</sup>, Mattia Fontana<sup>†\*</sup>, Serge Lemay<sup>‡</sup>, Klaus Mathwig<sup>#</sup>, Johannes Hohlbein<sup>†</sup>.  
DOI: 10.1101/201079 (\*equal contributions)

*Manuscript in preparation*



<sup>†</sup>Laboratory of Biophysics, Wageningen University & Research, Wageningen, The Netherlands

<sup>‡</sup>MESA+ Institute for Nanotechnology, University of Twente, Enschede, The Netherlands

<sup>#</sup>Groningen Research Institute of Pharmacy, University of Groningen, Groningen, The Netherlands

## Abstract

Single-molecule detection schemes offer powerful means to overcome static and dynamic heterogeneity inherent to complex samples. Probing chemical and biological interactions and reactions with high throughput and time resolution, however, remains challenging and often requires surface-immobilized entities. Here, utilizing camera-based fluorescence microscopy, we present glass-made nanofluidic devices in which fluorescently labelled molecules flow through nanochannels that confine their diffusional movement. The first design features an array of parallel nanochannels for high-throughput analysis of molecular species under equilibrium conditions allowing us to record 200.000 individual localization events in just 10 minutes. Using these localizations for single particle tracking, we were able to obtain accurate flow profiles including flow speeds and diffusion coefficients inside the channels. A second design featuring a T-shaped nanochannel enables precise mixing of two different species as well as the continuous observation of chemical reactions. We utilized the design to visualize enzymatically driven DNA synthesis in real time and at the single-molecule level. Based on our results, we are convinced that the versatility and performance of the nanofluidic devices will enable numerous applications in the life sciences.

## Introduction

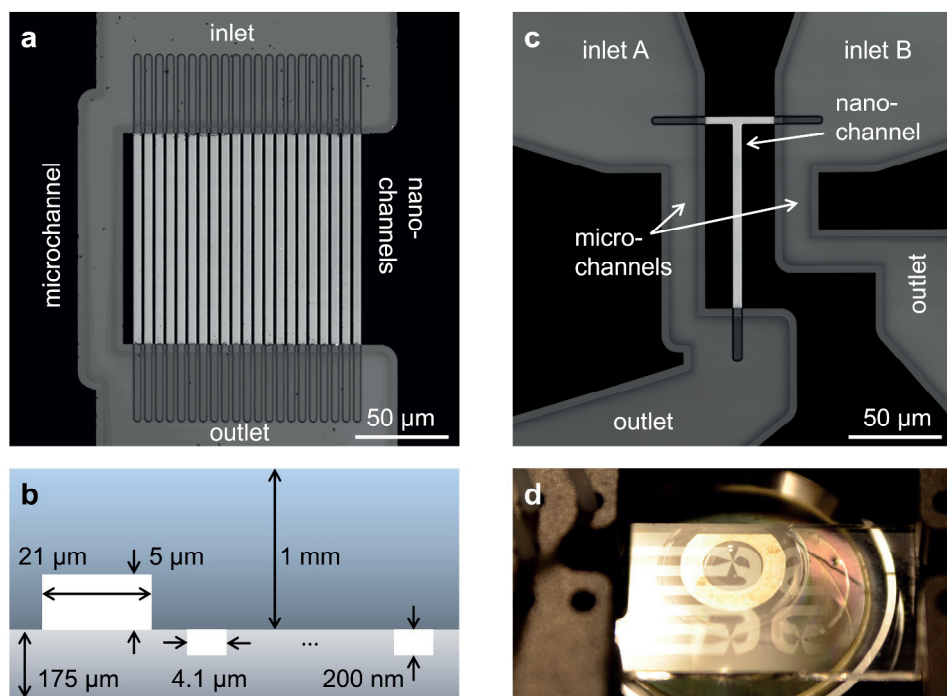
In many areas of the life sciences, single-molecule techniques are playing an increasingly important role in identifying and analyzing static or dynamic interactions of (bio-)molecules with high spatiotemporal resolution<sup>1-3</sup>. Among the large number of available frameworks<sup>4</sup>, single-molecule fluorescence detection (SMFD) schemes, such as confocal or camera-based total-internal-reflection fluorescence (TIRF) microscopy, are frequently chosen due to their simplicity, robustness and ease of use<sup>5</sup>. However, users often have to choose between high throughput or high time resolution<sup>6-8</sup>, many experiments require surface immobilization techniques to extend the observation time and it is difficult to access non-equilibrium conditions or follow reactions. Many creative solutions have been proposed to overcome these hurdles using fluidic platforms<sup>9,10</sup>. These include mixers for studying single-molecule kinetics<sup>11,12</sup>, titration devices<sup>13</sup>, devices to confine molecules in polymer nanochannels<sup>14</sup>, electroosmotic molecular traps<sup>15</sup>, or microfluidic droplets containing individual enzyme analytes<sup>16</sup>. The detection throughput can be increased by controlling the flow through the detection area; and, in the case of mixers, reactions can be triggered in front of the detection area. These fluidic platforms, despite their clear benefits, are however not yet widely used due to complex fabrication procedures, limited reusability and configurability or restrictions in integration into microscopy platforms. Here, we introduce a novel fluidic platform for SMFD consisting of nano-/microchannel devices fabricated entirely in glass. Our platform

uniquely combines several advantages: (1) Prolonged observation times of single analytes are achieved by geometrically confining the flow through nanochannels to a volume smaller than the excitation/detection focus along the long optical axis of a microscope. (2) High-throughput detection is possible by monitoring an array of nanochannels as well as by controlling the passage time through the nanochannels via the applied flow. (3) The devices are well-defined, robust and reusable. (4) The nanochannels can be functionalized to minimize, for example, non-specific adsorption of analytes. (5) The fluidic chips are compatible with a commercially available holder allowing a straightforward integration into a standard microscope stage as well as simple and reliable interfacing to tubing. (6) The flow rate can be controlled by cost-efficient syringe pumps. (7) Special nanochannel geometries can be used to continually mix and observe reactions at the single-molecule level. We first demonstrate high-throughput sensing and tracking of short DNA oligonucleotides in parallel nanochannels utilizing single-molecule Förster resonance energy transfer (smFRET) between a donor and an acceptor fluorophore attached to the DNA. As our method allows for continuous detection, several hundred thousand events can be combined to gather reliable single-molecule data as well as obtaining sub-micrometer resolved velocity profiles within the nanochannels. We then demonstrate the mixing geometry (T-junction) by monitoring the conformational equilibrium of a DNA hairpin before, during and after mixing with a buffer containing a high salt concentration that stabilizes the closed hairpin conformation. Finally, we validate the potential of the device for monitoring enzymatic reaction kinetics by observing DNA synthesis at the single-molecule level.

## **Materials and Methods**

### **Fluidic device fabrication**

Nano-/microfluidic devices were fabricated with a 45 mm × 15 mm footprint to be compatible with a commercially available chipholder Fluidic Connect PRO (Micronit Microtechnologies B.V., The Netherlands). Micronit also fabricated the devices. In brief, they consist of two thermally bonded borosilicate glass layers with photolithographically defined and wet-etched nano- and microchannels (Fig. 5.1). Microchannels and ports to connect to tubing are structured in the upper 1 mm thick glass, while nanochannels are etched into the bottom layer. Here, coverslip type D263 borosilicate glass (Schott) with a thickness of 175 μm was used enabling high-resolution fluorescence imaging with inverted oil-immersion microscope objectives with shortest working distances whilst minimizing undesirable autofluorescence.



**Figure 5.1.** Design of nanofluidic chips. *(a and c)* High-resolution confocal scans based on reflection of light. *(a)* The parallel channel design contains 21 straight nanochannels ( $l \times w \times h$ : 120 by 4.1 by 0.2 μm) and a microchannel ( $w \times h$ : 21 by 5 μm). *(b)* Schematic cross-section of the parallel channel array showing the dimensions (not to scale) of the microchannel etched into the top glass wafer and the nanochannels etched into the bottom glass wafer. *(c)* The design of the mixing device contains a single T-shaped nanochannel (horizontal part:  $l \times w \times h$ : 40 by 3.8 by 0.2 μm, vertical part  $l \times w \times h$ : 100 by 4.7 by 0.2 μm) in between two microchannels. *(d)* Picture of a chip with 4 mixing channels, one of which is centered on our TIRF objective. Chips fit in a Micronit chip holder for easy connection to tubes and pumps (see also Fig. S5.2).

### Fluidic flow control

Flow was driven by Pump 11 Pico Plus Elite (Harvard Apparatus, USA) syringe pumps. For high nanochannel flow rates (50 pL/min) the pump generates a pressure of about 200 kPa. Due to the large ratio of flows in PFC, the dead volume in the 1 mm wide and up to 2 cm long feeding microchannels is replaced almost instantaneously. In PFC, low flow rates are achieved by dividing the syringe flow rate  $Q$  into the microchannel and array of nanochannels according to the channels' hydraulic resistance, which is calculated as<sup>17,18</sup>



$$R_{\text{hyd}} \approx \frac{12\eta L}{1 - 0.630 \left(\frac{h}{w}\right)} \cdot \frac{1}{h^3 w} \text{ for } h < w. \quad (\text{eq. 5.1})$$

Here,  $\eta$  is the dynamic viscosity of 0.001 Pa s in water and  $h$ ,  $w$  and  $L$  are the height, width and length of a nanochannel, respectively. (A pressure  $\Delta p$  across a channel drops according to the Hagen Poiseuille Law  $Q = \Delta p / R_{\text{hyd}}$ , and for parallel resistances  $R_{\text{array}} = 1 / \sum R_{\text{hyd}, i}^{-1}$  applies.) Due to the cubic dependence  $h^3$ , the flow rate in each nanochannel is reduced by a factor of 60.000 compared to the microchannel with  $L = 140 \mu\text{m}$ ,  $w = 10 \mu\text{m}$ ,  $h = 5 \mu\text{m}$  compared to the applied syringe flow. Given the dimensions of the nanochannels, the flow is dominated by viscosity while inertia is negligible (Reynolds number  $< 10^{-4}$ , see Supporting Information Note 1); under these conditions laminar flow dictates the motion of the fluid and the equation for Poiseuille flow can be used to model the velocity field inside a nanochannel. In our design, the channel width is one order of magnitude larger than the height; as a consequence, the system resembles flow between two infinite parallel plates and the velocity is constant throughout most of the width (Supporting Information Note 5.2 and Fig. S5.5).

For the mixing device (Fig. 5.1c,d and Fig. S5.1b), two syringe pumps deliver flow to both feeding nanochannel inlets (each  $20 \mu\text{m}$  long,  $3.8 \mu\text{m}$  wide and  $200 \text{ nm}$  high). A single nanochannel ( $100 \mu\text{m}$  long,  $4.7 \mu\text{m}$  wide and  $200 \text{ nm}$  high) is positioned downstream of the junction. Two bypassing microchannels lead to an overall reduction factor of 40.000 of the nanochannel flow compared to the combined syringe flow. Each feeding nanochannels has a hydraulic resistance of 25% of the long nanochannel, thus, a backflow into another feeding channel is prevented for differences in the syringe pump rate of up to 25%. When mixing fluidics of different viscosities (e.g., 1 M aqueous NaCl has a 10% higher viscosity than water<sup>19</sup>), pump rates can be conveniently adjusted to compensate for the different effective hydraulic resistances and to ensure a 1:1 mixing ratio.

## DNA

Fluorescently labelled oligonucleotides were ordered from IBA, Germany. To construct the DNA hairpin, a 30-mer primer sequence (biotin-5'-CCT CAT TCT TCG TCC CAT TAC CAT ACA TCC-3') was annealed to a 75-mer hairpin sequence (5'-TGG ATT AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA TCC ATT GGA TGT ATG GTA ATG GGA CGA AGA ATG AGG-3'). The primer was internally labelled with ATTO647N at the -12 position; the hairpin was labelled with Cy3B at the 5' end. Gapped DNA, used to study KF polymerase binding, was constructed using the same primer sequence, annealed to a template strand (5'-CCA CGA AGC AGG CTC TAC TCT CTA AGG ATG TAT GGT AAT GGG ACG AAG AAT GAG G-3') and a downstream complementary strand (5'-TAG AGA GTA GAG CCT GCT TCG TGG-

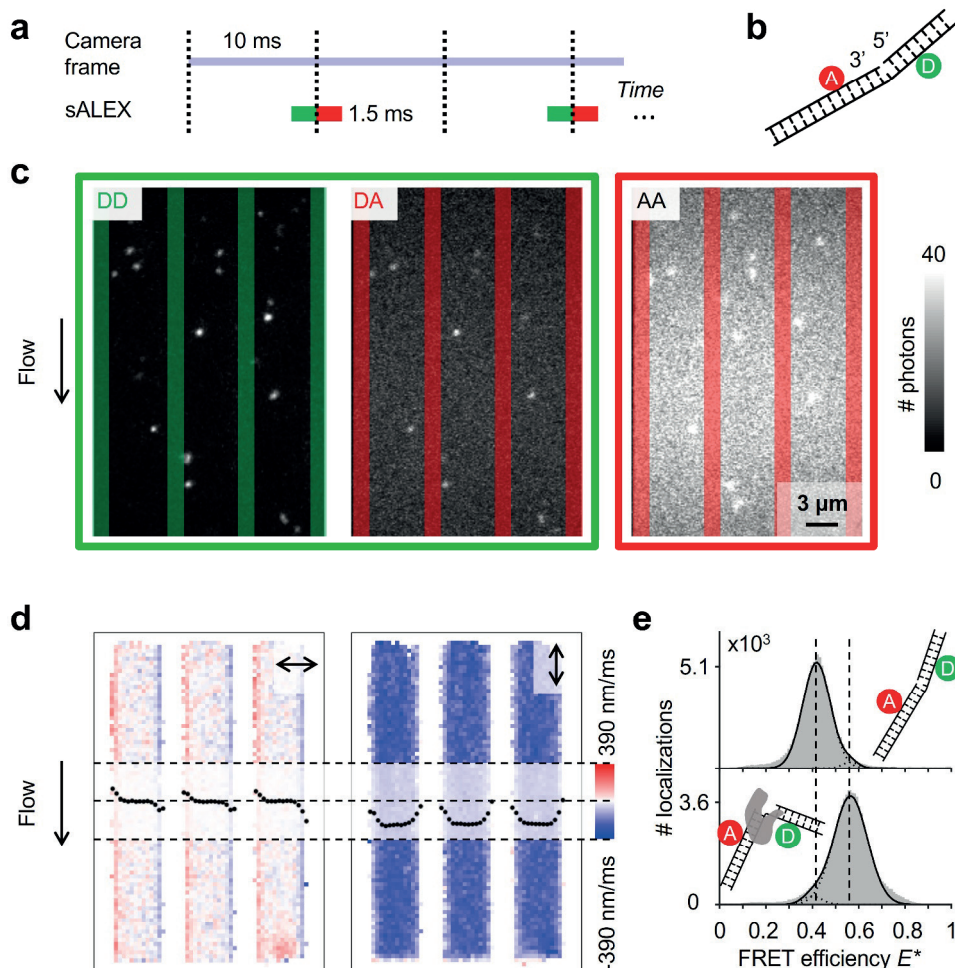
3'). The template strand was labelled with Cy3B at the +12 position. The DNA sensor used for the polymerization experiment consists of the same primer and template sequences, but with ATTO647N on the -7 position and Cy3B on the +25 position, respectively. dNTPs were ordered from Sigma-Aldrich / Merck, Germany.

### **Buffers**

DNA constructs (as well as DNA polymerases, if specified) were diluted in an imaging buffer containing 50 mM Tris HCl (pH7.5), 100 µg/mL BSA, 10 mM MgCl<sub>2</sub>, 5% glycerol, 1 mM DTT, 1 mM Trolox, 1% glucose oxidase/catalase and 1% glucose. Trolox is a triplet-state quencher and prevents fluorophore blinking. Glucose, glucose oxidase and catalase was used as an oxygen scavenger system to prevent premature photobleaching of fluorophores<sup>20,21</sup>. The concentration of gapped DNA was 1 nM and, if used, the concentration of KF was 10 nM. DNA hairpin concentrations were 500 pM in parallel channels and 1 nM in the mixing channel. (DNA hairpins were diluted in a similar imaging buffer without magnesium, but with additional NaCl as specified.) Prior to mixing, the concentration of DNA polymerization sensors was 1 nM; the concentration of KF was 5 nM, and the concentration of dNTPs was 200 µM each. For this polymerization experiment, we added neutravidin directly to the imaging buffer in a concentration of 0.6 µg/mL to block 5' end of the biotinylated DNA primer. We found this prevents the formation of a low  $E^*$  state, the cause of which is probably binding of multiple KF polymerases to the same DNA molecule. The imaging buffer was applied to a 100 µL syringe (ILS, Germany), which was then connected to the nanofluidic device using ethylene tetrafluoroethylene (ETFE) tubing (1/16" outer diameter, 0.010" inner diameter) (IDEX, USA).

### **Surface passivation**

To prevent non-specific adsorption in experiments with proteins, channels were passivated with PEG using a variation of a method described previously<sup>22</sup>. Burning the glass to remove organic contaminations was only performed for cover slips. The fluidic devices were first flushed and incubated (3 × 5 min) with a 1:50 (vol/vol) Vectabond:acetone solution. All subsequent washing and passivation steps were performed by flushing the channels with ~100 µL of the respective solutions. PEGylated channels were filled with PBS and stored in a humid chamber at 4°C.



**Figure 5.2.** Single-molecule detection and flow profiles inside (parallel) nanochannels. **(a)** Schematic of stroboscopic alternating laser excitation (sALEX) in which the excitation time is considerably shorter than the required acquisition time of a camera frame. **(b)** Schematic of a gapped DNA construct labelled with a donor and acceptor dye located on opposite sides of the one-nucleotide gap. **(c)** Typical (non-averaged) movie frames showing our gapped DNA construct ( $E^* \sim 0.4$ ) flowing through the parallel nanochannels at a pump rate of 40  $\mu\text{L/h}$  (Supporting Information Movie 2, watch online). Excitation colors are indicated by the surrounding boxes. For each molecule, photon counts of DD and DA after donor excitation are determined simultaneously. The photon counts of AA after acceptor excitation are collected during the next camera frame. Nanochannel boundaries are indicated with green bars (green detection channel) or red bars (red detection channel). Due to sALEX, motion blur is effectively suppressed and the back-to-back green and red excitation facilitates easy linking of AA to its corresponding signals DD and DA. **(d)** Flow velocimetry calculated from  $\sim 2 \times 10^5$  DA and

AA localization pairs perpendicular (left) and parallel (right) to the flow direction within the channels (binned  $3 \times 3$  pixels, see also Material and Methods and Figs. S5.3 and S5.4). The resulting velocity profiles are superimposed on each map. (e)  $E^*$  histograms (100 bins) of our gapped DNA construct, measured in the parallel channels for 5 minutes. Top: 1 nM DNA. Bottom: 1 nM DNA in presence of 10 nM *E.coli* DNA Polymerase I (Klenow Fragment). Dashed lines are added for visual guidance. For full  $E^*/S$  histograms see Figure S5.7.

### Single molecule detection

We used a home built TIRF microscope and a fiber-coupled laser engine (Omicron, Germany) equipped with lasers of four different wavelengths (405 nm, 473 nm, 561 nm, and 638 nm). Laser intensities were independently controlled by a home-written LabVIEW program. Divergent light from the fiber output is collimated ( $f = 30$  mm, Thorlabs, Germany) and focused by a second lens ( $f = 200$  mm, Thorlabs, Germany) into the backfocal plane of a 100x NA 1.49 objective (Nikon, Japan). A polychroic filter and a multi-bandpass filter (Chroma, USA) prevented laser light from entering the emission path. A tube lens focuses the emission on an aperture, which reduces the effective field of view to a rectangle. Next, the light is spectrally split into a blue, a green and a red emission channel. The three beams are focused on an Ixon Ultra 897 emCCD (Andor, UK) with 512 by 512 pixels (maximum acquisition rate: 56 Hz at full frame and 100 Hz after cropping the frame to 343 by 256 pixels). In our configuration 1 pixel on the camera corresponds to a length and width of 112 nm in the sample plane. We use a Rapid Automated Modular Microscope (RAMM) system as a stage holder (ASI, USA), combined with motorized  $x,y$ -scanning stage and a  $z$ -piezo for focusing.

Molecules were excited with 130 mW (561 nm and 638 nm lasers) measured after the fiber output within the fluidic devices and with 30 mW (561 nm) and 15 mW (638 nm) for the surface immobilized experiments. A stroboscopic alternating-laser excitation (sALEX)<sup>23</sup> scheme was used to reduce motion blur of diffusing molecules. Laser pulse widths were 1.5 ms (fluidic devices) and 3 ms (surface immobilized experiments) in a frame time of 10 ms. Green and red pulses were aligned back-to-back (Fig. 5.2a), so that particle movement between a green and a red frame is minimal. Shorter laser pulses, and the necessary higher laser powers, were found to cause rapid bleaching.

### Extracting emission intensities from movie frames

Particles were localized and tracked with a home-written variation on GaussStorm (Matlab)<sup>24</sup>. Time traces, histograms and binned maps were generated with custom-built software packages. The sALEX scheme effectively minimized motion blur and created mostly circular or elliptical intensity spots on the camera sensor. We applied a bandpass filter to find local intensity maxima before fitting the local maxima with elliptical 2D Gaussian functions from which we obtained the photon count as well as the position with sub-pixel accuracy<sup>24</sup>. After filtering of localizations based on their intensities, our algorithm links corresponding particles in the green and red channels

and calculates FRET efficiency  $E^*$  and stoichiometry ratio  $S$  for each molecule. A tracking algorithm is applied to track particles from frame to frame<sup>25</sup>.

### **FRET, alternating laser excitation (ALEX) and accurate FRET**

The apparent FRET efficiency  $E^*$  is calculated from the emission intensities of donor and acceptor after donor excitation (denoted  $DD$  and  $DA$ , respectively) for each molecule in each camera frame according to

$$E^* = DA / (DD + DA) \quad (\text{eq. 5.2})$$

To verify the presence of an acceptor fluorophore on the DNA and to obtain the additional information required for correcting  $E^*$  for spectral cross-talk and detection efficiencies, we applied alternating laser excitation (ALEX) in which every frame of donor excitation is followed by a direct excitation of the acceptor fluorophore using a second laser resulting in a third photon stream ( $AA$ ) for each molecule<sup>26,27</sup>. The detection of  $AA$  in addition to  $DD$  and  $DA$  allows for calculating the stoichiometry ratio  $S$ , defined as

$$S^{\text{raw}} = \frac{(DD + DA)}{(DD + DA + AA)} \quad (\text{eq. 5.3})$$

$S$  can be used to filter molecules: molecules with a stoichiometry close to 0 have no photoactive donor (e.g., because of donor bleaching), and molecules with a stoichiometry close to 1 have no photoactive acceptor. Depending on the ratio of the laser intensities for direct donor and acceptor excitation, a stoichiometry around 0.5 represents molecules having both a photoactive donor and acceptor that can undergo FRET. In this work, we show FRET data based on molecules that have stoichiometry values  $0.3 \leq S \leq 0.8$  (nanochannel data) or  $0.5 \leq S \leq 0.9$  (immobilized hairpins).

ALEX also allows for easy correction of the apparent FRET efficiency  $E^*$  to obtain an accurate FRET efficiency  $E^{\text{27}}$ . Corrections need to be made for 1) leakage of donor signal into the acceptor channel, 2) direct excitation of the acceptor when exciting the donor, and 3) quantum yields of the dyes and detection efficiencies of different colors on the camera sensor (summarized in the gamma factor). Leakage and direct excitation contributions can be deduced from the donor-only and acceptor-only peaks, respectively. Gamma will cause different FRET efficiencies to have slightly different stoichiometries; the contribution of gamma can therefore be deduced from the slope when plotting  $1/S$  against  $E$ .

### **Obtaining flow velocity profiles in parallel nanochannels**

The displacements were calculated through localization and tracking of single molecules between consecutive frames. The travelled distance was calculated from sub-pixel localizations. The time between displacements was the time between the

start of the green and the red excitation, respectively, i.e., 1.5 ms. Each calculated velocity was assigned to the starting position of the particle on the map. Then, the field of view was divided into regularly spaced bins of  $3 \times 3$  pixels, so a bin will contain all the velocities of the particles that started in it. To avoid visualizing bins with only a few data points, a threshold was applied to exclude bins with  $<10$  displacements from further analysis (Fig. S5.3). Excluded bins are located on the edges of the nanochannels.

## Results and Discussion

### Achieving parallel flow control

The device for high-throughput detection consists of an array of parallel nanochannels wet-etched into a first glass wafer (Fig. 5.1a,b, see also Material and Methods). Our array of nanochannels is connected in parallel via wafer-bonding with a bypassing larger microchannel etched into a second glass wafer such that the velocity of analytes in the nanochannel can be precisely tuned by the principle of parallel flow control (PFC) using a simple syringe pump<sup>28,29</sup>. In PFC, nanochannel flow rates as low as 0.5–50 pL/min (corresponding to flow velocities of 10–1000 nm/ms) are achieved by dividing the flow into the microchannel and the array of nanochannels (see Fig. S5.1 for equivalent circuit diagrams) such that the flow rate in each nanochannel is reduced by a factor of 60.000 compared to the microchannel (Material and Methods).

Mixing devices have been primarily demonstrated for confocal applications, with latest designs even enabling two consecutive mixing steps<sup>11,30</sup>. Here, for the first time, we use the PFC principle for a mixing geometry: Two syringe pumps deliver flow to two microchannels and two respective parallel feeding nanochannel inlets which merge to a single nanochannel at a T-junction (Fig. 5.1c,d and Supporting Information Movie 1, watch online). PFC allows stable mixing conditions: the flow rates of both syringe pumps can differ by up to 25% before a parasitic backflow into one feeding channel occurs (Material and Methods).

### Detecting fluorescence emission of moving particles inside nanochannels

To characterize our devices, we monitored fluorescently labelled DNA molecules and their respective distance-dependent smFRET signatures. The apparent FRET efficiency  $E^*$  is calculated for each molecule in each camera frame using photon streams after donor excitation emitted by the donor and the acceptor (*DD* and *DA* respectively, see Material and Methods). To verify the presence of the acceptor, we applied alternating-laser excitation (ALEX) in which every frame of donor excitation is followed by a direct excitation of the acceptor fluorophores creating photon streams of *AA* (Material and Methods)<sup>26,27</sup>. Since motion blur severely affected our ability to determine emission intensities of individual molecules inside the nanochannels at frame times as short as 10 ms, we opted for stroboscopic ALEX (sALEX)<sup>23</sup> in which



the molecules are only excited for a fraction of the camera exposure time (Fig. 5.2a). We further aligned the green and red laser pulses back-to-back to facilitate the linking of a particle's AA signal and position to its corresponding DD and DA signals from the previous frame.

### Obtaining velocity profiles in parallel nanochannels

We first measured 1 nM doubly labelled, gapped DNA (Fig. 5.2b) with a FRET efficiency of  $E^* \sim 0.4$  for 10 minutes at pump rates of 40  $\mu\text{L}/\text{h}$ . Upon donor excitation, both donor and acceptor intensities from individual DNA molecules are visible (Fig. 5.2c, left). Short excitation pulses (1.5 ms) lead to spots which are barely affected by motion blur and have an excellent signal-to-noise ratio allowing us to fit 2D Gaussians to obtain intensity numbers and sub-pixel localization accuracies for individual DNA molecules. The following direct excitation of the acceptor, provided information on the presence of the acceptor (Fig. 5.2c, right) and allowed us to calculate the displacements between the position of the molecule in the acceptor detection channel during donor excitation and during direct acceptor excitation.

Using  $2 \times 10^5$  displacements, we determined the particle velocities in vertical and horizontal direction and mapped the mean velocities back to the field of view; then, we constructed velocity profiles by averaging the mean velocities along the length of the channels (Fig. 5.2d). The velocity in the horizontal direction, perpendicular to the nanochannels' axis, is expected to be zero as no flow is applied along this direction and diffusion is random; this is true in the center of the channels. For positions close to the walls, however, spatial restrictions to diffuse further left or right are translated into a net movement from the walls toward the center. The vertical velocity plateaus at -240 nm/ms in the center of each nanochannel are in excellent agreement with the theoretical model (Figs. S5.4 and S5.5). Along the nanochannels, the flow velocity profiles at different pump rates show a linear relationship between pump rate and flow speed (Fig. S5.4) consistent with laminar Poiseuille flow (Fig. S5.5).

We compared the magnitude of molecular movement due to advective flow with movement due to diffusion. To obtain typical values for the mean flow speed and diffusivity, we fitted all vertical displacements with a Cumulative Distribution Function (CDF)<sup>31</sup> accounting for both flow and diffusivity, and applied it to a series of different pump rates (Fig. S5.6). We found an average diffusivity of  $D = 33 \pm 1 \mu\text{m}^2/\text{s}$ . The mean square displacement<sup>32</sup> due to diffusion can be calculated using  $\text{MSD} = 2D\Delta t$ , in which  $\Delta t$  is the time between the beginning of green and red excitation, respectively (1.5 ms, due to back-to-back illumination). Using the diffusion coefficient  $D$ , we found a MSD of  $0.1 \mu\text{m}^2$ , giving a standard deviation of  $\sim 300$  nm. CDF analysis also returned an average flow speed of 220 nm/ms, resulting in a flow-induced displacement of around 330 nm in 1.5 ms. These two displacements are in the same order of magnitude



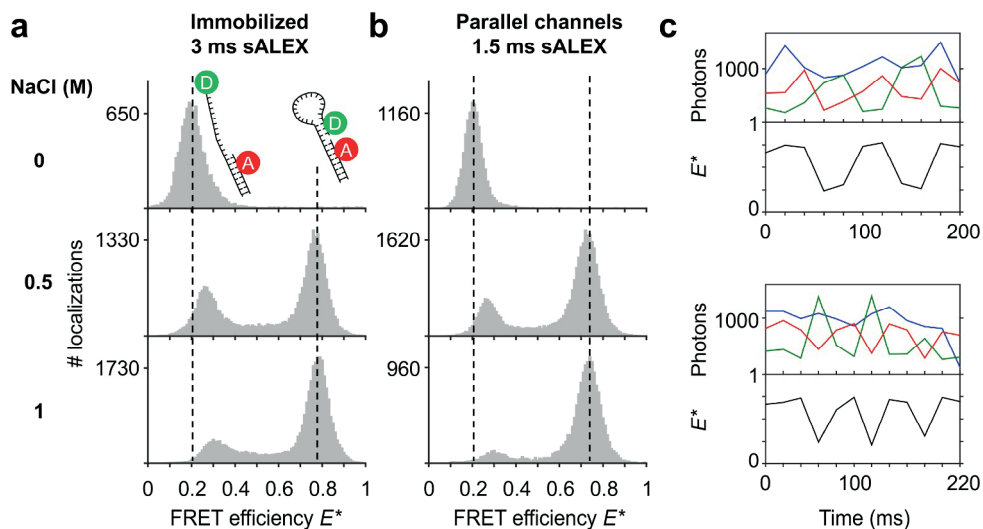
highlighting the fact that diffusion plays a major role at the time and length scales accessible in our experiments.

From our data, we compiled a FRET histogram with stoichiometry values  $0.3 \leq S \leq 0.8$  to select for doubly labelled DNA molecules (Fig. 5.2e). A single species is visible with a mean  $E^*$  of 0.4. To prove that our parallel device, after passivating the glass surface using PEG (see Material and Methods), is compatible with continuous enzyme detection, we added 10 nM of DNA polymerase I (Klenow fragment, KF) to the gapped DNA. We observed an increase in FRET efficiency from 0.4 to around 0.6, reflecting the shortened donor-acceptor distance upon KF binding and DNA bending (Fig. 5.2e).

As mentioned above, we obtained  $\sim 2 \times 10^5$  FRET data points within 10 minutes of measurement, with the current field of view on our setup ( $\sim 29 \mu\text{m}$  by  $19 \mu\text{m}$ ) allowing to monitor 3 out of 21 channels. For comparison, in diffusion-based confocal smFRET experiments, fluorescence bursts of single molecules passing the focus are collected with a similar time resolution ( $\sim 1\text{-}3$  ms), but the time between individual bursts must be kept long enough to avoid doubly occupancy of the focal volume<sup>7</sup>. Assuming a time of 500 ms to obtain a single FRET data point, 10 minutes yield only around 1000 FRET data points; considerably less than in our nanochannels.

### **Parallel nanochannels: resolving conformational dynamics of DNA hairpins at different salt concentrations**

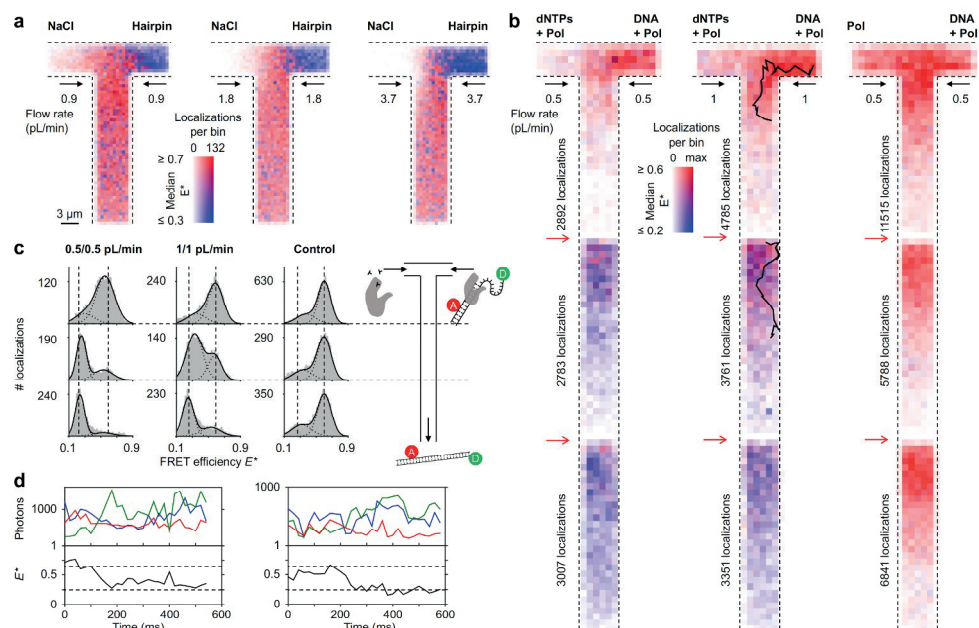
To probe the accessible dynamic range of our FRET measurements, we used DNA hairpins that can interconvert between an open (low FRET) and closed (high FRET) conformation. As reported before, the used DNA hairpin is mostly open at 0 M NaCl and mostly closed at 1 M NaCl, while showing fast opening and closing at an intermediate concentration of 0.5 M NaCl<sup>23</sup>. Our FRET histograms show that freely flowing DNA hairpins behave the same as immobilized ones, showing that we can access a broad FRET range (Fig. 5.3a,b). Here, the stroboscopic excitation scheme allows resolving the individual conformational states of “open” and “closed” despite the corresponding conformational dynamics being considerably faster than the frame rate of the camera. Individual time traces from freely flowing molecules show opening and closing of individual DNA hairpins in real time (Fig. 5.3c).



**Figure 5.3.** Single-molecule detection of DNA molecules immobilized (a) or freely flowing within parallel nanochannels (b). We analyzed the conformational equilibria of DNA hairpins as a function of the NaCl concentration and plotted FRET histograms (100 bins) derived from individual particle localizations. (a) DNA hairpins immobilized on a PEGylated coverslip are mostly open (low FRET) at  $[\text{NaCl}] = 0 \text{ M}$ ; the conformational equilibrium shifts towards the closed state (high FRET) with increasing salt concentrations. (b) The same DNA hairpin shows similar behavior when imaged in parallel nanochannels. When corrected for leakage, direct excitation and gamma, peak positions are the same as for the immobilized sample (Material and Methods and Supporting Information Table 5.1). For full  $E^*/S$  histograms see Figure S5.7. (c) Single-molecule time traces obtained from tracking individual molecules flowing through the channels at  $[\text{NaCl}] = 0.5 \text{ M}$ . The anti-correlation of the DD (green trace) and the DA (red trace) signal and the resulting FRET time traces (black trace) indicate conformational changes in the millisecond timescale of the DNA hairpin during its passage. The AA signal (blue trace) remains constant.

### Nanofluidic mixing: triggering DNA hairpin closing

We used the mixing device to shift the conformational equilibrium from mostly open to mostly closed. To this aim, we mixed a buffer containing 1 M NaCl from the left with a buffer containing 0 M NaCl and DNA hairpins from the right. We constructed a map of binned hairpin localizations and corresponding FRET efficiencies (Fig. 5.4a). For each bin, we plotted the median FRET efficiency albeit the full spatially resolved FRET information remained preserved. For each of the three flow rate combinations depicted, a sudden shift from low to high  $E^*$  is observed around the junction. These positions mark the onset of mixing. As flow rates are increased, fewer hairpins diffuse far into the salt channel and the onset of mixing shifts slightly more towards the junction. Apparently, particle displacement due to diffusion, as opposed to particle



**Figure 5.4.** Mixing studies. (a) Mixing DNA hairpin (from the right,  $[\text{NaCl}] = 0 \text{ M}$ ) with a high salt solution ( $[\text{NaCl}] = 1 \text{ M}$ ) at different flow rates (Supporting Information Movie 3, watch online). The figure shows binned maps ( $4 \times 4$  camera pixels per bin), in which color represents the median FRET efficiency and opacity indicates the number of localizations in a bin over the time span of the measurement. (b–d) DNA polymerization (Supporting Information Movie 4, watch online). (b) Binned maps ( $8 \times 8$  camera pixels per bin) of mixing (DNA sensors + KF) from the right with (KF + dNTPs) from the left at two different pump rates. A control (outer right) did not contain dNTPs. Maps contain data from adjacent fields of view (sections); the borders of these sections are indicated with red arrows. Color represents the median FRET efficiency; opacity indicates the number of localizations in a bin, normalized for each map separately (maximum 58 for 0.5/0.5 pL/min, maximum 66 for 1/1 pL/min and maximum 126 for the control). Fewer molecules are localized in the lower parts of the channel sections because of fluorophore bleaching. (c)  $E^*$  histograms for each section of the mixing channel, as indicated with the schematic on the right. (d) Time traces of two molecules from the 1/1 pL/min condition. Upper panels: fluorescence intensities of DD (green), DA (red) and AA (blue). Lower panel: FRET efficiency  $E^*$  (black). Both traces show a decrease in FRET efficiency over time, suggesting polymerization of the DNA sensor.

velocity due to applied flow, plays a large role in mixing. In fact, in the nanochannel junction, mixing is purely diffusion-controlled and the mixing area becomes more confined as flow speeds increase.

This high diffusivity is advantageous for applications for which rapid mixing is required<sup>33,34</sup> as turbulence does not occur in nanochannels due to very low Reynolds

numbers. By combining the reduction of flow rates using PFC with the observation of single molecules, we established a precise control over diffusive mixing such that reagents can be detected before, during and after mixing, i.e. without any dead time.

### **Nanofluidic mixing: triggering enzymatic catalysis of DNA synthesis**

To establish our mixing device as a powerful tool for non-equilibrium studies, we utilized a FRETbased DNA polymerization sensor containing a primer DNA, labelled with an acceptor fluorophore, annealed to a long template DNA in which the single stranded overhang (25 bases) is labelled with a donor fluorophore (see Material and Methods)<sup>35</sup>. Since the single strand is coiled, the FRET efficiency is high ( $E^* \sim 0.6$ ) before decreasing substantially upon addition of DNA polymerases and nucleotides (dNTPs: deoxyribonucleotide triphosphates) to  $\sim 0.2$ . Using the same construct, we previously found that full polymerization takes on average 1.6 s when immobilized on a surface<sup>35</sup>. In the nanofluidic device, we triggered the polymerization reaction by mixing (1 nM DNA + 5 nM KF) with (200  $\mu$ M dNTPs + 5 nM KF) in a 1:1 mixing ratio, using low flow rates of either 0.5 pL/min or 1 pL/min (Fig. 5.4b). Similar to the DNA hairpin data at low flow rates (Fig. 5.4a), some molecules transiently diffuse into the other nanochannel. Due to the long reaction time, we acquired movies using three different fields of view and stacked them together to follow the reaction profiles until completion (Fig. 5.4b,c). The maps show a gradual decrease from high FRET efficiency (unpolymerized DNA sensor) near the junction to low FRET efficiency (polymerized DNA sensor) towards the outlet indicating that polymerization is completed within the time the molecules spend in the nanochannel:  $\sim 3$  s at 1 pL/min and  $\sim 5$  s for 0.5 pL/min nanochannel flow rates. This is on the same time scale as DNA polymerization of surfaceimmobilized DNA sensors. We note that at these low flow speeds, the fluorophores attached to the DNA molecules tend to bleach while being in the field of view due to the high laser intensities required to image molecules and the presence of proteins (KF). Two time traces indicating DNA synthesis in real time are shown (Fig. 5.4d).

In our current design, the distance from junction to outlet is 100  $\mu$ m, corresponding to a residence time in the nanochannel of around 5 s using a flow rate of 0.5 pL/min and short DNA oligonucleotides. To gain access to further time points after mixing, designs using meandering channels could be implemented as demonstrated for confocal microscopy<sup>11,30</sup> or widefield microscopy<sup>36</sup>. Furthermore, our current field of view is cropped by a factor of two to ensure data acquisition at 100 Hz. With the use of faster cameras (sCMOS) and by reducing the overall magnification of the optical system (e.g., by replacing the 100x TIRF objective with a 60x objective), even more molecules could be simultaneously observed for a longer time over a larger area and with greater time resolution<sup>37</sup>.

## Conclusion

We pushed the boundaries of single-molecule detection under equilibrium and non-equilibrium conditions by introducing nanofluidic device designs that can be easily integrated in existing microscope setups: the first one with a series of parallel nanochannels for equilibrium studies, and the second with a single T-shaped nanochannel for non-equilibrium studies. For both cases, the core idea is the combination of camera-based single-molecule detection, which allows monitoring many fluorescent molecules in parallel, with constant replenishment of freely flowing molecules. Extended observation times are achieved by geometrical confinement of molecules, imposed by a nanochannel height of 200 nm. A stroboscopic excitation scheme resulted in near-circular emission spots on the camera sensor. Furthermore, in both designs we employed the concept of parallel flow control allowing us to work with simple syringe pumps whilst keeping the consumption of reagents low.

Our design using a parallel array of nanochannels is well suited for high-throughput measurements over extended periods of time. Moreover, longer tracks of single molecules flowing through the field of view can be observed by decreasing the sample concentration or decreasing the flow rate and increasing the viscosity of the buffer medium.

Using our nanochannel mixing devices, we accessed non-equilibrium conditions by mixing primarily open DNA hairpins from one inlet with a high-salt solution from a second inlet triggering hairpin closing, and showed that mixing is diffusive. Additionally, we observed polymerization of 25 bases on a DNA template by a DNA polymerase, illustrating that complex biological reactions can be followed in real time and in a continuous fashion.

We further note that the mixing device is conceptually equivalent to conventional stopped-flow instruments of the sort frequently used in the (bio-)chemical sciences to measure reaction kinetics of biomolecules, but with the unique benefit of allowing continuous observation of reactions. We therefore believe that the introduced nanofluidic designs represent a powerful platform for studying non-immobilized single molecules with high throughput and time resolution, under both equilibrium and non-equilibrium conditions.

## Acknowledgments

We thank Timo Wenzel and Ebru Acun for initial characterizations of the fluidic devices. We thank Adrie Westphal for help with the confocal scans. J.H. acknowledges support from a Marie Curie Career Integration Grant (#630992). S.G.L. acknowledges financial support from The Netherlands Organization for Scientific Research (NWO)

and the European Research Council (ERC) under Project 278801. M.F. acknowledges financial support from the Graduate School Experimental Plant Sciences (EPS), Wageningen University, under Project EPS3 3b 092. C.F. and M.F. contributed equally to this work.

## References

1. Smiley, R. D. & Hammes, G. G. Single Molecule Studies of Enzyme Mechanisms. *Chem. Rev.* **106**, 3080–3094 (2006).
2. Claessen, V. I. *et al.* Single-Biomolecule Kinetics: The Art of Studying a Single Enzyme. *Annu. Rev. Anal. Chem.* **3**, 319–340 (2010).
3. K  chler, A., Yoshimoto, M., Luginb  hl, S., Mavelli, F. & Walde, P. Enzymatic reactions in confined environments. *Nat. Nanotechnol.* **11**, 409 (2016).
4. Walter, N. G., Huang, C.-Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: How to use the modern single molecule toolkit. *Nat. Methods* **5**, 475–489 (2008).
5. Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in Single-Molecule Fluorescence Methods for Molecular Biology. *Annu. Rev. Biochem.* **77**, 51–76 (2008).
6. Hohlbein, J., Gryte, K., Heilemann, M. & Kapanidis, A. N. Surfing on a new wave of single-molecule fluorescence methods. *Phys. Biol.* **7**, 031001 (2010).
7. Santoso, Y. *et al.* Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 715–720 (2010).
8. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nat. Commun.* **4**, (2013).
9. Mathwig, K., Chi, Q., Lemay, S. G. & Rassaei, L. Handling and Sensing of Single Enzyme Molecules: From Fluorescence Detection towards Nanoscale Electrical Measurements. *ChemPhysChem* **17**, 452–457 (2016).
10. Banterle, N. & Lemke, E. A. Nanoscale devices for linkerless long-term single-molecule observation. *Curr. Opin. Biotechnol.* **39**, 105–112 (2016).
11. Wunderlich, B. *et al.* Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat. Protoc.* **8**, 1459–1474 (2013).
12. Gambin, Y. *et al.* Visualizing a one-way protein encounter complex by ultrafast single-molecule mixing. *Nat. Methods* **8**, 239–241 (2011).
13. Kim, S. *et al.* High-throughput single-molecule optofluidic analysis. *Nat. Methods* **8**, 242–245 (2011).
14. Tyagi, S. *et al.* Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nat. Methods* **11**, 297–300 (2014).
15. Lesoine, J. F., Venkataraman, P. A., Maloney, P. C., Dumont, M. E. & Novotny, L. Nanochannel-Based Single Molecule Recycling. *Nano Lett.* **12**, 3273–3278 (2012).



16. Arayanarakool, R., Shui, L., M. Kengen, S. W., Berg, A. van den & T. Eijkel, J. C. Single-enzyme analysis in a droplet-based micro- and nanofluidic system. *Lab. Chip* **13**, 1955–1962 (2013).
17. Bruus, H. Chapter 1 Governing Equations in Microfluidics. in *Microscale Acoustofluidics* 1–28 (Royal Society of Chemistry, 2014).
18. Mathwig, K. & Lemay, S. G. Mass transport in electrochemical nanogap sensors. *Electrochimica Acta* **112**, 943–949 (2013).
19. Lide, D. R. *CRC Handbook of Chemistry and Physics, 86th Edition*. (CRC Press, 2005).
20. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).
21. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *J. Am. Chem. Soc.* **131**, 5018–5019 (2009).
22. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res.* **43**, 5998–6008 (2015).
23. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17**, 27862–27872 (2015).
24. Holden, S. J. *et al.* Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophys. J.* **99**, 3102–3111 (2010).
25. Crocker, J. C. & Grier, D. G. Methods of Digital Video Microscopy for Colloidal Studies. *J. Colloid Interface Sci.* **179**, 298–310 (1996).
26. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).
27. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
28. Liang, H., Nam, W. J. & Fonash, S. J. A Novel Parallel Flow Control (PFC) System for Syringe-Driven Nanofluidics. in 281–283 (2008).
29. Mathwig, K., Mampallil, D., Kang, S. & Lemay, S. G. Electrical Cross-Correlation Spectroscopy: Measuring Picoliter-per-Minute Flows in Nanochannels. *Phys. Rev. Lett.* **109**, 118302 (2012).
30. Dingfelder, F. *et al.* Rapid Microfluidic Double-Jump Mixing Device for Single-Molecule Spectroscopy. *J. Am. Chem. Soc.* **139**, 6062–6065 (2017).
31. van den Wildenberg, S. M. J. L., Bollen, Y. J. M. & Peterman, E. J. G. How to quantify protein diffusion in the bacterial membrane. *Biopolymers* **95**, 312–321 (2011).
32. Chandrasekhar, S. Stochastic Problems in Physics and Astronomy. *Rev. Mod. Phys.* **15**, 1–89 (1943).
33. Liu, R. H. *et al.* Passive mixing in a three-dimensional serpentine microchannel. *J. Microelectromechanical Syst.* **9**, 190–197 (2000).



34. Yakhshi-Tafti, E., Cho, H. J. & Kumar, R. Diffusive mixing through velocity profile variation in microchannels. *Exp. Fluids* **50**, 535–545 (2011).
35. Fijen, C., Silva, A. M., Hochkoepler, A. & Hohlbein, J. A single-molecule FRET sensor for monitoring DNA synthesis in real time. *Phys. Chem. Chem. Phys.* **19**, 4222–4230 (2017).
36. Freitag, C. *et al.* Visualizing the entire DNA from a chromosome in a single frame. *Biomicrofluidics* **9**, 044114 (2015).
37. Juetten, M. F. *et al.* Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat. Methods* **13**, 341–344 (2016).
38. Reynolds, O. An Experimental Investigation of the Circumstances Which Determine Whether the Motion of Water Shall Be Direct or Sinuous, and of the Law of Resistance in Parallel Channels. *Philos. Trans. R. Soc. Lond.* **174**, 935–982 (1883).

## Supporting Information

### Supporting Note 5.1: Calculation of the Reynolds number

The Reynolds number is calculated using the following formula<sup>38</sup>:

$$Re = \frac{\rho u D_h}{\eta} \quad (\text{eq. S5.1})$$

in which  $\rho$  is the density of the liquid,  $u$  is the velocity,  $D_h$  is the hydraulic diameter and  $\eta$  is the dynamic viscosity. Using  $\rho = 1000 \text{ kg/m}^3$  (water),  $u = 0.2 \times 10^{-3} \text{ m/s}$  (for a pump rate of  $40 \text{ }\mu\text{L/h}$ ),  $D_h = 380 \text{ nm}$  ( $D_h = 4A/P$  in which  $A$  is the cross-sectional area and  $P$  is the wetted perimeter / circumference) and  $\eta = 1.002 \text{ mPa s}$  (water), we find a Reynolds number of  $7.6 \times 10^{-5}$ .

### Supporting Note 5.2: Poiseuille flow in rectangular nanochannels

The velocity for Poiseuille flow in our nanochannel geometry as derived by Bruus<sup>17</sup> can be calculated using

$$v_y(x, z) = \frac{4h^2 \Delta p}{\pi^3 \eta L} \sum_{n, \text{odd}} \frac{1}{n^3} \left[ 1 - \frac{\cosh\left(n\pi \frac{x}{h}\right)}{\cosh\left(n\pi \frac{w}{2h}\right)} \right] \sin\left(n\pi \frac{z}{h}\right). \quad (\text{eq. S5.2})$$

Here,  $\Delta p$  is the pressure difference between the two ends of the channels,  $\eta$  is the dynamic viscosity of the fluid, and  $h$ ,  $w$  and  $L$  are the height, width and length of a nanochannel, respectively. The resulting flow surface (Fig. S5.5a) shows a central velocity plateau along most of the width ( $x$ ) of the nanochannel whereas along the height ( $z$ ) a parabolic profile is present. While taking a movie of molecules flowing through the nanochannels we are acquiring the projection of the position of the particles in the  $x,y$ -plane; still, particle speed will also be a function of height. By calculating the displacements on the plane, we are neglecting the position of the molecule along  $z$  and the velocity profiles obtained in this way will therefore represent the velocity averaged along  $z$ :

$$v_y(x) = \frac{1}{h} \int_0^h v_y(x, z) dz \quad (\text{eq. S5.3})$$

Using this equation, it is possible to calculate the expected velocity profile as seen in our experiments (see Fig. S5.5b).

### Supporting Note 5.3: Cumulative Distribution Function (CDF) analysis

In the case of one-dimensional diffusion, the probability distribution for the position of the particle ( $\rho$ ) is a Gaussian distribution where the variance ( $\sigma^2$ , also called mean square displacement in context of diffusion) is a function of the diffusivity  $D$  of the particle<sup>32</sup>:

$$\sigma^2 = 2Dt \quad (\text{eq. S5.4})$$

The mean of this distribution ( $\mu$ ) is zero in case of simple diffusion but increases linearly with time in case of flow:

$$\mu = vt \quad (\text{eq. S5.5})$$

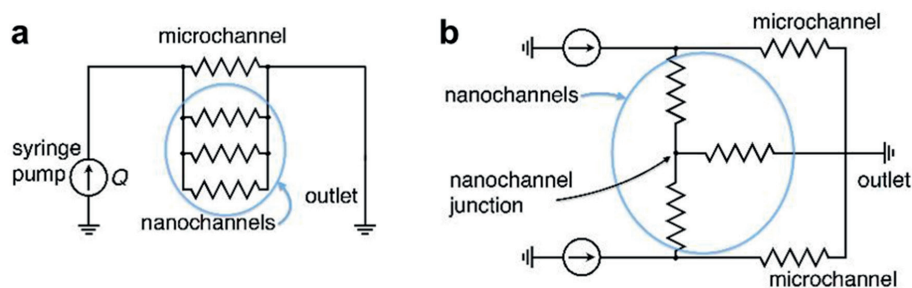
in which  $v$  is the flow speed. Particle displacements along the nanochannel axis ( $y$ ) are influenced by diffusion and flow; hence, the probability distribution for the position of a particle at time  $t$  is

$$\rho(y, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(y-vt)^2}{4Dt}} \quad (\text{eq. S5.6})$$

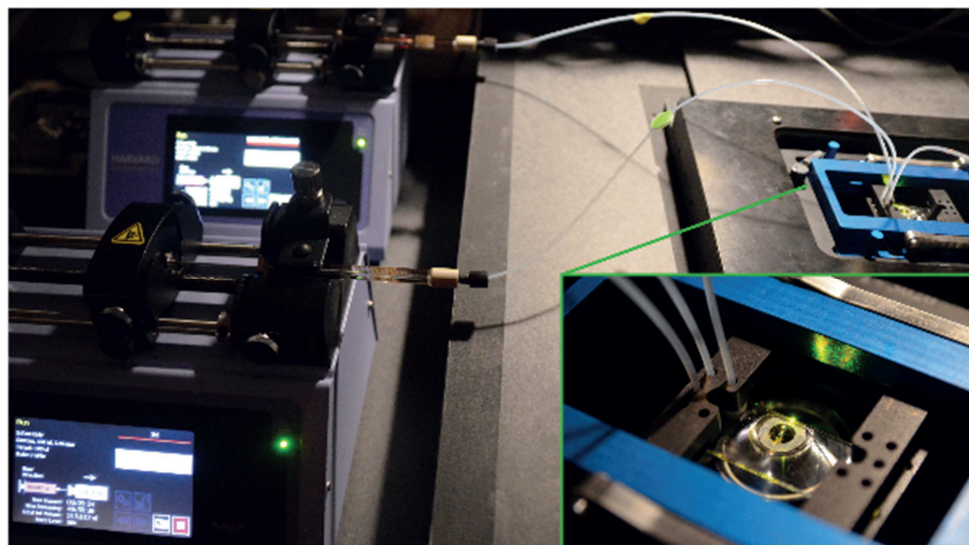
which can be integrated to obtain the Cumulative Distribution Function (CDF)<sup>31</sup> :

$$CDF(y, t) = \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{y - vt}{\sqrt{4Dt}}\right) \right] \quad (\text{eq. S5.7})$$

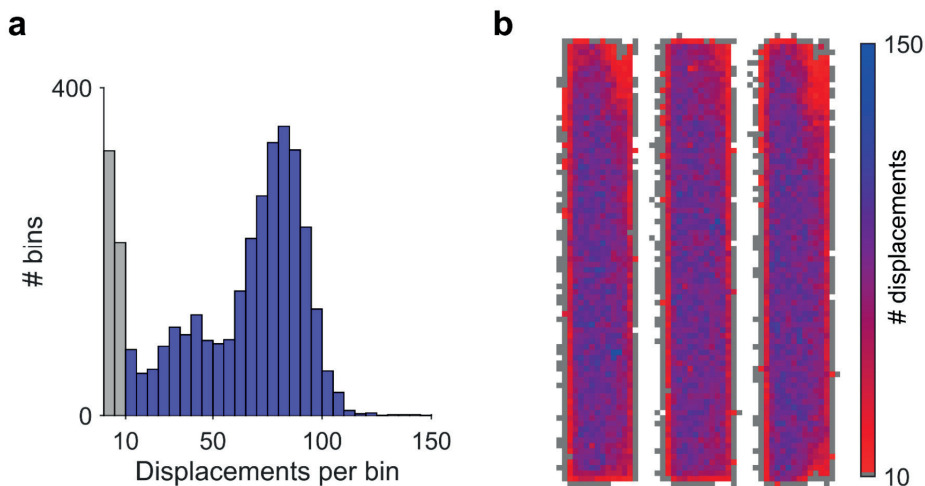
In Figure S5.6, the fit of experimental data to the CDF is shown for movies taken at different pump rates (10, 20 and 40)  $\mu\text{L/h}$ . The fit outputs for the flow speeds inside the channels are (-56, -114 and -221) nm/ms for the three respective pump rates; these values are in agreement with the experimental plateau values in the flow profiles (Fig. S5.4). The fits yields values for the diffusivity of (31, 32 and 34)  $\mu\text{m}^2/\text{s}$  for the three respective pump rates. The fit for the 40  $\mu\text{L/h}$  case is not as good as the other two; the reason is that, at this pump rate, a few particles can move outside the tracking window and are lost. (The tracking window throughout this analysis was set to 10 pixels; this means that we cannot see displacements longer than 1.1  $\mu\text{m}$ , which introduces a small bias in the analysis. In Figure S5.6c, the effect of this bias can be seen at displacements  $< -1 \mu\text{m}$ .)



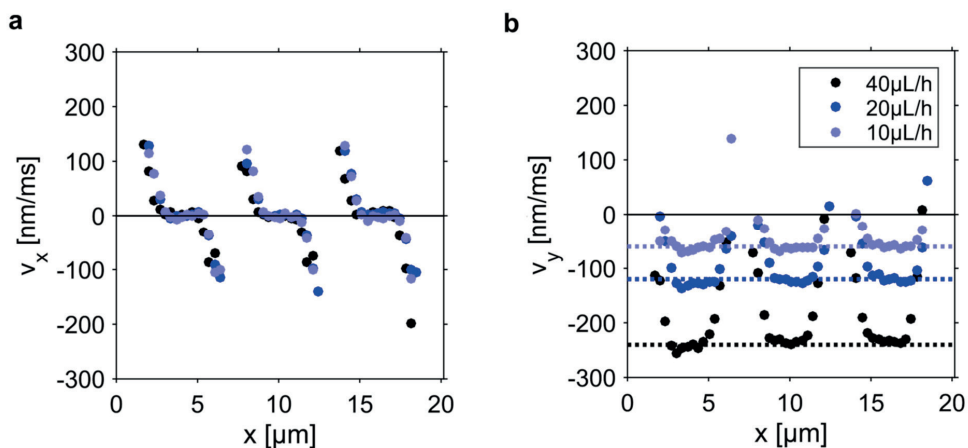
**Figure S5.1.** Achieving parallel flow control. Equivalent circuit diagram of the nano-/microfluidic channel geometries. Syringe pumps act as current sources delivering constant flow rates  $Q$  and channels oppose flow by a hydraulic resistance  $R_{hyd}$  (a) Parallel flow configuration of a bypassing microchannel and a nanochannel array. (b) Diagram of a micro-/nanofluidic mixing circuit. Here, reaction products are imaged downstream (right side) of the nanochannel junction.



**Figure S5.2.** Overview of the fluidic setup on a home-built microscope.

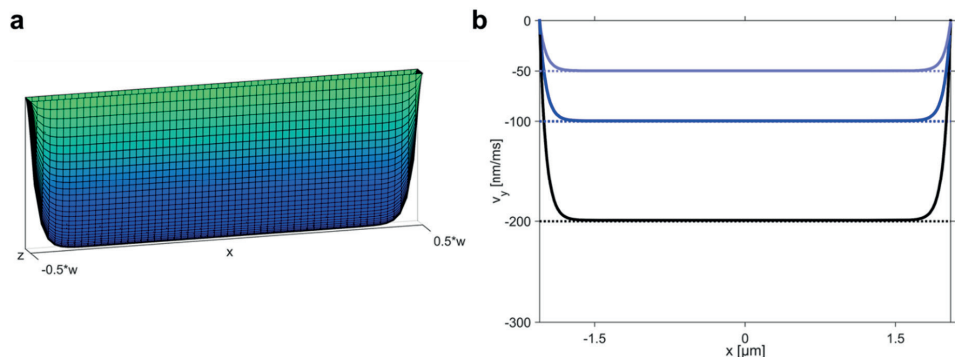


**Figure S5.3.** Number of displacements per bin and effect of the localization threshold. (a) Distribution of the number of displacements in the bins. Bins containing  $<10$  displacements were excluded from the analysis to reduce the error in the subsequent calculation of flow speeds. (b) Binned map showing the number of displacements at each location. Bins containing only a few displacements are mostly located at the edges of the nanochannels.

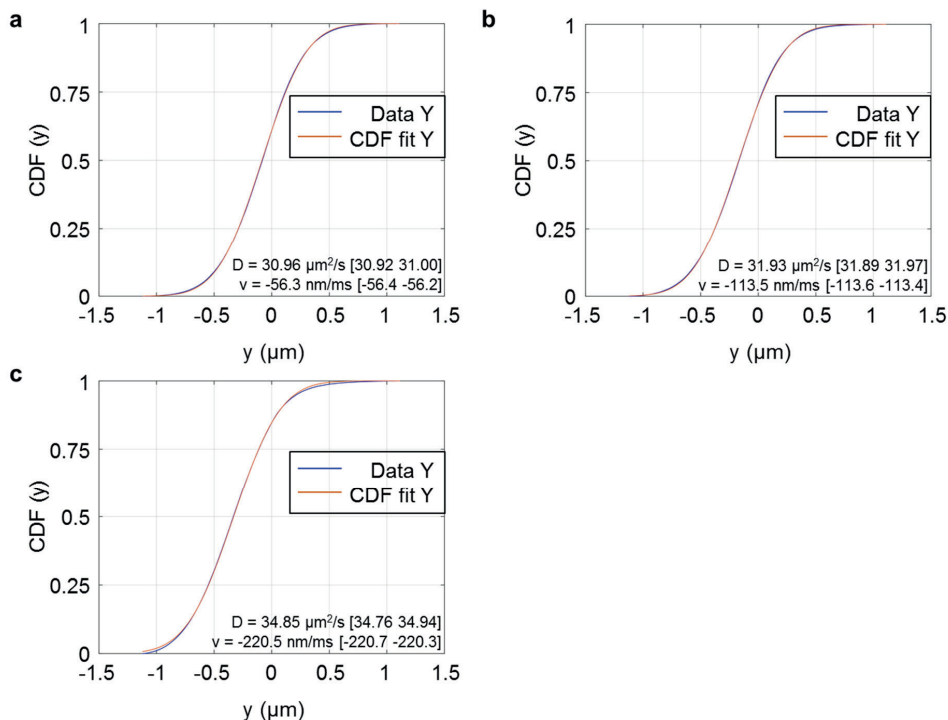


**Figure S5.4.** Experimental flow velocity profiles for different pump rates in three parallel nanochannels (40, 20 and 10)  $\mu\text{L/h}$  calculated starting from the binned map of the displacements. (a) Displacement along  $x$ : this direction is orthogonal to the applied flow thus the profiles are not influenced by the pump rate; i.e., they overlap with only minor differences on the most external values which are subjected to poor sampling (few tens of displacements per data-point). (b) Displacements along  $y$ : Reference lines are drawn at  $(-60, -120$  and  $-240)$   $\text{nm/ms}$ . The flow speed registered inside the nanochannels scales linearly with the pump rate

as expected for our system. All the profiles show a plateau for the velocity in the center of the channels (see also Fig. S5.5).

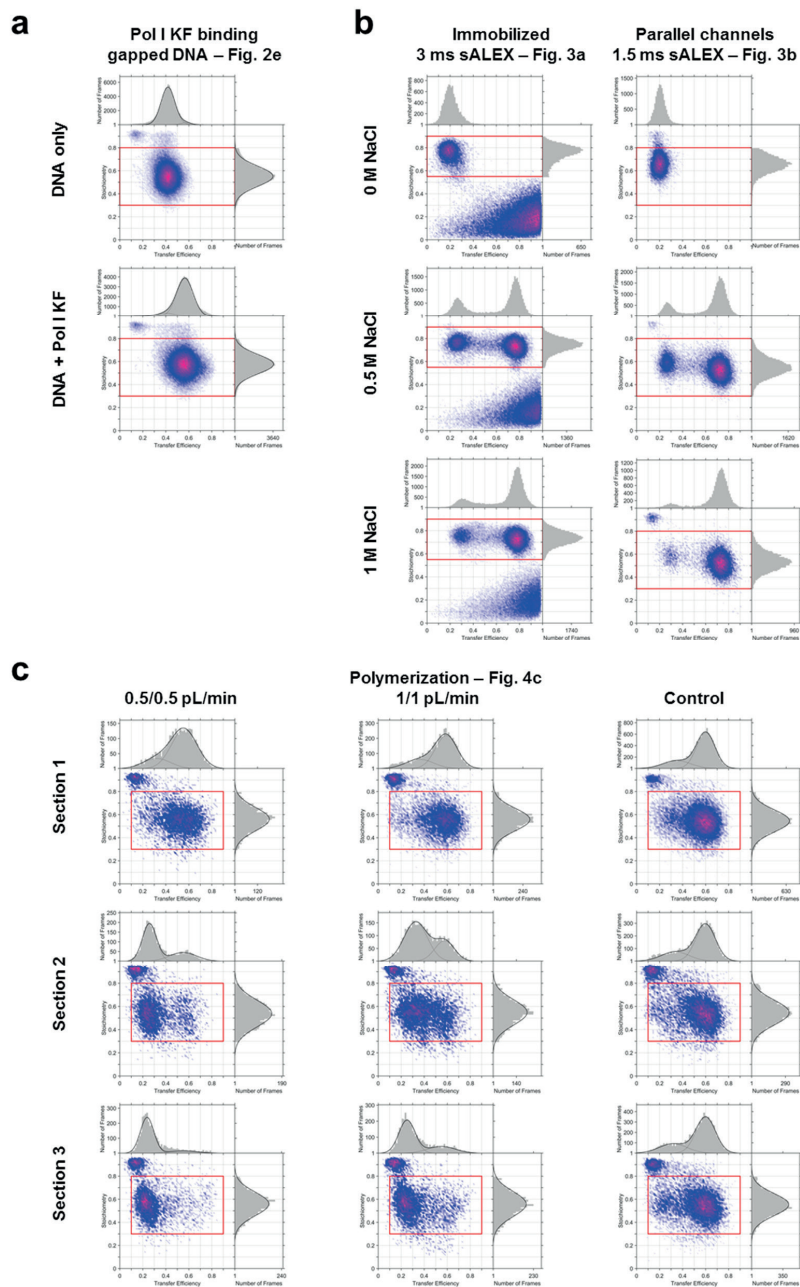


**Figure S5.5.** Theoretical 3D and 2D velocity profiles (see also Supporting Note 5.2). (a) Theoretical 3D velocity surface for Poiseuille flow in a nanochannel. The velocity surface was calculated using the dimensions of one parallel nanochannel. Along the height of the channels ( $z$ ,  $0.2\mu\text{m}$ ) the velocity profile is parabolic. Along most of the width ( $x$ , here  $4.1\mu\text{m}$ ) the flow speed is constant (parallel plate-like scenario). (b) Theoretical 2D velocity profiles for Poiseuille flow in a nanochannel. Theoretical velocity profiles calculated for different pump rates: (10, 20 and 40)  $\mu\text{L/h}$ ; the value of the plateau velocity ( $v_{\text{max}}$ ) scales linearly with the pump rate.



**Figure S5.6.** CDF fits of experimental data for different pump rates. Panels (a - c) show the experimental data and the CDF fit of pump rates of (10, 20 and 40)  $\mu\text{L/h}$  respectively. The fitted variables are reported as the mean followed by the 95% confidence interval. The fit in (c) is slightly worse than the fit in (a) and (b) as the higher flow speed increases the possibility that a particle would move out of the tracking window. This issue is visible at displacements  $< -1\ \mu\text{m}$  where the probability drops rapidly to zero introducing a small, and therefore negligible, bias. (see also Supporting Note 5.3).





**Figure S5.7.**  $E^*/S$  histograms of data shown in Figs 5.2e, 5.3a,b and 5.4c. Only data inside the red box is projected on the axes. FRET efficiency  $E^*$  is on the x-axis, stoichiometry  $S$  is on the y-axis.

**Supporting Table 5.1.** Accurate FRET for the DNA hairpin data shown in Figure 5.3. Comparison between immobilized (upper part) and nanochannel data (lower part). After corrections for leakage, direct excitation, gamma and beta, FRET efficiencies for immobilized and nanochannel data are similar.

	DNA	Uncorrected $E^*$				Corrected for Leakage + Direct Excitation				Corrected for gamma and beta			
		$E1$	$S1$	$E2$	$S2$	$E1$	$S1$	$E2$	$S2$	$E1$	$S1$	$E2$	$S2$
Immobilized	0 M NaCl	0.20	0.77	-	-	0.03	0.75	-	-	0.04	0.50	-	-
	0.5 M NaCl	0.27	0.77	0.77	0.74	0.13	0.74	0.75	0.72	0.15	0.50	0.78	0.50
	1 M NaCl	0.31	0.77	0.78	0.73	0.18	0.73	0.76	0.72	0.23	0.49	0.79	0.50
Parallel channels	0 M NaCl	0.20	0.66	-	-	-0.04	0.64	-	-	-0.02	0.56	-	-
	0.5 M NaCl	0.28	0.59	0.73	0.53	0.07	0.54	0.67	0.48	0.12	0.48	0.77	0.51
	1 M NaCl	0.30	0.59	0.73	0.53	0.09	0.53	0.68	0.48	0.17	0.48	0.78	0.50

**Supporting Movies** (watch online at <https://doi.org/10.1101/201079>)

#### Supporting Movie 1: Mixing two differently labelled fluorescent DNA strands – overlay

10 nM of a Cy3B-labelled DNA 55-mer (entering from the left) is mixed with 10 nM of an ATTO647N labelled DNA 30-mer (entering from the right). Both oligos were dissolved in PBS with an additional 1M NaCl. For this movie, the raw data was modified in ImageJ: the green and red channels (DD and AA signals) are merged, and colors were added. Flow rate from each side was 1.1 pL/min. Original image size corresponds to  $343 \times 256$  pixels. Display rate: 10 frames per second (original movie was taken at 100 frames per second). TIFF stack was compressed to JPEG: raw data is available upon request.

#### Supporting Movie 2: Fluorescently labelled gapped DNA flowing through parallel channels

1 nM of labelled gapped DNA flowing through parallel nanochannels. Pump rate was 40  $\mu$ L/h. Left panel: green detection channel. Right panel: red detection channel. Image size corresponds to  $343 \times 256$  pixels. Display rate: 10 frames per second (original movie was taken at 100 frames per second). TIFF stack was compressed to JPEG: raw data is available upon request.

### **Supporting Movie 3: Mixing DNA hairpins with a high-salt solution**

1 nM of labelled DNA hairpins (entering from the right) is mixed with a 1 M NaCl solution (entering from the left). Flow rate from each side was 3.7 pL/min. Left panel: green detection channel. Right panel: red detection channel. Image size corresponds to  $343 \times 256$  pixels. Display rate: 10 frames per second (original movie was taken at 100 frames per second). TIFF stack was compressed to JPEG: raw data is available upon request.

### **Supporting Movie 4: Triggering DNA polymerization inside a mixing channel**

1 nM of labelled DNA polymerization sensors (entering from the right together with 5 nM KF polymerase) is mixed with a 200  $\mu$ M dNTPs solution (entering from the left together with 5 nM KF polymerase). Flow rate from each side was 0.5 pL/min. Left panel: green detection channel. Right panel: red detection channel. Image size corresponds to  $343 \times 256$  pixels. Display rate: 10 frames per second (original movie was taken at 100 frames per second). TIFF stack was compressed to JPEG: raw data is available upon request.





## General Discussion



## General Discussion

In this thesis, we have explored single-molecule FRET as a tool to study DNA and DNA polymerases *in vitro*. We aimed to complement already existing biochemical and single-molecule approaches by developing powerful assays with high throughput, with which we are able to study polymerase activity and conformational dynamics. More specifically, we focused on questions regarding polymerase processivity and fidelity: relevant questions that studies on the ensemble level struggle to answer because of averaging effects in time and space. Single-molecule techniques have developed rapidly over the past 15 years, because of their ability to visualize or probe (bio)molecular kinetics without ensemble averaging. It remains a challenge, however, to monitor a large number of single molecules simultaneously, which is required to increase throughput and improve statistics, especially with force-based methods like optical tweezers or atomic force microscopy. TIRF microscopy, as applied in this thesis, achieves considerable higher throughput by offering parallelized detection of fluorescent molecules in a large field of view. We chose to combine this with FRET: a phenomenon in which energy is non-radiatively transferred between two fluorophores. The efficiency of energy transfer is sensitive to distance changes in the 2-10 nm range<sup>1</sup>, a region in which many intermolecular interactions and conformational changes take place. Despite considerable advances made by super-resolution microscopy to bring down the effective diffraction limit to lower than  $\sim\lambda/2$ , it cannot reliably cover this low-nanometre region.

As we discussed in chapter 2, DNA structures (such as hairpins or Holliday junctions) were among the first molecular species studied with single-molecule techniques. This is mainly due to the structural features of the DNA helix, which is stable, well-characterized and easy to modify. DNA processing enzymes were a logical next target for single-molecule studies. For our work, we made extensive use of the technical framework laid out by the earlier studies described in chapter 2, especially the alternating laser excitation scheme<sup>2,3</sup> and FRET-restrained structural modelling<sup>4</sup>.

## Probing DNA processivity and speed using a simple DNA sensor

DNA polymerization speed has traditionally been determined using single turnover experiments and gel-shift assays. With the advance of single-molecule methods, several assays have been developed to determine polymerization speed using force measurements or fluorescently labelled polymerases on fluorescently labelled DNA<sup>5-9</sup>. These approaches are inventive, and shine new light on variations in speed, sequence dependency and pausing motifs. We decided to develop a simple, widely applicable polymerization assay based on a fluorescent DNA sensor: a short DNA



primer-template labelled with a FRET dye pair (chapter 3). Our sensor is designed such that FRET efficiency decreases as the flexible template is incorporated in a more rigid double-helical structure. The design with both FRET labels on the DNA makes the assay compatible with a broad range of unlabelled polymerases. As a proof of concept, we tested three polymerases with different origin and function: *E. coli* DNA polymerase I Klenow Fragment (KF, involved in synthesis of short stretches of DNA), human Polymerase beta (POLB, which has a role in DNA repair) and the  $\alpha$  subunit of bacterial Polymerase III (POLIII $\alpha$ , part of the DNA replication machinery). We were able to trigger full polymerization of the template by adding each of these polymerases together with nucleotides to our surface-immobilized DNA sensors. For KF and POLIII $\alpha$ , we could distinguish multiple steps during polymerization, indicating that for these two polymerases processivity on our 25 nucleotide template is low. Their polymerization speeds, determined from our measurements, are consistent with literature values<sup>7</sup>. Notably, POLIII $\alpha$  is much slower than the POLIII holoenzyme, which can reach speeds of 1000 nucleotides per second (nt / s)<sup>8-10</sup>. Therefore, other POLIII subunits are thought to speed up the reaction in the holoenzyme. POLB is noticeably slower than KF and POLIII $\alpha$ , showing only a gradual decrease in FRET efficiency over time. This behaviour is explained by the enzyme's low processivity, and its preference for gapped DNA over a recessed primer-template<sup>11</sup>. We were able to temporarily stop polymerization by adding only a subset of nucleotides: in the template we used, the addition of a single nucleotide at position +3 is visible in FRET efficiency histograms. This effective "resolution" can be influenced by choosing different labelling positions that create inter-dye distances in the FRET-sensitive region around the Förster radius  $R_0$ . We showed a comparison between sensors labelled at positions +25/-7 and +12/-12: because the +25 position is farther on the template than the +12 position, the +25/-7 sensor can report on polymerization of more bases. However, the proximity of the -7 position to the 3' end of the primer may interfere with initial polymerase binding: the apparent affinity of KF for the +12/-12 sensor was 3-4 times higher than for the +25/-7 sensor. Because of the relative simplicity of our assay, we used it to test our nanofluidic mixing device. Finding "full" polymerization traces of single sensors turned out to be challenging, mainly due to premature fluorophore bleaching, but the overall time for the polymerase to finish the reaction was well within the time molecules spent inside the channel (~5 s), which in turn is on the same time scale as the surface-immobilized measurements. Although we specifically looked at polymerization at the single-molecule level, we see possibilities for wide application of this sensor in ensemble assays that use a fluorescence readout. An assay very similar to ours, but aimed at ensemble studies, was recently published by Sharma *et al.*<sup>12</sup>. They used it to screen potential inhibitors of HIV-1 reverse transcriptase in an implementation that makes the assay immediately applicable for the pharmaceutical industry, using stopped-flow and plate readers. Like

us, they found that dye placement can affect the observed polymerase speed, and the sensors should therefore be carefully designed. As a follow-up, we plan to use our assay to test two artificial polymerases that are based on a truncated version of Klenow Fragment<sup>13</sup>, to help characterize its smallest catalytic unit both at the ensemble level (stopped-flow) and single-molecule level (nanofluidic mixing).

## **DNA polymerase beta: fingers closing as a fidelity checkpoint**

We took a closer look at the mechanism behind nucleotide selection and decided to take POLB as an example (chapter 4). POLB is involved in the base excision repair pathway, where it takes care of incorporating the correct base into a gapped, damaged DNA substrate. From previous work on KF, it had become clear that movement of the polymerase fingers domain (“fingers closing”) is an important fidelity checkpoint<sup>14,15</sup>. We aimed to study this conformational change by using a version of POLB that carries an acceptor label on the fingers (prepared by the lab of dr. Joann Sweasy, Yale University School of Medicine), while providing an immobile position on the DNA primer with a FRET donor. This approach enabled us to monitor fingers closing as a change in FRET efficiency, and to extract opening and closing rates as a function of nucleotide concentration. Interestingly, we found evidence that fingers close only partially when an incorrect nucleotide is bound, similar to the behaviour of the fingers domain of KF. This implies that nucleotides are indeed already “screened” during fingers closing, and rejected if they are not complementary or have the wrong backbone structure (non-complementary dNTPs versus complementary rNTPs). DNA binding times, however, do not become shorter when incorrect nucleotides are supplied, and are even increased at nucleotide concentrations  $> 1$  mM. This is fundamentally different from a mechanism observed for KF, in which a wrong nucleotide increases the dissociation rate and therefore shortens the binding time<sup>15</sup>. We hypothesize that the increased complex stability observed for POLB shortens the time needed for DNA repair. Despite our success with a singly labelled version of POLB and a labelled DNA construct, a more direct approach of measuring these small changes of the fingers domain would see both fluorophores on the polymerase: one label on the fingers, and a reference label on an immobile part of the polymerase. We previously tested a version of POLB with a single label at position T93C, in the “thumb” region of the polymerase, and found that full fingers closing only caused a slight change in  $E^*$  ( $\Delta E^* = 0.03$ , data not shown in this thesis). Position T93C may therefore be a good candidate for this reference label. An assay with a doubly labelled protein would still require labelled DNA molecules for initial localization, so non-specific adsorption can be distinguished from substrate binding. The label on the DNA can then be bleached off early in the experiment to not interfere with the FRET signal. This doubly labelled version is currently under construction, and will hopefully shed more light on the fidelity checkpoint that we identified. On a side note, we did not

observe sharp bending of the DNA substrate upon POLB binding (as is predicted by crystal structures 1BPX and 1BPY). Whether substrate bending is dependent on the length of the DNA, the sequence or the presence of fluorophores remains an open question.

## **Go with the flow: increasing the throughput of single-molecule experiments**

As we described in this thesis, TIRF microscopy allows for monitoring >100 molecules simultaneously for tens of seconds when combined with surface immobilization of the target molecules. Surface immobilization, however, has several disadvantages: molecules to be immobilized need an additional tag and parts of the molecule close to the surface may become inaccessible for large binding partners. Moreover, reactions at non-equilibrium conditions are hard to tune (for example, our polymerization assay can only be performed once per well). For studying non-immobilized molecules, one had to revert to confocal microscopy. Special mixing devices for studies under non-equilibrium conditions have been realized for this type of microscopy<sup>16,17</sup>, but one has to sacrifice observation times of single molecules and overall throughput. In the past, several alternative methods for long observation of non-immobilized single molecules have been presented, e.g. the ABEL trap and SWIFT<sup>18,19</sup>. The latter is of particular interest, since it allows detection of multiple single molecules in a parallelized manner using camera-based detection, enabling high throughput. However, accessing non-equilibrium conditions was not possible. In chapter 5, we presented two novel nanofluidic devices that allow to work at high throughput with non-immobilized single molecules. One design contains an array of parallel channels (to be used for equilibrium studies), and the other contains a T-shaped mixing channel (which can be used to trigger reactions in very small volumes). In these devices, fluorescent molecules move through an array of channels with a height of 200 nm, which restricts their movement along the optical axis of the microscopy, ensuring that they do not move out of focus. Molecules are detected and tracked with home-built software. While designed for high throughput experiments, individual molecules can be tracked for up to 50 frames at a slightly lower throughput. Our devices are made entirely out of glass and are user-friendly in their implementation. Additionally, we implemented parallel flow control, which reduces the pressure needed to pump molecules through the channels by including a far larger, parallel microchannel in the designs. To test the possibilities of our devices, we flowed DNA constructs labelled with FRET dye pairs through the channels. As expected, we measured similar FRET efficiencies for molecules inside the channels and molecules that were immobilized on a PEGylated coverslip. Flow in channels with such small dimensions is expected to be laminar. Based on sub-pixel localizations of molecules over a time span of 1.5 ms, we were able to construct flow profiles inside the channels. Indeed, we observed the profile

associated with laminar flow through a rectangular channel. The distribution of movements also contained information on the flow velocity in relation to diffusion: a fit with a cumulative distribution function revealed that motion due to flow and motion due to diffusion are of the same order of magnitude. This property is especially important for our T-shaped mixing channels, since this so-called “diffusive mixing” ensures that mixing already takes place at the junction. This phenomenon is well illustrated when a DNA hairpin is mixed with a high salt solution: upon mixing, the equilibrium of the hairpin shifts from mostly open to mostly closed, which we observed as a sudden change in FRET efficiency at the junction. In a next step, we tried to trigger DNA polymerization inside the channels. Although we did observe polymerization, long exposures of the dyes to the lasers (because of low flow speeds) and the presence of proteins induced strong bleaching. We believe, however, that precise tuning of a biochemical reaction at this scale is an important advancement for the field. A possible approach to counteract bleaching, to be tested in the future, could include attaching a large biomolecule to the DNA sensor to decrease the motion due to diffusion, which would in turn allow a longer excitation time and lower laser powers.

## **A note on the use of FRET as a molecular ruler**

Throughout this thesis, we frequently used FRET to calculate inter-dye distances. Though the application of correction procedures is straightforward, a note needs to be made with regard to the validity of the underlying assumptions<sup>4,20</sup>. FRET is a consequence of a dipole-dipole interaction, and therefore depends on the orientation of both fluorophores. As explained in the General Introduction, it is often assumed that those fluorophores can rotate freely, and the dipole orientation factor  $\kappa^2$  averages out to  $2/3$ . If both fluorophores are inhibited in their rotational freedom,  $\kappa^2 = 2/3$  is not a valid assumption anymore, and the measured FRET efficiency is affected. The FPS software that we used to calculate the accessible volumes does not consider any chemical or physical interactions between the fluorophore and the molecular surface that may cause rotational restrictions: the dye is assumed not to “stick”, and all positions are sampled equally. Molecular Dynamics simulations, however, have shown that interactions between fluorophores and DNA or proteins are well possible: cyanine dyes are known for their ability to undergo stacking interactions with bases in a DNA helix<sup>21,22</sup>, and charged fluorophores like ATTO647N may be attracted to or repelled from charged (protein) surfaces. In practice, fluorescence anisotropy measurements can be used to test for interactions that inhibit the rotational freedom (rotational relaxation should be faster than energy transfer). We did not perform any anisotropy measurements in this thesis, thereby assuming that at least 1 fluorophore of the FRET pair can rotate freely. For the relatively simple structures of DNA double helices (like our polymerized sensors in chapter 3), we could easily compare the

obtained distances against a model. Therefore, we gave additional anisotropy measurements less priority. In chapter 4, however, we question whether we can use the conventional crystal structures 1BPX and 1BPY for modelling our POLB-DNA complex. Additional anisotropy measurements may reveal interactions of the dyes with the protein, and are therefore a good next step before strong claims on the validity of the crystal structure are made. In general, when large-scale structural modelling needs to be done, anisotropy measurements are absolutely necessary.

For a more detailed discussion on other parts of the FRET correction procedure, amongst others the uncertainties associated with distance determination, please refer to the worldwide benchmark study by Hellenkamp *et al.*<sup>20</sup>.

## A brief look into the future of single-molecule FRET

A major limitation of the emCCD camera that we used is the limited acquisition rate (56 Hz on 512 x 512 pixels). We applied a stroboscopic excitation scheme to visualize dynamics at faster timescales than continuous illumination with an emCCD camera would allow<sup>23</sup>. Since a few years, however, faster sCMOS cameras have appeared on the market. These cameras can reach high acquisition rates on full frame (100 Hz on 2048 x 2048 pixels), making them an attractive choice to study fast biochemical reactions<sup>24</sup>. With the demand for increased throughput and time resolution come challenges to handle large amounts of single-molecule data and to improve fluorophore photostability. Powerful computers and new algorithms are needed for fast (ideally real-time) data analysis. A new localization algorithm based on Phasor has recently been developed, achieving fast localization rates (3 million per second) with comparable accuracy to established methods like Gauss-MLE<sup>25</sup>. An example of the other challenge (insufficient photostability) comes from our DNA polymerization assay in the T-shaped nanochannel, which shows severe premature dye bleaching. Improvements in photostability have been achieved by covalently attaching a triplet state quencher to a dye molecule, so transitions from a dark state back to a fluorescent state do not rely on random collisions<sup>26,27</sup>. Further optimization of photostability, however, remains a priority before technological advances can be used to their full extent.

While considerable progress has been made to push the boundaries of single-molecule techniques, it is easy to overlook the day-to-day challenges which are encountered when setting up a single-molecule experiment. Method-wise, improvements in the field can not only be made in higher time resolution and throughput, but also in the basic, but essential, area of surface passivation. For the work in this thesis, we used PEG to reduce non-specific surface interactions for almost all single-molecule experiments. The use of PEG layers, however, is not ideal. The main limitation is the variation in anti-fouling properties among batches and over time. Furthermore,

experiments with labelled proteins show that non-specific adsorption can become a problem at higher protein concentrations (e.g. fluorescently labelled POLB showed strong sticking at 50 nM). Recent studies to improve surface passivation have come up with optimized PEGylation protocols<sup>28</sup>, and with completely different approaches using for example combinations of silanes and surfactants<sup>29</sup> or peptoid chains<sup>30</sup>. Another promising candidate are zwitterionic polymers, which have applications in biosensors and are known for their excellent anti-fouling properties<sup>31</sup>. Recently, functionalization of these polymers with a biotin moiety was reported<sup>32</sup>, which would facilitate easy integration with our existing immobilization procedure that uses biotin-neutravidin-biotin linkages. We have started promising pilot experiments in collaboration with the Laboratory of Organic Chemistry (WUR). The current challenge is larger scale synthesis of these functionalized zwitterionic coatings, which at the moment requires a time-consuming surface-initiated polymerization method.

Apart from increased time resolutions and higher throughput, it would be interesting to see more single-molecule experiments carried out *in vivo*, bridging the gap with super-resolution microscopy. Fluorescent proteins like GFP have proven very useful for super-resolution microscopy in living cells, but structural and kinetic work needs integration with bright, small organic fluorophores. One of the approaches is transfection of labelled biomolecules into living cells. Recent studies used electroporation to internalize fluorophore-labelled DNA in bacteria<sup>33–35</sup>, which was successful and allowed the researchers to retrieve the expected FRET efficiencies from measurements at the single-cell level. Another study used microinjection, which allows for more control over the amount of internalized species<sup>36</sup>. A second approach aims to label proteins *in vivo*. These reactions should not interfere with any other reaction in the cell, and are therefore called “bio-orthogonal reactions”. Early developments were the introduction of the SNAP- and CLIP-tags, which use a mutant version of the (small) hAGT enzyme coupled to a protein of interest. hAGT can then covalently link a fluorophore, supplied as an O<sup>6</sup>-benzylguanine derivative, to itself and thus to the protein of interest<sup>37,38</sup>. Introduction of tetracysteine tags in proteins has also proven successful for labelling with cell-permeable biarsenical dyes<sup>39,40</sup>. Importantly, both these labelling procedures are intracellular. Since then, efforts have been made to decrease the size of the tag<sup>41</sup>. Intracellular labelling was not always possible because of cell permeability problems or questions regarding labelling specificity or cytotoxicity<sup>41</sup>. The development of shorter tags, however, did lead to an extracellular FRET assay on dynamics of single HIV-envelope trimers in intact virions<sup>42</sup>. In 2017, Sakin *et al.* introduced non-canonical amino acids<sup>43</sup> into the protein of interest, which can then react to a synthetic dye through click-chemistry. Labelling of the modified protein (also here an HIV-envelope protein) was successful, but took place at the plasma membrane. We conclude that the development of intracellular single-molecule assays with bright fluorophores remains challenging, which is not



surprising given the complex environment in the cell. However small, exciting steps are being made.

Many biological studies would benefit from a molecule-by-molecule approach, and therefore single-molecule techniques are rapidly becoming more popular throughout the life sciences. Previous work proves that a wide variety of nucleic acid-nucleic acid interactions (Holliday junctions<sup>44</sup>, riboswitches<sup>45</sup>), protein-nucleic acid interactions (DNA polymerases, chromatin structure<sup>46</sup>) and protein-protein interactions (molecular motors<sup>47</sup>, protein folding and modifications<sup>48</sup>) can all be investigated in great detail. In this thesis, we explored and expanded the possibilities of fluorescence-based single-molecule experiments: we visualized stepwise DNA synthesis and found evidence for a fidelity checkpoint in POLB, while our nanofluidic devices open possibilities to measure non-immobilized molecules with increased time resolution and throughput.

## References

1. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55–75 (1948).
2. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).
3. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
4. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).
5. Christian, T. D., Romano, L. J. & Rueda, D. Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc. Natl. Acad. Sci.* **106**, 21109–21114 (2009).
6. Schwartz, J. J. & Quake, S. R. Single molecule measurement of the ‘speed limit’ of DNA polymerase. *Proc. Natl. Acad. Sci.* **106**, 20294–20299 (2009).
7. Maier, B., Bensimon, D. & Croquette, V. Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12002–12007 (2000).
8. Tanner, N. A. *et al.* Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat. Struct. Mol. Biol.* **15**, nsmb.1381 (2008).
9. Tanner, N. A. *et al.* Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Res.* **37**, e27–e27 (2009).
10. Pham, T. M. *et al.* A single-molecule approach to DNA replication in *Escherichia coli* cells demonstrated that DNA polymerase III is a major determinant of fork speed. *Mol. Microbiol.* **90**, 584–596 (2013).



11. Prasad, R., Beard, W. A. & Wilson, S. H. Studies of gapped DNA substrate binding by mammalian DNA polymerase beta. Dependence on 5'-phosphate group. *J. Biol. Chem.* **269**, 18096–18101 (1994).
12. Sharma, K. K., Przybilla, F., Restle, T., Godet, J. & Mély, Y. FRET-based assay to screen inhibitors of HIV-1 reverse transcriptase and nucleocapsid protein. *Nucleic Acids Res.* **44**, e74 (2016).
13. Martina, C. E., Lapenta, F., Montón Silva, A. & Hochkoeppler, A. HoLaMa: A Klenow sub-fragment lacking the 3'-5' exonuclease domain. *Arch. Biochem. Biophys.* **575**, 46–53 (2015).
14. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nat. Commun.* **4**, (2013).
15. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res.* gkv547 (2015). doi:10.1093/nar/gkv547
16. Dingfelder, F. *et al.* Rapid Microfluidic Double-Jump Mixing Device for Single-Molecule Spectroscopy. *J. Am. Chem. Soc.* **139**, 6062–6065 (2017).
17. Wunderlich, B. *et al.* Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat. Protoc.* **8**, 1459–1474 (2013).
18. Cohen, A. E. & Moerner, W. E. Suppressing Brownian motion of individual biomolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4362–4365 (2006).
19. Tyagi, S. *et al.* Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nat. Methods* **11**, 297–300 (2014).
20. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements - a worldwide benchmark study. *ArXiv171003807 Phys. Q-Bio* (2017).
21. Dolgikh, E., Roitberg, A. E. & Krause, J. L. Fluorescence resonance energy transfer in dye-labeled DNA. *J. Photochem. Photobiol. Chem.* **190**, 321–327 (2007).
22. Spiriti, J., Binder, J. K., Levitus, M. & van der Vaart, A. Cy3-DNA Stacking Interactions Strongly Depend on the Identity of the Terminal Basepair. *Biophys. J.* **100**, 1049–1057 (2011).
23. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17**, 27862–27872 (2015).
24. Juetten, M. F. *et al.* Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat. Methods* **13**, 341–344 (2016).
25. Martens, K., Bader, A. N., Baas, S., Rieger, B. & Hohlbein, J. Phasor based single-molecule localization microscopy in 3D (pSMLM-3D): an algorithm for MHz localization rates using standard CPUs. *bioRxiv* 191957 (2017). doi:10.1101/191957

26. Altman, R. B. *et al.* Cyanine fluorophore derivatives with enhanced photostability. *Nat. Methods* **9**, 68–71 (2012).
27. Tinnefeld, P. & Cordes, T. ‘Self-healing’ dyes: intramolecular stabilization of organic fluorophores. *Nat. Methods* **9**, 426–427 (2012).
28. Chandradoss, S. D. *et al.* Surface Passivation for Single-molecule Protein Studies. *J. Vis. Exp. JoVE* (2014). doi:10.3791/50549
29. Hua, B. *et al.* An improved surface passivation method for single-molecule studies. *Nat. Methods* **11**, 1233–1236 (2014).
30. Ryu, J. Y. *et al.* New Antifouling Platform Characterized by Single-Molecule Imaging. *ACS Appl. Mater. Interfaces* **6**, 3553–3558 (2014).
31. Schlenoff, J. B. Zwitteration: Coating Surfaces with Zwitterionic Functionality to Reduce Nonspecific Adsorption. *Langmuir* **30**, 9625–9636 (2014).
32. Lange, S. C., van Andel, E., Smulders, M. M. J. & Zuilhof, H. Efficient and Tunable Three-Dimensional Functionalization of Fully Zwitterionic Antifouling Surface Coatings. *Langmuir* **32**, 10199–10205 (2016).
33. Crawford, R. *et al.* Long-Lived Intracellular Single-Molecule Fluorescence Using Electroporated Molecules. *Biophys. J.* **105**, 2439–2450 (2013).
34. Plochowitz, A., Crawford, R. & Kapanidis, A. N. Characterization of organic fluorophores for in vivo FRET studies based on electroporated molecules. *Phys. Chem. Chem. Phys.* **16**, 12688–12694 (2014).
35. Aigrain, L., Sustarsic, M., Crawford, R., Plochowitz, A. & Kapanidis, A. N. Internalization and Observation of Fluorescent Biomolecules in Living Microorganisms via Electroporation. *J. Vis. Exp. JoVE* (2015). doi:10.3791/52208
36. König, I. *et al.* Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat. Methods* **12**, 773–779 (2015).
37. Juillerat, A. *et al.* Directed Evolution of O6-Alkylguanine-DNA Alkyltransferase for Efficient Labeling of Fusion Proteins with Small Molecules In Vivo. *Chem. Biol.* **10**, 313–317 (2003).
38. Gautier, A. *et al.* An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chem. Biol.* **15**, 128–136 (2008).
39. Griffin, B. A., Adams, S. R. & Tsien, R. Y. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **281**, 269–272 (1998).
40. Gaietta, G. *et al.* Multicolor and Electron Microscopic Imaging of Connexin Trafficking. *Science* **296**, 503–507 (2002).
41. Crivat, G. & Taraska, J. W. Imaging proteins inside cells with fluorescent tags. *Trends Biotechnol.* **30**, 8–16 (2012).
42. Munro, J. B. *et al.* Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. *Science* **346**, 759–763 (2014).

43. Sakin, V. *et al.* A Versatile Tool for Live-Cell Imaging and Super-Resolution Nanoscopy Studies of HIV-1 Env Distribution and Mobility. *Cell Chem. Biol.* **24**, 635–645.e5 (2017).
44. McKinney, S. A., Déclais, A.-C., Lilley, D. M. J. & Ha, T. Structural dynamics of individual Holliday junctions. *Nat. Struct. Mol. Biol.* **10**, 93–97 (2003).
45. Savinov, A., Perez, C. F. & Block, S. M. Single-molecule Studies of Riboswitch Folding. *Biochim. Biophys. Acta* **1839**, 1030–1045 (2014).
46. Cuvier, O. & Fierz, B. Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells. *Nat. Rev. Genet.* **18**, 457–472 (2017).
47. Veigel, C. & Schmidt, C. F. Moving into the cell: single-molecule studies of molecular motors in complex environments. *Nat. Rev. Mol. Cell Biol.* **12**, 163–176 (2011).
48. Bustamante, C. J., Kaiser, C. M., Maillard, R. A., Goldman, D. H. & Wilson, C. A. M. Mechanisms of Cellular Proteostasis: Insights from Single-Molecule Approaches. *Annu. Rev. Biophys.* **43**, 119–140 (2014).





## Summary



This thesis focuses on the characterization of DNA polymerases with single-molecule techniques. More specifically, I aimed to study polymerase processivity and fidelity-related conformational changes using assays based on Förster Resonance Energy Transfer (FRET) on a total internal reflection fluorescence (TIRF) microscope.

Chapter 2 reviews some of the recent applications of single-molecule FRET (smFRET) to study DNA and DNA binding proteins, in particular DNA polymerases. The chapter begins with an introduction of FRET, employed to measure distance changes in the 1-10 nm region, and introduces the two most common fluorescence-based implementations of single-molecule techniques: confocal microscopy and TIRF microscopy. The chapter concludes with a short discussion on FRET-based structural modelling, parts of which are applied in practice later in this thesis.

In chapter 3, I report the development of a short, fluorescently labelled DNA sensor to probe DNA polymerization at the single-molecule level. The sensor is a simple primer-template combination labelled with donor and acceptor fluorophores suitable for FRET. The advantage of this assay is that polymerases do not need to be labelled with any fluorophore. I show that the FRET efficiency of the sensors changes significantly upon polymerization of the 25 nucleotide template, and I present time traces showing polymerization of single sensors by three different polymerases (*E. coli* DNA Polymerase I (KF), human Polymerase Beta (POLB) and the  $\alpha$  subunit of bacterial Polymerase III (POLIII $\alpha$ )). Based on these traces, I can measure polymerase speed and pausing: KF and POLIII $\alpha$  extended the primer in  $\sim$ 1.0-1.5 s, but POLB was far slower (tens of seconds). I foresee applications for these sensors in the single-molecule field, where they can be used to characterize the processivity of other polymerases, but also for ensemble experiments in which native polymerases need to be tested for activity.

I take a closer look at POLB in chapter 4. This polymerase is involved in DNA repair, and I address the question whether resolving the conformational dynamics of the enzyme can shed new light on fidelity-related mechanisms. Previous work on both KF and POLB showed that the polymerase “fingers” domain binds a nucleotide and subsequently transfers it to the active site (a conformational change known as “fingers closing”). For KF, it was shown that the fingers domain does not entirely close when non-complementary nucleotides are present, suggesting that nucleotides are screened for complementarity with the templating base during fingers closing. To see whether POLB employs a similar mechanism, I designed an smFRET assay with an immobile donor fluorophore on the DNA primer and an acceptor fluorophore on the fingers domain. Using this approach, I can visualize fingers closing in the presence of the correct nucleotide in single POLB-DNA complexes. Incorrect nucleotides (non-complementary dGTPs and complementary rUTPs) did not induce fingers closing. Instead, we observed a slight shift in the mean FRET efficiency of the open



conformation (from  $E^* \approx 0.55$  to  $E^* \approx 0.62$ ), while a fully closed conformation corresponds to  $E^* \approx 0.75$ . I find evidence for a partially closed, fidelity-related conformation of the fingers subdomain. Simultaneously, I find that high concentrations of incorrect nucleotides (1 mM and 3 mM) stabilize the POLB-DNA complex by lowering the POLB dissociation rate. In contrast, for KF, a destabilizing effect was shown previously. The mechanism behind this stabilization remains unknown, but I hypothesize that with the abundance of incorrect nucleotides in the cell, DNA repair is much faster if high levels of these nucleotides do not promote dissociation.

In chapter 5, I introduce novel nanofluidic devices for high-throughput single-molecule imaging. These devices are completely made of glass. I present two designs: one design with a parallel array of nanochannels for equilibrium studies, and another with a single, T-shaped nanochannel for mixing studies allowing access to non-equilibrium conditions. A channel height of 200 nm confines movement of the molecules such that they do not move out of focus. With the implementation of parallel flow control, the devices can be driven with conventional syringe pumps. I achieve a high temporal resolution on our emCCD camera due to stroboscopic excitation (1.5 ms excitation in 10 ms frame time). I show that we can track single molecules at low concentrations for extended periods of time. The track length depends on the flow speed, but ranges from several frames to tens of frames. Moreover, at higher concentrations, I achieve hundreds of thousands of localizations within 10 minutes. These localizations allowed me to construct flow profiles, which confirms that the flow in the nanochannels is laminar. I also calculate that, at low flow rates and with the small DNA molecules I used, motion due to flow is of the same order of magnitude as motion due to diffusion. I illustrate this concept by mixing DNA hairpins in a primarily open configuration with a high-salt solution in the mixing channel: the FRET signature of the hairpins changes abruptly towards an equilibrium of primarily closed DNA hairpins. After fine-tuning the conditions, this so-called “diffusive” mixing is employed to trigger single-molecule reactions: I successfully polymerize my previously described DNA sensor inside the channel. I believe that these nanofluidic devices are a promising platform for studying non-immobilized single molecules at high throughput and high temporal resolution.



## About the author





## Acknowledgments

My name may be the only one on the front cover, but this thesis (and science in general) is always a team effort. I have had an amazing four years working on this thesis, and I would like to take the opportunity to acknowledge the people that were involved.

First of all, I would like to thank my day-to-day supervisor, dr. Johannes Hohlbein. He gave me the chance to start working on a truly exciting, but challenging project: starting up single-molecule research in Wageningen. His guidance and advice have been really helpful. Thanks for giving me the opportunity to take part in this exciting journey!

Herbert, thanks for your support as my promotor and creating a great research environment in the lab.

For most of the research, I have worked closely with very talented scientists from other labs: Alejandro Montón Silva and Alejandro Hochkoeppler for the polymerization sensor, Klaus Mathwig and Serge Lemay for the nanochannels and Mariam Mahmoud, Jamie Towle-Weicksel and Joann Sweasy for polymerase beta. Science is more fun when you build bridges to different labs and disciplines. Thanks for the work you have done and the insights you have shared. It has been a pleasure working with you!

Of course there are my close colleagues in the single-molecule group: Shazia Farooq, Mattia Fontana and Koen Martens. We shared the same sense of nerdy humour (or is that just me trying too hard to be funny?). Guys, I had a great time! Tatiana, Yashar, Julia, Donny, Ahmad, Ana, Alena, Olga, Daan, Fugui, Shanthi, Evgenia and Caner: thanks for joining me in this adventure of so many fluorescent lifetimes!

The people of the Laboratory of Biophysics (BIP) are of a special kind. They are all on the same wavelength and therefore get excited easily, especially by laser talk. Thanks for great discussions about philosophical topics like life, science and stroopwafels. My special thanks go to John Philippi and Arjen for their technical support, and Cor for helping me out with the lab supplies. Rob, Frank, Edo, Netty, Emilie, Elena, Henk, Pieter and John van Duynhoven: you make BIP an awesome place to work.

I would also like to thank my students: Timo Wenzel, Ebru Acun and Rebecca Kaup. You did a wonderful job, and I wish you all the best for the rest of your careers.

Finally, I would like to thank my friends, my parents and my sister who have supported me over the years. Whether it was a relaxing board game night, a one-star-horror-movie night or just some good beers: it is great to come home and realize I have such great friends and such a supportive family. Thank you!

## Curriculum Vitae

Carel Fijen was born on July 7, 1989 in Purmerend, the Netherlands. He finished secondary school (VWO gymnasium) in 2007. Because of his great interest in chemistry and biology, he continued his education at Wageningen University, where he studied Molecular Life Sciences. This interdisciplinary programme combines chemistry, biology and physics. His thesis work at the Laboratory of Microbiology focused on finding DNA aptamers for haemoglobin and ectoine; during his internship at Montana State University (U.S.A.), he developed nanoreactors based on bacteriophage protein cages, which resulted in a publication in the Journal of Materials Chemistry B. He obtained his MSc degree in early 2013. Eager to continue in science and intrigued by DNA repair, he started his PhD in the group of dr. Johannes Hohlbein, at the Laboratory of Biophysics. In his free time, he likes to write short stories and go outdoors with his camera.







## Overview of completed training activities

### Discipline specific activities

Biophysics @ FOM	FOM	Veldhoven, NL	2013-2017
Picoquant Single Molecule Workshop	Picoquant	Berlin, DE	2014
EMBL conference: Seeing is Believing	EMBL	Heidelberg, DE	2015
Advanced course Biomolecular Modelling	VU	Amsterdam, NL	2015
VLAG course Advanced Chemistry	VLAG	Wageningen, NL	2016

### General courses

Reviewing a Scientific Paper	WGS	Wageningen, NL	2013
VLAG PhD week	VLAG	Baarlo, NL	2014
Scientific Publishing	WGS	Wageningen, NL	2014
Techniques for Writing and Presenting a Scientific Paper	WGS	Wageningen, NL	2014
High Impact Writing	WGS/WIAS	Wageningen, NL	2015
Writing Grant Proposals	Wageningen Into Languages	Wageningen, NL	2016

### Optional courses and activities

Course Programming in Python	Wageningen University & Research	Wageningen, NL	2013
Writing a Scientific Review			2013
Group meetings and colloquia at the Laboratory of Biophysics		Wageningen, NL	2013-2017

Layout and cover design by Carel Fijen.

Research presented in this thesis was performed at the Laboratory of Biophysics, Wageningen University & Research, The Netherlands.

This research received funding from a Marie Curie Career Integration Grant [630992] and the Laboratory of Biophysics.

Financial support from the Laboratory of Biophysics for printing this thesis is gratefully acknowledged.

Printed by: Digiforce | | ProefschriftMaken



## Propositions

1. Diffusion-based confocal detection of single molecules can be fully replaced by the use of nanofluidic devices implemented on a TIRF microscope.  
(this thesis)
2. Human DNA Polymerase beta adopts a partially closed fingers conformation to check for nucleotide compatibility with the DNA template.  
(this thesis)
3. A program to develop good teaching skills should be an obligatory part of the PhD program at our university.
4. Raw data is only as valuable as the (wo)manpower available to analyse it.
5. The right to organize a referendum is an essential part of a democracy.
6. The abundant publicity for the dangers of climate change does not contribute to establishing a greener society.

Propositions belonging to the thesis, entitled

“DNA polymerases at work: single-molecule observations of DNA synthesis in real time”

Carel Fijen

Wageningen, 21 March 2018