Chapter 4: Immobilization of proteins for single-molecule studies

4.1 Introduction

Single-molecule fluorescence measurements are generally performed using either diffusing molecules or molecules immobilized on a surface. While diffusion experiments are relatively straightforward to perform, they do not permit the observation of individual molecules for long periods of time. In most cases, extended time trajectories of molecules are obtained by visualising immobilized molecules using TIRF microscopy (134).

Numerous immobilization strategies have been developed to study biological systems using TIRF microscopy (155–157). The most straightforward approach that has proved useful in studies of DNA-binding proteins has been to immobilize substrate DNAs directly to a surface and then add the DNA-binding protein of interest to the observation chamber (158, 159). However, in some instances this approach can be difficult to implement. For example, our group has experienced considerable difficulties in immobilizing RNAP-DNA complexes using this approach due to the apparent instability of complexes once deposited onto the surface. In our experience, the vast majority of RNAP-DNA complexes immobilized in this way dissociate rapidly and leave DNA alone attached to the surface. In order to circumvent this problem, we explored whether alternative immobilization strategies could improve our yield of immobilized complexes. In this chapter, we develop protocols to immobilize proteins using a vesicle-encapsulation approach, and an antibody immobilization approach.

The ultimate goal of the method development work described in this chapter was to establish working protocols for the vesicle and antibody (Ab) immobilization techniques that could go on to be used to immobilize RNAP-DNA complexes for single-molecule investigations. To do this, we optimised our immobilization protocols using DNA Polymerase I Klenow Fragment (KF) as a model system. We took this approach because the conformational dynamics of KF has been extensively characterised in

solution, and therefore provides a benchmark against which to compare results obtained using our new immobilization strategies. Accordingly, we now provide a brief introduction into KF conformational dynamics.

Subtilisin enzymatically cleaves DNA polymerase I (Pol) to produce KF, which retains DNA polymerisation activity (160). The molecular structures of Pol, Pol-DNA binary complex, and Pol-DNA-deoxyribonucleotide (dNTP) ternary complex have all been solved (161–164), and have revealed that KF resembles a half-open right hand with a cleft formed by the fingers, thumb, and palm subdomains (Fig. 4.1). These structural studies led to initial suggestions that the fingers subdomain closes over the polymerase active site when the binary complex binds to cognate dNTP. Our group has since demonstrated the occupancies of the fingers-open and fingers-closed conformations under a variety of conditions in solution by performing single-molecule experiments using KF labelled at the fingers and thumb subdomains with donor and acceptor dyes, respectively (139, 140).

In this chapter, we demonstrate that immobilizing KF using the vesicle and antibody techniques does not influence the occupancies of the fingers-open and fingers-closed conformations. Furthermore, we show that these immobilization strategies allow us to observe KF dynamics directly.

4.2 Materials and Methods

4.2.1 DNA and protein preparation

DNA substrate was synthesised and HPLC purified by IBA Life Sciences (Germany). KF was expressed, purified, and fluorescently labelled by researchers in the Grindley laboratory (Yale University, USA) as described in (139, 140).

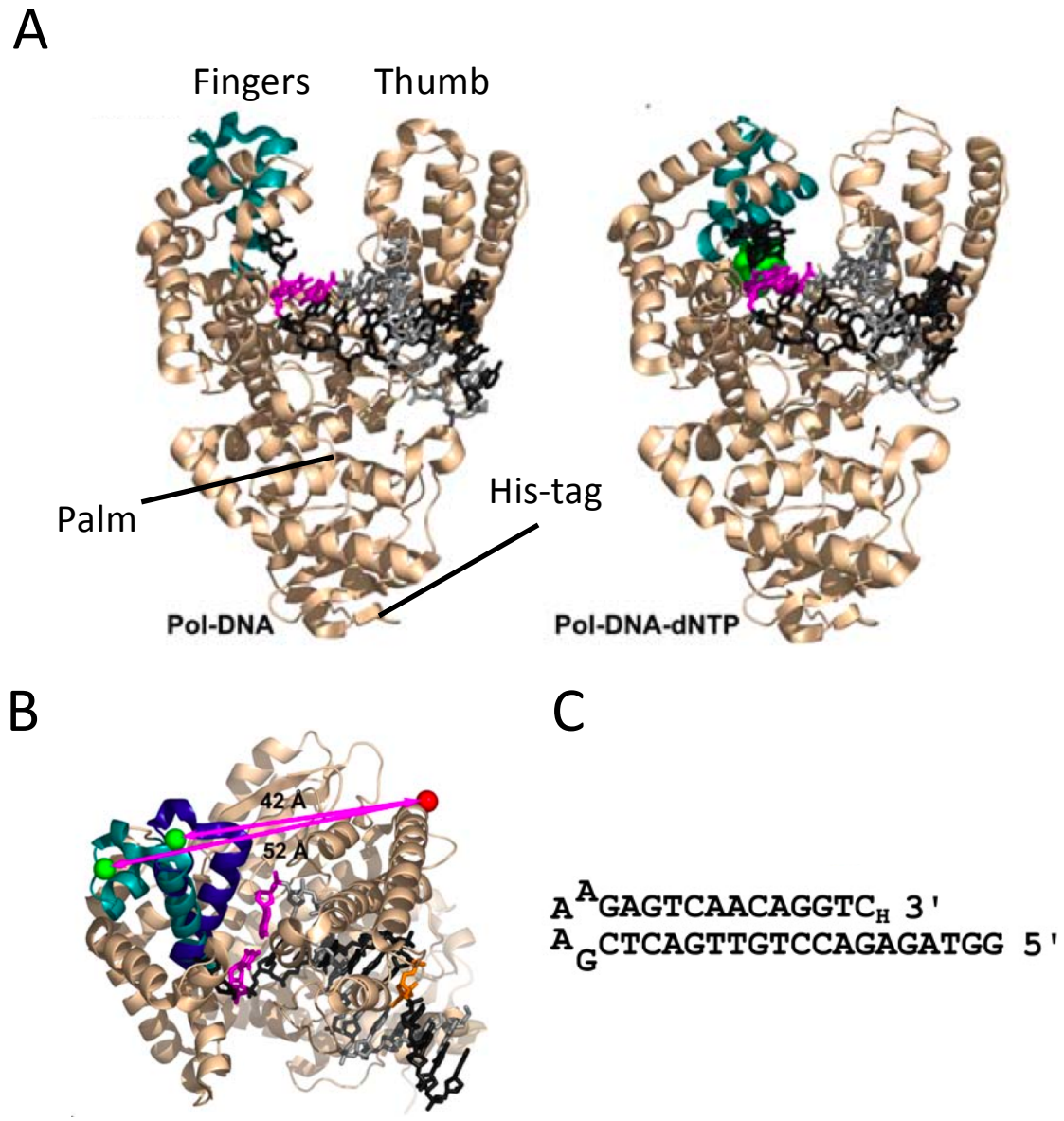

Figure 4.1. Open and closed conformations of the KF fingers subdomain. (A) Open (Pol-DNA binary complex; left) and closed (Pol-DNA-dNTP ternary complex; right) conformations of Pol (PDB files 1L3U and 1LV5, respectively). The fingers subdomain is shown in light blue. DNA is shown as a stick model. The terminal base is in magenta, and the complementary dNTP is in green. (B) Superposition of the open and closed fingers conformations viewed from the top. The fingers subdomain is in light blue in the open conformation, and in dark blue in the closed conformation. The donor and acceptor fluorophore labelling sites are shown in green and red, respectively. (C) KF substrate DNA. The dideoxy nucleotide (3'-H) prevents dNTPs from being incorporated into the 3'end. Figures from (139).

4.2.2 Vesicle encapsulation of KF

The extrusion process, as described in (165), was used to produce unilamellar vesicles. Here, 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(cap biotinyl) (16:0 biotinyl cap PE) was mixed with egg L-α-phosphatidylcholine (egg PC) at a 2:100 molar ratio in chloroform. The lipids were sourced from Avanti Polar Lipids, Inc. The chloroform was evaporated using a stream of nitrogen gas and the resulting 2.5 mg lipid film left in a desiccator overnight to remove residual chloroform. The lipid film was hydrated with 200 µL of the hydration buffer containing the fluorescent molecule of interest at a concentration of 200 nM. For KF encapsulation, the hydration buffer consisted of 50 mM Tris (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, 5% (vol/vol) glycerol, 100 µg/mL bovine serum albumin (BSA), and 2 mM Trolox (UV treated). The concentration of the fluorescent molecule of interest was chosen to ensure the encapsulation of, on average, less than one molecule per vesicle ($\lambda = 0.5$) according to:

$$\lambda = \left(\frac{4\pi R^3}{3}\right) \Big/ \left(\frac{10^{24}}{NC}\right)$$

Where λ is the average number of molecules per vesicle, R is the radius of the vesicle in µm, N is Avogadro's constant, and C is the concentration needed, in nM, of the molecule to be encapsulated (165). This was followed by 41 extrusion passes through a polycarbonate membrane with 200 nm diameter pores, using an Avanti Mini Extruder according to the manufacturer's instructions, which generated 200 nm diameter vesicles that were stored at 4°C and used on the same day.

Unliganded KF was encapsulated as described above. KF-DNA binary complex was first formed *in vitro* by incubating 200 nM KF with 3 µM template-primer (TP) DNA (see Fig. 4.1C for DNA sequence). The complex was then encapsulated as described above. The concentration of the TP DNA was expected to give, on average, 7.5 molecules per vesicle, which was expected to maximise the likelihood that most vesicles that contained KF in their interior also contained at least one DNA molecule. KF-DNA-dNTP ternary complex was encapsulated as described in section 4.2.4.

The vesicle-encapsulation protocol resulted in a mixture of vesicles containing the molecule of interest, and unencapsulated molecules outside of the vesicles. We separated vesicles from unencapsulated molecules using size-exclusion chromatography. Here, a glass column was loaded with ~25 mL of Sepharose 4B agarose slurry (Sigma) that was allowed to settle by gravity at room temperature for at least 5 h resulting in an ~20 mL resin. The resin was washed with ~100 mL of Milli-Q water and equilibrated by a passage of at least 80 mL of the hydration buffer. The vesicle sample was then added to the column and eluted using gravity flow. The eluate was collected in 10-drop fractions (~500 µL), and the resin washed after use with Milli-Q water and stored in 20% ethanol at 4°C. Fractions containing vesicles were easy to identify by their milky appearance.

4.2.3 Vesicle immobilization and imaging

For all immobilization approaches, we passivated the imaging surface as follows: glass coverslips (25x65 nm, thickness number 1, VWR) were burned in a furnace, heated at a rate of 90 K/hr, held at 500°C for 1 hour, and cooled at 50 K/hr. Burned coverslip surfaces were then incubated with a 1:50 mixture of Vectabond (Vector Labs) to acetone (vol/vol) for 5 mins. Surfaces were subsequently washed with MilliQ water and dried using a stream of nitrogen gas. Silicone gaskets (Grace Bio-labs) were placed on each dried slide and each gasket incubated with 30 µL of 200 mg/mL NHS-PEG (PEG-SVA MW 5000 Da, Laysan Bio) and 2.5 mg/mL NHS-biotin-PEG (PEG-SVA MW 5000 Da, Laysan Bio) in 50 mM MOPS (pH 7.5) buffer. After a three-hour incubation, the wells were washed with the relevant imaging buffer (without the oxygen scavenging system).

The vesicle fractions were diluted to a concentration of ~100 pM prior to immobilization. The sample was then added to a NeutrAvidin-coated glass slide surface and binding monitored using TIRF microscopy. Once ~60 fluorescent molecules were bound, excess vesicles were washed away using the hydration buffer. The immobilized vesicles were then imaged using TIRF microscopy in an imaging buffer that consisted of the hydration buffer plus an oxygen scavenging system consisting of 1 mg/mL glucose

oxidase, 40 μg/mL catalase, 1.4% w/v D-glucose. Laser powers used in experiments are provided in relevant figure captions.

4.2.4 Porous vesicle formation

The protocol used to generate porous vesicles was based on that described in (166). Unilamellar vesicles were formed and immobilized as described above. We then incubated the immobilized vesicles with 30 μM of the beta-pore-forming toxin α-hemolysin (a gift from Hagen Bayley, Oxford University, UK) in hydration buffer for 15 mins. α-hemolysin molecules not incorporated into the lipid bilayer of immobilized vesicles were washed away using hydration buffer. Under these conditions, >95% of immobilized vesicles containing single fluorescent molecules exhibited the expected FRET change when 1 mM cognate dNTP was added. No further improvements were achieved using higher concentrations of α-hemolysin. In fact, the number of vesicles in the field-of-view decreased considerably at higher α-hemolysin concentrations, presumably because too many pores were formed in the vesicle membrane, which compromised the structural integrity of the vesicles.

4.2.5 Antibody-immobilization of KF and imaging

A 10 nM solution of biotinylated penta-His antibody (Qiagen) was incubated for 10 min on a surface coated with NeutrAvidin (Thermo Scientific) to allow binding. After the incubation period, excess unbound antibodies were washed away, and a 500 pM solution of KF incubated on the surface for <5 min followed by washing to remove excess unbound KF. Data acquisition was performed in the same imaging buffer described in section 4.2.3. Where indicated, 3 μM TP DNA was added to form KF-DNA binary complex, 3 μM TP DNA and 1 mM dTTP was added to form KF-DNA-dNTP ternary complex.

4.2.6 Instrumentation

Single-molecule TIRF experiments were performed on a custom built objective-type TIRF microscope. A green (532 nm; Cobolt Samba) and a red (635 nm; Cube Coherent) laser were combined using a dichroic mirror and then coupled into a fibre optic cable. The fibre output was focussed into the back focal plane of the objective (100× oil-immersion, numerical aperture 1.4, Olympus) and displaced off the optical axis such that the laser light was incident at the slide-solution interface at an angle greater than the critical angle, creating an evanescent wave. ALEX (150) was implemented by directly modulating the two lasers. Data was acquired at different exposure times depending on the experiment; these are given in relevant figure captions. For most experiments, two sets of incident laser powers were used: 1) 8 mW (320 W/cm^2) and 4 mW (160 W/cm^2) for the 532 nm and 635 nm lasers, respectively, and 2) 1 mW (40 W/cm^2) and 0.5 mW (20 W/cm^2) for the 532 nm and 635 nm lasers, respectively. All laser power measurements were done in continuous-wave mode and evanescent wave intensity was not meaured directly. Fluorescence emission was collected from the objective and separated from the excitation light by a dichroic mirror (545 nm/650 nm, Semrock) and cleanup filters (545 nm LP, Chroma; and 633/25 nm notch filter, Semrock). The emission signal was focussed on a slit to crop the image and then spectrally separated, using a dichroic (630 nm DRLP; Omega), into donor and acceptor emission channels, which were focused side by side onto an EMCCD camera (Andor iXon 897). The EMCDD was set to an EM gain of 300, which corresponds to an approximate real gain of 4.55 counts per photon. With our incident laser intensities and EM gain settings single molecules displayed a brightness of between 700 and 1000 counts. A continuous reflective interface feedback focus system (CRIFF, Applied Science Instrumentation) was used to ensure focus stability during experiments. The CRIFF reduces focal drift by using the totally internally reflected beam as a reference; the reflected beam is detected by a position-sensitive detector, which monitors changes in distance between the specimen and objective lens; when there is focal drift, the CRIFF alters the focus of the microscope.

All data acquisition was performed using Alternating Laser Excitation (ALEX) (150). Here, the immobilized sample was exposed to 532 nm laser light, and emitted photons imaged on an EMCCD camera and collected for a set time period. At the end of this period, the 532 nm laser excitation was stopped, at which point the sample was exposed to 635 nm laser excitation and the EMCDD camera triggered to begin a new exposure interval. At the end of another time period, the 635nm laser excitation was stopped. The duration of the exposure intervals was the same for both 532 nm and 635 nm laser excitations. At the end of both intervals, one alternating laser excitation period is completed thus defining the temporal resolution of the data. This excitation cycle was repeated for the duration of the acquisition. In our setup, the alternation of the excitation lasers is coupled with each camera exposure interval. This external triggering is controlled by custom-built software written in LabView by Justin Pinkney.

4.2.7 Data analysis and visualisation

The fluorescence intensities of fluorescent spots were extracted from the EMCCD images using image analysis software called *twoTone* (167). Fluorescent spots were detected in an automated fashion. Here, a high pass spatial filter was used to remove noise, followed by convolution with a Gaussian mask similar in size to our fluorescent spots, followed by the identification of above-threshold pixels. The centroids of the fluorescent spots were then localized to high precision by fitting the detected molecules with an elliptical 2D Gaussian using ordinary-least-squares minimization (167). 55% of molecules passed selection criteria and were taken forward for further analysis.

The apparent FRET efficiency (E*) and stoichiometry (S) values were calculated from the fluorescence emission as outlined in chapter 3. Vesicle encapsulation is a stochastic process and therefore results in vesicles containing between 0 and 4 fluorescent molecules. Because we were only interested in vesicles containing one fluorescent molecule, we monitored the fluorescence intensities until photobleaching occurred. Only fluorescent spots exhibiting single photobleaching steps were selected

for further analysis. FRET efficiency (E*) histograms and time trajectories were visualised using *Seneca* software built by Kristofer Gryte.

By eye, we manually classified intensity time trajectories of individual molecules into groups. Static open fingers signals were defined as time trajectories showing a stable E*~0.45 signal with no fluctuations to higher FRET states during our observation periods (~50 sec for 100-200 ms temporal resolution experiments; ~5 sec for 5-10 ms temporal resolution experiments). Static closed fingers signals exhibited a stable E*~0.65 signal with no fluctuations to lower FRET states during our observation periods. Dynamic signals were defined as time trajectories showing any anti-correlated intensity changes in the donor and acceptor fluorescence during our observation periods. The tendency for individual dynamic molecules to preferentially sample either the open or closed fingers conformations was determined from the E* histogram of individual molecules. Fluorescent spots exhibiting confounding photophysical fluctuations were excluded from analysis.

4.3 Results and discussion

4.3.1 Immobilization of vesicle-encapsulated KF and comparison with solution data

An immobilization strategy involving the encapsulation of molecules inside of lipid vesicles, which are subsequently immobilized on a surface was recently developed (165, 168, 169). If the vesicles carry fluorescent molecules in their interior, the encapsulated molecules can be monitored using TIRF microscopy (Fig. 4.2A). This section describes how we applied this technique to immobilize KF, and how we assessed whether the technique influenced the KF fingers conformation.

A glass slide surface was passivated with a 1:100 biotin-PEG:PEG mixture to minimise non-specific adsorption. The passivated-surface was then incubated with NeutrAvidin, which bound to the biotin moieties on the biotin-PEG. 200 nm diameter vesicles containing unliganded KF in their interior were then immobilized on the surface. In a separate experiment, vesicles containing KF-DNA binary complex were immobilized

on the surface. The vesicles consisted of a mixture of biotinylated and non-biotinylated lipid, which enabled interactions between the NeutrAvidin anchored to the surface and the biotin moieties on the vesicle exterior. Unbound vesicles were washed away and the surface then illuminated using TIRF microscopy in combination with ALEX. We acquired data at a temporal resolution of 200 ms. Control experiments in which vesicles were added to surfaces lacking NeutrAvidin showed little non-specific adsorption after a washing step (Fig. 4.2B–C).

Vesicle encapsulation is a random process that results in some vesicles containing multiple fluorescent molecules. These can be seen in TIRF images as anomalously bright point-spread-functions (PSF; Fig. 4.2B), and in fluorescence intensity time trajectories that show multiple photobleaching steps (Fig. 4.2D). We collected data over 5 fields-of-view (FOV) and selected fluorescent spots for analysis according to the following criteria: 1) bleached in a single step. 2) No donor photoblinking where photoblinking is defined as a minimum of four frames when fluorescent intensity reaches zero (although fluctuations were permitted). 3) No acceptor photoblinking where photoblinking is defined as a minimum of four frames when fluorescent intensity reaches zero (we allowed for different ATTO647N fluorescence states (170)). 4) Fluorescence intensities within a limited range between 200 and 285 photons per frame. We then plotted data from all frames of selected molecules into FRET efficiency (E*) histograms[2].

The E* histograms for vesicle-encapsulated KF and vesicle-encapsulated KF-DNA binary complex are shown in Fig. 4.3B next to solution confocal data (Fig. 4.3A) for comparison purposes. There are only slight differences between the solution-based and surface based data sets. Both show two populations centred at E*~0.45 and E*~0.65, which we assign to the fingers open and closed conformations (139, 140). The relative proportions of the FRET populations are similar between the two data sets. We therefore conclude that vesicle immobilization has little, if any, effect on the fingers conformation. The histograms also show that vesicle encapsulation does not interfere with the DNA binding activity of KF.

[2] The first five frames were excluded from the histogram to allow for laser stabilization.

4.3.2 KF fingers opening and closing

The time trajectory data for unliganded KF revealed transitions between E*=0.45 and E*=0.65[3]. This observation confirms earlier proposals that the KF fingers subdomain undergoes opening and closing transitions in the unliganded state (139). Interestingly, we observed considerable inter-molecular heterogeneity in the immobilized unliganded KF population. KF behaviours were divided into the following groups: i) predominately open (30% of molecules), in which there were no obvious transitions to the closed conformation (Fig. 4.4A, top left); ii) predominately closed (15% of molecules), in which transitions to the open conformation were very fast and difficult to observe directly with our temporal resolution of 100 ms (Fig. 4.4A, top right); and iii) dynamic (55% of molecules), in which the KF fingers subdomain readily underwent opening and closing transitions (Fig. 4.4A, bottom panels). KF-DNA binary complexes were also found to exhibit heterogeneous behaviour; 85% of molecules exhibited fingers conformational dynamics in which occupancy of the open conformation was favoured (Fig. 4.4B), and only 15% of molecules showed a stable fingers-open conformation. However, it should be noted that since our FRET assay does not permit us to monitor DNA binding directly, we cannot exclude the possibility that the FRET transitions for the KF-DNA complex are due to short periods when KF has dissociated from DNA while inside the vesicle.

Our results demonstrate that the vesicle immobilization approach can be used to immobilize KF, and that it does not affect the ability of the protein to bind to DNA. Furthermore, individual time trajectories of vesicle-immobilized KF reveal intriguing behaviours that merit further investigation. However, in the context of this method development chapter, we conclude that the vesicle immobilization technique has sufficient potential to be applied to RNAP-DNA complexes. We note however, that in order to monitor the transcription reaction inside of vesicles in real-time, we would need to deliver substrate NTPs to the vesicle interior. The development of a technique that would allow us to do this is described in the next section.

---

[3] To observe clear FRET transitions for unliganded KF, we performed additional experiments at a temporal resolution of 100 ms. A temporal resolution of 200 ms was sufficient to observe transitions for KF-DNA binary complex.

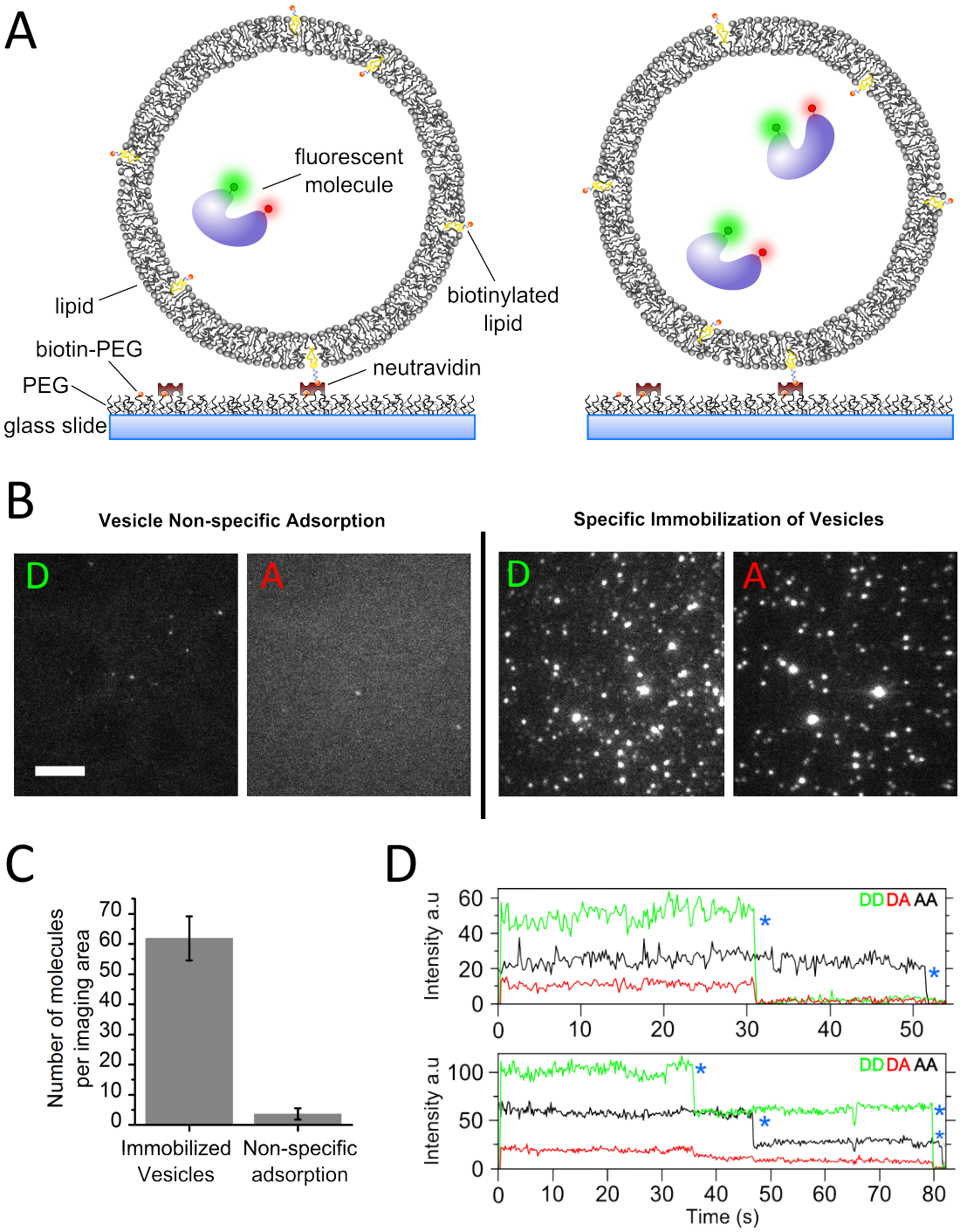

Figure 4.2. Vesicle immobilization of KF. (A) Labelled schematic of immobilized vesicles. Left-hand figure shows one molecule encapsulated; right-hand figure shows two molecules encapsulated. (B) Example images of TIRF fields-of-view. Left-hand figure shows vesicle non-specific sticking to a PEG-coated surface; right-hand figure shows specific immobilization via biotin-NeutrAvidin interactions. Donor channels are labelled with a green 'D'; acceptor channels are labelled with a red 'A'. Scale bar is 5 µm. (C) Bar graph with average number of fluorescent molecules per image. Error bars show standard deviation of the mean across 5 images. (D) Example time trajectories of bleaching events. The top panel shows a single bleaching event for each fluorophore. The bottom panel shows two bleaching events for each fluorophore. Bleaching steps are marked with blue asterisks.

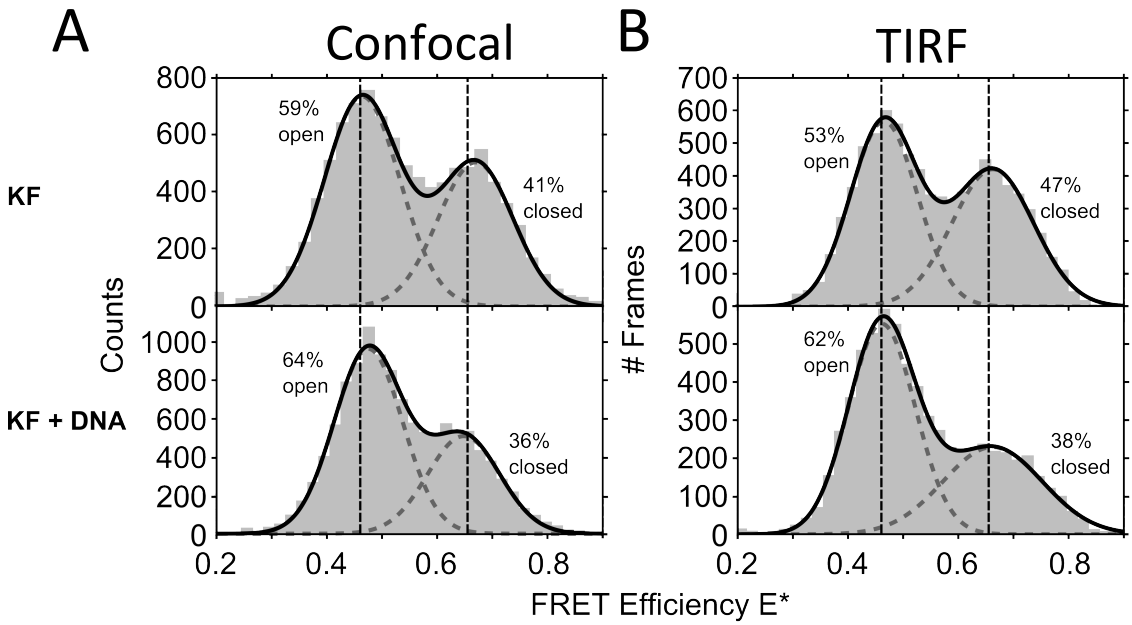

Figure 4.3. Vesicle-immobilized KF FRET histograms. (A) Stacked E* histograms of KF in solution. (B) Stacked FRET histograms of immobilized vesicle-encapsulated KF. Temporal resolution is 200 ms, excitation powers are 1 mW for the 532 nm laser, and 0.5 mW for the 635 nm laser. In both stacked histograms, the top panel is unliganded KF and the bottom panel is KF-DNA binary complex. For confocal data, 6806 bursts contributed to the top panel hitogram, 7313 bursts contributed to the bottom histogram. For TIRF data, 112 molecules contributed to the top panel histogram, 98 molecules contributed to the bottom panel histogram. All histograms are fit with a 2 Gaussian profile.

4.3.3 Permeabilising the vesicle lipid membrane

Controlling the availability of NTPs to the vesicle interior is not straightforward using the vesicle-immobilization approach since the vesicle lipid bilayer acts as an impermeable barrier to small charged molecules. To enable the delivery of small molecules to the vesicle interior, we permeabilised the vesicle lipid membranes using a bacterial pore-forming toxin, as described in (166). This permeabilisation method involved incubating vesicles with α-hemolysin from the bacterium *Staphylococcus aureus*. α-hemolysin monomers are soluble in solution and spontaneously insert into lipid membranes where they form heptameric assemblies that comprise a membrane channel or pore (Fig. 4.5) (171). The narrowest region of the pore is 1.8 nm in diameter, which is sufficiently wide to permit the passage of small molecules including organic molecules less than 3kD in molecular weight (172).

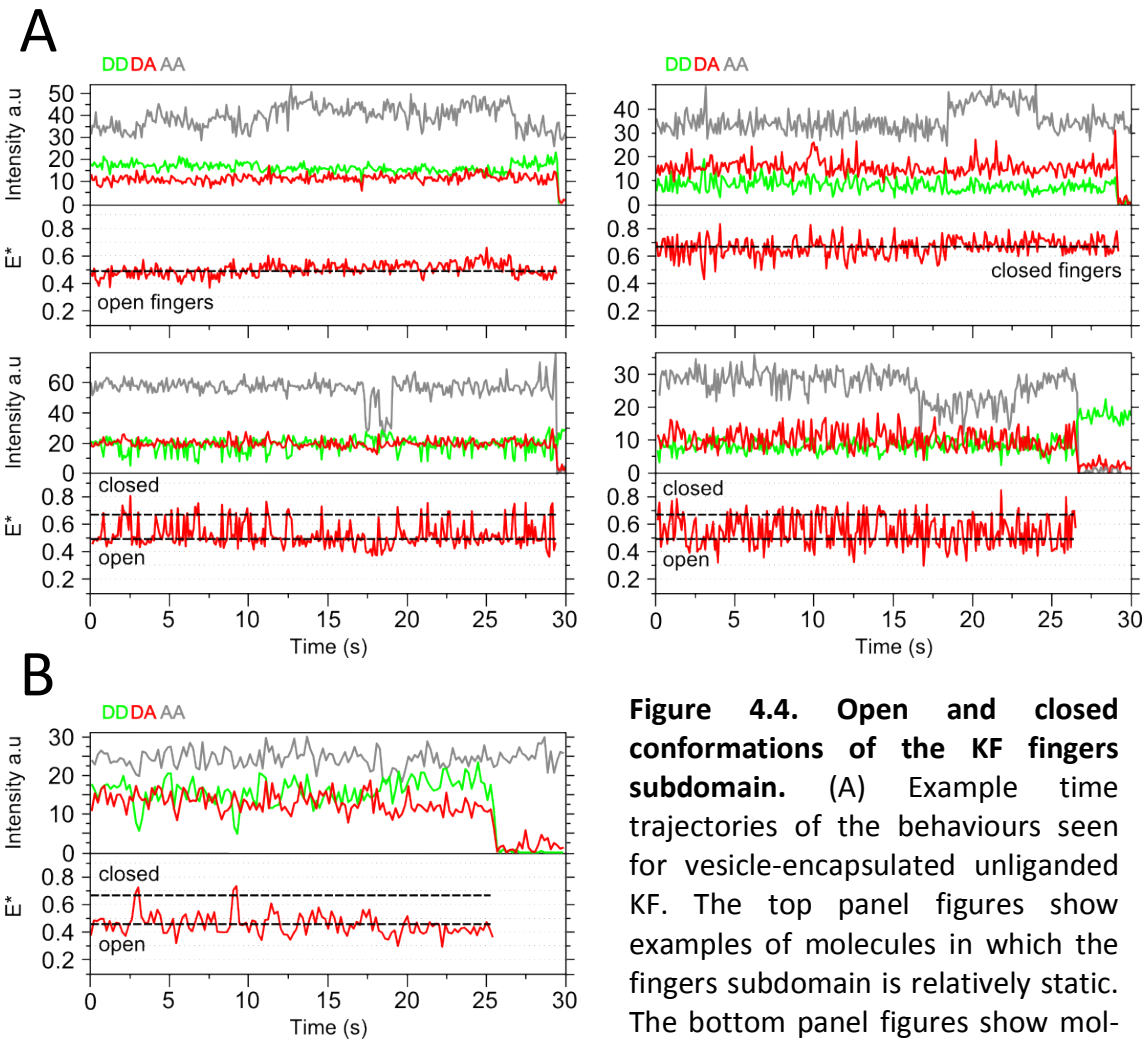

Figure 4.4. Open and closed conformations of the KF fingers subdomain. (A) Example time trajectories of the behaviours seen for vesicle-encapsulated unliganded KF. The top panel figures show examples of molecules in which the fingers subdomain is relatively static. The bottom panel figures show mol--ecules that undergo clear fingers opening and closing transitions. Temporal resolution is 100 ms; 532 nm excitation is 1.5 mW, 635 nm excitation is 1 mW (B) Example time trajectory of the most common behaviour observed for vesicle-encapsulated KF-DNA binary complex. Temporal resolution is 200 ms, and excitation powers are the same as in Fig. 4.3. All figures show a bleaching event. Approximate FRET values corresponding to the open and closed fingers conformations are marked with dashed lines. The top panels of all figures show donor under donor excitation (DD; green line), the acceptor under acceptor excitation (AA; grey line), and the acceptor under donor excitation (DA; red line). The E* trajectory is shown as a red line. This colour scheme is used throughout this thesis.

First, we encapsulated the KF-DNA binary complex as usual using a DNA substrate containing a dideoxy 3'-end that renders it non-extendable (Fig. 4.1C). This permits the binding of the complementary dNTP to the binary complex, which in turn induces fingers closure, but prevents catalysis (139, 140). We immobilized the vesicles containing binary complex and incubated them with α-hemolysin monomers (see *Materials and Methods*). After a 15 min incubation period, we washed away

unincorporated α-hemolysin and added dTTP to the observation chamber; dTTP is the deoxynucleotide complementary to the substrate DNA. We selected molecules to generate an E* histogram as described in section 4.3.1. Again, we compared our surface-based TIRF data with solution confocal data (Fig. 4.6A). And again, we could see only small differences between the two data sets. Both show that upon binding to dTTP the fingers subdomain closes, resulting in a FRET histogram with a large population centred at E*~0.65, and only a small population centred at E*~0.45.

Additionally, our time trajectory data showed that ~10% of molecules exhibited a stable fingers-closed conformation with no obvious structural transitions to the open state, while the remainder showed dynamic behaviour with the equilibrium lying very much towards the closed conformation (Fig. 4.6B)[4]. Our results clearly demonstrate our ability to form porous vesicles *in situ* and deliver small molecules to the vesicle interior, which is absolutely necessary for studying polymerisation reactions using the vesicle-immobilization technique.

During the optimization of the vesicle immobilization protocol, it became apparent that the technique was time-consuming, technically not straightforward, and inefficient (many molecules were not encapsulated). This limited our ability to increase the temporal resolution beyond 100 ms and obtain sufficiently high yields of immobilized complexes that could be taken forward for further analysis. We therefore decided to explore an alternative immobilization strategy involving the use of antibodies that specifically recognise the His-tag of recombinant proteins.

4.3.4 Antibody-immobilization of KF

A protein immobilization technique involving the use of antibodies was recently reported (137). Here, an antibody (Ab) is tethered to a surface and then incubated with the molecule of interest to which it binds (Fig. 4.7A). This interaction stably anchors the molecule of interest on the surface, and has so far been used to monitor the

[4] It is possible that our stable molecules exhibit fingers opening transitions that occur faster than our temporal resolution of 200 ms.

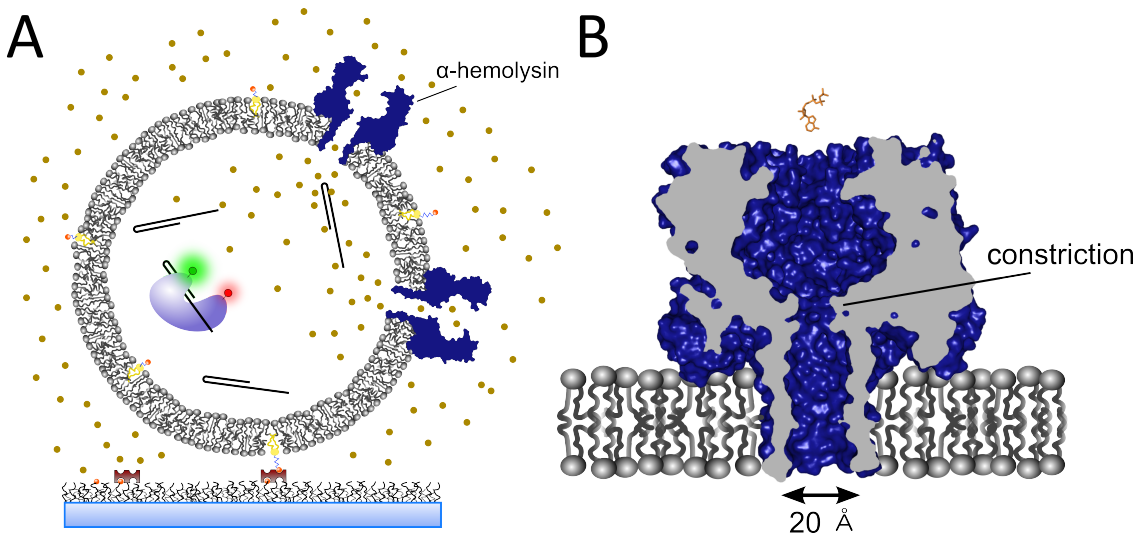

Figure 4.5. Schematic illustrations of porous vesicles. (A) Cartoon of an immobilized porous vesicle. The vesicle contains one fluorescent KF molecule and multiple DNA molecules. The α-hemolysin pores allow diffusion of small molecules across the membrane. dNTPs are represented as gold circles. (B) Cutaway structure of alpha-hemolysin showing the pore internal channel (PDB 7AHL) and an ATP molecule in orange for scale comparison taken from PDB 1TWA.

composition of oligomeric assemblies, to monitor biochemical activity, and to investigate conformational heterogeneity in an ADP-binding protein (137, 173, 174). This section describes how we used this Ab-immobilization technique to immobilize unliganded KF, KF-DNA binary complex, and KF-DNA-dNTP ternary complex. We compare our results with earlier vesicle-immobilization results.

The immobilization protocol we developed was modified slightly from (137) and is described in *Materials and Methods*. Briefly, we passivated a glass slide surface with a 1:100 biotin-PEG:PEG mixture to minimise non-specific adsorption, and then incubated the surface with NeutrAvidin. Unbound NeutrAvidin was washed away and biotinylated penta-His Ab incubated with the surface-bound NeutrAvidin for 10 mins. Unbound Ab was then washed away and 500 pM of KF with an N-terminal His-tag incubated on the surface until ~60 molecules were immobilized (His-tag location is shown in Fig. 4.1A). This generally took less than 5 mins. Control experiments in which KF was added to a NeutrAvidin coated surface without Ab showed minimal non-specific adsorption of KF even after a 15 min incubation period (Fig. 4.7B). This showed that the vast majority of the immobilized fluorescent molecules detected were due to the specific immobilization

of KF via the penta-His Ab. To immobilize KF-DNA binary complex using this technique, we simply immobilized KF as described above, and then added substrate DNA to the imaging buffer at a concentration of 3 µM. To form the KF-DNA-dNTP ternary complex, we added 1 mM of dTTP to the immobilized KF-DNA binary complex. Experiments were performed at a temporal resolution of 10 ms to try and capture clear KF fingers conformational changes[5]. Data from 40 fields-of-view were used to select molecules and generate E* histograms as described in section 4.3.1 (Fig. 4.8B).

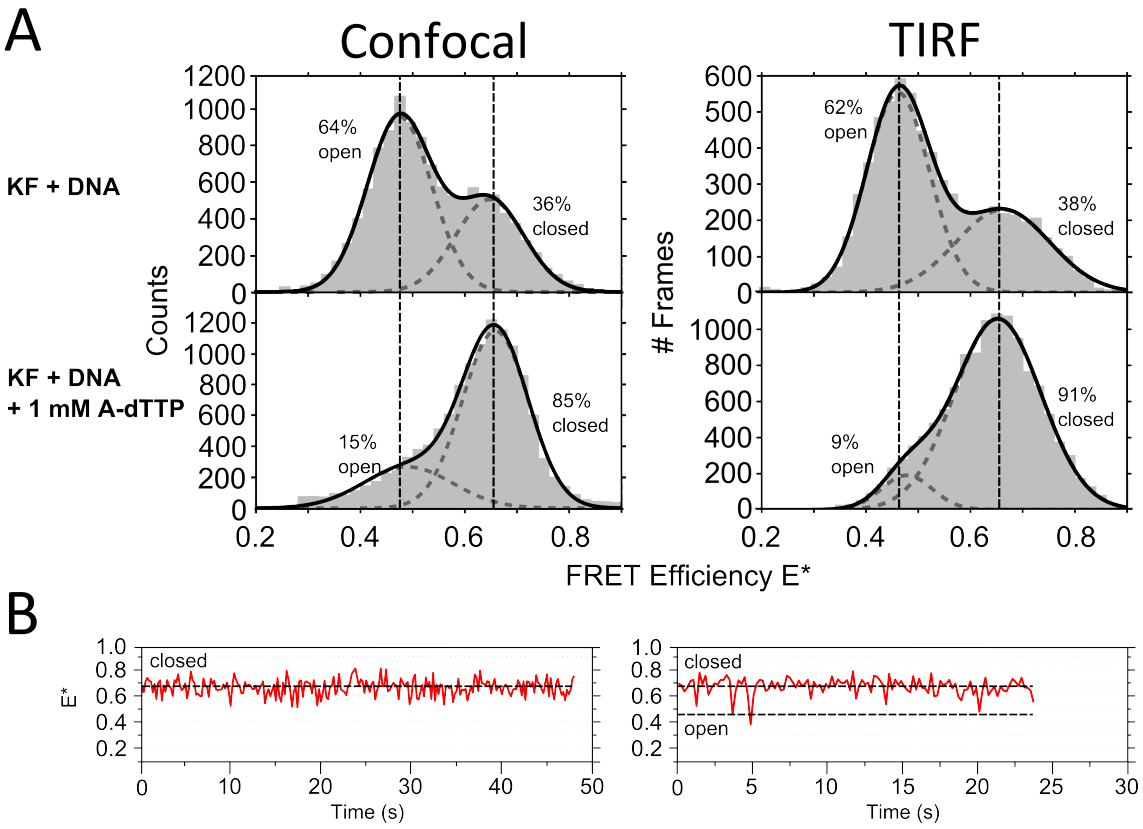

Figure 4.6. KF-DNA-dNTP ternary complex formation inside of porous vesicles. (A) Stacked FRET histograms of KF-DNA binary complex (top) and of KF-DNA-dNTP ternary complex (bottom). The left-hand figure shows solution confocal data, the right-hand figure shows surface TIRF data of vesicle-encapsulated molecules. The vesicle ternary complex is formed *in situ* by permeabilising the immobilized vesicles containing KF-DNA binary complex and adding 1 mM dTTP to the observation chamber. For confocal data, 7313 bursts contributed to the top histogram and 7217 contributed to the bottom histogram. For the TIRF data, 105 molecules contributed to the top histogram and 119 molecules to the bottom histogram. (B) Example time trajectories of vesicle-immobilized KF-DNA-dNTP ternary complex. Temporal resolution is 200 ms; excitation powers are the same as in Fig. 4.3.

[5] Experiments at a temporal resolution of 5 ms were required to capture the dynamics of unliganded KF.

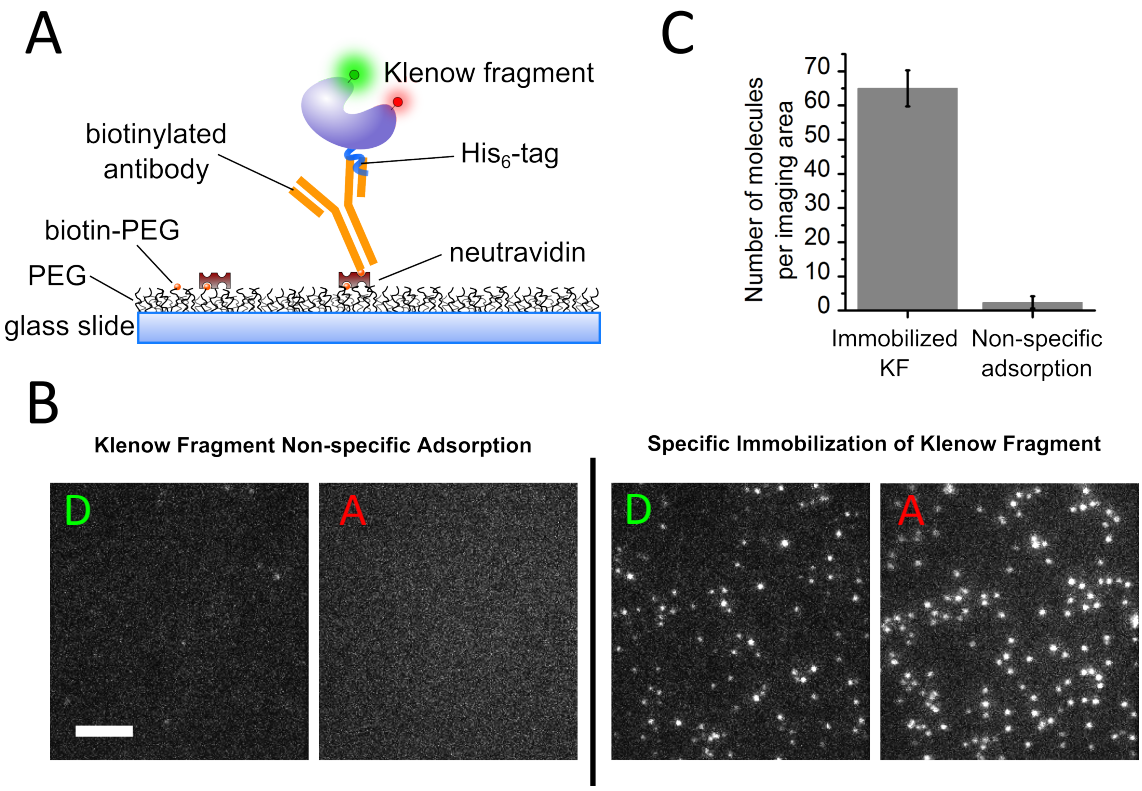

Figure 4.7. Antibody-immobilization of KF. (A) Schematic of the Ab-immobilization approach. (B) Example images of TIRF fields-of-view. Left hand panel shows non-specific sticking of KF to the PEG-coated surface, right hand panel shows specific immobilization of KF via NeutrAvidin-biotin-antibody interactions. The donor channels are labelled with a green 'D', the acceptor channels are labelled with a red 'A'. Scale bar is 5 µm. (C) Bar graph with the average number of fluorescent molecules per imaging area. Error bars show standard deviation of the mean across 5 images.

We compared TIRF data with solution data (Fig. 4.8A)[6], and found little difference between the two datasets, which indicates that the Ab-immobilization strategy does not interfere with the KF fingers conformation or the ability of KF to bind to DNA and dNTP. Our time trajectory data revealed approximately the same behaviour as that observed in our vesicle-immobilization experiments. The unliganded KF fingers subdomain conformation behaviour was classified into the following groups: i) predominately open (20% of molecules) (Fig. 4.9A, left); ii) predominately closed (20% of molecules) (Fig. 4.9A, right); and iii) dynamic (60% of molecules) (Fig. 4.9B). Interestingly, dynamic molecules displayed different tendencies to occupy the open and

[6] For Ab experiments, we used an alternative KF preparation to that used for vesicle encapsulation experiments. The two preparations differ slightly in the ratio of the low and high E* populations due to different labelling strategies (see (139, 140)). The general trend of the E* histograms remains unchanged.

closed conformations, which can be seen in the example time trajectories and E* histograms shown in Fig. 4.9B. The majority of binary complex molecules (80%) exhibited fingers conformational dynamics (Fig. 4.9C), while only 20% displayed a stable fingers-open conformation (E*=0.45) with no obvious transitions to higher FRET states. Only 25% of ternary complex molecules adopted a stable closed conformation (E*=0.65), while 75% showed dynamic behaviour in which occupancy of the closed conformation was very much favoured (Fig. 4.9D). These results demonstrate our ability to immobilize proteins and protein-DNA complexes using the penta-His Ab approach.

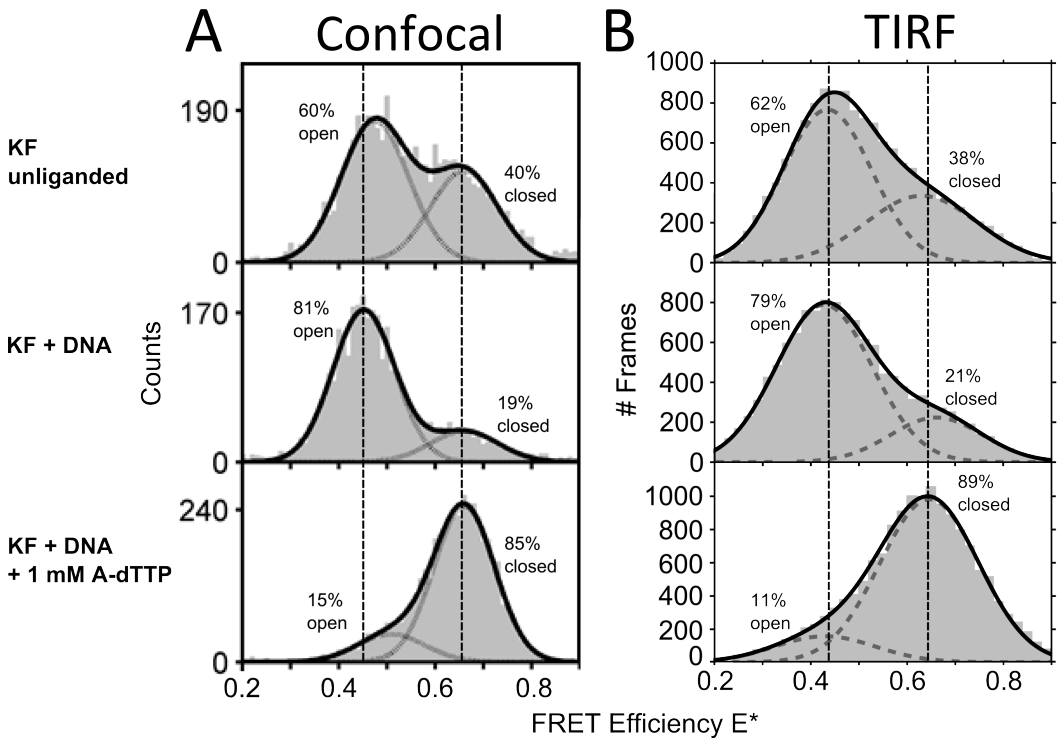

Figure 4.8. Ab-immobilized KF FRET histograms. (A) Stacked FRET histograms of Klenow fragment in solution (figure from (140)). (B) Stacked FRET histograms of Ab-immobilized KF. Data taken at a temporal resolution of 10 ms; 6 mW 532 nm excitation, 4 mW 635 nm excitation. The increased width of the Ab-TIRF data histograms compared with vesicle-TIRF data histograms (Fig. 4.3 and Fig. 4.6) is due to the increased noise at this faster temporal resolution. For TIRF data, 47 molecules contributed to the top histogram, 49 molecules contributed to the middle histogram, and 45 molecules contributed to the bottom histogram.

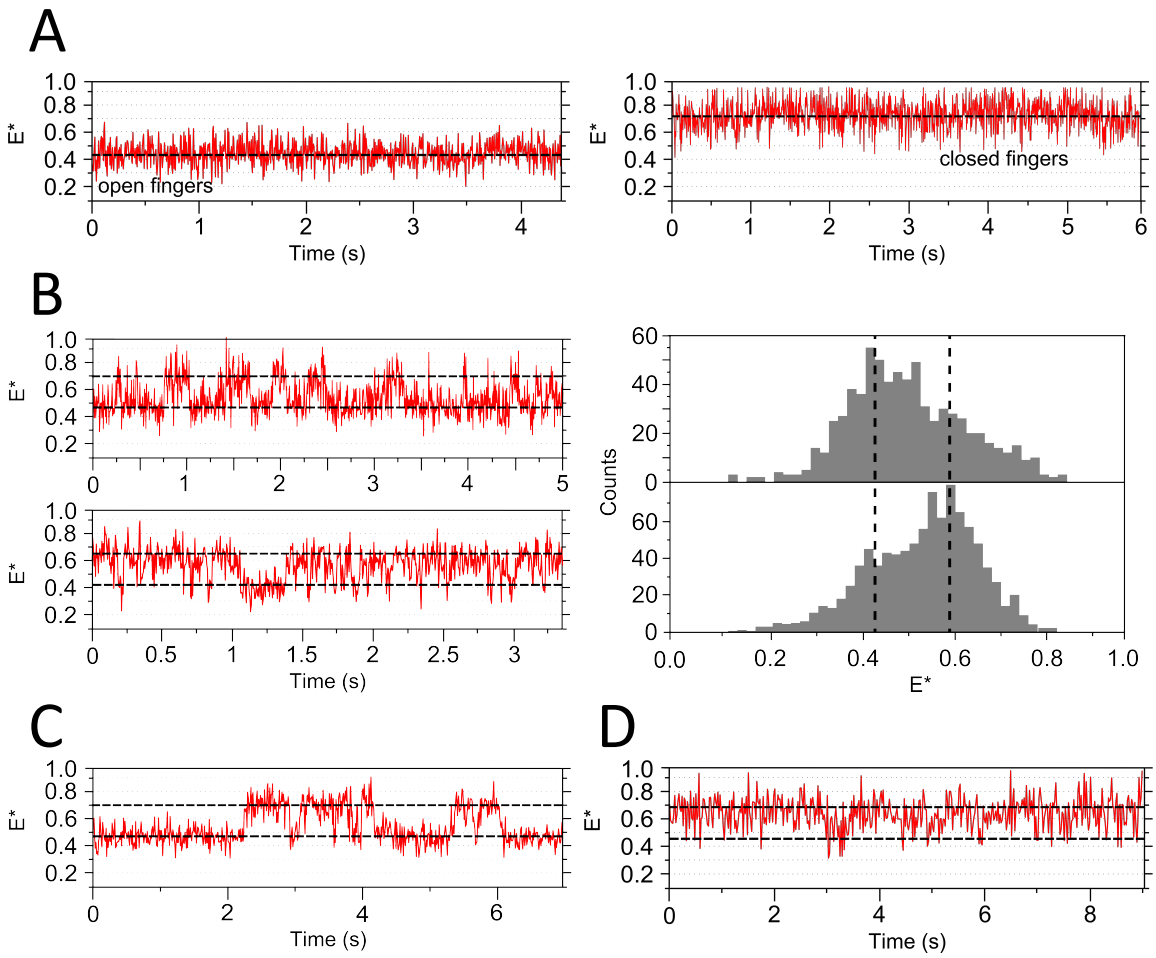

Figure 4.9. Example time trajectories of Ab-immobilized KF. (A) Unliganded KF. Left panel shows static open behaviour, right panel sows static closed behaviour. (B) Unliganded KF molecules undergoing clear fingers opening and closing transitions. The time trajectory projections for the two example molecules are shown on the right. Data shown in A and B was taken at a temporal resolution of 5 ms; 8mW 532 nm excitation, 5 mW 635 nm excitation. (C) Antibody-immobilized KF with substrate DNA in the observation chamber (KF-DNA binary complex). (D) Antibody-immobilized KF with substrate DNA and complementary dNTP (KF-DNA-dNTP ternary complex) in the observation chamber. Data shown in C and D was taken at a temporal resolution of 10 ms with the same excitation powers in Fig. 4.8. Dashed lines indicate the approximate FRET states of the open and closed fingers conformations.

4.4 Conclusions

In this chapter we presented results that demonstrate our ability to immobilize biomolecules using two independent techniques. The first technique involved the encapsulation of KF inside of vesicles that were immobilized. The second technique involved immobilizing KF using an Ab specific for protein His-tags. The two

immobilization techniques did not appear to interfere with the KF fingers subdomain conformation as judged by the similarity of surface-based E* histograms to solution-based E* histograms under a variety of conditions. Furthermore, the two techniques permitted us to obtain intriguing KF time trajectory data that showed considerable inter-molecular heterogeneity. The behaviour of unliganded KF varied to the extent that some molecules displayed a very stable open-fingers conformation, others displayed a stable closed-fingers conformation, and a third group showed considerable conformational dynamics. It is not clear what the origin of this heterogeneity is; however, the fact that it is observed using both immobilization approaches, which rely on very different principles, suggests that the heterogeneity is, at the very least, not an artefact of how the protein is immobilized. Future work will look at the possible functional significance of these observations.

Comparing the two immobilization techniques, we found that the Ab approach was more straightforward to implement and enabled the accumulation of large amounts of data at a very fast temporal resolution (up to 5 ms). The vesicle immobilization approach on the other hand was technically very demanding, and required many steps before we could even begin TIRF microscopy experiments. Furthermore, a considerable amount of data filtering time was required to eliminate vesicles containing multiple fluorescent molecules. Although the vesicle approach can be argued to mimic *in vivo* conditions by acting like a pseudo-cell, the practical difficulties in using the technique meant that in subsequent chapters the Ab-immobilization technique is used as the method of choice. We note, however, that the availability of two working protocols to immobilize transcription complexes means that we can use the vesicle immobilization technique to test the observations found by the easier-to-implement Ab immobilization technique.

4.5 Contributions

I performed all work described in this chapter except for the following contributions made by others: Kristofer Gryte (Oxford University, UK) wrote custom-

built software used to produce FRET histogram and time trajectory figures. The Grindley laboratory (Yale University, USA) provided KF samples. α-hemolysin was a gift from Hagen Bayley (Oxford University, UK).

Chapter 5: RNAP transcription bubble dynamics

5.1 Introduction

Transcription initiation begins with RNAP holoenzyme binding to promoter DNA to yield the RNAP-DNA closed complex (RP_C), which subsequently undergoes large conformational changes that trigger the melting of 12-14 base pairs (bp) of DNA around the transcription start site. This yields the catalytically competent RNAP-DNA open complex RP_O (see chapter 2 for a detailed description of transcription initiation).

Despite major advances in our understanding of the RP_O formation mechanism, a complete RP_O structure has not yet emerged. Recent investigations suggest that RP_O exhibits structural flexibility and inter-molecular heterogeneity that may explain why the complex has remained inaccessible to structural methods. For example, a solution-based single-molecule FRET (smFRET) study has produced evidence suggesting that the template (T) DNA in RP_O exhibits dynamic heterogeneity (175). Furthermore, surface-based smFRET studies of the mitochondrial and eukaryotic transcription initiation machinery have revealed large-scale DNA conformational dynamics in these initiation complexes (176, 177). While studies finding structural flexibility are suggestive, little is actually known about DNA conformational dynamics in bacterial transcription initiation complexes.

Researchers in our group have attempted to gain new insight into the conformational changes that occur during bacterial transcription initiation using single-molecule TIRF microscopy. These attempts have involved the immobilization of pre-formed RNAP-DNA initiation complexes via a biotin moiety at one end of the promoter DNA. However, progress has been slow due to a low yield of surface-immobilized RNAP-promoter DNA complexes. In our experience, RNAP tends to dissociate from the complex during the immobilization step, which leaves promoter DNA alone attached to the surface.

In chapter 4 of this thesis, we demonstrated our ability to immobilize proteins and protein-DNA complexes using two different immobilization approaches, one involving vesicles, and the other involving an antibody (Ab) specific for protein His-tags. In this chapter we use these immobilization methods to anchor RNAP-promoter DNA complexes to a surface for smFRET TIRF measurements. We perform experiments that address questions regarding transcription bubble structural heterogeneity and dynamics in RNAP-promoter DNA complexes. Our results reveal that the transcription bubble undergoes large-scale structural transitions between closed and open conformations. Our results also reveal the presence of two long-lived heparin-resistant complexes with the transcription bubble in conformations distinct from that expected for fully formed RP_O. Our results also provide insight into how σ^{70} region 3.2 (σR3.2) and the β' Jaw influence the transcription bubble conformation and dynamics.

5.2 Materials and Methods

5.2.1 DNA and protein preparation

Oligonucleotides labelled with Cy3B and ATTO647N were purchased from IBA Life Sciences (Germany), and gel-purified by Nicole Robb. The purified single-stranded DNA (ssDNA) were annealed in annealing buffer consisting of 50mM Tris-HCl pH 8.0, 1mM EDTA, and 500 mM NaCl. See Fig. 5.1 for DNA sequences and labelling schemes.

Wild-type (WT) RNAP core from *E. coli* was prepared as described previously (178). Briefly, BL21-AI chemically competent cells (Life Technologies) transformed with pVS10 plasmid encoding the five core RNAP subunits ($\alpha_I\alpha_{II}\beta\beta'\omega$) (179) were inoculated into 30 mL of LB medium with ampicillin at 100 mg/L and grown with agitation at 37°C. When OD_{600}=0.7 was reached, 0.6 mL of the starter culture was added to 600 mL of LB medium supplemented with both carbenicillin at 100 mg/L and an Overnight Expression Autoinduction system (Novagen) to co-express the RNAP subunits. The culture was grown at 37°C for 16 hours, and the cells then harvested by centrifugation. The cells were lysed using a French pressure cell press, and the cleared lysate combined with Ni-

NTA slurry, washed, and then eluted with imidazole. RNAP containing fractions were pooled and loaded onto a 5 mL HiTrap Heparin column (GE Healthcare, UK) and purified using an ÄKTA FPLC (GE Healthcare, UK) by eluting with a linear NaCl gradient. RNAP containing fractions were pooled, dialysed, and then loaded onto a MonoQ column (GE Healthcare, UK). Washing and elution was carried out as described for the HiTrap column except this time with a gentler gradient in order to isolate core from holoenzyme (180). RNAP containing fractions were dialysed into RNAP storage buffer and stored at -20°C. RNAP storage buffer consisted of 50% glycerol (vol/vol), 10 mM Tris pH 7.5 at 4 °C, 100 mM NaCl, 0.1 mM DTT, and 0.1 mM EDTA.

WT σ^{70} and mutant σ^{70} lacking residues 513-519 (Δ3.2) were purified by our collaborator, Dr. Konstantin Brodolin (CPBS, France) as described in (80, 181). Our collaborator Dr. Sivaramesh Wigneshweraraj (Imperial College London, UK) purified ΔJaw RNAP core as described in (41). WT, ΔJaw, and Δ3.2 holoenzymes were prepared by incubating 50 nM RNAP core with 250 nM σ^{70} for 30 min at 33 °C.

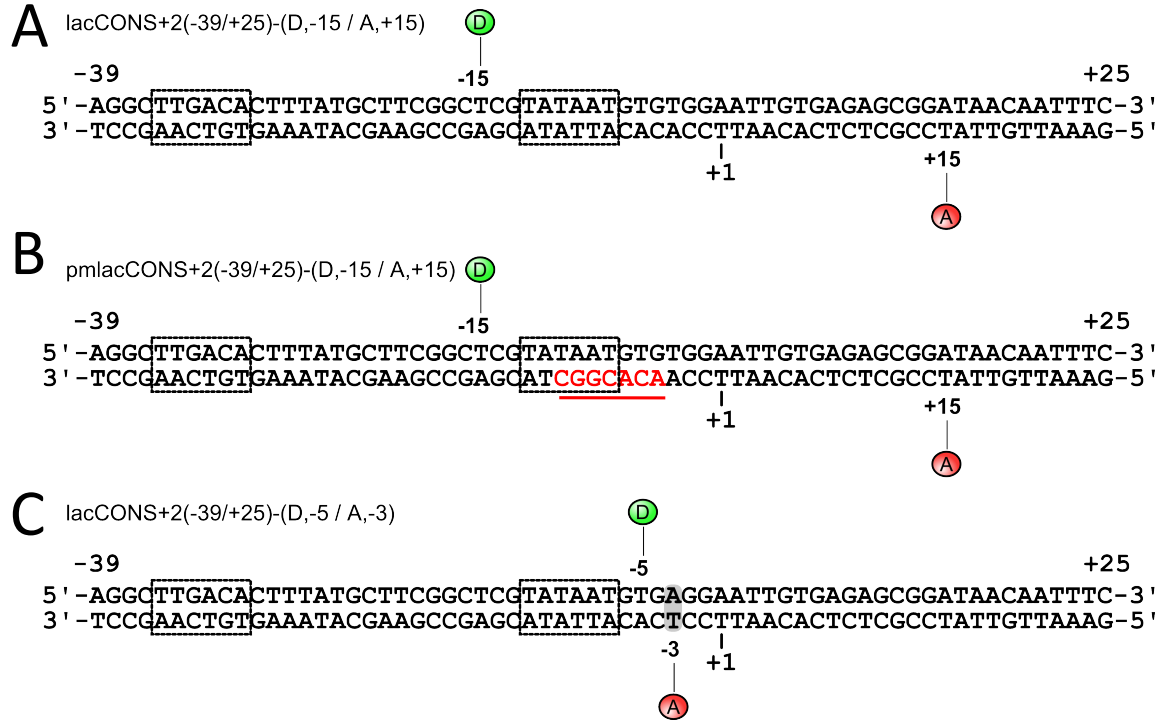

Figure 5.1. Sequences and labelling positions of DNA. -35 and -10 promoter elements are shown with boxes. The transcription start site (+1) is marked, as are the fluorophore labelling sites. (A) Fully complementary promoter DNA. (B) Pre-melted promoter DNA with the mismatch region shown and underlined in red. (C) 'Bubble ruler' DNA. The -3 position has an inversion mutation to permit the labelling of the -3 position. The donor is Cy3B, the acceptor is ATTO647N.

5.2.2 RP_O formation

RP_O was formed as described in (6, 182). Briefly, 10 nM of the promoter DNA fragment was incubated with 50 nM *E. coli* RNAP holoenzyme in T8 buffer (50 mM Tris-HCl (pH 8.0), 100 mM KCl, 10 mM $MgCl_2$, 100 µg/mL BSA, 1 mM DTT, 5% glycerol) to give a final volume of 20 µL. The mixture was incubated at 37°C for 15 minutes, after which 1 mg/mL heparin sepharose (GE Healthcare) was added to disrupt non-specific RNAP-promoter DNA complexes. The mixture was incubated at 37°C for 30 seconds and then centrifuged to remove sepharose beads. 13 µL of the supernatant was removed and transferred to a pre-warmed eppendorf tube and incubated for a further 20 min at 37°C. Our ability to form RP_O efficiently was demonstrated using an electrophoretic mobility shift assay. Here, 0.5 µL of 50% glycerol was added to 5 µL of the RP_O mixture and run on a 1% native TBE agarose gel at 21°C (Fig. 5.2).

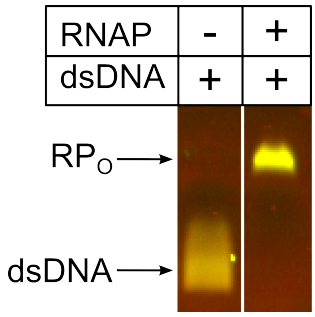

Figure 5.2. RP_O formation *in vitro*. Electrophoretic mobility shift assay showing retardation of the fluorescent DNA signal after incubation with RNAP as described in section 5.2.2. The image shows an overlay of donor emission upon donor excitation and acceptor emission upon donor excitation.

5.2.3 Vesicle encapsulation and immobilization of RP_O

RP_O was formed according to the steps described in section 5.2.2 except with a slight modification in concentrations to give 200 nM of both RNAP holoenzyme and promoter DNA in T8 (RNAP hydration buffer). The complexes were then encapsulated inside of 200 nm diameter vesicles as described in chapter 4. Vesicles were kept on ice, and samples incubated at 37°C for 10 minutes prior to immobilization in an attempt to re-form RP_O that might have dissociated during the encapsulation process. Samples were used on the same day they were formed. Vesicles were purified using size-exclusion chromatography and immobilized as described in chapter 4.

5.2.4 Antibody-immobilization of RP_O

A PEG-passivated glass surface was functionalized with penta-His antibody (Ab) as described in chapter 4. Then, a 1 nM solution of pre-formed RP_O was incubated on the surface for 5 min. Binding was monitored until ~60 molecules were deposited on the field of view surface (600 μm^2). Excess unbound complexes were washed away.

5.2.5 Instrumentation

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

5.2.6 Single-molecule experiments

Once RP_O was immobilized using either the vesicle or Ab approaches, imaging buffer containing 50 mM Tris-HCl (pH 8.0), 100 mM KCl, 10 mM $MgCl_2$, 100 µg/mL BSA, 1 mM DTT, 5% glycerol and 2 mM UV-treated Trolox was added to the observation chamber. An oxygen scavenging system described in chapter 4 was added to the observation chamber just before it was sealed using a glass coverslip as a lid. Data of vesicle-immobilized RNAP-DNA complexes was acquired at a temporal resolution of 200 ms and with excitation powers of 1 mW (532 nm laser) and 0.5 mW (635 nm laser). Data of Ab-immobilized RNAP-DNA complexes was acquired at a temporal resolution of 20 ms, and with excitation powers of 5 mW (532 nm laser) and 2 mW (635 nm laser). All experiments were performed at 21°C unless indicated otherwise.

5.2.7 Data analysis and visualisation

Fluorescence intensities were extracted from images and uncorrected FRET efficiencies (E*) calculated as described in chapter 4. We manually inspected intensity time trajectories and selected 100 molecules for analysis according to the following

criteria: 1) bleached in a single step after at least 105 frames. 2) Circular PSF across DD, DA, and AA channels. 3) No donor photoblinking (although fluctuations permitted). 4) No acceptor photoblinking although we allowed for different states (see chapter 4). 5) Fluorescence intensities within a limited range. 6) No defocusing. 7) No nearby molecules as measured using a nearest-neighbour criterion. All E* histograms were constructed using 250 frames of data from each of our 100 selected molecules. 40% of spots passed all our selection criteria.

We manually classified the intensity time trajectories of individual molecules into groups. Static closed bubble signals were defined as time trajectories showing a stable E*~0.20 signal with no fluctuations to higher FRET states during our observation period. Static open bubble signals exhibited a stable E*~0.45 signal with no fluctuations to lower FRET states. Dynamic bubble signals were defined as time trajectories showing any anti-correlated E* changes during our observation periods. Molecules exhibiting confounding photophysical fluctuations were excluded from analysis.

5.2.8 HMM Analysis of dynamic smFRET time trajectories

HMM analysis was performed on time trajectories displaying dynamic FRET fluctuations using custom-written MATLAB software (184, 185). Each time trajectory was fitted with 1 to 3 states and the best model (number of states) was selected automatically using maximum evidence criteria (184, 185). The best model selected was invariably a 2 state model. The states extracted by HMM were categorized according to their FRET efficiency and the states that preceded them; low FRET efficiency states were classified as "closed bubble states", and high FRET efficiency states were classified as "open bubble states". The dwell time distributions for each state category were fitted with either a single or double exponential decay function using least squares fitting. The single exponential decay function used was $F(t) = Aexp\left({}^{-t}/_{\tau}\right)$ with A and τ the fitted parameters. The double exponential decay function used was $F(t) = Aexp\left({}^{-t}/_{\tau_1}\right) + Bexp\left({}^{-t}/_{\tau_2}\right)$ with A, B, τ_1, and τ_2 the fitted parameters. The fits were used to extract the mean open (τ_{open}) and closed (τ_{closed}) dwell times, respectively.

5.3 Results

5.3.1 Vesicle immobilization of RNAP-promoter DNA complexes

In RP_O, 12-14 bp of DNA surrounding the transcription start site is melted to form the transcription bubble. Our group previously performed a solution-based smFRET study of RNAP-promoter DNA complexes utilizing a fluorophore FRET pair at positions -15 and +15 to report on the transcription bubble conformation (175). Here, the opening of the transcription bubble brings the donor and acceptor dyes closer together and results in an increase in FRET. This study found that a mixture of free promoter DNA and RP_O generated an E* histogram with a bimodal FRET distribution. The low FRET population (E*~0.2) was assigned to free promoter DNA (closed transcription bubble) while the high FRET population (E*~0.4) was assigned to RP_O (open transcription bubble). The RP_O population was found to be wider than the shot-noise prediction, which led us to hypothesize that the excess width could be due to conformational heterogeneity. Unfortunately, the short observation window (~1 ms) characteristic of diffusion studies hindered further investigations as to the heterogeneity's origins. In this chapter, we extended this study by performing smFRET TIRF experiments on immobilized RNAP-promoter DNA complexes using the vesicle and Ab immobilization approaches (see chapter 4) (Fig. 5.3). We begin by presenting results of vesicle-immobilized RNAP-promoter DNA complexes.

We formed RP_O *in vitro* (see *Materials and Methods*) using the lacCONS+2 promoter DNA fragment described in (175) and shown in Fig. 5.1A. We then encapsulated the protein-DNA complex inside 200 nm diameter vesicles and incubated the vesicles at 37°C for 10 mins prior to immobilization (see *Materials and Methods*). Once the vesicles were immobilized, we acquired data at a temporal resolution of 200 ms. As a control, we also encapsulated promoter DNA alone inside of vesicles. We selected molecules according to the criteria described in *Materials and Methods* section 5.2.7, and generated E* histograms for both promoter DNA alone and RNAP-promoter

DNA complexes inside of vesicles, which we compared to solution-based E* histograms (Fig. 5.4).

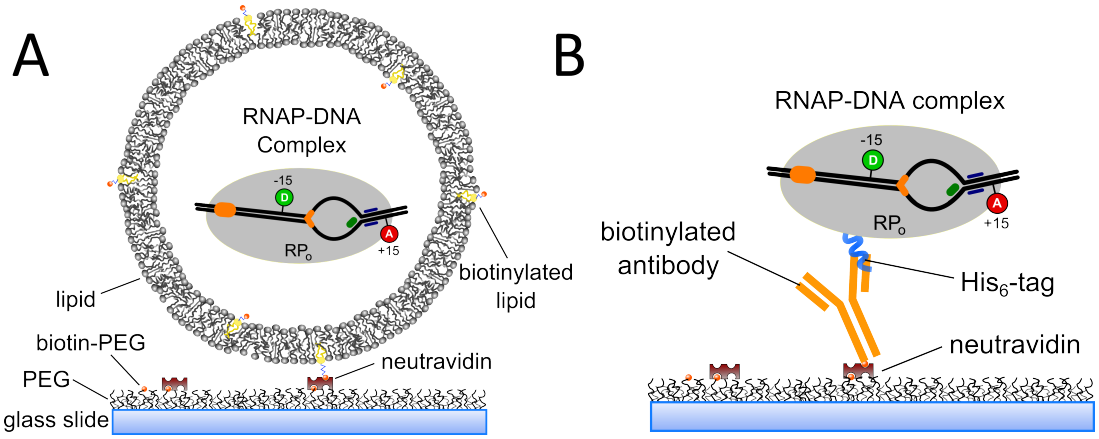

Figure 5.3. Vesicle- and Ab-immobilization of RP_O. Cartoon schematics of (A) the vesicle-immobilization approach, (B) the Ab-immobilization approach used to immobilize RP_O in this chapter. The -35 and -10 promoter elements are shown in orange. The RNAP active site is marked with a green oval. The downstream DNA binding cleft is shown in blue.

The histograms for the two datasets resemble each other; the promoter DNA E* histograms show a single population centred at E*~0.2, while the RNAP-promoter DNA complex E* histograms show at least two populations. The first is centred at E*~0.2 while the second is centred at E*~0.45 (Fig. 5.4).

Our time trajectory data showed that promoter DNA alone exhibits a static FRET signal centred at E*~0.2 (Fig. 5.5A). By contrast, the time trajectories of vesicle-encapsulated complexes reveal a variety of behaviours. We identified molecules that adopt a stable high FRET state in which no clear FRET fluctuations can be observed (Fig. 5.5B). We also observed molecules that exhibit a stable low FRET state similar to that seen for the vesicle-encapsulated DNA control. However, since RNAP is not fluorescently labelled we cannot comment as to whether these vesicles contain RNAP in their interior. We identified molecules that undergo clear FRET fluctuations between the E*~0.20 and E*~0.45 states (Fig. 5.5C), and also observed molecules that adopt a stable E*~0.35 state (Fig. 5.5D). The E*~0.35 state is in between that expected for a fully open and fully closed transcription bubble, which indicates that the transcription bubble here may be in a partially open conformation. Alternatively, the intermediate FRET state may be due

to unresolved FRET transitions that occur on a faster timescale than our temporal resolution. An analysis of the time trajectory data using Mean Variance Analysis supports the hypothesis that fast transitions are partially responsible for the E*~0.35 state (186).

Our results demonstrate that we are able to encapsulate RNAP-promoter DNA complexes within vesicles and use this tool to monitor molecules for many seconds (~60 sec). The fact that the TIRF and confocal E* histograms resemble each other suggests that the vesicles have not significantly affected the behaviour of the complexes. However, small differences in the E* distributions (Fig. 5.4) do raise some concerns that the physical manipulation exerted during the encapsulation process may have slightly perturbed the RNAP-promoter DNA complexes. Furthermore, there was uncertainty as to whether the E*~0.20 population was due to vesicles lacking RNAP in their interior. These uncertainties combined with the fact that the encapsulation process was technically very demanding and time-consuming led us to immobilize RNAP-DNA complexes using the Ab technique described in chapter 4.

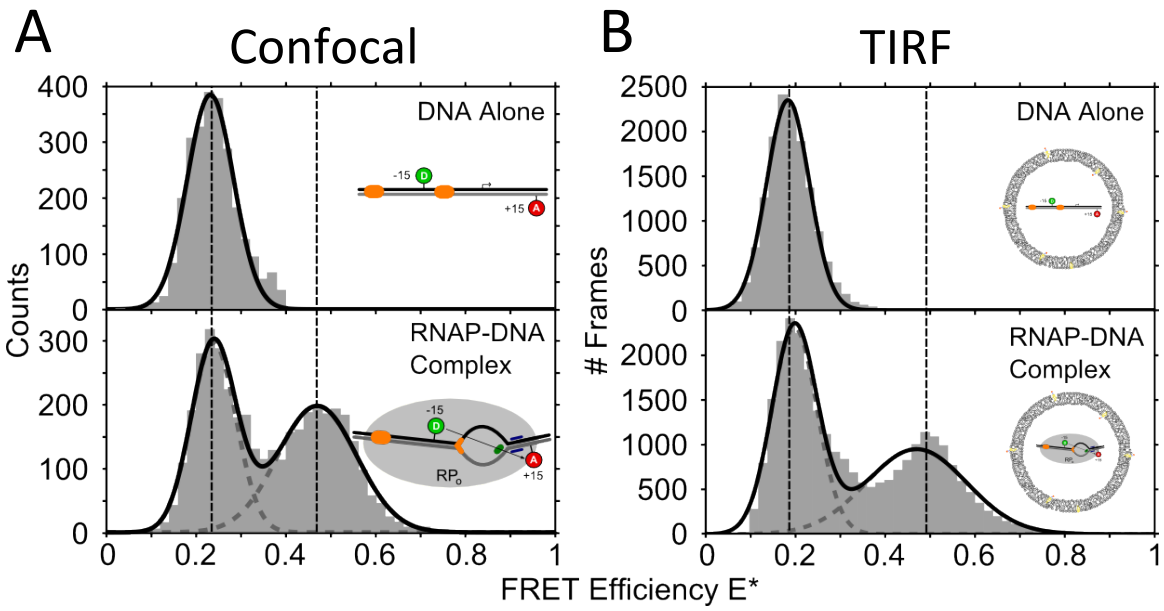

Figure 5.4. Vesicle-immobilized RNAP-DNA complexes. (A) Stacked FRET histograms of solution confocal data of free lacCONS+2 promoter DNA (top; 4372 bursts), and of a mixture of free promoter DNA and RNAP-promoter DNA complexes (bottom; 4993 bursts). (B) Stacked FRET histograms of vesicle-encapsulated lacCONS+2 promoter DNA alone (top; 100 molecules); and of vesicle-encapsulated RNAP-DNA complexes (bottom; 100 molecules). TIRF data at 200 ms temporal resolution. Black lines show Gaussian fits.

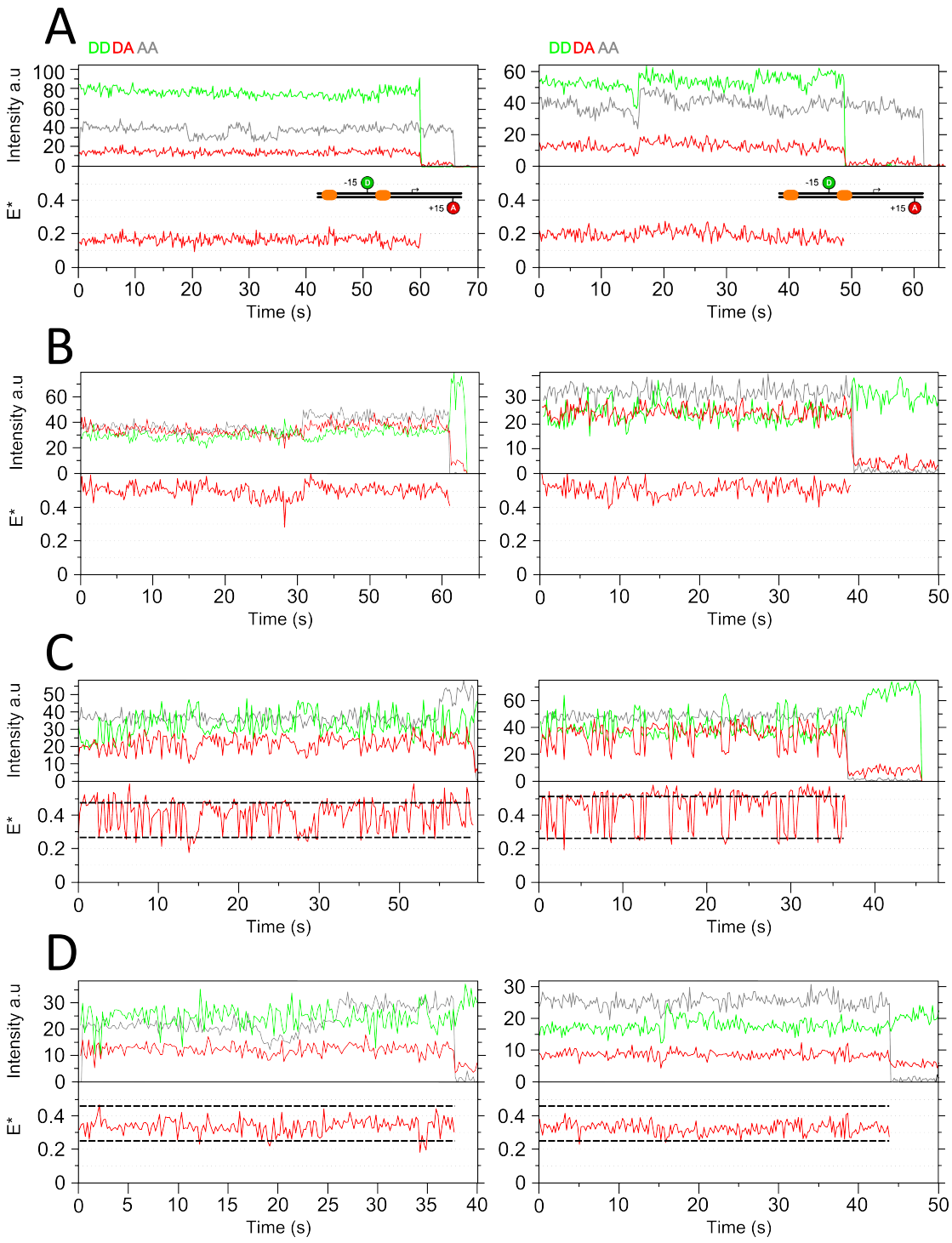

Figure 5.5. Example time trajectories of vesicle-immobilized RNAP-DNA complexes. (A) lacCONS+2 promoter DNA alone in vesicles. Vesicle-immobilized RNAP-DNA complexes exhibiting (B) stable high FRET states, (C) dynamic transitions in which the transcription bubble appears to fluctuate between two conformations, and (D) a stable intermediate FRET value. The highest and lowest FRET states are marked by dashed guidelines. A double-stranded DNA schematic is shown inset in A. Temporal resolution is 200 ms.

5.3.2 Antibody-immobilization of RNAP-DNA complexes

We formed RP_O *in vitro* and immobilized the complex using the Ab-immobilization approach as described in *Materials and Methods*. As a control, we also immobilized biotinylated promoter DNA alone via biotin-NeutrAvidin interactions. We acquired data at a temporal resolution of 20 ms, selected molecules for analysis, and generated E* histograms according to the criteria described in *Materials and Methods*. We then compared the TIRF E* histograms with solution confocal E* histograms (Fig. 5.6).

As expected, the biotin-immobilized promoter DNA E* histogram showed a single population (E*~0.20). The Ab-immobilized RNAP-DNA complex E* histogram showed two populations; a smaller population was centred at E*~0.20 and a more abundant population was centred at E*~0.45 (Fig. 5.6B). We expected to observe the E*=0.45 population and assigned it to an RNAP-promoter DNA complex with the transcription bubble in an open conformation. However, because we immobilized complexes via the protein, we did not expect to observe a population with the same FRET signature as DNA alone (E*~0.20). Since in our immobilization strategy any free DNA can only be due to non-specific adsorption, and since control measurements showed that non-specific adsorption was negligible[7], we conclude that the E*=0.20 population we observe cannot be due to free promoter DNA non-specifically adsorbed to the surface. Furthermore, since this population shows a similar FRET signature to DNA alone, we infer that the promoter DNA transcription bubble is closed in this complex.

We investigated the E*=0.20 population further by repeating the experiment described above with 100 µg/mL heparin in the observation chamber. Heparin is a non-specific DNA competitor that should remove DNA from any RNAP-DNA complexes that may have reverted to the heparin-sensitive RP_C state while attached to the surface (187). Heparin was found to have little effect on the FRET histogram (Fig. 5.6C), which

[7] ~1 molecule per 100 immobilized complexes (average over 5 slides).

suggests that the E*=0.20 population corresponds to an intermediate that occurs downstream of the heparin-sensitive RP_C state.

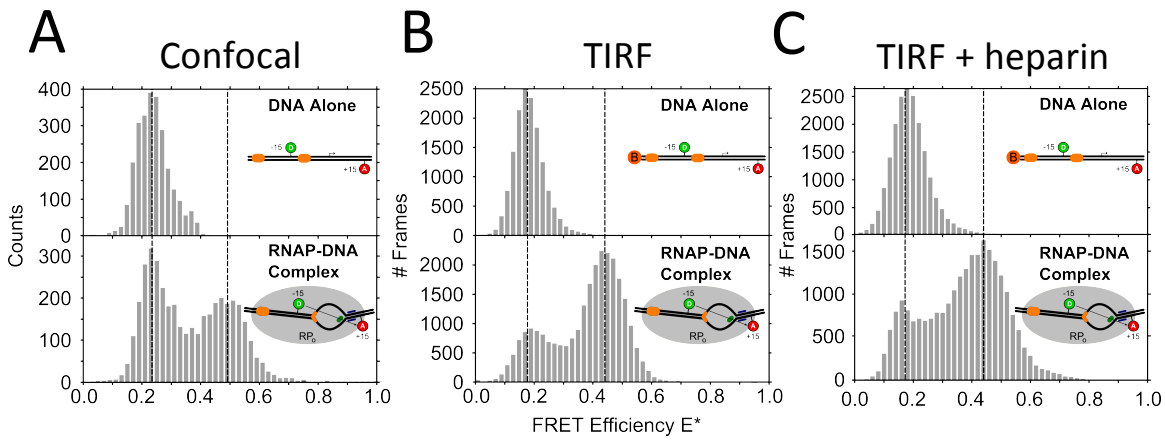

Figure 5.6. Ab-immobilized RNAP-promoter DNA complex E* histograms. (A) Stacked FRET histograms of promoter DNA (top; 4372 bursts) and RNAP-promoter DNA complexes (bottom; 4993 bursts) in solution. (B) Stacked FRET histograms of biotin-immobilized promoter DNA (top) and Ab-immobilized RNAP-promoter DNA complexes (bottom). Biotin is shown as an orange sphere labelled 'B' (C) Stacked FRET histograms of biotin-immobilized promoter DNA (top) and Ab-immobilized RNAP-DNA complexes in imaging buffer supplemented with 100 µg/mL heparin (bottom). All TIRF FRET histograms show data for 100 molecules.

Next, we inspected 100 individual time trajectories of immobilized RNAP-promoter DNA complexes (Fig. 5.7), and found the same general behaviours seen for our earlier vesicle-immobilization data (Fig. 5.5). Specifically, we found 50% of molecules exhibited a stable E*~0.45 state (Fig. 5.7B), 20% exhibited a stable E*~0.20 state (Fig. 5.7B), 25% displayed dynamic FRET fluctuations (Fig. 5.7C), and 5% displayed a stable E*~0.35 state (Fig. 5.7D). This finding supports the notion that the heterogeneity identified earlier in this chapter is not due to the immobilization strategy but may be an inherent feature of RNAP.

We assigned the stable E*~0.45 state as an RNAP-promoter DNA complex with the transcription bubble fully open, and the stable E*~0.20 state as a structural intermediate with the transcription bubble in a stably closed conformation. We assigned molecules exhibiting dynamic FRET fluctuations as complexes in which the transcription bubble undergoes opening and closing structural transitions, and we hypothesized the existence of an RNAP-Promoter DNA complex in which the transcription bubble

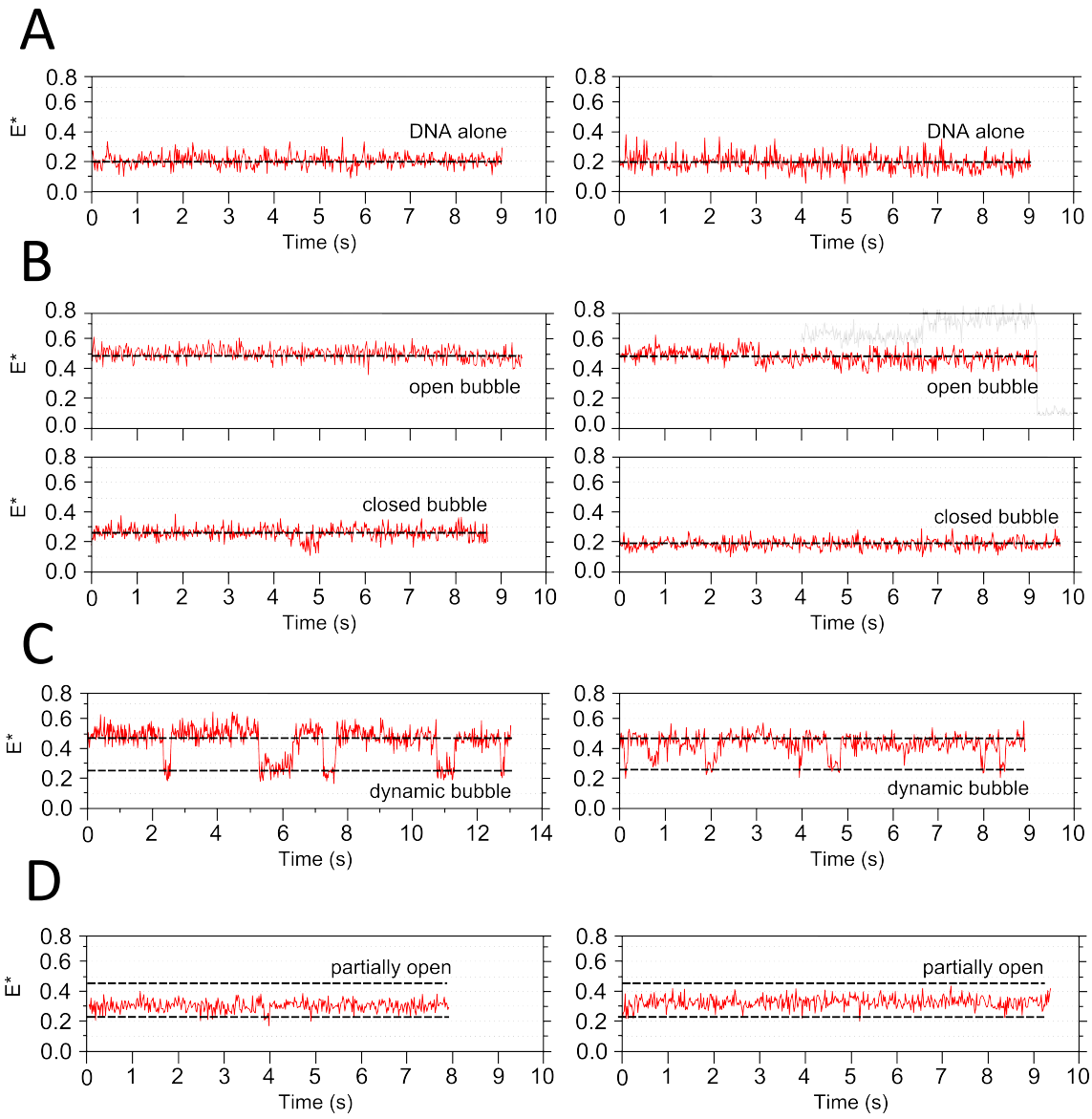

Figure 5.7. Example time trajectories of Ab-immobilized RNAP-promoter DNA complexes. (A) Biotin-immobilized promoter DNA. (B) Top figures show RNAP-promoter DNA complexes exhibiting a stable high FRET state. The bottom figures show RNAP-DNA complexes exhibiting a stable low FRET state. We assign these as open and closed transcription bubbles, respectively. (C) RNAP-DNA complex exhibiting dynamic transitions between open and closed bubble conformations. (D) RNAP-DNA complexes with a stable intermediate FRET value between the expected FRET values for open and closed transcription bubbles. Data was acquired at a temporal resolution of 20 ms. The FRET signal ends abruptly due to bleaching events. Approximate positions of high and low FRET states are shown with dashed lines.

occupies a partially open conformation that is neither fully closed nor fully open. This hypothesis is supported by time trajectory data showing clear transitions between either E*~0.45 and E*~0.35, or between E*~0.20 and E*~0.35 states (Fig. 5.8A). The FRET transitions are anti-correlated, suggesting that they are real events. A final

observation worth noting is that the E*~0.20 and E*~0.35 populations remained present when we repeated experiments at 37°C (Fig. 5.9).

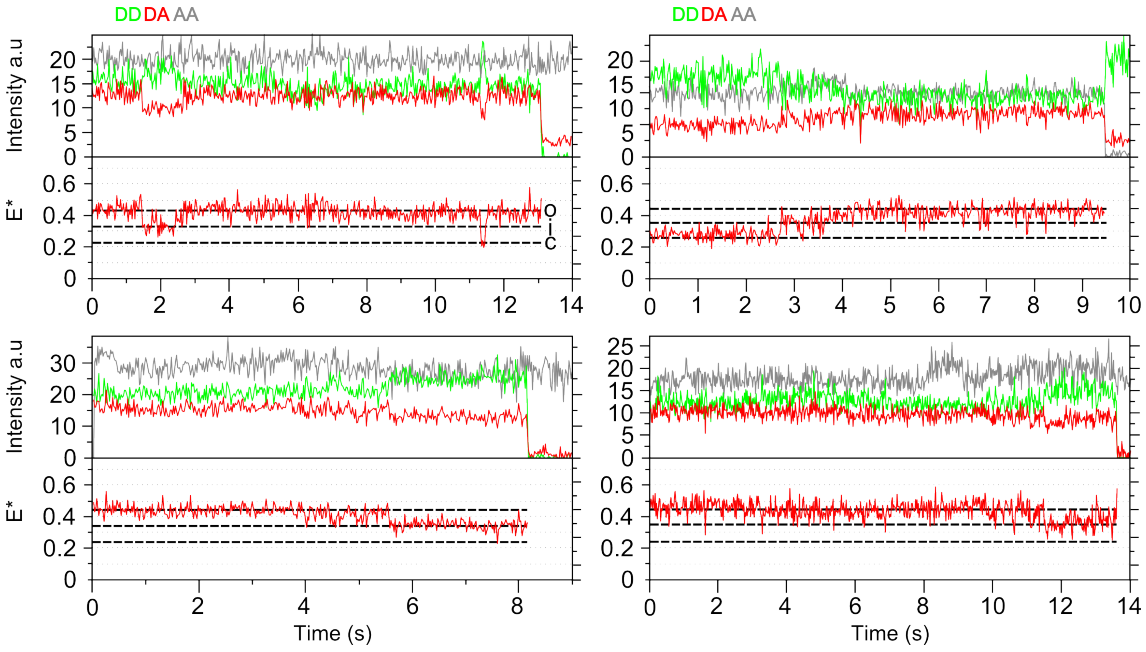

Figure 5.8. A partially open transcription bubble conformation. Example time trajectories of Ab-immobilized RNAP-promoter DNA complexes displaying transitions between E*~0.35 and E*~0.45, or between E*~0.35 and E*~0.20. Fluorescence intensity time trajectories are shown in the top panels of figure to re-assure the reader that the FRET changes are anti-correlated and not due to fluorophore photophysical aberrations. The dashed lines indicate the approximate FRET values of the closed (c), intermediate (i), and open (o) conformations.

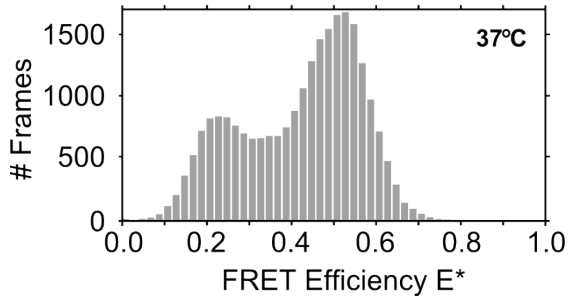

Figure 5.9. Effect of temperature on the RNAP-DNA complex E* distribution. E* histogram of 100 Ab-immobilized RNAP-promoter DNA complexes at 37°C. Temporal resolution is 20 ms. Note that the shape of the distribution is largely unchanged compared with Fig. 5.6B.

5.3.3 Investigating the origin of transcription bubble dynamics: DNA determinants

Several time trajectories (25%) showed dynamic transitions between E*~0.23 and E*~0.45 (Fig. 5.7C). To provide another vantage point to monitor transcription bubble dynamics, we moved the fluorophores to within the transcription bubble by

employing a -5/-3 labelled promoter DNA fragment (Fig. 5.1C). This DNA fragment is referred to as a 'bubble ruler' and has been previously used to monitor bubble opening (175, 188). When the transcription bubble is closed, contact-mediated quenching precludes fluorescence, then as the two DNA strands separate during RP_O formation, fluorescence is restored, which results in an E*>0.5 population (Fig. 5.10A) (175). We performed experiments using the bubble ruler and identified molecules exhibiting dynamic behaviour in which the FRET signal fluctuates between a quenched FRET state and a FRET state centred at E*=0.6 (Fig. 5.10B). As the bubble ruler DNA directly reports on the proximity of the template and non-template strands in the transcription bubble, the results support our conclusion that the FRET fluctuations we observed in Fig. 5.7C correspond to structural transitions in which the transcription bubble opens and closes. More detailed structural interpretations of the bubble ruler data are not possible due to the possible influence of the protein environment on the dye fluorescence signals.

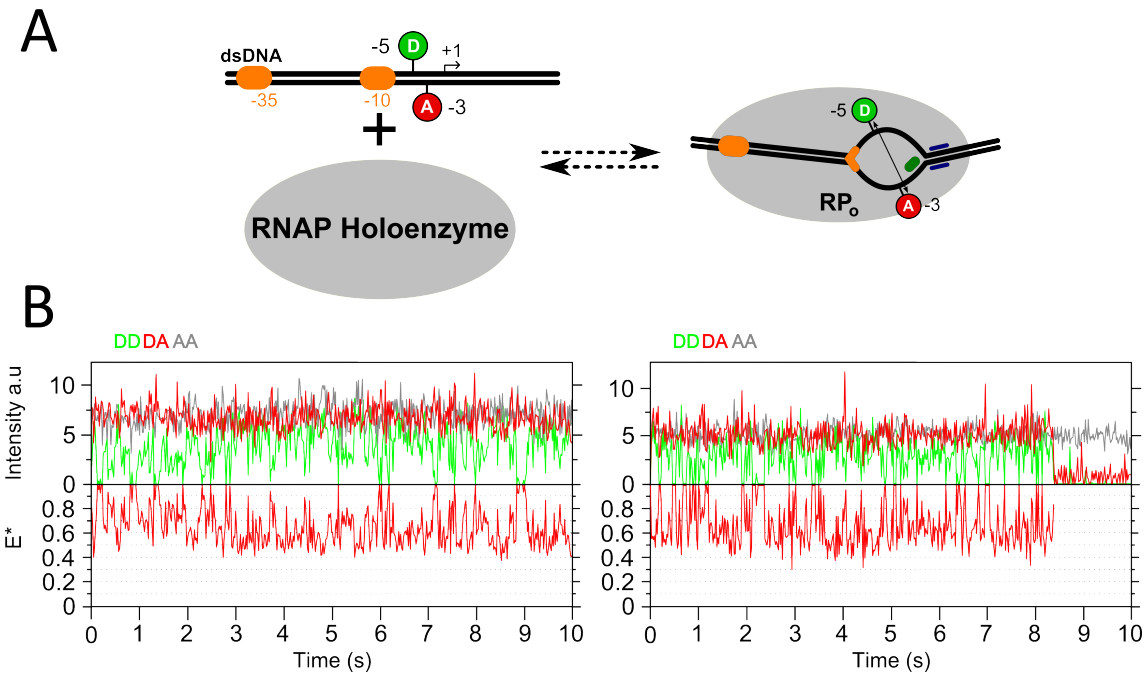

Figure 5.10. Bubble ruler DNA results. (A) Cartoon schematic showing the formation of RP_O using the bubble ruler DNA. The donor and acceptor fluorophores are close together in dsDNA and move further apart in RP_O. The positions of the active site (green oval), the binding cleft (blue), the -10 element, and the -35 element are shown. Figure adapted from (175). (B) Example time trajectories of complexes formed using the bubble ruler. Clear transitions occur between a high FRET state (E*~0.6) and a quenched state. Temporal resolution is 20 ms.

One possible explanation for the observed dynamics is that a fraction of RNAP molecules are incapable of stable promoter opening. We tested this hypothesis by performing smFRET TIRF experiments as described above, except this time using a pre-melted promoter DNA fragment (pmDNA) to form RNAP-DNA complexes (see Fig. 5.1B). By placing mismatches within the transcription bubble, we mimicked bubble nucleation and biased the reaction toward RP_O (189, 190). If transcription bubble dynamics is due to unstable promoter opening, we would expect the mismatch to reduce the size of the dynamic subpopulation. We would also expect the size of the static E*=0.45 population to increase because of the improved stability of the open transcription bubble.

The RNAP-promoter pmDNA E* histogram showed two populations that are approximately equal in size (Fig. 5.11). The first is centred at E*=0.25, and the second at E*=0.45. It was unexpected that the E*~0.20 population would be larger than that seen for regular RNAP-promoter dsDNA complexes. However, an inspection of individual time trajectories revealed that the increase in the size of the E*~0.20 population was due to an increase in the proportion of molecules exhibiting bubble dynamics (60% for pmDNA compared with 25% for dsDNA) and not due to an increase in the proportion of molecules adopting a static E*~0.20 state. However, molecules adopting a stable E*~0.20 state remained present even with a mismatch in the transcription bubble region. These observations suggest that bubble dynamics is not due to unstable promoter opening, and raises the intriguing possibility that the bubble dynamics may be an inherent feature of RP_O with some functional role in transcription initiation.

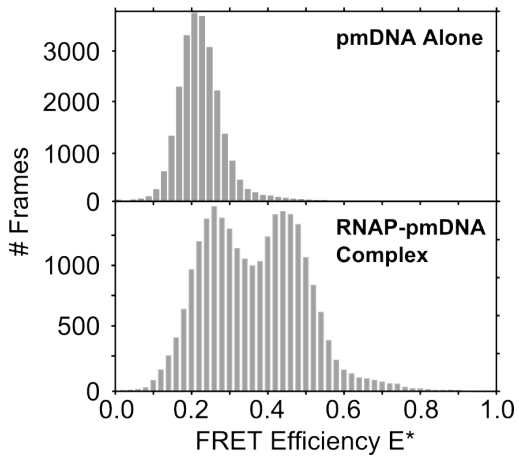

Figure 5.11. Effect of the transcription bubble mismatch on the RNAP-DNA complex E* distribution. Stacked E* histograms of 100 biotin-immobilized promoter pmDNA (top) and 100 Ab-immobilized RNAP-promoter pmDNA complexes (bottom). Temporal resolution is 20 ms. The RNAP-dsDNA complex E* histogram is shown in Fig. 5.6B.

5.3.4 Investigating the origin of transcription bubble dynamics: Protein determinants

Next, we sought to better understand how specific RNAP structural regions that are known to play important roles in transcription initiation influence transcription bubble dynamics. We investigated the σ^{70} unstructured subdomain σR3.2 (σR3.2 finger) and the β' jaw subdomain (Fig. 5.12).

The crystal structure of a *T. thermophilus* RNAP-fork junction DNA complex recently showed that the σR3.2 finger forms a loop that protrudes towards the RNAP active centre, where it comes into close proximity with the template strand (Fig. 5.12) (72). A current hypothesis is that interactions between the σR3.2 finger and the melted template strand are involved in the strand's correct positioning at the active site (73, 191). The β' jaw on the other hand is thought to position the NT strand correctly in the DNA binding cleft and stabilise RP_O by assembling on the downstream promoter DNA late in the RP_O formation pathway. In fact, deleting the jaw was shown to reduce the lifetime of the open complex at various promoters (192), and to stabilise an RP_O intermediate at the λP_R promoter (41). We hypothesized that the absence of either the σR3.2 finger or the β' jaw might destabilise the transcription bubble open conformation.

To test our hypothesis, we first performed experiments with a σ^{70} subunit variant lacking amino acids 513–519, corresponding to the loop tip (Δ3.2). We performed TIRF experiments using the fully complementary and pre-melted promoter DNAs and compared the E* histograms to those of wt RNAP; the histograms are shown in Fig. 5.13. We found that the E* distribution of Δ3.2 RNAP-promoter dsDNA complexes was similar to that of wt RNAP-promoter dsDNA complexes; however, there were marked differences between the FRET distributions of Δ3.2 and wt complexes formed using pmDNA. The major difference between the two was that the two FRET peaks for Δ3.2 complexes were closer together and skewed towards the open-bubble conformation (Fig. 5.13A) compared to the corresponding histogram for wt RNAP (Fig. 5.11). Our time trajectory data revealed that the difference in FRET distributions was largely due to faster bubble dynamics in Δ3.2 complexes relative to wt complexes (Fig. 5.13B).

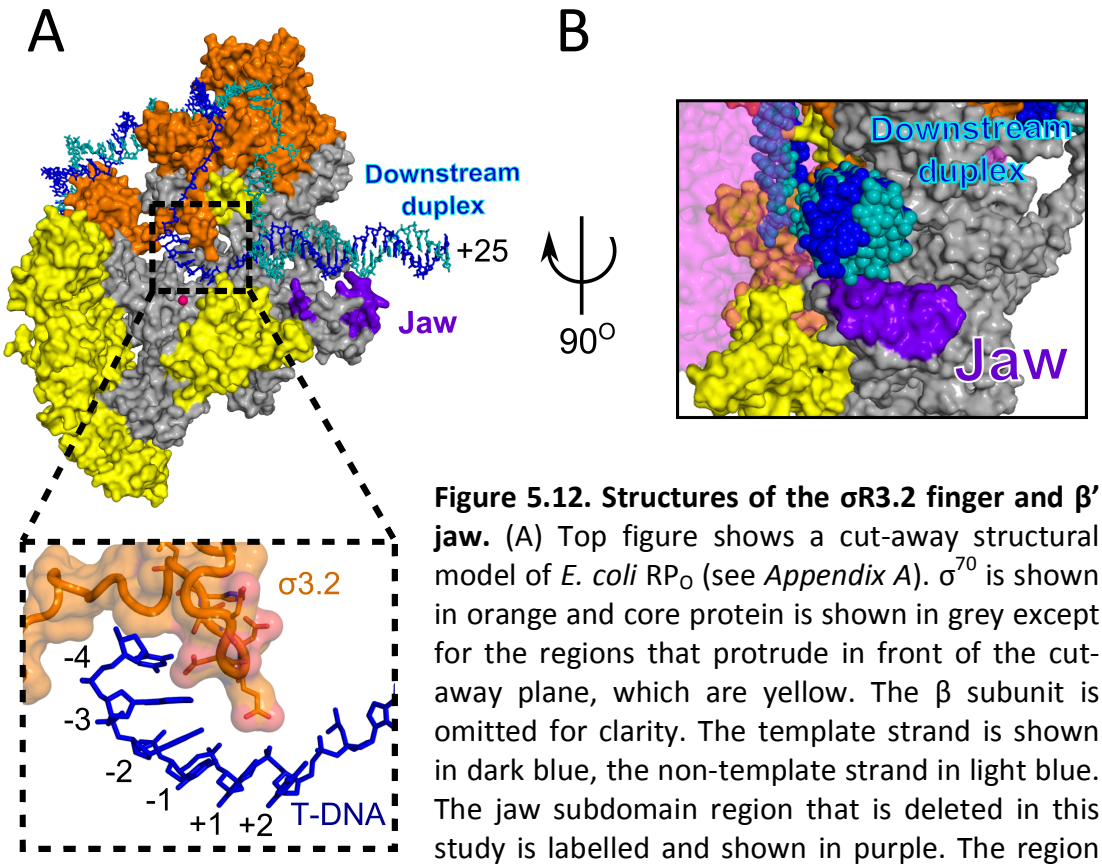

Figure 5.12. Structures of the σR3.2 finger and β' jaw. (A) Top figure shows a cut-away structural model of *E. coli* RP_O (see *Appendix A*). σ^{70} is shown in orange and core protein is shown in grey except for the regions that protrude in front of the cut-away plane, which are yellow. The β subunit is omitted for clarity. The template strand is shown in dark blue, the non-template strand in light blue. The jaw subdomain region that is deleted in this study is labelled and shown in purple. The region in the box is a zoomed view of the RNAP active site that is marked in the full RP_O model and shows the interactions made between the σR3.2 finger and the transcription bubble template strand. The side chains of σR3.2 residues deleted in this study are shown. (B) Zoomed view of the open complex after it has been rotated 90° in the direction of the curly arrow. This view shows the close contacts made between the jaw and downstream DNA. The β subunit is shown partially transparent and coloured pink.

In addition, we observed an increase in the proportion of molecules exhibiting a stable E*=0.35 state (20% for Δ3.2 pmDNA complexes compared with 10% for wt pmDNA complexes). Collectively, our results indicate that the σR3.2 finger influences the conformational dynamics of the transcription bubble.

Next, we performed smFRET experiments using RNAP containing a β' subunit variant lacking the jaw subdomain (amino acids 1149-1190; ΔJaw). The E* histograms for the ΔJaw complexes formed using dsDNA and pmDNA are shown in Fig. 5.14. The ΔJaw histograms show only small differences compared with the wt histograms. Firstly, the ΔJaw RNAP-promoter dsDNA complex E* histogram (Fig. 5.14A) shows an intermediate FRET population (E*~0.35) that is much more prominent than that seen in the wt RNAP histogram (Fig. 5.6B). Secondly, the ΔJaw RNAP-promoter pmDNA time