Studying σ^{70}-finger displacement during initial transcription using single-molecule FRET

Anna Wang

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

The Queen's College

University of Oxford

Hilary Term 2023

Disclaimer

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

- Double-labelled σ^{70}-factor protein, double-labelled hexahistidine-tagged RNAP holoenzyme, single-labelled σ^{70}-factor protein, single-labelled hexahistidine-tagged RNAP holoenzyme and hexahistidine-tagged RNAP holoenzyme was prepared by Abhishek Mazumder.
- *In vitro* transcription gels shown in Fig. 3.3 in Chapter 3.3 were performed by Zakia Morichaud and Konstantin Brodolin.
- The dual-colour TIRF microscope setup for single-molecule FRET measurements described in Chapter 2.4 was built by Ludovic Le Reste and Johannes Hohlbein.
- The confocal microscope setup for single-molecule FRET measurements described in Chapter 2.4 was built by Geraint Evans and Johannes Hohlbein.

Anna Wang
April 2023

Abstract

Bacterial RNA Polymerases (RNAPs) bind to transcription initiation protein factors called σ-factors to start DNA sequence-specific transcription. Within σ-factors lies a highly conserved structural module, the 'σ-finger', a loop that that resides very close to the 'heart' of transcription, the active-centre of RNAP. The σ-finger is implicated in the pre-organisation of template DNA and the synthesis of the first short RNAs. The σ-finger also blocks entry of the nascent RNA to the RNA-exit channel of the RNAP and must be displaced to allow entry into transcription elongation. Despite structural studies, σ-finger conformational changes during late transcription initiation are still unknown. To uncover the dynamic conformational landscape and mechanism of the *E. coli* σ-finger during initial transcription and promoter escape, this thesis uses a new single-molecule FRET (smFRET) ruler. The results show that the σ-finger is displaced from its position inside the active site cleft, before promoter escape and after synthesis of RNA lengths that are highly dependent on the sequence of the promoter DNA used. Additionally, the chemical moiety at 5'-end of RNA, which are used in different modes of transcription, was also found to influence the point of σ-finger displacement. Real-time smFRET measurements revealed the presence of significant heterogeneity in the timing of σ-finger displacement and show that different initial conformations of the σ-finger are linked to significantly different kinetics in transcription initiation and promoter escape.

This thesis identifies different mechanisms of σ-finger displacement that influence the kinetics of initial transcription and have the potential to impact gene regulation in bacteria. Since archaeal and eukaryotic transcription systems contain σ-finger-like structural modules, these mechanisms may be general and apply to all kingdoms of life.

Acknowledgements

I am hugely grateful to my supervisor Achilles Kapanidis for his unreserved support, mentoring and supervision over these past years, through thick and thin, and always generous.

I would like to thank everyone in the Kapanidis group, past and present, for your help and for making me feel so welcome first as a summer student and then as I started my DPhil journey. Thank you especially to Abhishek Mazumder for all your help and patience in teaching a physicist biology, and all the laughs along the way – I am very grateful. Thank you to Hafez El Sayyed for your charisma, Oliver Pambos for humouring my cheese sandwich plans, Christof Hepp for generally keeping me sane, Heesoo Uhm for being my desk neighbour, Stelios Chatzimichail in advance for the Popeyes next year (#teamNotGandalf), Piers Turner for being an AMR celebrity, Jagadish Prasad Hazra for your generosity, Nicole Robb for all your help, Rebecca Andrews for your advice and kindness when I first joined, Andrew McMahon for always knowing the right form I need to fill in, Nicolas Shiaelis for the laughs, Aleksander Zagajewski for the nice chats, Rasched Haidari for your gym masterclasses, Mirjam Kummerlin for the nice socials, Emma Lalande for your eccentric stories, Alison Farrar for all the lovely tea and Qing Zhao for my secret santa present.

I am grateful to the EPSRC for the funding.

To my parents: thank you for your love and support.

To Edmund Derby: thank you for your love and care.

Anna Wang
April 2023

Publications

- Wang, A., Mazumder, A., Kapanidis, A. N. (2023) Displacement of the σ^{70}-finger in initial transcription is highly heterogeneous and promoter-dependent. bioRxiv. https://doi.org/10.1101/2023.06.10.544452. *Manuscript submitted*.

- Ward, E. N., Hecker, L., Christensen, C. N., Lamb, J. R., Lu, M., Mascheroni, L., Chung, C. W., Wang, A., Rowlands, C. J., Schierle, G. S. K., and Kaminski, C. F. (2022). Machine learning assisted interferometric structured illumination microscopy for dynamic biological imaging. *Nature Communications*, *13*(1), 7836.

- Mazumder, A., Wang, A., Uhm, H., Ebright, R. H., and Kapanidis, A. N. (2021). RNA polymerase clamp conformational dynamics: Long-lived states and modulation by crowding, cations, and nonspecific DNA binding. *Nucleic Acids Research*, *49*(5), 2790–2802.

Talks

- σ^{70}-finger displacement during initial transcription: a single-molecule analysis. UK RNA Polymerase workshop, Newcastle, April 2022.

- Conformational control of the mechanism of promoter escape by RNA Polymerase. Single-molecule Biophysics Meeting, Les Houches, March 2022.

- Single-molecule FRET analysis of key protein conformational changes during promoter escape by RNA Polymerase. Biological Physics Seminar, Oxford, November 2021.

- Conformational changes of the σ^{70}-finger during late initial transcription, UK RNA Polymerase workshop, online, March 2021.

- Single-molecule FRET analysis of key protein conformational changes during promoter escape by RNA Polymerase, Biophysical Society Meeting, online, February 2021.

- Observing protein movements during initial transcription. Institute of Physics Biological Physics Group Early Career meeting, online, December 2020.

Contents

Chapter 1: Introduction

The goal of this thesis is to show how characterisation of the structure-function relationship of proteins and nucleic acids at the single-molecule level can clarify biological mechanisms. Single-molecule studies of molecular biology are important as biological molecules and cellular systems are highly complex, and ensemble studies are often unable to unpick the structural and kinetic heterogeneity that exists in chemically identical molecules.

The problem addressed is part of the question of how transcription initiation is regulated in bacteria. This topic has been studied extensively both by ensemble and single-molecule assays and much is known about the different factors affecting sequence-specific transcription initiation of RNA Polymerase (RNAP) from promoters, either through RNAP- or promoter-centred mechanisms (reviewed in Browning and Busby, 2016). One particular RNAP-centred regulatory point of interest that remains coarsely defined is the role of a structural module called the 'σ-finger', which is thought to undergo a large conformational change during promoter escape, when RNAPs containing short RNAs advance from the promoter and make the transition from transcription initiation to elongation.

The single-molecule approach to characterise the σ-finger is similar in principle to structural studies measuring its conformation at varying points of initial transcription. Structural studies produce three dimensional images with atomic resolution, however are strictly limited to static perspectives and generally cannot distinguish between different reaction pathways. This leaves behind unresolved pictures of the σ-finger's mechanism.

The single-molecule FRET (smFRET) characterisation of the σ-finger presented in this thesis makes a number of different observations that, pieced together with ensemble structural and biochemical results, forms a mechanism by which transcription initiation is regulated under varying physiological conditions. As equivalents of the σ-finger are pervasive for cellular transcription in all kingdoms of life, such mechanisms are also of interest outside of prokaryotic transcription.

Overview

This chapter introduces the problem and provides the relevant general concepts in transcription initiation and single-molecule methodology. Chapter 2 describes the materials and methods used in the whole thesis. In Chapter 3, a smFRET assay using a double-labelled σ^{70}-factor construct to observe σ-finger movements during transcription initiation is developed and characterised. In Chapter 4, this double-labelled σ^{70}-factor construct is used to measure the conformation of the σ-finger in complexes containing a consensus DNA promoter sequence. In Chapter 5, these experiments are extended to two naturally occurring DNA promoter sequences. Interpretations of results and future work are found in Chapter 6, and final conclusions in Chapter 7.

1.1 Transcription initiation

First proposed by Crick in 1958, the central dogma of molecular biology describes the transfer of genetic information between molecules of DNA, RNA and proteins (Crick, 1958). Transcription describes the process by which information stored in DNA is copied to synthesise RNA. The most highly regulated stage of this process is transcription

initiation, where the enzyme RNA Polymerase (RNAP) associates with a protein σ-factor to initiate transcription at specific sequences (promoters) of DNA.

The RNAP core enzyme is composed of five subunits: α1, α2, β, β' and ω. Structures reveal these to be arranged in a crab claw-like shape with β and β' subunits making up two 'pincers' (Darst, 2001). At the hinge of these 'pincers' lies the active-centre cleft, where the main transcriptional activity of RNA synthesis takes place. A secondary channel connected to the active-centre cleft provides access to incoming ribonucleotides (RNA bases) and the RNA exit channel accommodates the nascent RNA chain (Fig. 1.1).

To initiate transcription from *E. coli* promoters, RNAP core binds with a σ-factor to form RNAP holoenzyme (Fig. 1.2A-B; Feklístov et al., 2014). Different σ-factors specify for different promoters, providing a means by which to select for different sets of genes. The majority of σ-factors belong to the σ^{70}-family and are structurally related to a 'housekeeping' σ-factor (σ^{70} in *E.coli*) responsible for the majority of transcription. This family is further divided into four groups based on the degree to which they resemble the σ^{70}-factor with Group 1 containing 'housekeeping' σ-factors such as σ^{70}, and Groups 2-4 containing alternative σ-factors which are selected for by stress signals (Feklístov et al., 2014). The second family of σ-factors, the σ^{54} family, are unrelated by structure or sequence to the σ^{70} family and are also involved in a variety of stress responses (Danson et al., 2019). This thesis will focus on transcription from the 'housekeeping' σ^{70}-factor (*rpoD* gene product) in *E.coli*.

Structural studies of the σ^{70}-factor and σ^{70}-RNAP holoenzyme show that five main domains; σR1.1, σR2, σR3, σR3.2 and σR4 are conserved in all σ^{70} proteins (Campbell et

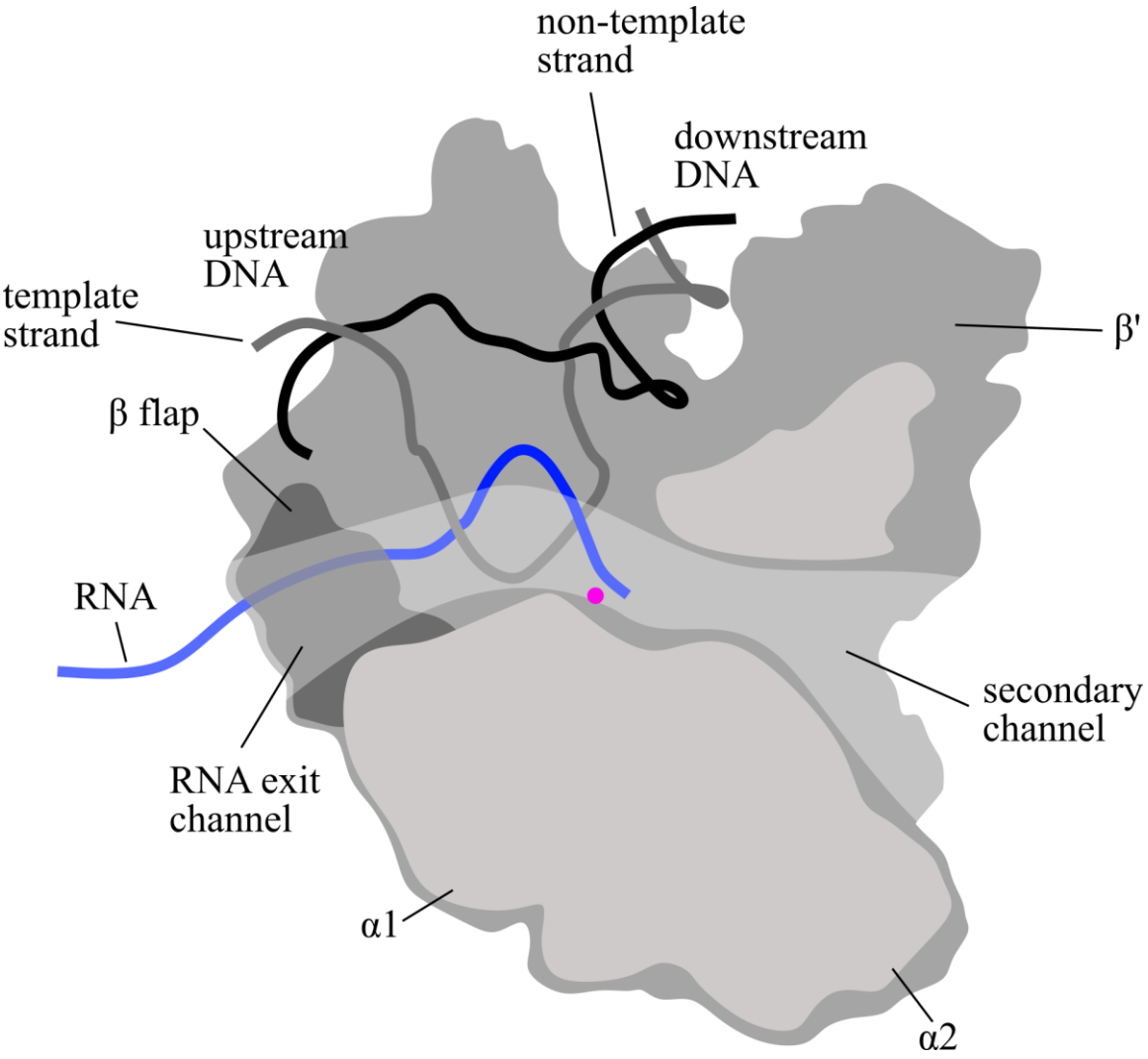

Figure 1.1: Cross-sectional schematic of bacterial RNAP. RNAP core enzyme is coloured grey with areas of the structure that were sliced in the plane of the page light coloured (apart from the β flap, which is dark). The template and non-template DNA strands are coloured dark grey and black respectively, and the RNA is blue. The active site Mg^{2+} is shown as a magenta circle. The secondary and RNA exit channels are highlighted in white. Created with BioRender.com.

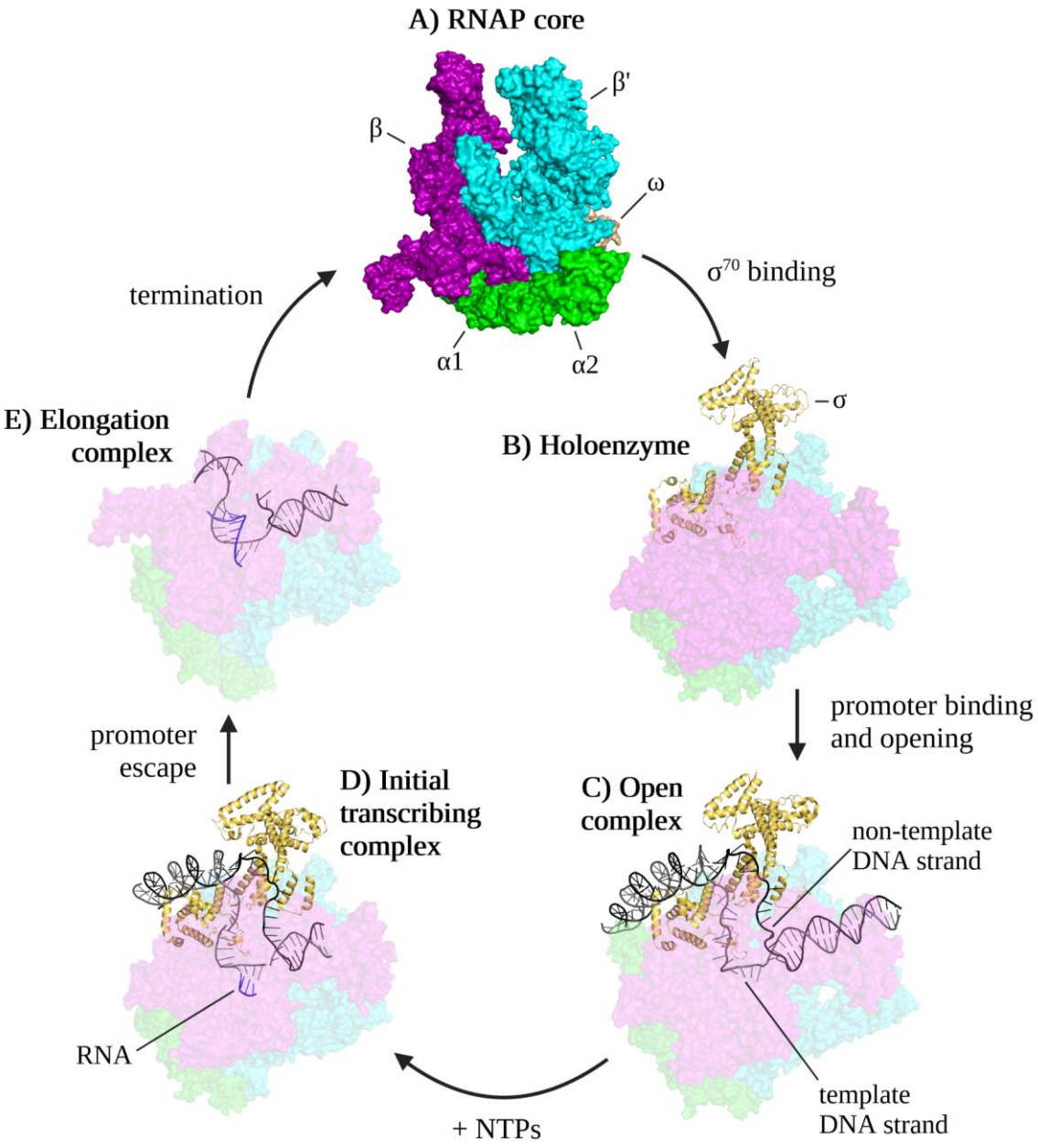

Figure 1.2: Structures of *E. coli* RNAP complexes during transcription. **(A)** RNAP Core (7MKP; Qayyum et al., 2021). **(B)** RNAP holoenzyme (6C9Y; Narayanan et al., 2018). **(C)** RNAP open complex (7MKD; Saecker et al., 2021). **(D)** RNAP initial transcribing complex with a 4-nt RNA transcript (4YLN; Zuo and Steitz, 2015). **(E)** RNAP elongation complex (7MKO; Qayyum et al., 2021). All structures shown are solved crystal or cryo-EM structures. Created with BioRender.com.

al., 2002). Interactions between σ^{70} and RNAP core spread open and expose sequence specific DNA binding regions in the σ-factor (Callaci et al., 1998, 1999), explaining observations that free σ cannot bind to DNA (Burgess and Travers, 1970; Wu et al., 1975). These interactions occur between σR2 and the β' pincer, σR3 and the base of the β flap and σR4 with the tip of the β flap (Murakami et al., 2002). The remaining conserved regions, σR1.1 and σR3.2, act as negatively-charged nucleic acid mimics – σR1.1 is positioned within the active-centre cleft in RNAP holoenzyme and acts as a 'gatekeeper' to prevent non-specific DNA association (Vuthoori et al., 2001; Wilson and Dombroski, 1997) and σR3.2 or the 'σ-finger' forms part of the linker between σR3 and σR4 occupying the path of the nascent RNA will take during initial transcription (Y. Zhang et al., 2012; Zuo and Steitz, 2015). The role of the σ-finger is the main focus of this thesis and will be discussed throughout.

Once bound, the σ^{70}-factor guides RNAP holoenzyme to form sequence specific interactions with the promoter, which is characterised by two main DNA sequence motifs upstream of the transcription start site: the -35 (5'-TTGACA-3') and -10 (5'-TATAAT-3'), recognised by σR4 and σR2 respectively (numbers refer to the position on the DNA, consensus sequences are in brackets) (Dickson et al., 1975; Shimada et al., 2014). More specifically, structural studies show that a helix-turn-helix motif in σR4.2 makes contacts with the -35 hexamer and the -11 and -7 non-template bases of the -10 hexamer are flipped into protein pockets of σR2.3 (Y. Zhang et al., 2012). These interactions are shown in Fig. 1.3A. Several isomerisation events, investigated by structural, biochemical and single-molecule methods (Chen et al., 2020; Mazumder et al., 2021; Mazumder and Kapanidis, 2019; Ruff, Record, et al., 2015), then take place to form the RNAP open complex (RPo)

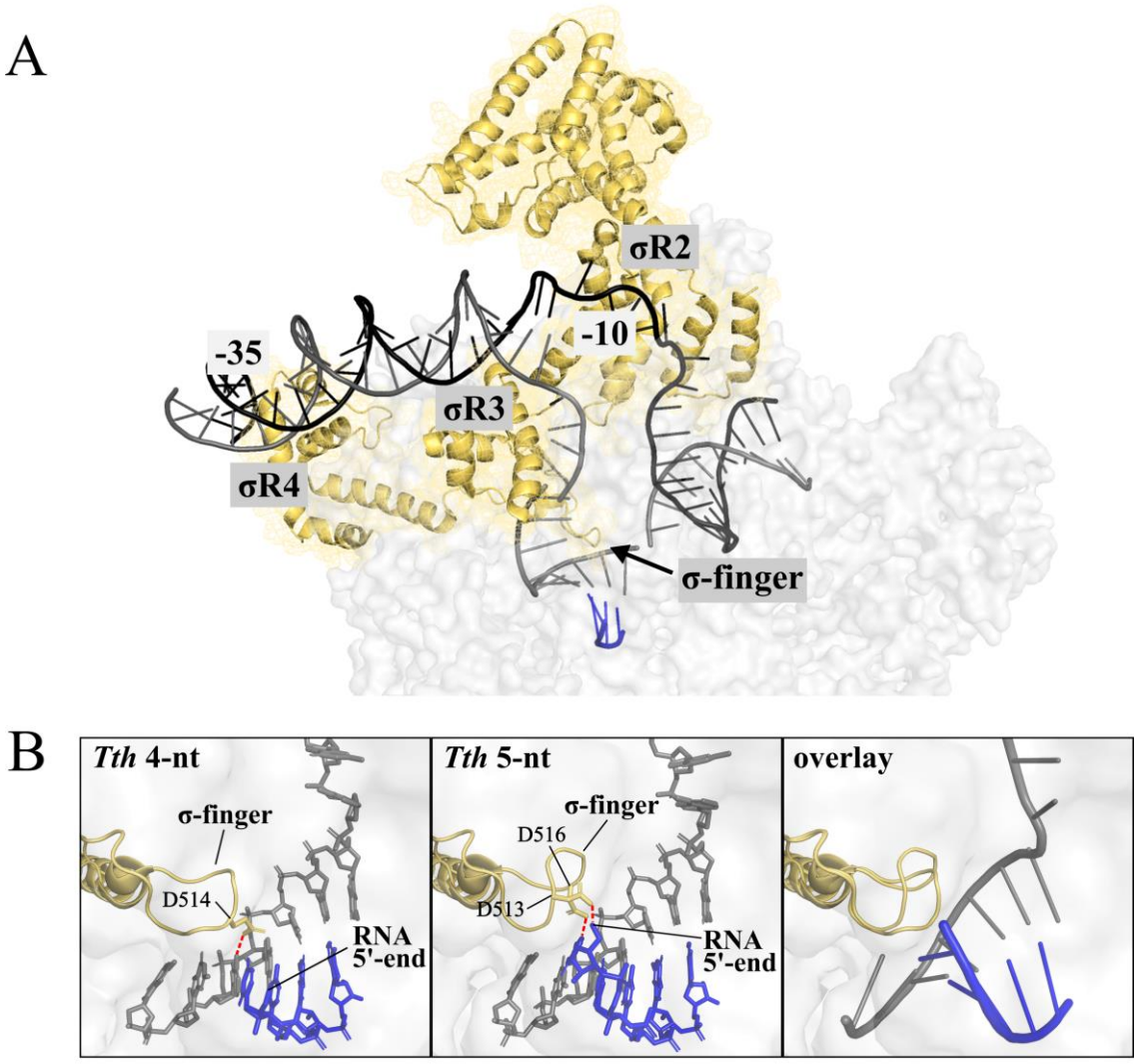

Figure 1.3: Crystal structures of RNAP initial transcribing complexes. **(A)** Initial transcribing complex of *E. coli*. The σ-factor is coloured yellow-orange, RNA is blue, template DNA is grey, non-template DNA is black, RNAP core is grey. (PDB 4YLN; Zuo and Steitz, 2015). **(B)** Initial transcribing complex of *Thermus Thermophilus* (*Tth*) containing a 4-mer RNA (PDB 6KQE; Li et al., 2020) and a 5-mer RNA (PDB 6KQF; Li et al., 2020). The overlay shows an alignment of the two structures and shows the tip of the σ-finger folding back over itself in response to the growing RNA chain. Colours as in (A) with H-bonds between the σ-finger and DNA or RNA as red dashed lines. Non-template DNA is hidden for clarity. (PDB 6KQE and 6KQF; Li et al., 2020).

in which ~13bp of the DNA is unwound and separated to form the transcription bubble (Fig. 1.2C). During this process, structural studies show that the σ-finger makes contacts with DNA template bases (-4 and -3), pre-organising the ssDNA into a helical conformation (Y. Zhang et al., 2012). Biochemical assays show that this indirectly facilitates the binding of initial NTPs and formation of the first phosphodiester bond (Kulbachinskiy and Mustaev, 2006; Pupov et al., 2014).

Next steps of initial transcription involve initiation of RNA synthesis where following each nucleotide addition cycle, the nascent RNA transcript either remains bound as a DNA-RNA hybrid and a relative translocation of RNAP active-centre by 1bp takes place by DNA scrunching in which downstream DNA is pulled into the RNAP active-centre (Fig. 1.2D; Kapanidis et al., 2006; Revyakin et al., 2006) – the productive pathway, or the nascently formed RNA product is released and the complex returns to RPo – the abortive pathway (Carpousis and Gralla, 1980; Gralla et al., 1980; Hsu, 2009). In the productive pathway, the nascent RNA chain typically grows up to around 9- to 11-nt until energy stored in the scrunched DNA is used to break previously established RNAP-promoter contacts and promoter escape into elongation occurs. The exact length of the RNA required for promoter escape depends on the promoter (Henderson et al., 2017).

The σ-finger is implicated to play a crucial role in initial transcription and promoter escape. Made up of thirteen amino acid residues (510 to 522) from the σR3/4 linker, it forms a hairpin-like loop extruding towards and close to the RNAP active-site, blocking the RNA exit channel (Murakami et al., 2002; Y. Zhang et al., 2012; Zuo and Steitz, 2015). When the RNA reaches 4-nt in length, the σ-finger interacts with the DNA-RNA hybrid through H-bonds and van der Waals interactions, likely as a stabilising influence (Li et al., 2020).

However, the σ-finger also provides a steric barrier to the growing RNA chain due to its placement along the RNA exit channel, with steric collision of the σ-finger and 5'-end of the nascent RNA estimated from structures to occur when the nascent RNA is 5- or 6-nt long (Murakami et al., 2002). For further RNA extension and therefore promoter escape, the σ-finger must be displaced from the exit channel (Li et al., 2020; Murakami et al., 2002; Shin et al., 2021; Winkelman et al., 2015). This is supported by biochemical studies showing that mutations of the σ-finger residues alter the abortive initiation profile to bias for longer transcripts, whilst deleting the σ-finger altogether eliminates all abortive initiation but causes defects in promoter escape (Cashel et al., 2003; Kulbachinskiy and Mustaev, 2006; Murakami et al., 2002; Pupov et al., 2014). Similar changes in the abortive initiation profile were observed for alternative σ-factors, which contain a structurally homologous module to the σ^{70}-finger (Oguienko et al., 2021). Single-molecule studies reveal that after a 6-mer RNA chain has been synthesised, the on pathway complex may enter a paused state for ~20s (Duchi et al., 2016; Dulin et al., 2018). The pause is greatly decreased in complexes with deleted σ-finger residues, indicating the σ-finger to be one of the pause lifetime-determining factors and highlighting its regulatory role in RNA synthesis.

Displacement of the σ-finger in the productive pathway has never been observed directly but recent structures come close, showing that before displacement, the tip of the finger gradually folds back over itself away from the active-site in a stepwise response (Fig. 1.3B; Li et al., 2020). However, full displacement required for promoter escape is still not reported and these structures were unable to capture kinetic information – providing motivation for a single-molecule study of the details and mechanism of σ-finger

displacement explored in this thesis. Additionally, modules functionally similar to the σ-finger exist in all domains of life. These modules interact with the template strand DNA and block the path of the nascent RNA, likely operating in a similar fashion to the prokaryotic σ-finger (Li et al., 2020). In archaeal RNAP this module is the TFB zinc ribbon and CSB (Fig. 1.4A; Renfrow et al., 2004) in eukaryotic RNAP I, the Rrn7 zinc ribbon and B-reader (Fig, 1.4B; Engel et al., 2017; Han et al., 2017); in eukaryotic RNAP II, the TFIIB zinc ribbon and B-reader (Fig.1.4C; He et al., 2016; Kostrewa et al., 2009; Liu et al., 2010; Plaschka et al., 2016); and in eukaryotic RNAP III, the Brf1 zinc ribbon (Fig, 1.4D; Abascal-Palacios et al., 2018; Vorländer et al., 2018). Mechanistic insight provided by a single-molecule study on the prokaryotic σ-finger could therefore also support studies of archaeal and eukaryotic transcription.

Structural studies estimate that the RNA chain will emerge from the end of the exit channel at a length of 16-nt. Aside from the σ-finger, this nucleotide length also implies the displacement of σR4 from its original holoenzyme position and suggests that the 5'-end of RNA must compete with both the σ-finger and σR4 to progress to elongation (Nickels et al., 2005). This is supported by only rare observations of abortive transcripts above 16 nucleotides in length (Hsu, 2002; Vo et al., 2003).

1.2 Single molecule studies of transcription

The first visualisations of bacterial transcription on the single-molecule level came from a tethered particle motion study that monitored the motions of gold particles attached to DNA as it was being transcribed by T7 RNAP (Schafer et al., 1991). Since then, mechanistic

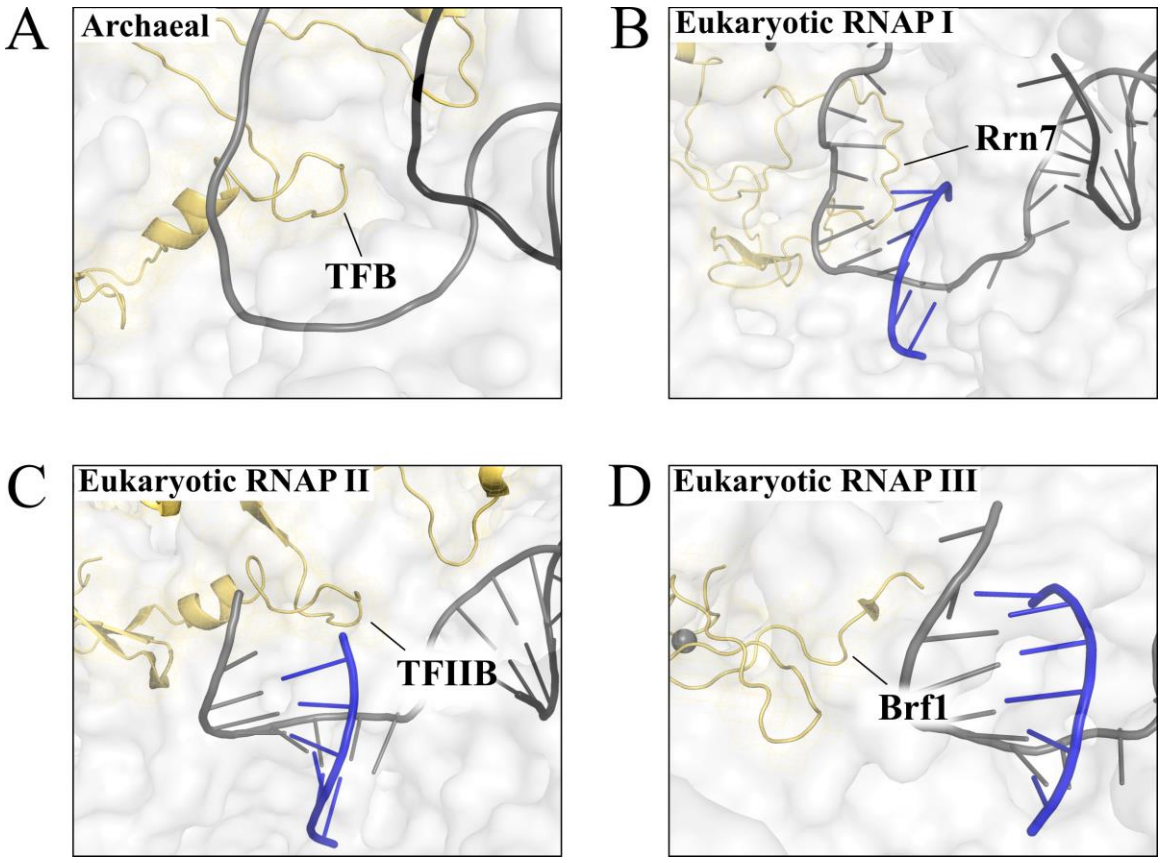

Figure 1.4: Structures of archaeal and eukaryotic RNAP complexes with modules equivalent to the prokaryotic σ-finger labelled. **(A)** Archaeal RNAP. Model from Nagy et al. (2015). **(B)** Eukaryotic RNAP I (PDB 5W65; Han et al., 2017). **(C)** Eukaryotic RNAP II (PDB 4BBS; Sainsbury et al., 2013). **(D)** Eukaryotic RNAP III (PDB 6F41; Vorländer et al., 2018).

models in transcription have been greatly advanced by single-molecule techniques, especially single-molecule fluorescence and DNA nano-manipulation.

Fluorescent-based assays directly visualise transcriptional mechanisms through fluorescent labelling and detection strategies. Colocalisation single-molecule spectroscopy (CoSMoS) and DNA curtain assays involve fluorescent labelling of both RNAP and DNA to study the RNAP promoter-search mechanism (Cha and Lee, 2021; Friedman et al., 2013; Gourse and

Landick, 2012; F. Wang et al., 2013). Single-molecule FRET (smFRET), the method used in this thesis, acts as a 'spectroscopic ruler' between two fluorescent molecules, one donor and one acceptor, for nanoscale distances (~2-10 nm) and is extremely well-suited to detect the structure-function relationships underlying transcription (Ha et al., 1996). Labelling systems typically involve either double-labelled DNA (DNA-DNA), double-labelled RNAP (RNAP-RNAP) or a single-label on both DNA and RNAP (RNAP-DNA). Important advances using smFRET relevant to *E. coli* RNAP include the identification of DNA scrunching in initial transcription using a combination of RNAP-DNA and DNA-DNA labelling schemes (Kapanidis et al., 2006), the discovery of transcription initiation pausing using DNA-DNA labelling schemes to interrogate highly stable scrunched intermediates (Duchi et al., 2016; Dulin et al., 2018) and the finding of σ^{70} retention during transcription elongation using RNAP-DNA labelling (Kapanidis et al., 2005; Mukhopadhyay et al., 2001), to name a few. Similar smFRET assays using DNA-DNA labelling have also observed DNA scrunching and the complete pathway from initiation to elongation in T7 RNAPs (Koh et al., 2018; Tang et al., 2008, 2009). Difficulties in obtaining structures of eukaryotic RNAPs, due to their large size and diverse composition from the substantial number of transcription factors involved, led to the development of the Nano-Positioning System (NPS; Michaelis and Treutlein, 2013; Muschielok et al., 2008; Muschielok and Michaelis, 2011). In combination with smFRET, NPS triangulates the position of an 'antenna' dye labelled at an unknown position using three or more 'satellite' dyes labelled at known positions – and has enabled the course of nucleic acids in eukaryotic Pol II open and elongation complexes to be determined (Andrecka et al., 2008, 2009; Muschielok et al., 2008; Treutlein et al., 2012). Archaeal RNAP complexes have

also been analysed in this fashion (Grohmann et al., 2011; Michaelis and Treutlein, 2013; Nagy et al., 2015).

Recently, protein induced fluorescence enhancement (PIFE) of cyanine dyes, typically Cy3 and Cy5, has been exploited to probe distances <3nm on the single-molecule level (smPIFE), which FRET interactions are insensitive to (Hwang et al., 2011; Hwang and Myong, 2014; Ploetz et al., 2023). Sorokina et al. (2009) used the high sensitivity of smPIFE to distinguish between different RNAP complexes with closed and open transcription bubbles and Koh et al. (2018) combined smFRET and smPIFE schemes in one assay to monitor RNAP molecules during the entirety of transcription initiation including abortive initiation cycles and promoter escape with near base-pair resolution – both studies using T7 RNAP. In addition to smPIFE, Cy3 labelled in the region of the transcription bubble has been reported to exhibit a ~two-fold increase in fluorescence intensity upon RPo formation (Feklístov et al., 2017; Ko and Heyduk, 2014; Koh et al., 2018). Mazumder et al. (2021) adapted this for single-molecule unwinding-induced fluorescence enhancement (smUIFE) studies to complement a smFRET assay monitoring the *E. coli* RNAP clamp and propose a mechanism for open complex formation. To directly monitor RNA synthesis, fluorescently-modified NTPs, FISH-like and CoSMoS strategies have also been developed (Friedman and Gelles, 2012; Gueroui et al., 2002; Z. Zhang et al., 2014).

Complementary single-molecule approaches to fluorescence methods include force-based assays such as optical and magnetic tweezers. Optical tweezers allow the measurement of RNAP translocation over large-scale distances along DNA beyond the FRET regime and

have therefore mainly been used to study transcription elongation (Heller et al., 2014; Ma et al., 2013; Neuman et al., 2003; M. D. Wang et al., 1998; Yin et al., 1995). Optical tweezers have also been combined with fluorescence assays to monitor RNAP interactions with DNA (Harada et al., 1999). Magnetic tweezers can study DNA topologies such as supercoiling and have been used to show DNA unwinding by RNAP with ~1bp resolution (Revyakin et al., 2004) and, independently to smFRET, identify the DNA scrunching mechanism of initial transcription (Revyakin et al., 2006).

This thesis focusses on using smFRET to underpin structure-function relationships and deduce mechanisms involved in regulating transcription initiation.

1.3 Single-molecule FRET (smFRET)

smFRET involves one donor and one acceptor labelled on the same complex (Ha et al., 1996).

When excited by a laser, the donor molecule transitions from the ground state into an excited state. After vibration relaxation, there are two major mechanisms by which the donor can directly return to the ground state: spontaneous emission and intramolecular non-radiative resonance energy transfer. The latter, FRET, only occurs if the energy required to excite a nearby acceptor molecule exactly corresponds to that of these mechanisms and couples the electronic states of the donor and acceptor molecules via a dipole-dipole interaction (Fig. 1.5). The probability of transfer depends on the physical separation of the two fluorophores. Using Fermi's golden rule, this is calculated per donor molecule as

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6}\ , \qquad 1.1$$

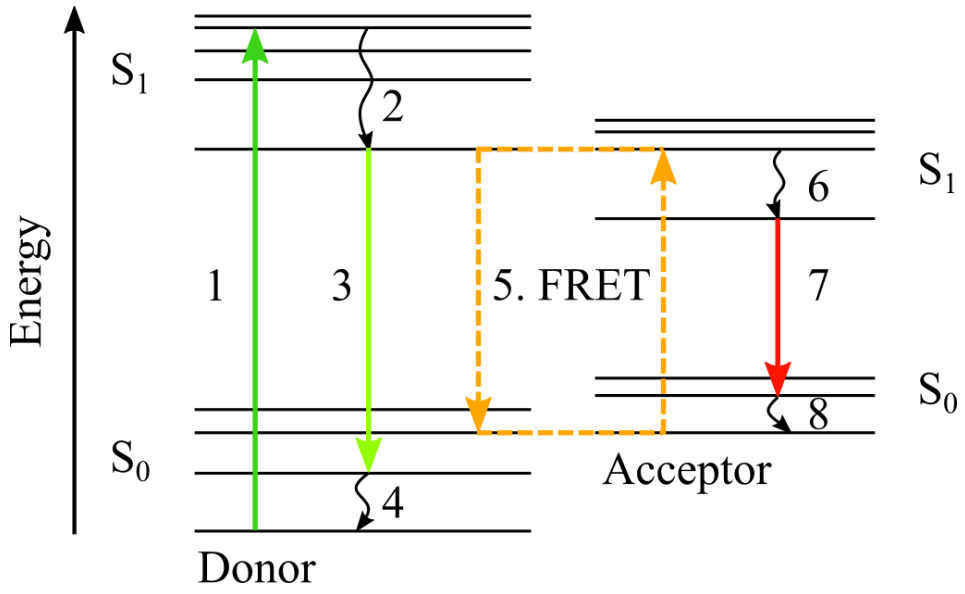

Figure 1.5: Energy level transitions of the donor and acceptor molecules. 1) direct excitation of the donor to an excited state, 2) vibrational relaxation of the donor, 3) spontaneous emission of the donor, 4) vibrational relaxation of the donor back to the ground state, 5) FRET, 6) vibrational relaxation of the acceptor, 7) spontaneous emission of the acceptor and 8) vibrational relaxation of the acceptor back to the ground state. Ground and excited levels are labelled S_0 and S_1 respectively.

where R is the donor-acceptor separation (Förster, 1948). The Förster radius, R_0, is the separation between the donor and acceptor for which energy transfer is equally likely as spontaneous emission. It can be calculated for each dye pair as

$$R_0 = \frac{9000 \cdot ln10 \cdot \varphi_D \cdot \langle \kappa^2 \rangle}{128\,\pi^5 \cdot n^4 \cdot N_A} \cdot \int_0^\infty f_D(\lambda)\varepsilon_A(\lambda)d\lambda\,. \qquad 1.2$$

φ_D is the donor quantum yield, n the refractive index of the medium and N_A is Avogadro's constant. κ^2 describes the relative orientation of the donor and acceptor dipoles moments and is calculated as $\langle \kappa^2 \rangle = 2/3$ for freely rotating fluorophores. $f_D(\lambda)$ is the donor emission spectrum and $\varepsilon_A(\lambda)$ the acceptor absorption spectrum. The FRET pair used in this thesis, Cy3B-Alexa647, has a Förster radius of ~60Å (Duchi et al., 2018).

Whilst the FRET efficiency only requires direct excitation of the donor molecule, an alternating laser excitation (ALEX) scheme has been developed to report on the labelling of individual molecules (Kapanidis et al., 2004). The scheme requires the alternation of two lasers, one each for direct excitation of donor and acceptor, from which three intensities are extracted: the donor emission upon donor excitation (I_{DD}), the acceptor emission upon donor excitation (I_{DA}) and the acceptor emission upon acceptor excitation (I_{AA}). A non-detector corrected stoichiometry ratio S between these intensities is calculated as

$$S = \frac{I_{DA} + I_{DD}}{I_{DA} + I_{DD} + I_{AA}} \tag{1.3}$$

and the expression used to calculate and the apparent FRET efficiency E* is written as

$$E^* = \frac{I_{DA}}{I_{DA} + I_{DD}}. \tag{1.4}$$

E* ≈ 1 and S ≈ 0 denotes acceptor-only labelling and E* ≈ 0 and S ≈ 1 denotes donor-only labelling. Molecules of interest are labelled with one donor and one acceptor. They are selected for by mid-range S ratios and single donor and acceptor photobleaching steps, which are the hallmark of single fluorophores.

The values reported in this thesis are apparent FRET efficiencies (E*) and stoichiometries (S) which are not corrected for leakage of donor fluorescence into the acceptor channel and direct excitation of the acceptor at the donor excitation wavelength.

To make smFRET observations, this thesis uses confocal and total internal reflection fluorescence (TIRF) microscopy (Axelrod et al., 1983). Confocal microscopes use a

focused excitation laser, a pinhole in the emission path to remove background fluorescence and avalanche photodiodes (APDs) for emission detection from freely diffusing molecules. TIRF microscopes detect surface-immobilised molecules and reject background fluorescence by using an evanescent wave to excite fluorophores. This wave is generated at the slide-sample interface when the excitation angle, α, is greater than the critical angle required for total internal reflection given by

$$\alpha > sin^{-1}\left(\frac{n_{water}}{n_{glass}}\right) \approx 63° \quad 1.5$$

where $n_{water} = 1.33$ and $n_{glass} = 1.5$. The evanescent field generated has a penetration depth, d ~300nm, calculated by

$$d = \frac{\lambda}{4\pi}\left(n_{glass}^2 \, sin^2 \, \alpha - n_{water}^2\right)^{-1/2} \quad 1.6$$

where λ is the wavelength of incident light (Axelrod et al., 1983). To detect the resulting fluorescence, TIRF microscopes typically use an EMCCD.

Instead of directly transitioning to the ground state after excitation, fluorophores may also enter a triplet state, which decreases the rate of fluorescence and increases the probability of photobleaching. Triplet quenchers such as Trolox are therefore used in smFRET assays to reduce triplet state lifetimes (Rasnik et al., 2006).

The following chapter provides more detail on the materials and methods used in this thesis.

Chapter 2: Materials and Methods

2.1 Accessible volume calculations

Accessible volume estimations were carried out using FRET-restrained positioning and screening (FPS) software (Kalinin et al., 2012). This uses structural modelling to geometrically assess the spatial distribution for each dye to estimate distances and FRET efficiencies.

To assess dye labelling positions on the σ-finger, PDB files 6KQD, 6KQE and 6KQF of Tth-σ^A initial transcribing complexes were used. The labelled σ^A residue was 321 Ile (which corresponds to σ^{70} residue 511). For FRET distance estimations, PDB files 4YLN, 7KHB, 7MKD and 7MKE of *E.coli*-σ^{70} transcription complexes were used with labelling at σ^{70} residues 366 Ser and 511 Ile.

To model dye distributions, amino acid side chains on σ^A residues 321 and 174 (6KQD, 6KQE and 6KQF) and on σ^{70} residues 511 and 366 (4YLN, 7KHB, 7MKD and 7MKE), were deleted in all structures. The dye attachment point used was Cα. Dye parameters for accessible volume measurements were as follows: Alexa647 maleimide (linker length = 21Å, linker width = 4.5Å and dye radii = 11.0Å, 4.7Å, 1.5Å) and Cy3B maleimide (linker length = 9.1Å, linker width = 4.5Å and dye radii = 7.7Å, 2.5Å, 1.3Å). Average distances and corresponding FRET values were then calculated using a Förster radius of 60Å.

2.2 Preparation of reagents

σ^{70} and RNAP holoenzyme

The double-labelled (DL) σ^{70}, single-labelled σ^{70}, DL and single-labelled hexahistidine-

tagged RNAP holoenzymes and non-labelled hexahistidine-tagged holoenzyme were prepared by Dr. Abhishek Mazumder.

Double cysteine modified σ^{70} derivatives were prepared as follows: single colonies of E. coli strain BL21(DE3) (Millipore) were co-transformed with a pGEMD (-Cys) derivative encoding two cysteine residues at positions 366 and 511 constructed from plasmid pGEMD (-Cys)46 by use of site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit, Agilent) to replace codons 366 and 511 by a codon encoding cysteine residue, and used to inoculate 20ml LB broth containing 100μg/ml ampicillin. Cultures were incubated 16h at 37°C with shaking. Culture aliquots (2x10ml) were used to inoculate LB broth (2x1L) containing 100μg/ml ampicillin. Cultures were incubated at 37°C with shaking until OD600 = 0.7; IPTG was added to 1mM. Cultures were further incubated for 4h at 37°C with shaking. Cells were harvested by centrifugation (5,000g; 20 minutes at 4˚C), re-suspended in 50ml lysis buffer (40mM Tris–HCl pH 7.9, 300mM NaCl, 1mM EDTA, one protease inhibitor cocktail tablet, and 0.2% deoxycholate), and lysed by emulsification (Emulsiflex-C5; Avestin, Inc., Ottawa, Canada). Inclusion bodies containing σ^{70} derivatives were isolated by centrifugation (10,000g; 20 minutes at 4°C), washed with 20ml lysis buffer containing 0.2mg/ml lysozyme and 0.5% Triton X-100, and washed with 20ml lysis buffer containing 0.5% Triton X-100 and 1 mM DTT (with each wash step involving sonication 2 × 1 minute at 4°C in wash buffer, incubation 10 minutes at 4°C in wash buffer, and centrifugation – 10,000g; 20 minutes at 4°C). Washed inclusion bodies containing σ^{70} derivatives were solubilised in 40ml 6M guanidine–HCl, 50mM Tris–HCl (pH 7.9), 10mM MgCl2, 10μM $ZnCl_2$, 1mM EDTA, 10mM DTT, and 10% glycerol, and

dialyzed against 2L TGED (20mM Tris–HCl pH 7.9, 0.1mM EDTA, 0.1mM DTT, and 5% glycerol) containing 0.2M NaCl (20h at 4°C; two changes of buffer). The sample was centrifuged (10,000g; 20 minutes at 4°C) to remove particulates and applied to a Mono-Q HR 10/10 column (Amersham-Pharmacia Biotech, Piscataway, NJ) pre-equilibrated in the same buffer. The column was washed with 16ml of the pre-equilibration buffer and eluted in 2ml fractions of a 160ml linear gradient of 200–600mM NaCl in TGED (with σ^{70} derivatives typically eluting at ~360mM NaCl in TGED). Fractions containing the σ^{70} derivative were identified by SDS–PAGE and Coomassie staining and are pooled. Pooled fractions were concentrated in storage buffer (20mM Tris-HCl, pH 7.9, 200mM NaCl, 0.1mM EDTA, 1mM TCEP, and 5% glycerol) using 10kDa MWCO Amicon Ultra-15 centrifugal ultrafilters, and stored in aliquots at -80°C.

Fluorescent-probe-labelled, DL σ^{70} derivatives were prepared as follows: A reaction mixture containing 10μM σ^{70} derivative (with cysteines at position 366 and 511), 500μM Cy3B maleimide and 500μM Alexa647 maleimide in 0.5ml buffer B (100mM potassium phosphate buffer pH 8.0, 50mM NaCl, 1mM EDTA, 5mM TCEP, and 2% dimethylformamide) was incubated 4h on ice, subjected to 5 cycles of buffer exchange, dilution with 5ml buffer B without dimethyl formamide, followed by concentration to 0.5ml using 10kDa MWCO Amicon Ultra-15 centrifugal ultrafilters (EMD Millipore), and stored in aliquots at -80°C.

Labelling efficiencies were determined from UV/Vis-absorbance measurements as 88% and 75% for Cy3B and Alexa647 respectively. These products are stochastically labelled and therefore contain a mixture of σ^{70} derivatives labelled at residues 366 and 511 with

either two Cy3B fluorophores, two Alexa647 fluorophores, one Cy3B fluorophore labelled at residue 366 and one Alexa647 fluorophore labelled at 511, or one Cy3B fluorophore labelled at residue 511 and one Alexa647 fluorophore labelled at residue 366.

Single-labelled σ^{70} was prepared in an analogous fashion to the DL σ^{70} but with only one cysteine mutation introduced into the σ^{70}-PGEMD (-cys) construct at residue 511. The resulting σ^{70}-511C was labelled with Cy3B to form single-labelled σ^{70}.

Labelled hexahistididine-tagged holoenzyme (DL holoenzyme or single-labelled holoenzyme) was formed by incubating RNAP core with DL σ^{70} or single-labelled σ^{70} in a 1:2 ratio in KG7 buffer (40mM HEPES-NaOH, pH 7.0, 100mM potassium glutamate, 10mM $MgCl_2$, 1mM dithiothreitol and 5% glycerol) at 30°C for 30 mins, and then on ice for 12 hours.

Non-labelled hexahistidine-tagged *E.coli* RNAP holoenzyme was prepared as described in Mazumder et al. (2021).

Nucleic acids

DNA oligos were purchased from Biomers and IDT. Oligo sequences and modifications are stated in Table 1.

Double-stranded DNA was annealed by heating non-template oligos and template oligos (in the ratio of 1:1.1 when the non-template oligo is biotinylated and in the ratio 1:1 for nonbiotinylated oligos) in buffer (10mM-Tris-HCl, pH 7.9 and 0.2M NaCl) for 5 minutes at 95°C. The solution was then cooled to 25°C in 2°C steps with 1 minute per step in a thermal cycler (Applied Biosystems).

lacCONS+12C construct for measurements in Figs. 4.2, 4.3, 4.4 (RPitc2, RPitc4, RPitc7, RPitc11), 4.5, 4.6 (RPitc2, RPitc4, RPitc11)	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGTGAGAGCGGATAACAATTTC-3’
Template strand	5- GAAATTGTTATCCGCTCTCACAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+6A+10C construct for measurements in Figs. 4.4 (RPitc5, RPitc9), 4.6 (RPitc9)	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGAGAGCGCGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGCGCTCTCAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+6C construct for measurements in Fig. 4.6 (RPitc5)	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGCGAGAGCGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGCTCTCGCAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT -3’
lacCONS+7C construct for measurements in Figs. 4.4 (RPitc6), 4.6 (RPitc6)	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGTCAGAGCGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGCTCTGACAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+8C construct for measurements in Fig. 4.6 (RPitc7)	

Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGTGCGAGCGGATAACAATTTC-3’
Template strand	5’-GAAATTGTTATCCGCTCGCACAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+9C construct for measurements in Figs. 4.4 (RPitc8), 4.6 (RPitc8)	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGTGACAGCGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGCTGTCACAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+11C construct for measurements in Figs. 4.4 (RPitc10), 4.6 (RPitc10)	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGTGAGACCGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGGTCTCACAATTCCACACATTAT ACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+15C construct for measurements in Figs. 4.4 (RDe14), 4.6 (RDe14), 4.9, 4.10, 4.11, 5.9B, 5.9C	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATA ATGTGTGGAATTGTGAGGAGGACGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGTCCTCCTCACAATTCCACACAT TATACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+12C (-15NT/+20T) construct for measurements in Figs. 4.7	
Non-template strand	5’-AGGCTTGACACTTTATGCTTCGGC(T/Cy3B)CGTATA ATGTGTGGAATTGTGAGAGCGGATAACAATTTC-3’

Template strand	5’- GAAAT(T/Atto647N)GTTATCCGCTCTCACAATTCCACACATTATACGAGCCGAAGCATAAAGTGTCAAGCCT-3’
lacCONS+12C (-25T) construct for measurements in Fig. 4.12	
Non-template strand	5’-(biotin)AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGAGCGGATAACAATTTC-3’
Template strand	5’- GAAATTGTTATCCGCTCTCACAATTCCACACATTATACGAGCCGAAGCA(T/Atto647N)AAAGTGTCAAGCCT-3’
rrnB P1 construct for measurements in Figs. 5.5, 5.7, 5.8A, 5.9D	
Non-template strand	5’-(biotin)CTCTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACAC-3’
Template strand	5’- GTGTTTGCCGTTGTTCCGTGTCAGTGGTGGCGCATTATA GGGAGTTATTCCGGCCTGACAAGAG-3’
rrnB P1 (-15NT/+20T) construct for measurements in Figs. 5.6, 5.8B, 5.8C	
Non-template strand	5’-CTCTTGTCAGGCCGGAATAACTCC(T/Cy3B)TATAATGCGCCACCACTGACACGGAACAACGGCAAACAC-3’
Template strand	5’- GTGT(T/Atto647N)TGCCGTTGTTCCGTGTCAGTGGTGGCGCATTATAAGGAGTTATTCCGGCCTGACAAGAG-3’
pR construct for measurements in Figs. 5.3, 5.4, 5.9A	
Non-template strand	5’-(biotin)ATCTATCACCGCAAGGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCTCTGGCGGTGATAATGGTTGCATGTAGTAAGGAGGTGGTATGGAAT-3’

Template strand	5’- ATTCCATACCACCTCCTTACTACATGCAACCATTAT CACCGCCAGAGGTAAAATAGTCAACACGCACGGTGTT AGATATTTATCCCTTGCGGTGATAGAT-3’

Table 1: DNA oligos used for each experiment

Small molecules

NTPs (New England Biolabs) and dinucleotides (Trilink, Biolog LSI GmbH) were diluted in nuclease-free water (Ambion) to a final concentration of 100mM or 10mM and stored at -20°C.

RPo

For experiments monitoring the σ-finger, RPo was formed by incubating 15nM DL or single-labelled RNAP holoenzyme with 10nM biotinylated promoter DNA in KG7 buffer for 20 minutes at 37°C, with the exception of experiments in Chapter 5.5 where incubation was performed at 22°C.

For experiments monitoring the DNA transcription bubble, concentrations of 10nM RNAP holoenzyme and 30nM double-labelled promoter DNA were used.

Incubation was followed by heparin challenge using 250μg/ml of heparin sepharose and further incubation for 1 minute, after which the mixture was centrifuged and supernatant collected.

2.3 Chapter 3: smFRET using confocal microscopy

Data collection and microscopy

A custom epifluorescence confocal scanning ALEX microscope described in Robb et al. (2013) was used. Green (532nm) and red (638 nm) lasers modulated at 10kHz were coupled into a single-mode fibre, collimated and coupled into a 60x oil-immersion objective (NA 1.35, Olympus). The fluorescence emission was separated from the excitation by a dichroic mirror, focussed onto a 200µm pinhole and spectrally split by two filters (585DF70, 650LP) onto two avalanche photodiodes (SPC-AQR-14).

Data were collected by adding ~100pM of DL σ^{70} diluted in KG7 buffer to a No. 1.5 thickness coverslips (0.13-0.16mm). Laser powers of 125µW (532nm laser) and 55µW (638nm laser) were used.

Data analysis

Fluorescent bursts using the threshold criteria $10 < I_{DD} < 100$, $I_{DA} > 3$ and $I_{AA} > 50$ where I_{DD} is the donor emission upon donor excitation, I_{DA} is the acceptor emission upon donor excitation and I_{AA} is the acceptor emission upon acceptor excitation, were used to calculate E* and S values using equations 1.3 and 1.4 from Chapter 1.2 in the Seneca software package written by Kristopher Gryte. Values were binned, plotted and gaussian fits made in Python.

2.4 Chapter 4 and 5: smFRET using TIRF-ALEX

Observation chamber preparation

Imaging chambers were formed using silicone gaskets (6mm diameter, Grace Bio-Labs)

on No. 1.5 coverslips. The glass floor was then passivated with biotin-PEG (with the ratio 40:1 of PEG-SVA:biotin-PEG MW 5000, Laysan Bio. Inc.) and functionalised with NeutrAvidin (Sigma Aldrich) as in Mazumder et al. (2021) (Fig. 2.1A).

For experiments monitoring the DNA transcription bubble and the σ-finger in RNAP holoenzyme only, the glass was additionally treated with biotinylated anti-hexahistidine monoclonal antibody (Pena-His Biotin Conjugate; Qiagen) as in Duchi et al. (2016) (Fig. 2.1B).

RPo was added to the observation well at a concentration of 100pM and incubated at 22°C for 30-60s for immobilisation. Wells were then washed with 3x50ul KG7 to remove excess RPo molecules.

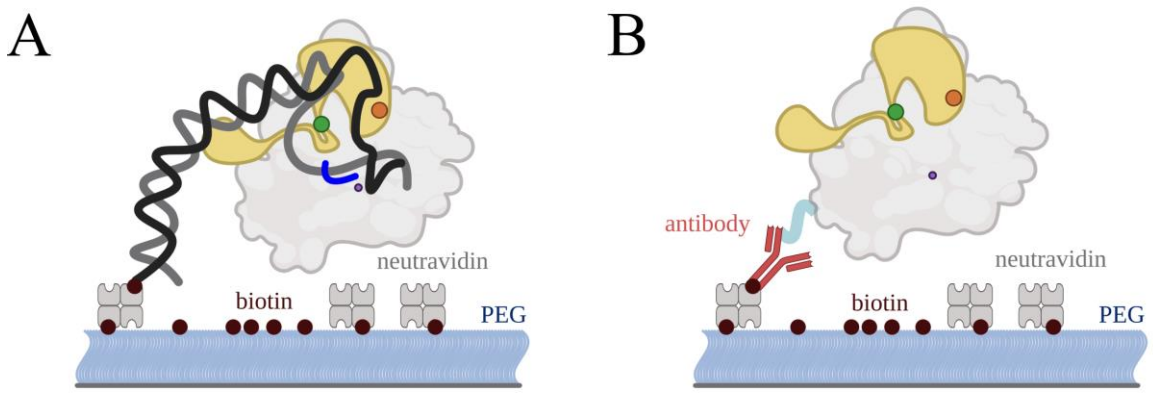

Figure 2.1: Surface-immobilisation strategies used. **(A)** Immobilisation via biotinylated DNA for initial transcription complexes containing DL σ^{70} **(B)** Immobilisation via a hexahistidine-tag on RNAP core. Created with BioRender.com.

Endpoint transcription assays

For observations of RPo, imaging buffer (KG7 buffer containing 2mM Trolox (Sigma-Aldrich), 1mg/ml glucose oxidase (Sigma-Aldrich), 40 ug/ml catalase (from bovine liver C30; Sigma-Aldrich), and 8mM D-glucose) was added to the observation well and data acquired.

For observations of RPitc2, imaging buffer additionally containing 500µM of dinucleotides was added to the observation well and data acquired.

For observations of artificially stalled initial transcribing complexes above RPitc2, a specified combination of dinucleotides (500µM) and NTPs (100µM) were added to the observation well and incubated for 5 minutes. Wells were washed with 5x50ul KG7 buffer and imaging buffer added. Data were then acquired.

Specifically for measurements of the lacCONS DNA transcription bubble in Chapter 4.5, data were acquired for RPitc11 as above, immediately after which wells were washed with 5x50ul KG7 and a 50ul aliquot of 1mM GTP/CTP incubated into the well for 5 minutes at 22°C to form RDe14. The well was then washed with 5x50ul KG7 buffer and data acquired for RDe14.

All transcription reactions and measurements were performed at 22°C using ALEX-TIRF. Data were collected for 50s at a frame rate of 100ms/frame (50ms ALEX) using laser powers of 0.7mW (532nm laser) and 0.3mW (635nm laser) to monitor the σ-finger and laser powers of 0.6mW (532nm laser) and 0.2mW (635nm laser) to monitor the DNA transcription bubble.

Real-time transcription assays

For real-time experiments in Chapter 4 involving the lacCONS promoter, RPo was immobilised and imaging buffer supplemented with 625µM dinucleotide (ApA or pppApA) added to the well and incubated for 2 minutes. Data acquisition was started and GTP/UTP/ATP added added into the well and mixed to give final concentrations of 500µM dinucleotide and 1mM NTPs.

For real-time experiments in Chapter 5 involving the pR promoter, experiments were performed in a similar fashion and imaging buffer added to the well. Data acquisition was then started and ApU and GTP/UTP/ATP/ mixed into the well to give final concentrations of 500µM ApU and 1mM NTPs

For real-time experiments in Chapter 5 involving the rrnB P1 promoter, experiments were performed in a similar fashion and imaging buffer added to the well. Data acquisition was then started and ApC and GTP/UTP/ATP/CTP mixed into the well to give final concentrations of 500µM ApU and 1mM NTPs.

All transcription reactions and measurements were performed at 22°C using ALEX-TIRF. Data were collected for 8min 20s at a frame rate of 400ms/frame (200ms ALEX) using laser powers of 190µW (532nm laser) and 80µW (635nm laser).

Objective-type TIRF Microscope

smFRET experimental data were collected using a custom-built objective-type TIRF microscope as described in Holden et al. (2010) (Fig. 2.2). Light from green (532nm, Cobolt) and red (635nm, Coherent) lasers were combined using a dichroic mirror and coupled into single-mode fibre. The output was focused onto the back focal plane of a 100x

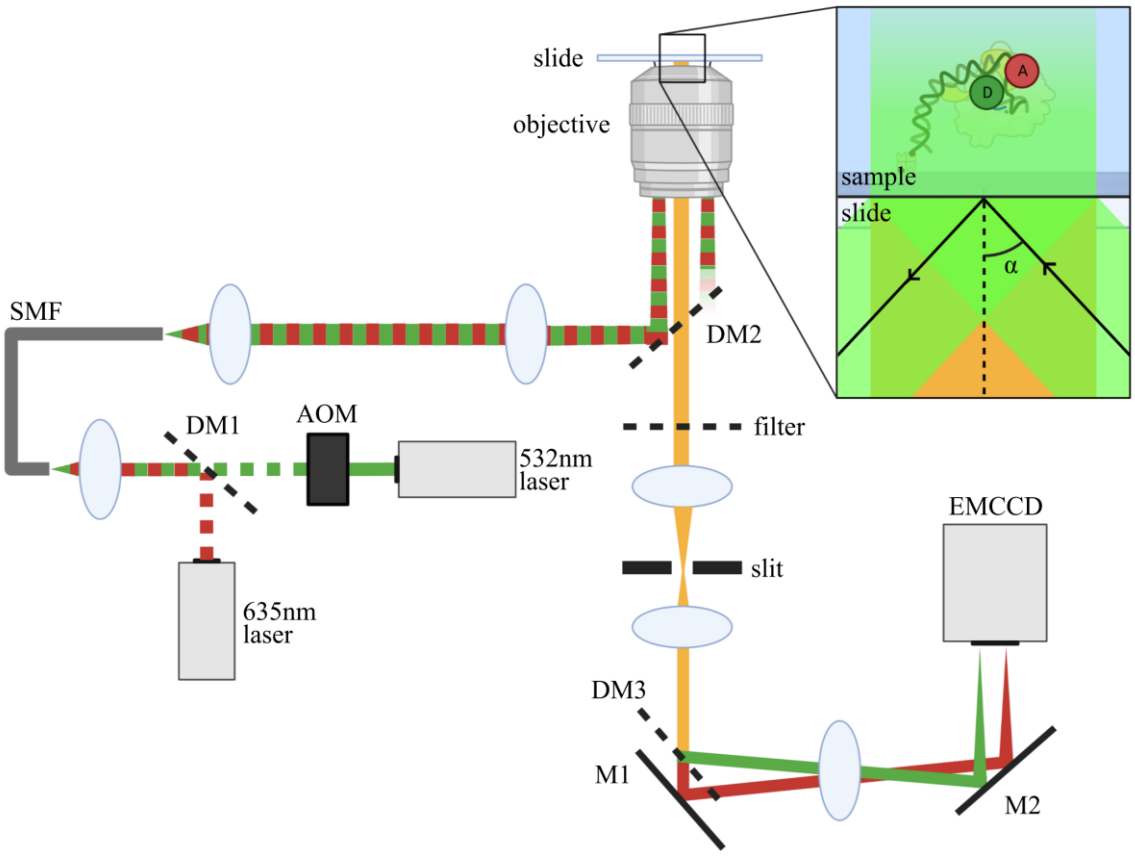

Figure 2.2: Schematic of the TIRF microscope used. Light from 532nm and 635nm lasers are combined using a dichroic mirror (DM1), coupled into a single-mode fibre (SMF), collimated and focussed onto the sample at incident angle (α) that allows for total internal reflection. Emission light is filtered (DM2 and filter), separated into red and green channels with a dichroic mirror (DM3) and imaged onto an EMCCD camera. M1 and M2 are mirrors. Created with BioRender.com.

oil-immersion objective (NA 1.4, Olympus) and laterally displaced off the optical axis such that the incident angle at the slide exceeded the critical angle and created an evanescent illumination wave. ALEX was achieved through an acousto-optical modulator (1205C, Isomet). A dichroic mirror (545nm/650nm, Semrock) and emission filters (545nm LP Chroma and 633/25nm notch, Semrock) separated the fluorescence emission from the excitation. The fluorescence emission was then focussed on a slit and spectrally separated

using a dichroic mirror (630nm DLRP, Omega) into two channels focussed onto an EMCCD (iXon 897, Andor). A motorised x/y-scanning stage (MS-2000, ASI) was used to mount the observation chambers.

Data analysis

Background corrected fluorescence-emission intensity vs. time trajectories were extracted using software package Twotone-ALEX (Holden et al., 2010). Intensity vs time trajectories were manually inspected to exclude trajectories exhibiting I_{DD}<100 or >1500 counts, I_{AA}<100 or >1000 counts, multi-step donor or acceptor photo-bleaching and donor or acceptor photo-blinking. Selected trajectories were then used to calculate E* and S trajectories using equations 1.3 and 1.4 in Chapter 1.2.

S values were used to distinguish between species containing both donor and acceptor fluorophores and species containing only one of either donor or acceptor fluorophore. E* histograms for species containing both donor and acceptor fluorophores were built and fitted to gaussian distributions using Origin (Origin Lab) to provide subpopulation percentages and mean E* values.

Errors in the low FRET 'displaced' σ-finger population percentage were calculated by finding the standard error in the mean of 'displaced' σ-finger subpopulation percentages between individual experimental datasets (or between three subsets of when the experiment was only performed once).

To identify dynamics and their time scales, E*-time trajectories were analysed using Hidden Markov Modelling (HMM) as implemented in the software package ebFRET (van de Meent et al., 2014). By maximising the lower bound, a three-state model was chosen to

fit dynamic E*-time trajectories monitoring the σ-finger. The three-state classified raw E* values were then plotted and fitted to gaussian distributions in Origin to extract population percentages and a mean E* value for each state. Dynamic molecules were defined as those with three or more transitions between different states. Dwell times were extracted and plotted as histograms and fit either to a single exponential or peaked exponential, as stated, in Origin.

Errors in the percentage of dynamic molecules were calculated by finding the standard error in the mean of the percentage of dynamic molecules between individual experimental datasets (or between three subsets of when the experiment was only performed once).

For real time experiments, all E*-time trajectories were analysed initially using HMM and then categorised into four groups as described in Chapter 4.6. Class-I and Class-II molecules, in which σ-finger displacement was observed, were then analysed using a four-state HMM, and Class-III molecules, which exhibited transitions between different FRET states, were analysed using a three-state HMM.

Displacement dwell times for Class-I and Class-II molecules were calculated by subtracting the estimated time of nucleotide addition from the time of σ-finger displacement, which was taken from HMM analysis. The time of nucleotide addition was estimated using traces that exhibited a sudden brief drop in intensity in both donor and acceptor channels due to the injection of nucleotides using a pipette. The times for σ-finger displacement were then plotted as dwell-time histograms, fitted to double exponential decays in Origin and the two time constants extracted. The σ-finger conformation before nucleotide addition was extracted from E*-time trajectories exhibiting σ-finger displacement (Class-I and Class-II). These traces were classified into either exhibiting

‘short’ or ‘long’ finger displacement using a boundary calculated from the time at which the two components of the double exponential decay intersect. The E* values before the estimated time of nucleotide addition was then extracted and a histogram produced from binned values. These were then fit to single gaussian distributions in Origin.

Chapter 3: σ-finger detection

3.1 Introduction

Structural studies offer high resolution static snapshots of the σ-finger conformation (Murakami et al., 2002; Y. Zhang et al., 2012; Zuo and Steitz, 2015). They show the σ-finger to be a small looped element (13 amino acids) positioned close to the active centre of RNAP, interacting with the DNA template strand at bases -3 and -4 (Fig. 3.1). This positioning had two important implications as was discussed in Chapter 1.1. First, template strand contacts suggested that the σ-finger pre-organises the DNA template strand in RPo formation. This is supported by biochemical assays (Kulbachinskiy and Mustaev, 2006; Pupov et al., 2014) and more recently structures, which show an entirely disordered template when the contacting σ-factor residues were deleted (Li et al., 2020). Second, the positioning of the σ-finger blocks the RNAP RNA exit channel. It was therefore theorised that for productive transcription, the σ-finger would have to be displaced from this position before or during initial transcription (Li et al., 2020; Murakami et al., 2002; Shin et al., 2021; Winkelman et al., 2015). Li et al. (2020) provided further insight, reporting structures showing the tip of the σ-finger folding back away from the active-site in a stepwise response to the growing RNA (Chapter 1 Fig. 1.2B). They suggest that the σ-finger serves as a 'protein spring' which stores energy required for the breakage of RNAP-promoter and RNAP-σ interactions during promoter escape, which acts in addition to the storage of energy by DNA scrunching. Whilst Li et al. (2020) propose that the growing nascent RNA provides the sole driving force for σ-finger folding, a cross-linking study by Winkelman et al. (2015) suggested that changing interactions between the DNA template strand, the σ-finger and RNAP resulting from DNA scrunching could also partially supply the force

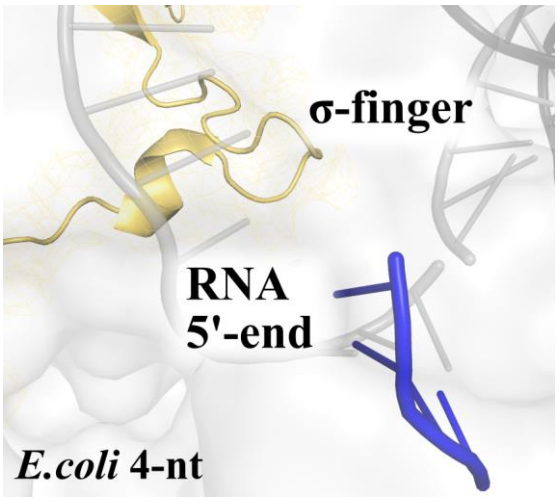

Figure 3.1: Initial transcribing complex of *E. coli* showing the σ-finger conformation. The σ-factor is coloured yellow-orange, RNA is blue, template DNA is grey, non-template DNA is black, RNAP core is grey. (PDB 4YLN; Zuo and Steitz, 2015).

required for displacement. These effects would not have been observed in the structural study by Li et al. (2020) either because the structures do not contain a sufficiently long RNA chain for full σ-finger displacement or due to the use of nucleic acid scaffolds which eliminated DNA scrunching, an important property of an initial transcribing complex.

Full displacement of the σ-finger away from the RNAP active-centre cleft in the context of a full promoter remains unobserved and as a result, the mechanism of σ-finger displacement during active transcription is unknown.

Current structural methods are unable to observe displacement of the σ-finger due to its flexibility and general difficulties with obtaining late initial transcription structures. smFRET provides an alternative means, able to define conformation dynamics and heterogeneity from early to late stages of transcription initiation of this flexible module (Chapter 1.3). These advantages have already been exploited to investigate other

transcription time points, reviewed in Chapter 1.2, such as the mechanism of open complex formation (Mazumder et al., 2021), the nucleotide addition cycle (Mazumder et al., 2020), DNA scrunching in transcription initiation (Kapanidis et al., 2006), transcription initiation pausing (Duchi et al., 2016; Dulin et al., 2018), promoter escape (Margeat et al., 2006) and σ-factor retention in elongation (Kapanidis et al., 2005).

Chapter summary

In this chapter, the double-labelled mutant of σ^{70}-factor required for smFRET studies of the σ-finger was designed, prepared and characterised. Initial confocal and TIRF-ALEX experiments agreed strongly with published structures and two new conformations of the σ-finger were found, providing new information about the heterogenous landscape of σ-finger conformations and dynamics.

3.2 σ-finger detection

smFRET was chosen to track σ-finger movements. This involved the attachment of two fluorescent probes, one at the σ-finger and the other at a known stationary point in initial transcription. In order to minimise the interference of the probes with σ-finger and RNAP function, accessible volume (AV) modelling using FRET Positioning and Screening (FPS) software of Cy3B was performed on high resolution crystal structures of initial transcribing complexes containing 3-, 4- and 5-nt of RNA (Fig. 3.2A; Kalinin et al., 2012; Li et al., 2020). This shows that attaching Cy3B at σ^{70} residue 511 (with a point mutation from Ile to Cys) should not interfere with interactions between the σ-finger and nascent RNA. The second probe attachment pointwas chosen to be σ^{70} residue 366 (with a point mutation from Ser to Cys) as this is known to be stable during initial transcription (Kapanidis et al.,

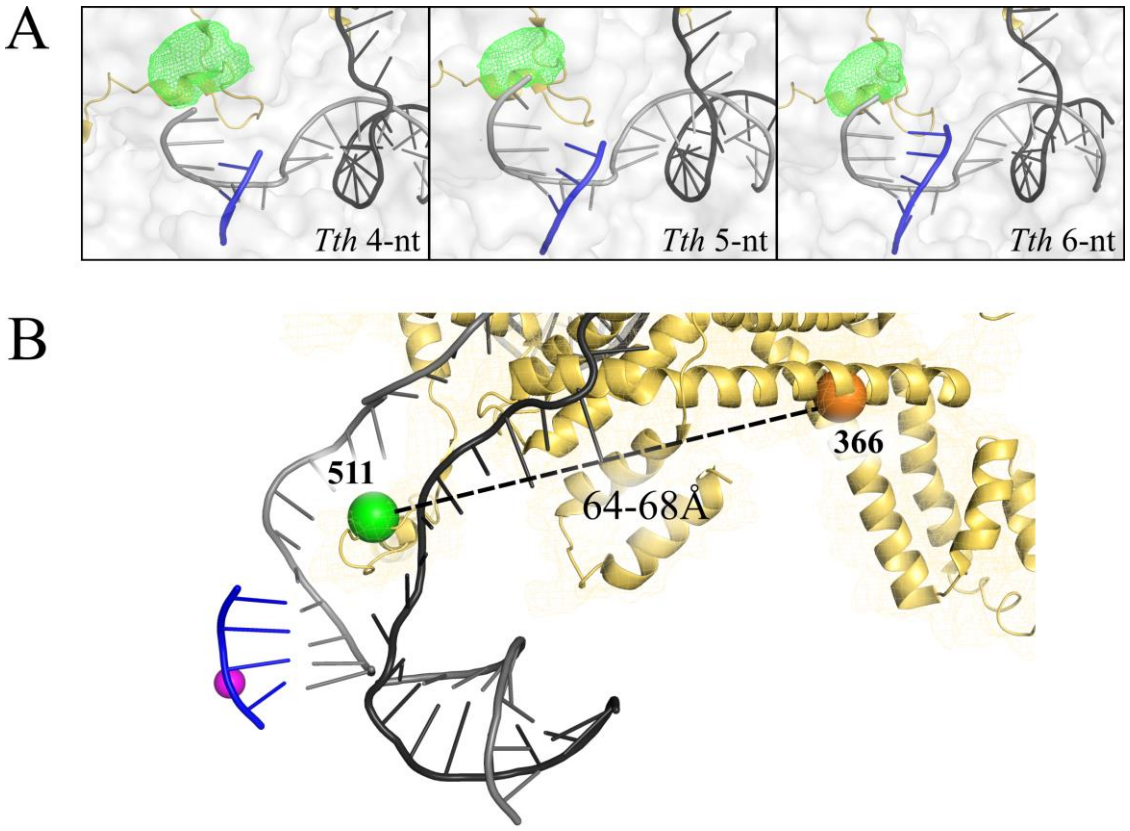

Figure 3.2: smFRET assay for detecting σ-finger movements in solution. **(A)** AV modelling of Cy3B attached at σ^{70}-factor residue 511, the base of the σ-finger. Cy3B accessible volume is green and other colours as in Fig. 3.1. (PDB 6KQD, 6KQE and 6KQE; Li et al., 2020). **(B)** smFRET dye labelling scheme with one probe at σ^{70}-factor residue 511 and the other probe at σ^{70}-factor residue 366 (green and dark orange spheres). (PDB 4YLN; Zuo and Steitz, 2015).

2005). The distance between Cy3B and Alexa647 attached at σ^{70} residues 366 and 511 was estimated to be 64-68Å using AV modelling on PDB structures 6C9Y, 4YLN, 7KHB, 7MKD and 7MKE. This corresponded to estimated FRET efficiencies E ~ 0.3-0.4 (Fig. 3.2B).

The designed double-labelled σ^{70} (DL σ^{70}) construct with fluorescent probes at σ^{70} residues 366 and 511 used in this thesis was prepared by Dr. Abhishek Mazumder (Department of Physics, Univeristy of Oxford) using stochastic labelling.

To experimentally determine whether labelling affected RNAP function, RiboGreen and *in vitro* transcription assays were performed. The RiboGreen measurements were performed by Dr. Abhishek Mazumder and showed that the transcriptional activity of complexes containing DL σ^{70} was >80% of the transcription activity of complexes containing unmodified σ^{70}. Similarly, the *in vitro* transcription measurements in Fig. 3.3, taken by Dr. Zakia Morichaud and Dr. Konstantin Brodolin (Institut de Recherche en Infectiologie de Montpellier, Université Montpellier – CNRS), show that the initial transcription profile of complexes containing DL σ^{70} is similar to that of those containing unmodified σ^{70}.

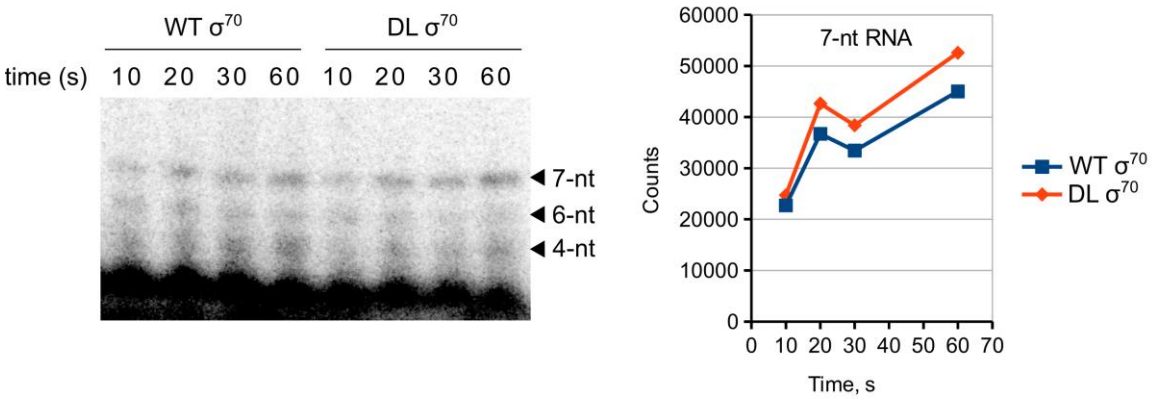

Figure 3.3: *In vitro* transcription activity for RPitc7 in complexes containing unmodified σ^{70} (WT σ^{70}) and double-labelled σ^{70} (DL σ^{70}) over 60s. RNAs were made under RPitc7 conditions (RPo + 500µM ApA + 100µM UTP + 100µM GTP) and [α-^{32}P]-UTP used. The graph shows quantification of the 7-nt band for WT and DL σ^{70}.

3.3 smFRET characterisation of the σ-finger in freely diffusing σ^{70}

To find the initial conformation of the σ-finger in freely diffusing σ^{70}-factor proteins, confocal ALEX spectroscopy on the DL σ^{70} construct was performed (Fig. 3.4; Materials and Methods 2.3). The apparent FRET efficiency (E*) burst data, with donor only and

acceptor only signals filtered out, displays a bimodal distribution with one major population at E* ~ 0.2 (66%) and one minor population at E* ~ 0.4 (34%).

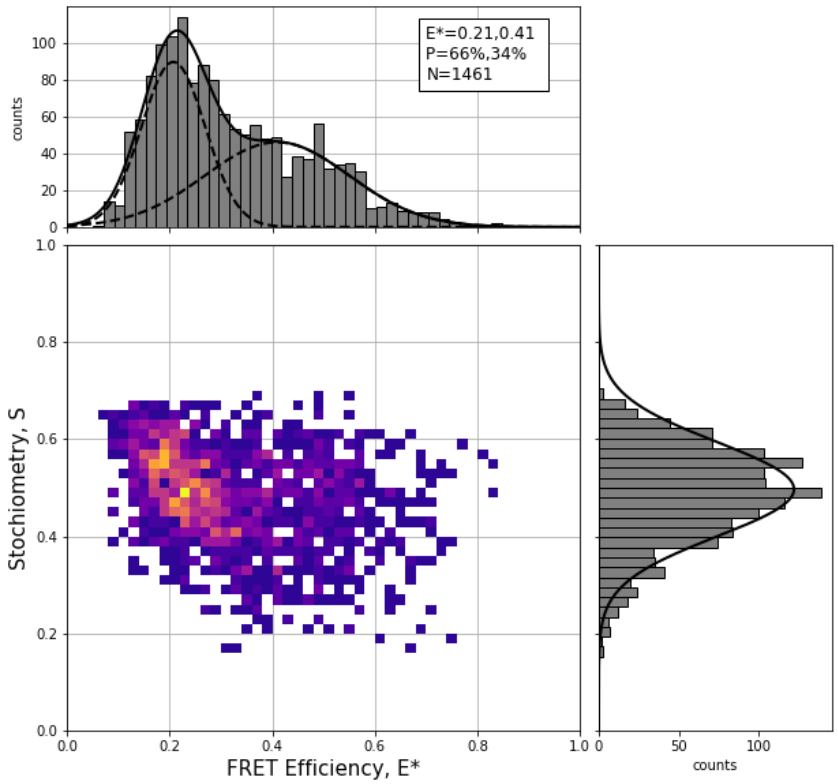

Figure 3.4: Conformational state of the σ-finger in freely diffusing σ^{70} labelled at residues 366 and 511 with Cy3B and Alexa647. Donor only and acceptor only signals are filtered out.

3.4 σ-finger conformation in RNAP holoenzyme

In the next stage of transcription, free σ^{70} associates with RNAP core to form RNAP holoenzyme. To investigate the σ-finger conformation in this complex, DL holoenzyme was formed and observed using TIRF-ALEX by immobilisation of complexes via a hexahistidine-tag onto glass slides functionalised with a penta-His antibody (Materials and

Methods 2.4). Fig. 3.5 shows controls for specific immobilisation and example fields of view. Example fluorescence and E* traces with HMM fitting are shown in Fig. 3.6B.

Fig. 3.6C shows the overall E* histogram to be multimodal with populations of mean E* ~ 0.15 (10%), ~ 0.37 (72%) and ~ 0.69 (18%). The major subpopulation at E* ~ 0.37 is in strong agreement to AV modelling of fluorescent probes on the *E. coli* RNAP holoenzyme structure (PDB 6C9Y), E ~ 0.35. The two minor subpopulations, E* ~ 0.15 and E* ~ 0.69, are new and show previously unobserved conformations of the σ-finger. HMM identified 28% of molecules to be dynamic and exhibit transitions between different three FRET states centred around E* values of ~ 0.22 (28%), ~ 0.45 (54%), and ~ 0.70 (18%) with dwell times of 0.5-0.8s (Fig. 3.6D).

Comparing Fig. 3.6B and Fig. 3.4 shows that there is a large change in average E* as σ^{70} binds to RNAP core since the dominant subpopulation shifts from E*~ 0.21 (66%) to E*~ 0.37 (72%). These results are consistent with observations showing a large conformational change in σ^{70} during this process to unmask DNA binding domains (Callaci et al., 1998, 1999).

3.5 Conclusions

A double-labelled construct suitable for studying the σ-finger during transcription using smFRET was prepared and initial experiments performed.

- Introduction of fluorescent probes at the base of the σ-finger (σ^{70} residue 511) did not alter RNAP function.
- The σ-finger conformation in freely diffusing σ^{70} shows two distinct FRET subpopulations, E* ~ 0.2 (66%) and E* ~ 0.4 (34%).

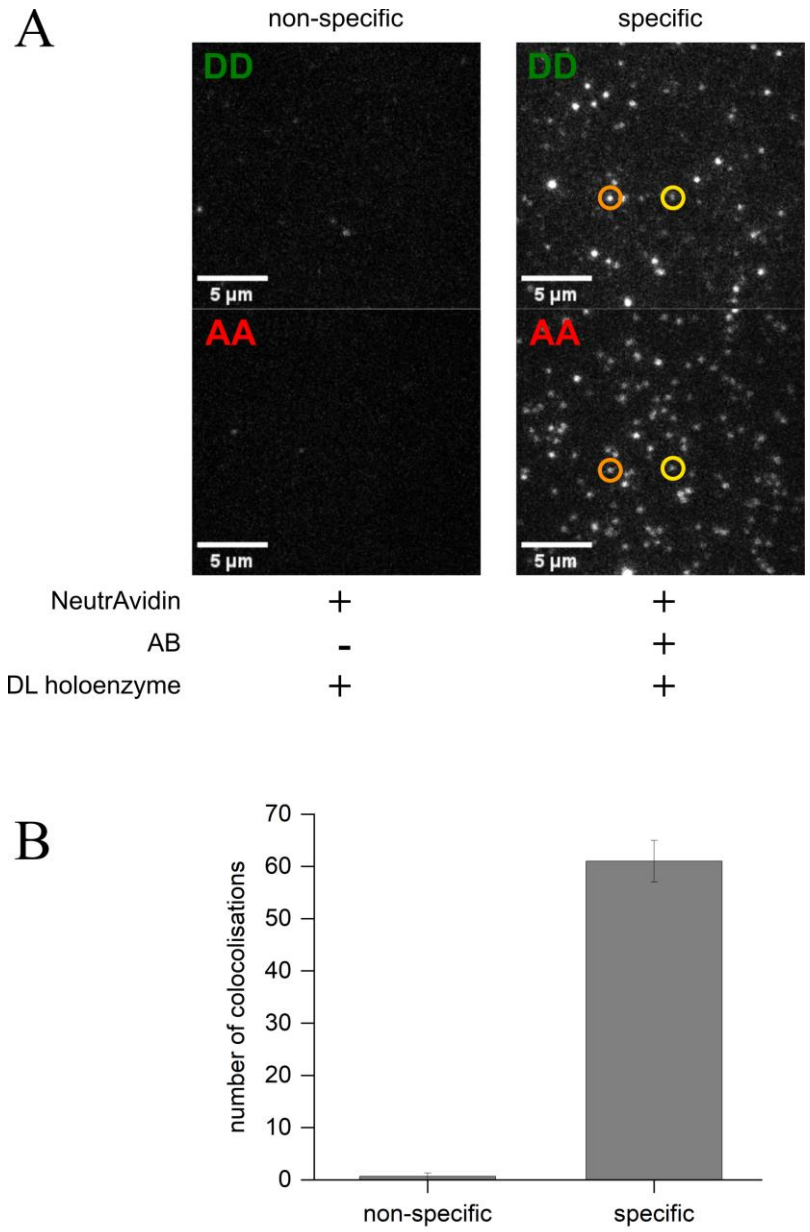

Figure 3.5: Controls for specific immobilisation. **(A)** Example fields of view when DL holoenzyme was non-specifically (left) or specifically (right) immobilised to PEG-passivated glass slides in the absence or presence of penta-His antibody (AB). Orange and yellow circles highlight pairs of donor and acceptor molecules that colocalised. **(B)** Comparison of the mean number of colocalisation counts across 3 fields of view between the donor and acceptor channels for non-specific and specific immobilisation of DL holoenzyme. Error bars show the standard error of the mean.

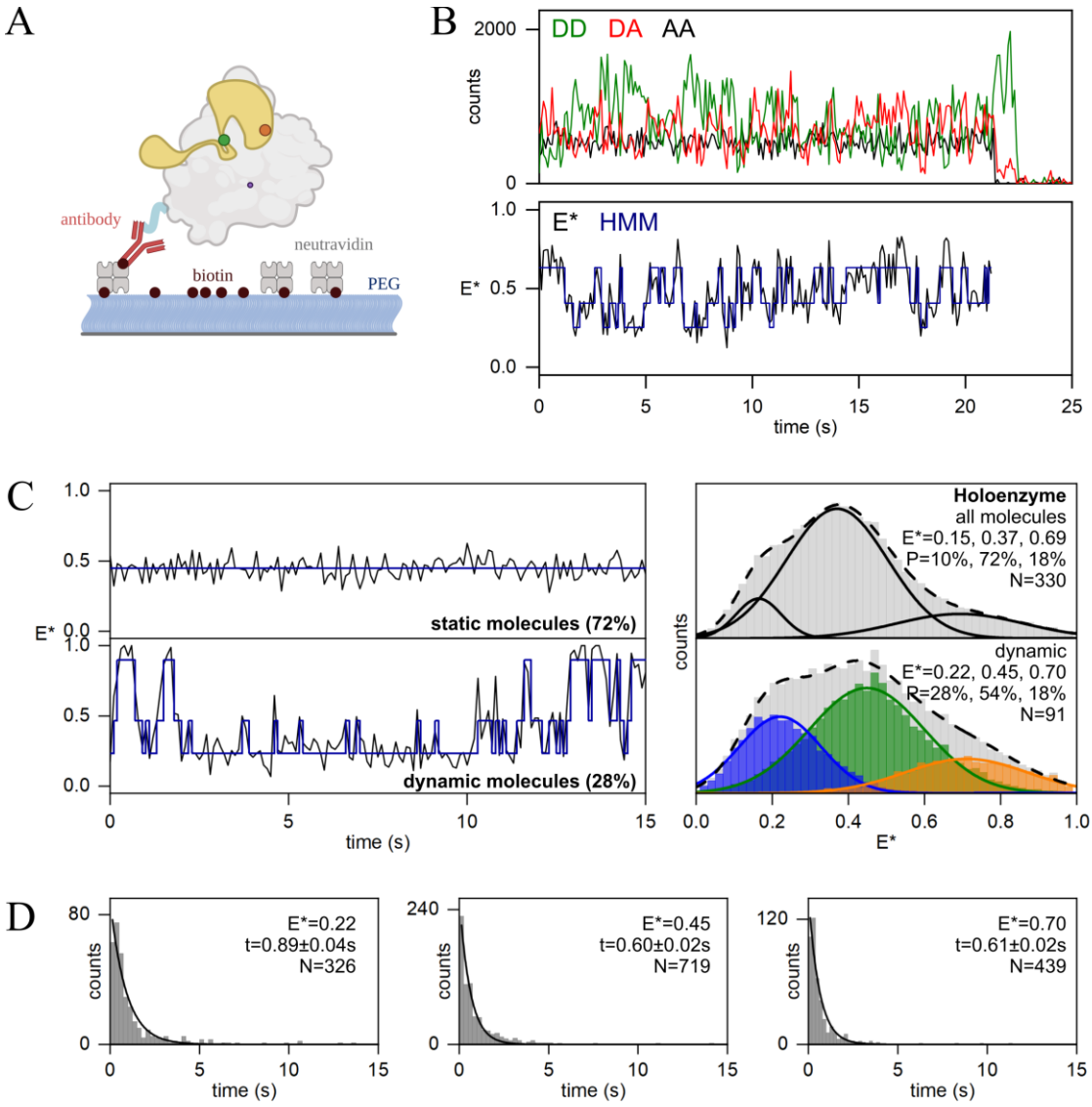

Figure 3.6: Conformation of the σ-finger in RNAP holoenzyme. **(A)** Schematic of the surface immobilisation via a hexahistidine-tag on RNAP core method used for RNAP holoenzyme measurements. **(B)** Example fluorescence and FRET traces with HMM fitting. Donor and acceptor traces both show single-step bleaching indicative of a single molecules. **(C)** Example E*-time trajectories of static and dynamic molecules and E* histograms of DL holoenzyme for all and only dynamic molecules. **(D)** Dwell time histograms (0.25s bin for the left and middle panels, and 0.2s bins for the right panel) of each of the three states found using HMM.

- The σ-finger conformation in RNAP holoenzyme shifts towards intermediate FRET and displays three interconverting subpopulations, E* ~ 0.15 (10%), ~ 0.37 (72%) and ~ 0.69 (18%).

Whilst structural studies of RNAP holoenzyme show only one conformation of the σ-finger, these smFRET measurements in solution also identify two additional conformations, showing that at this stage of transcription, the σ-finger is in fact mobile.

The smFRET assay designed here can therefore provide new information about the mechanism of the σ-finger during initial transcription.

Chapter 4: Conformation of the σ-finger with a consensus promoter, lacCONS

4.1 Introduction

This chapter uses a consensus promoter, lacCONS, to assess σ-finger movements during primer-dependent and primer-independent transcription initiation.

Consensus promoter

Bacterial RNAP core associates with a σ-factor to perform transcription from specific promoters, a typical schematic of which is shown in Fig. 4.1A. In σ^{70}-dependent transcription, the activity of these promoters mainly depends on two sequence motifs: the -10 and -35 hexamers, which have consensus sequences of 5'-TTGACA-3' and 5'-TATAAT-3' respectively (Dickson et al., 1975; Shimada et al., 2014). These hexamers are separated by a non-conserved spacer region with consensus length of 17bp and typical range of 16-19bp that is thought to aid correct positioning of the conserved regions when binding to RNAP (McClure et al., 1983). Interactions of these elements with the σ-factor have been described previously (Chapter 1. 1) and are shown in Fig. 4.1C for the consensus promoter. Additionally, the promoter strength is also influenced by the 'UP' element upstream of the -35 region (Ross et al., 1993), the discriminator (consensus 5'-GGG-3'; Haugen et al., 2008) and core recognition element (Y. Zhang et al., 2012) – both downstream of the -10 region and, in some promoters, an extended -10 element (consensus 5'-TGTG-3'; Barne et al., 1997). Most naturally occurring promoters have a sub-maximal activity (Lisser and Margalit, 1993) which is modulated by different combinations of consensus/non-consensus elements in each promoter (Browning and Busby, 2016).

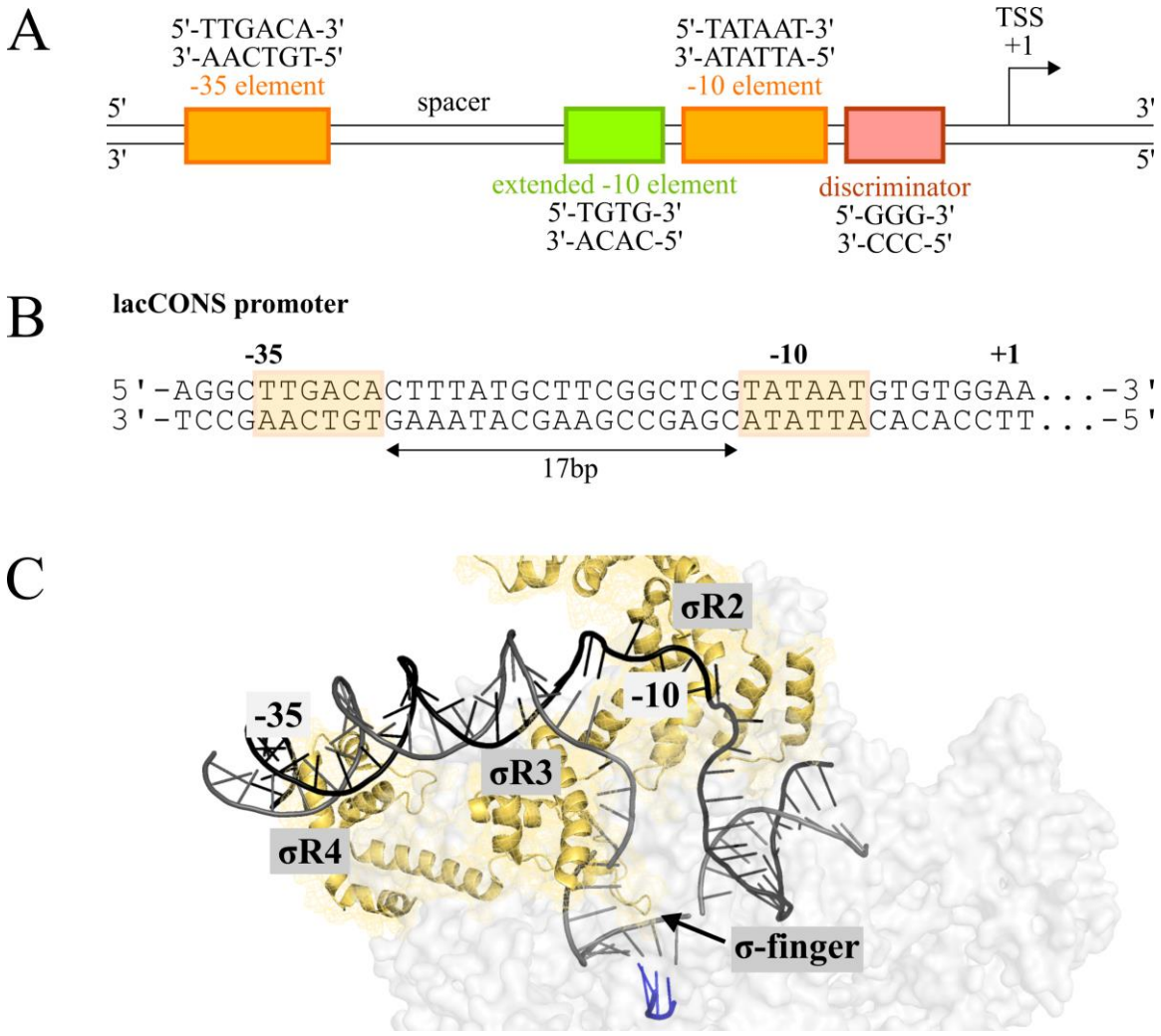

Figure 4.1: Schematic of bacterial promoters. **(A)** Diagram of a bacterial promoter with -25, extended -10 (ext. -10), -10, discriminator (dis.) and transcription start site (TSS) labelled and their respective consensus sequences. Created with BioRender.com. **(B)** The lacCONS promoter with -35 and -10 regions highlighted. **(C)** Structure of an initial transcribing complex containing the lacCONS consensus promoter and a 4-mer RNA. Interactions of the promoter regions with the σ-factor are labelled. Colours as in Fig. 1.3. (PDB 4YLN; Zuo and Steitz, 2015).

The lacCONS promoter used in this chapter is shown in Fig. 4.1B. This derivative of the *lac* promoter has several σ^{70} consensus features such as a consensus –35 element, a consensus –10 element, and a consensus −35/–10 spacer length. RPo formed using this

promoter contains a non-scrunched 13bp DNA transcription bubble with start site positioned at seven bases downstream of the -10 region. The lacCONS promoter has been used extensively to study transcription mechanisms such as RNAP clamp motions (Duchi et al., 2018; Mazumder et al., 2021), transcription start-site selection (Winkelman, Vvedenskaya, et al., 2016), DNA scrunching in initial transcription (Kapanidis et al., 2006), transcription initiation pausing (Duchi et al., 2016; Lerner et al., 2016; Winkelman et al., 2020) and promoter escape (G. Wang et al., 2016). Whilst not all promoters necessarily share the same mechanisms, most of those listed have also since been confirmed in other promoters. For example, relevant to the σ-finger, Duchi et al. (2016) identified a transcription initiation pause for the lacCONS promoter after synthesis of 6-mer RNA, which was greatly reduced in RNAP mutants containing a seven residue deletion of the σ-finger. This initiation pause was also found in promoters containing different -10, -35/-10 spacer and -35 promoter elements (Bauer et al., 2016), and found to be dependent on the initial transcribed sequence (Winkelman et al., 2020) – with both studies verifying the σ-finger's role as pause lifetime determinant.

The well-studied lacCONS promoter therefore provides a good starting point to define σ-finger conformations during initial transcription.

Primer-dependent and primer-independent transcription

Transcription in *E. coli* can be initiated by both primer-dependent and primer-independent mechanisms (Winkelman et al., 2021). During primer-dependent transcription, a 'nanoRNA' containing a 5'-hydroxyl (5'-OH) end is used to initiate RNA synthesis during stationary-phase growth (Vvedenskaya et al., 2012). The nanoRNA used is predominantly a dinucleotide primer (Skalenko et al., 2021) and modulates the expression of promoters

involved in biofilm formation and stress (Nickels, 2012). Many single-molecule assays have used primer-dependent initiation to understand transcription mechanisms (Duchi et al., 2016, 2018; Dulin et al., 2018; Kapanidis et al., 2005, 2006; Lerner et al., 2016; Margeat et al., 2006), however, the majority of transcription occurs via primer-independent transcription using an initiating NTP with a 5'-triphosphate (5'-ppp) end. Studying the primer-independent mode of transcription is especially relevant to the σ-finger since the 5'-ppp has a greater steric bulk and charge (-4 charge and ~163Å^3 solvent accessible volume; Li et al., 2020) compared to the 5'-OH group (0 charge and ~7Å^3 solvent accessible volume; Li et al., 2020) which may affect movements of the negatively charged σ-finger.

Chapter summary

In this chapter, the DL σ^{70} construct described in Chapter 3 was used to assess the σ-finger conformation in transcription initiation with the consensus lacCONS promoter. Three main questions were asked:

1) At what point in transcription is the σ-finger displaced?

2) How does displacement differ in primer-dependent transcription compared to primer-independent transcription?

3) What are the timescales for σ-finger displacement?

The results show that for the lacCONS promoter in primer-dependent transcription, the σ-finger is displaced upon addition of the 10th RNA nucleotide with heterogeneous timescales (13s, 88%; 180s, 12%). However, during primer-independent transcription, this displacement happens much earlier, on the addition of the 5th nucleotide with faster

timescales (1.4s, 94%; 70s, 6%). This is the first-time full displacement of the σ-finger has been observed, and the displacement is found to occur before promoter escape. In addition, both primer-dependent and -independent transcription results show conformational heterogeneity amongst RPitc2 molecules formed with the lacCONS promoter that display an order of magnitude difference in σ-finger displacement kinetics.

4.2 RPo

The DL holoenzyme construct designed in Chapter 3 was incubated with biotinylated lacCONS promoter DNA (Materials and Methods Table 1) to form heparin resistant RPo molecules. These were observed using TIRF-ALEX by immobilisation onto glass slides functionalised with NeutrAvidin (Fig. 4.2A-B; Materials and Methods 2.4).

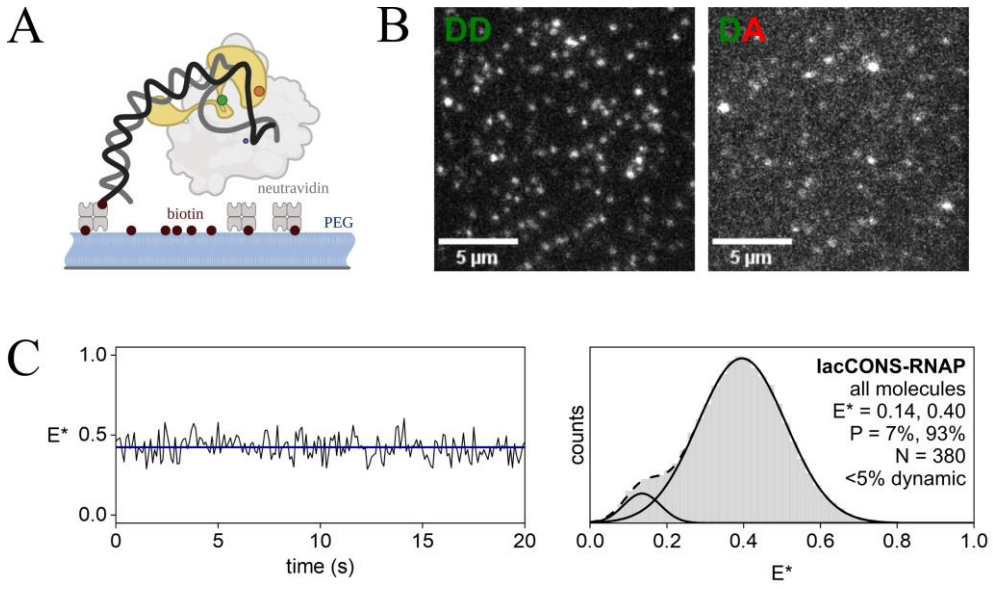

Figure 4.2: Conformation of the σ-finger in RPo. **(A)** Schematic of the surface immobilisation via biotinylated DNA for RPo. **(B)** Example fields of view of donor (DD) and FRET (DA) channels when DL RPo was immobilised to PEG-passivated glass slides functionalised with NeutrAvidin. **(C)** E*-time trajectory and histogram for the σ-finger conformation in RPo molecules formed with the lacCONS promoter.

Fig. 4.2C shows the overall E* histogram to be multimodal with two subpopulations at E* ~ 0.14 (7%) and E* ~ 0.40 (93%). The major population (E* ~ 0.40) strongly agrees with AV calculations on high resolution structures of RNAP-promoter complexes with a similar promoter DNA, E ~ 0.34 (67±9Å, PDB 4YLN) – and was therefore attributed to the conformation of the 'in-cleft' σ-finger conformation that contacts the DNA template strand as visualised in structural studies. The minor population (E* ~ 0.14) represents a structural state where the σ-finger is positioned away from σR2 and the static probe.

4.3 σ-finger displacement in primer-dependent transcription

The σ-finger conformation at different points of transcription in primer-dependent transcription was assessed by adding ApA dinucleotide primer along with subsets of RNA nucleotides to RPo molecules prepared in the previous section 4.2 using lacCONS constructs (Materials and Methods Table 1). The use of ApA and withholding of CTP lead to the RNA synthesis of RNAs with a 5'-OH group up to a desired length (2-nt, 4- to 11-nt and 14-nt; Fig.4.3A). Nucleotides were washed out of solution before imaging (Materials and Methods 2.4).

Fig. 4.3B-C shows smFRET data for RPitc2. The E* histogram is multimodal with a major population at E* ~ 0.38 and minor population with E* ~ 0.14. HMM identified a subpopulation of molecules (18%) that interconverted between three states E* ~ 0.22 (17%), 0.37 (54%) and 0.59 (29%) with time scales of 0.6-1.1s (Fig. 4.3B). These FRET states and kinetics are similar to those of the DL holoenzyme and show the flexibility of the σ-finger in these complexes.

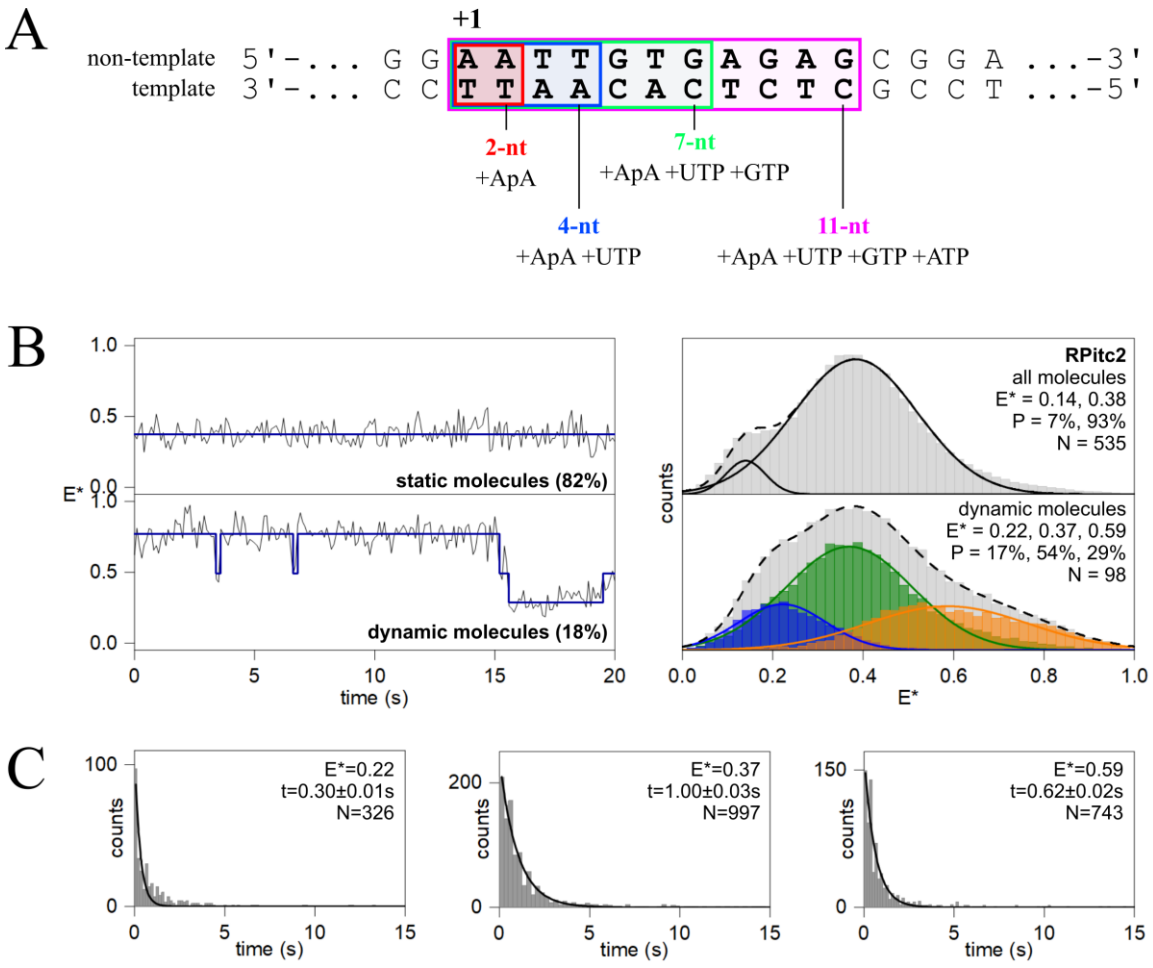

Figure 4.3: Experimental approach and conformation of the σ-finger in RPitc2. **(A)** Schematic of experimental design with different subsets of NTPs leading to different maximum RNA lengths. **(B-C)** Conformation of the σ-finger in RPitc2. **(B)** E*-time trajectories of static and dynamic molecules and E* histograms of RPitc2 for all and only dynamic molecules. **(C)** Dwell time histograms (0.15s bins for the left and right panels, and 0.25s bins for the middle panel) for the three states found using HMM.

Fig. 4.4A shows E* histograms for molecules containing increasing lengths of RNA (up to 14nt). For complexes containing RNA up to 9-nt in length, the E* histograms are similar to those obtained with RPo (section 4.2) and almost only exhibit the 'in-cleft' σ-finger conformation at E* ~ 0.40 (96-100%). The percentage of molecules that HMM identified as being dynamic for each of these complexes is shown in Fig. 4.4C and shows that unlike

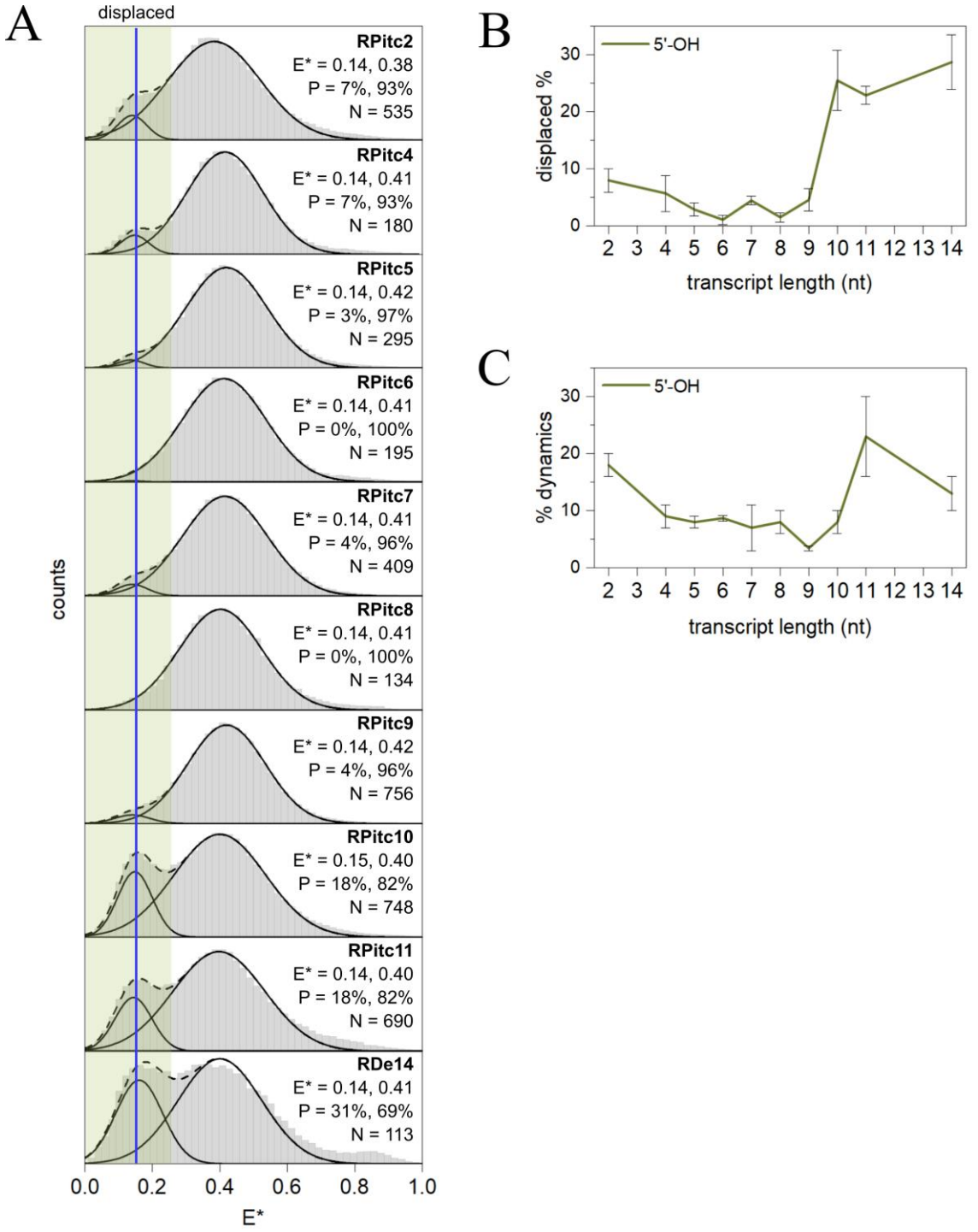

Figure 4.4: Conformation of the σ-finger in transcription initiation complexes involving the lacCONS promoter and 2- to 14-nt of RNA (RPitc2-RDe14) with a 5'-OH end. **(A)** E* histograms. Estimations of mean E* values and abundances are shown for the displaced (blue line, green highlighting) and 'in-cleft' populations. **(B)** Occupation probabilities of the displaced σ-finger conformation and associated errors calculated from the standard error of the mean in different datasets. **(C)** Percentage of dynamic molecules and associated errors calculated from the standard error of the mean in different datasets.

RPitc2, few complexes containing RNA from 4- to 9-nt in length (RPitc4-9) exhibited dynamics (<10%). This shows the formation of a structural state in early transcription initiation that stabilises the base of the σ-finger (Li et al., 2020), where the probe is attached.

For RPitc10 complexes, the E* histogram shows an increased abundance in the low FRET subpopulation (E* ~ 0.15; 19%). This shows that for a significant number of molecules, synthesis of a 10-mer RNA results in the displacement of the σ-finger away from the 'in-cleft' position to one positioned away from σR2 (Fig. 4.4B). This low FRET subpopulation was therefore assigned to the 'displaced' σ-finger conformation.

Similarly to RPitc10, The E* histogram for RPitc11 also shows a significant abundance of the 'displaced' σ-finger conformation (E* ~ 0.14; 18%). However Fig. 4.4C shows that unlike RPitc10, HMM identified a much higher fraction of dynamic molecules (21%) – which interconverted between three states E* ~ 0.23 (25%), 0.42 (47%) and 0.64 (28%) with time scales of 0.5-1.1s (Fig. 4.5). The increase in dynamics may have possible two origins:

1) Loss of σ-finger contacts with the DNA-RNA hybrid, or
2) Transcription bubble 'scrunching-unscrunching' and entire σ-finger-DNA-RNA hybrid motions.

If the first case were true, and upon breakage of σ-finger-DNA-RNA hybrid contacts, the σ-finger becomes more mobile, kinetics would resemble those observed in early transcription initiation complexes such as RPitc2. However, dwell times for the low FRET state (E*~ 0.23), shown in Fig. 4.5B, follow a peaked distribution (t_1 = 0.9s, t_2 = 0.5s)

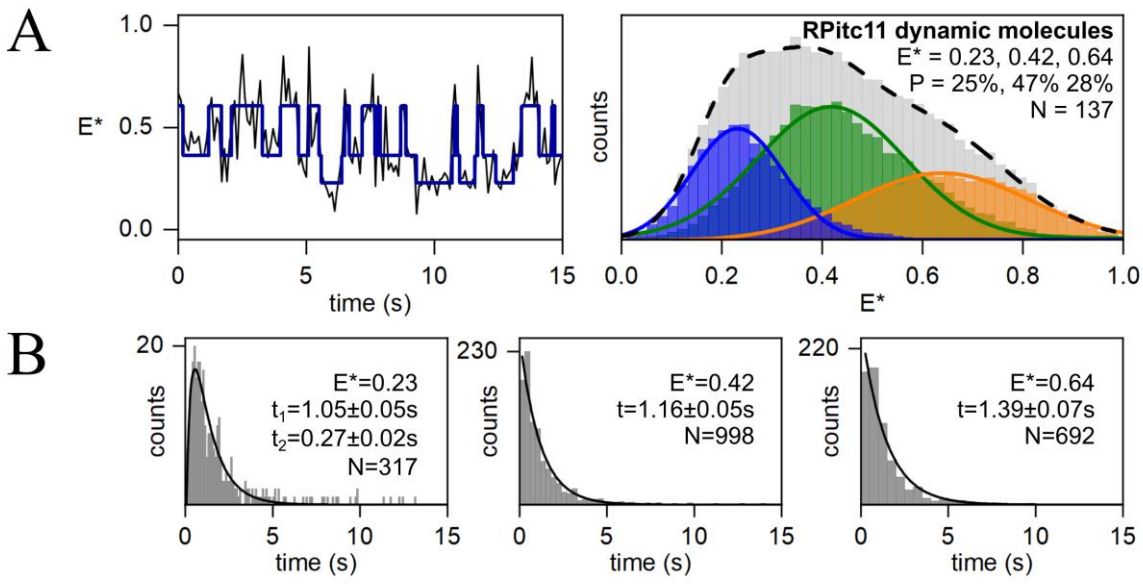

Figure 4.5: Conformation of the σ-finger in dynamic molecules of RPitc11 **(A)** E*-time trajectory of a dynamic molecule and histogram. **(B)** Dwell time histograms (0.1s bin for the left panel, 0.3s bin for the middle panel and 0.5s bin for the right panel) for the three states found using HMM.

rather than a single-exponential distribution observed in RPitc2 (Fig. 4.3B). The peaked distribution indicates a multistep process. Here, this may be two steps – first, the 5'-end of RNA moves out of the exit channel, and second, σ-finger movement out of the low FRET state to other conformations. The timescales of σ-finger motions in RPitc11 closely resemble 'scrunching-unscrunching' rates of the DNA transcription bubble previously found for complexes in late initial transcription (Dulin et al., 2018) and taken together, make the second case of entire σ-finger-DNA-RNA hybrid motions more likely.

When the RNA is extended up to 14-nt in length (RDe14), the E* histogram shows a significant abundance of the 'displaced' σ-finger conformation (E* ~ 0.14; 31%) but with a decrease in the fraction of dynamic molecules to 13% (Fig. 4.4). This behaviour is

consistent with an early elongation complex that has progressed out of late initial transcription and promoter escape but which still retains σ^{70} (Kapanidis et al., 2005).

4.4 σ-finger displacement in primer-independent transcription

In order to assess the role played by the functional group on the 5'-end of RNA in σ-finger displacement, similar measurements to those in the previous section but without ApA dinucleotide primer were performed. This resulted in primer-independent transcription – synthesis of RNAs with a 5'-triphosphate (5'-ppp) rather than a 5'-hydroxyl (5'-OH). Since the triphosphate group has a greater steric bulk and charge compared to the hydroxyl group, motions of the negatively charged σ-finger may be affected.

Fig. 4.6A shows E* histograms for complexes containing increasing lengths of 5'-ppp RNA (2-14nt). The E* distributions for complexes synthesising up to 4-nt of RNA (ppp-RPitc2 and ppp-RPitc4) are similar to those involved in primer-dependent transcription. However, the E* histogram of complexes containing 5-nt of RNA shows an increased abundance of the low FRET σ-finger conformation at E*~ 0.14 (14%; Fig. 4.6B) compared to the primer-dependent case (3%; Fig. 4.4B). E* histograms of complexes ppp-RPitc6 and beyond show a similarly abundant E*~ 0.14 subpopulation (18-28%) and complexes up to ppp-RPitc11, except for one, display a considerable fraction of dynamic molecules (14-26%; Fig. 4.6C). Dwell times and distribution shapes for ppp-RPitc11 (0.3-1.1s) are similar to RPitc11 in primer-dependent transcription. The E* histogram for ppp-RDe14 is similar to the primer-dependent case and shows both a large low FRET σ-finger subpopulation (E*~ 0.14, 31%) and a reduced fraction of dynamic molecules (4%).

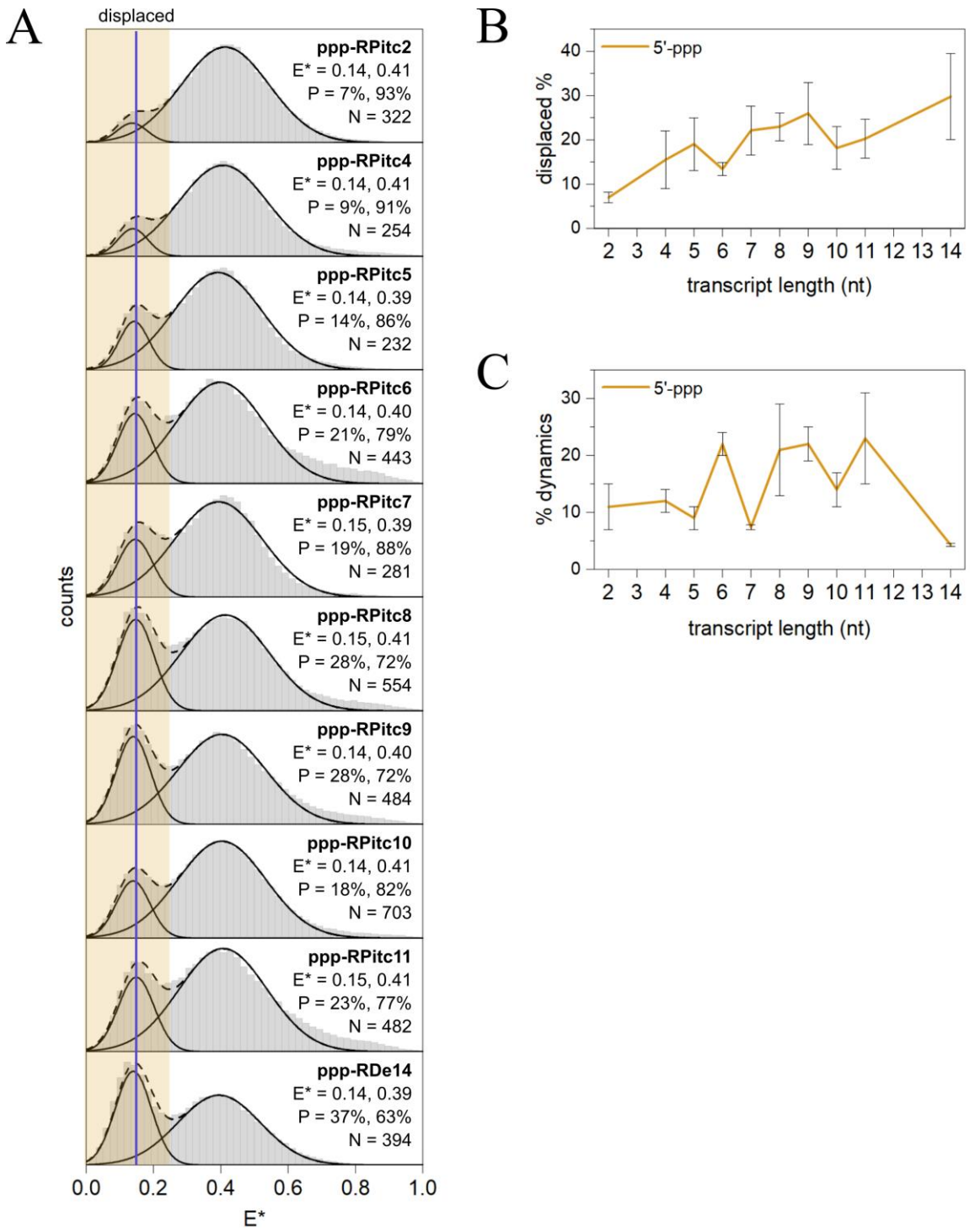

Figure 4.6: Conformation of the σ-finger in transcription initiation complexes involving the lacCONS promoter and 2- to 14-nt of RNA (RPitc2-RDe14) with a 5'-ppp end. **(A)** E* histograms. Estimations of mean E* values and abundances are shown for the displaced (blue line, green highlighting) and 'in-cleft' populations. **(B)** Occupation probabilities of the displaced σ-finger conformation and associated errors calculated from the standard error of the mean in different datasets. **(C)** Percentage of dynamic molecules and associated errors calculated from the standard error of the mean in different datasets.

4.5 σ-finger displacement and promoter escape

The fraction of dynamic molecules reduced upon formation of complexes containing up to 14-nt of RNA (RDe14 and ppp-RDe14) with the lacCONS promoter. To test if this kinetic behaviour was due to progression past promoter escape into an early stage σ^{70}-retaining elongation complex, a previously reported smFRET ruler was used to measure the status of the transcription bubble – which 'unscrunches' during promoter escape. A double-labelled lacCONS promoter, with probes at positions -15 of the non-template DNA and +20 of the template DNA (Materials and Methods Table 1), was incubated with an unlabelled hexahistidine-tagged RNAP holoenzyme and observed using TIRF-ALEX (Materials and Methods 2.4). Measurement of the transcription bubble status for complexes containing 11-nt and 14-nt of RNA were then taken.

Fig. 4.7 shows E* histograms for these measurements. The E* histogram for RPitc11 shows a large subpopulation centred around E* ~ 0.45. This subpopulation corresponds to the scrunched DNA bubble characteristic of late initial transcribing complexes, as established by previous single-molecule studies (Fig. 4.7A shows a schematic; Duchi et al., 2016). However, this population is not present in the E* histogram for RDe14, which shows that upon RNA extension to 14-nt, the scrunched DNA bubble disappears. The point of escape for the lacCONS promoter therefore lies between formation of a 11-mer and 14-mer RNA. The small subpopulation of complexes in the RDe14 E* histogram at E* ~ 0.45 may be either transcriptionally inactive or in abortive initiation so therefore do not procced to elongation.

Previous sections showed that the σ-finger becomes displaced before formation of an 11-mer RNA (upon 10-mer for 5'-OH RNA and 5-mer for 5'-ppp RNA). Taken together, the

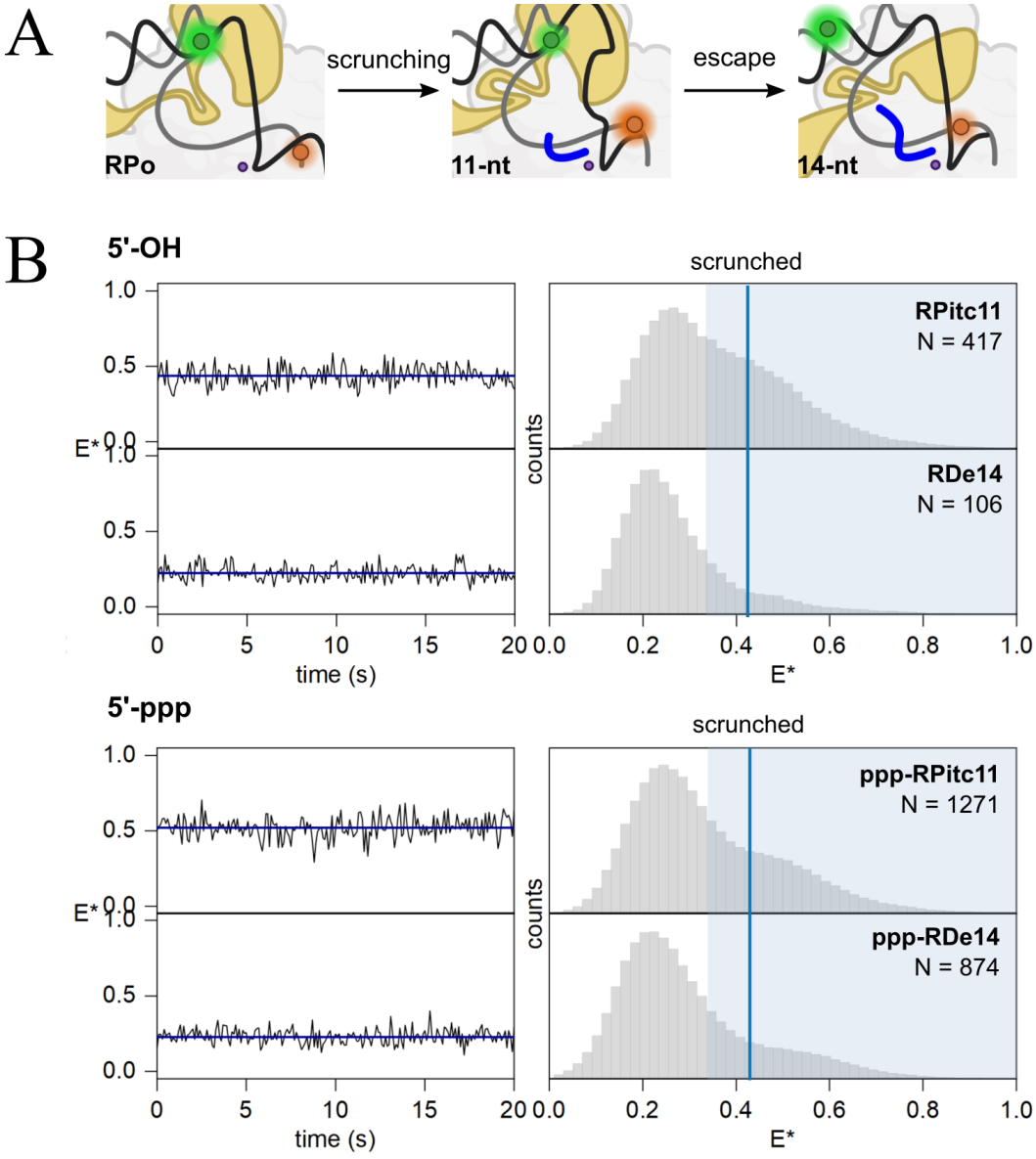

Figure 4.7: Conformation of the DNA transcription bubble during transcription initiation and prompter escape. **(A)** Schematic of experiment to monitor the DNA transcription bubble with probes at -15 (non-template DNA; black) and +20 (template DNA; grey). DNA scrunching due to initial transcription increases FRET and unscrunching due to promoter escape leads to a FRET decrease. **(B)** E*-time trajectories and histograms for the conformation of the transcription bubble in RPitc11 and RDe14 where the RNA as a 5'-OH end (left) and a 5'-ppp end (right). Created with BioRender.com.

results show that the base of σ-finger is displaced before promoter escape for the lacCONS promoter.

4.6 Observations of σ-finger displacement in actively transcribing complexes

The timescale of σ-finger displacement was found by taking E* measurements of complexes involved in real-time synthesis of RNA up to promoter escape. RPo complexes to monitor the σ-finger were immobilised onto functionalised glass slides and incubated with 500μM ApA or pppApA to form RPitc2. To observe the σ-finger for a longer period of time, the movie was recorded using 400ms frametime (200ms ALEX). During this, a subset of NTPs (1mM GTP, UTP and ATP) added to direct transcription up to formation of a 14-mer RNA. E*-time trajectories were extracted from the movies and were fitted using HMM to find the time between nucleotide addition and σ-finger displacement (Fig. 4.8; Materials and Methods 2.4).

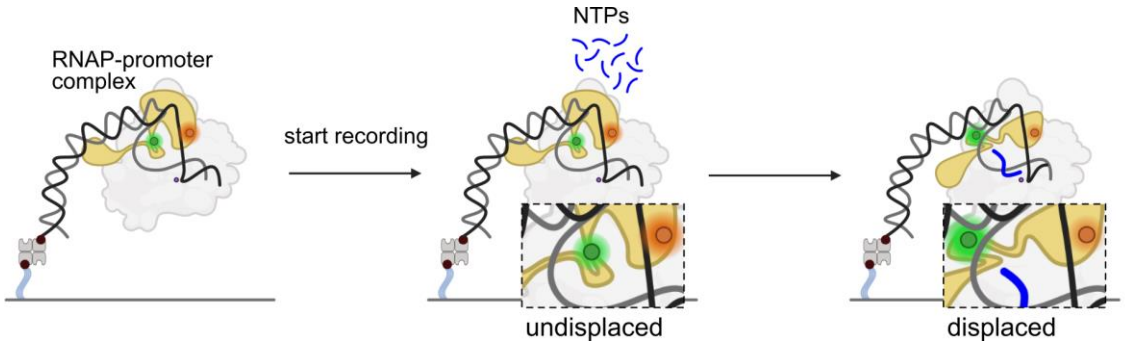

Figure 4.8: Experiment design of real-time measurements of the σ-finger conformation showing fluorescent probe positions before and after displacement. Created with BioRender.com.

For both primer-dependent and primer-independent transcription, E*-time trajectories showed four classes of molecules:

Class-I Molecules exhibiting a FRET efficiency decrease to a stable ‘displaced’ σ-finger conformation , E*~ 0.15, on a short timescale

Class-II Molecules exhibiting a FRET efficiency decrease to a stable ‘displaced’ σ-finger conformation, E*~ 0.15, on a long timescale

Class-III Molecules exhibiting multiple transitions between FRET states but no FRET efficiency decrease to a stable ‘displaced’ σ-finger conformation

Class-IV Molecules exhibiting no FRET efficiency transitions

Example E*-time trajectories are shown in Fig. 4.9A for primer-dependent and Fig. 4.9B for primer-independent cases.

Both smFRET datasets for RDe14 and ppp-RDe14 monitoring the σ-finger exhibited a low fraction of dynamic molecules, 13% and 4% respectively. This means that for at least some Class-III molecules, the σ-finger is displaced at later time points and is missed due to premature ending of the E* trajectory due to photobleaching. Class-IV molecules likely consist of inactive complexes or those stuck in early initial transcription.

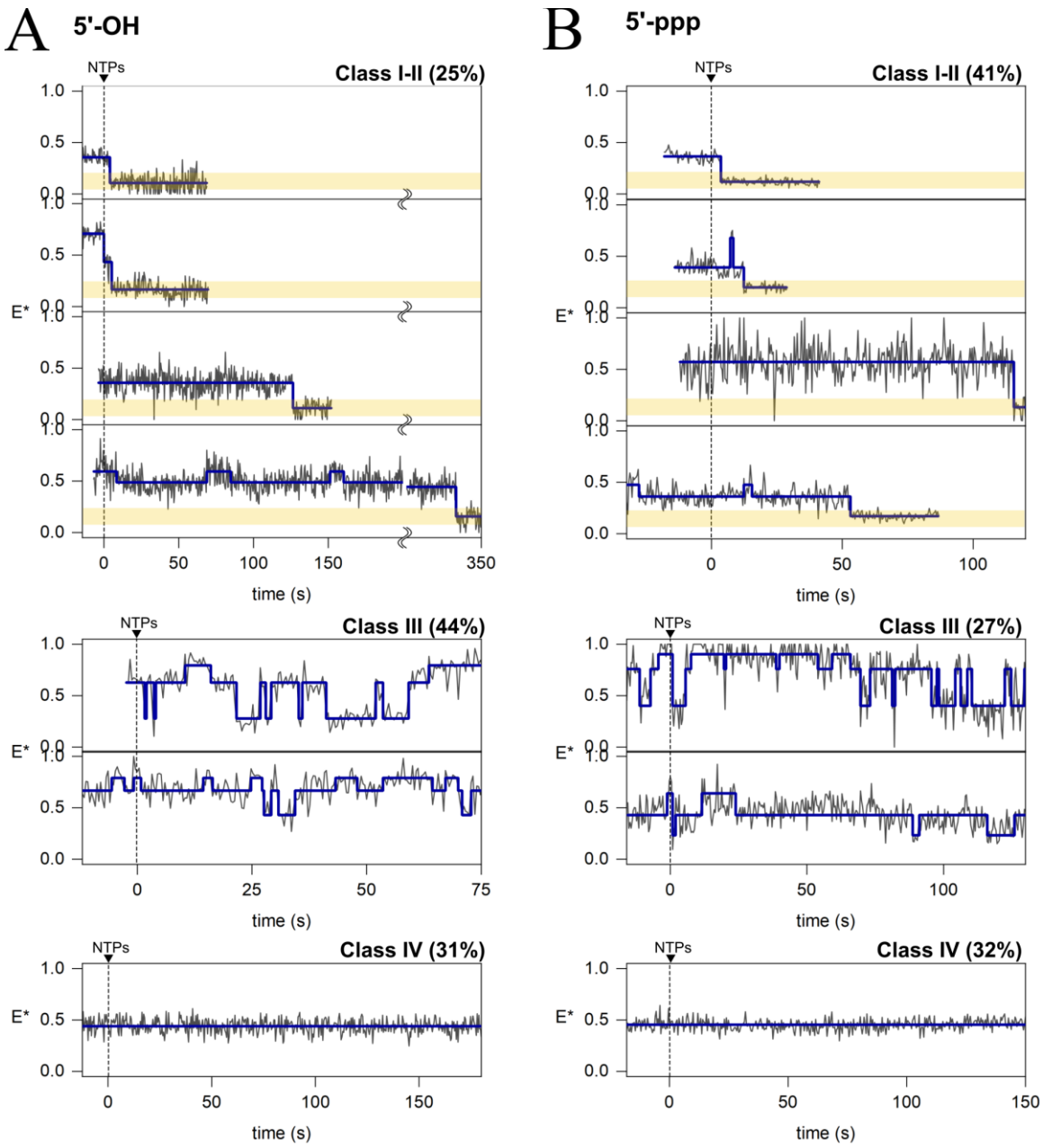

Figure 4.9: Example E*-time trajectories for each Class I-IV obtained in real-time experiments tracking the σ-finger conformation for complexes containing the lacCONS promoter and RNA with a **(A)** 5'-OH end. **(B)** 5'-ppp end. Yellow-orange highlighting shows the displaced σ-finger E* state.

Primer-dependent transcription

During real-time RNA synthesis of molecules in primer-dependent transcription, 24% of molecules belonged to Class-I and Class-II, 44% to Class-III and 31% to Class-IV.

In order to find the σ-finger displacement time, t_d, Class-I and Class-II molecules were pooled and the time between nucleotide addition and the E* step decrease corresponding to σ-finger displacement calculated (Fig. 4.10A).

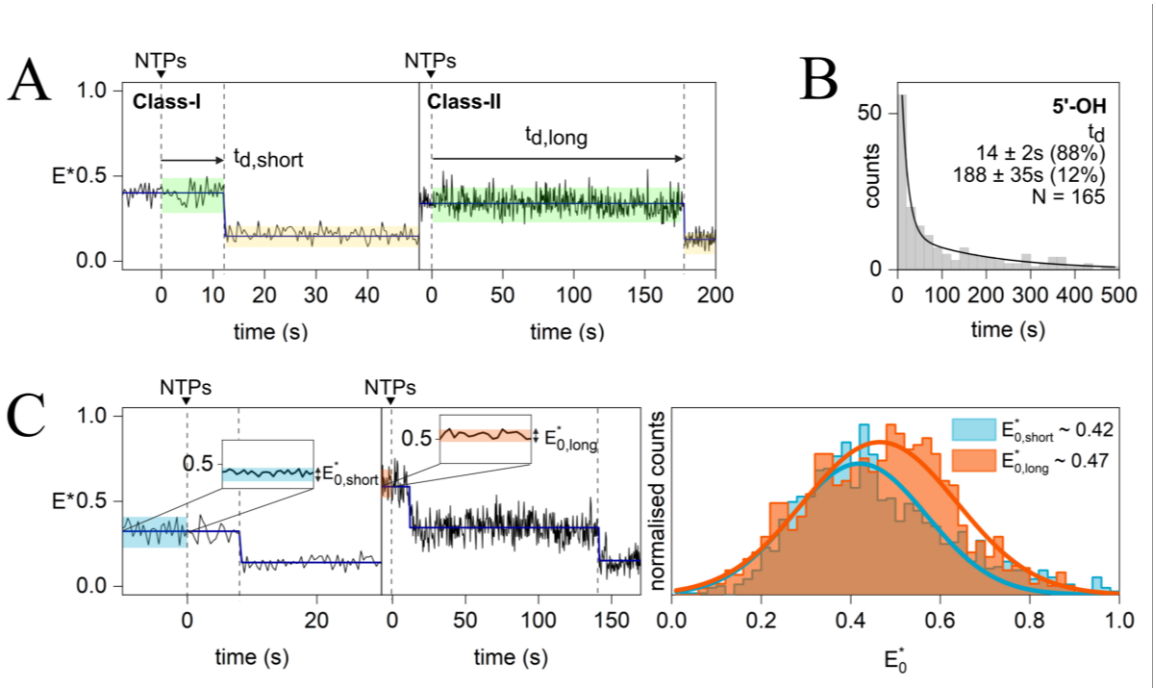

Figure 4.10: Real-time measurements of the σ-finger conformation in complexes containing the lacCONS promoter and RNA with a 5'-OH end. **(A)** Example E*-time trajectories of Class-I and Class-II molecules showing σ-finger displacement on two timescales. Green highlighting shows FRET states between σ-finger displacement and NTP addition, and yellow-orange highlighting shows the FRET state after displacement **(B)** Displacement time histogram (grey; bin size 20s) fitted to a double-exponential decay (black line). **(C)** FRET state before NTP addition (E^*_0). (*left*) Example E*-time trajectories with E^*_0 values highlighted for trajectories exhibiting a short displacement time (blue) and trajectories exhibiting a long displacement time (orange). (*right*) E^*_0 histograms for trajectories involving a short displacement time (blue) and trajectories showing a long displacement time (orange). Created with BioRender.com.

Fig. 4.10B shows the binned displacement times fitted with a biexponential decay curve with two displacement times; $t_{d,\ short}$ ~ 14s (~90% of events) and $t_{d,\ long}$ ~ 190s (~10% of events). The two timescales for σ-finger displacement are an order-of-magnitude apart. To test whether this originates from conformational differences in RPitc2, the E* values before nucleotide addition were extracted and separated into two groups; 'fast-displacers' if t_d<31s and 'slow-displacers' if t_d>31s (Materials and Methods 2.4). Fig. 4.10C shows that the two groups exhibit remarkably distinct E* histograms centred around E*~0.42 for the 'fast-displacers' and E*~0.47 for the 'slow-displacers'. Changing the cut-off time for 'slow-displacers' and 'fast-displacers' did not affect this result.

Primer-independent

During real-time RNA synthesis of molecules in primer-independent transcription, 41% of molecules belonged to Class-I and Class-II, 27% to Class-III and 32% to Class-IV.

Analysis of real-time measurements were carried out as for primer-dependent transcription. Example traces of Class-I and Class-II molecules used to calculate σ-finger displacement times are shown in Fig. 4.11A. Fig. 4.11B shows binned σ-finger displacement times fitted with a biexponential decay with displacement times $t_{d,\ short}$ ~ 1.4s (~97% of events) and $t_{d,\ long}$ ~ 70s (~3% of events) in the case of primer-independent transcription (synthesis of RNA with 5'-ppp). These timescales are faster than that of primer-dependent transcription and, combined with results showing earlier σ-finger displacement in primer-independent transcription, mean that the interaction between the 5'-ppp end of RNA and σ-finger lowers the energy barrier for displacement – due to the triphosphate group's greater steric bulk and negative charge that acts in opposition to acidic σ-finger residues (D513, D514, E515 and D516) compared with the 5'-OH end.

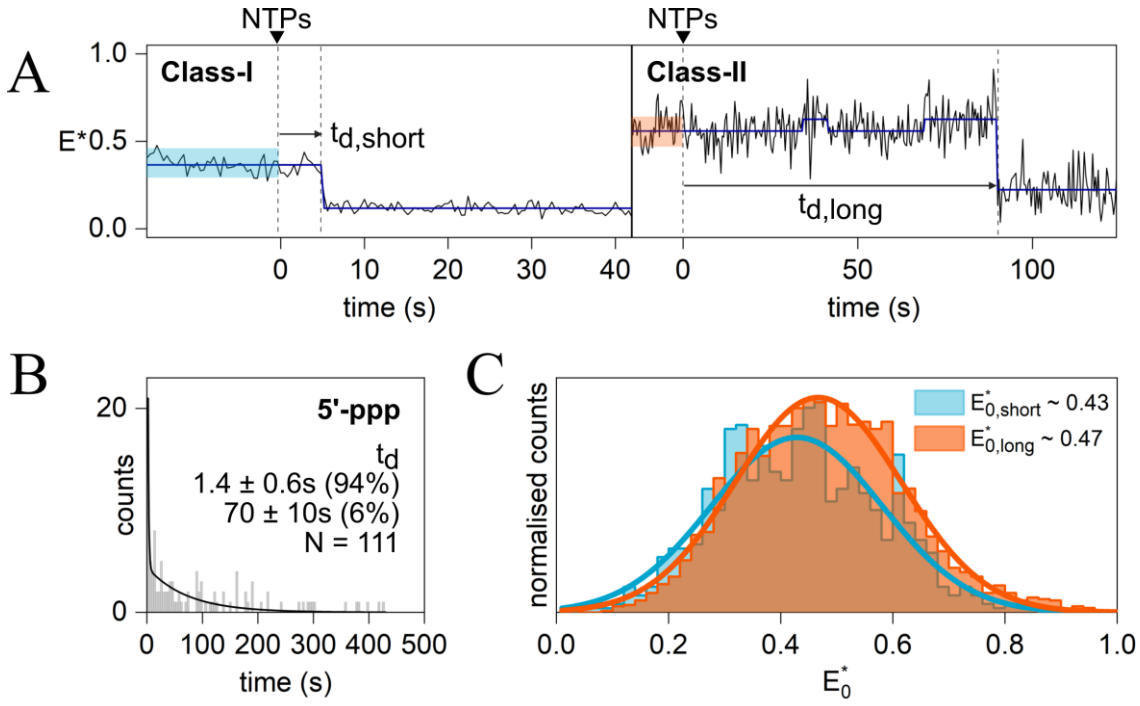

Figure 4.11: Real-time measurements of the σ-finger conformation in complexes containing the lacCONS promoter and RNA with a 5'-ppp end. **(A)** Example E*-time trajectories of Class-I and Class-II molecules showing σ-finger displacement on two timescales. E^*_0 values before NTP addition are highlighted blue (Class-I) and orange (Class-II). **(B)** Displacement time histogram (grey; bin size 4s) fitted to a double-exponential decay (black line). **(C)** E^*_0 histograms for trajectories involving a short displacement time (blue) and trajectories showing a long displacement time (orange). Created with BioRender.com.

Fig. 4.11C shows E* histograms comparing the σ-finger conformation before nucleotide addition in 'fast-displacers' (t_d<4s) and 'slow-displacers' (t_d>4s) during primer-independent transcription. Similarly to Fig. 4.9D, the two histograms are distinct with the 'fast-displacers' exhibiting a RPitc2 σ-finger conformation centred around E*~ 0.43 and 'slow-displacers' exhibiting E*~ 0.47. Changing the cut-off time for 'slow-displacers' and 'fast-displacers' did not affect this result.

4.7 Conformation of the σ-finger relative to DNA upstream of the transcription bubble

Structural constraints suggest that the σ-finger moves towards DNA upstream of the transcription bubble when displaced. In order to assess movements of the σ-finger relative to DNA, an alternative smFRET ruler was used – Cy3B was positioned at the base of the σ-finger (σ^{70} residue 511), as before, and Atto647N was positioned at -25 of the template DNA (Materials and Methods 2.2). The distance between these probes for the 'in-cleft' conformation of the σ-finger is estimated to be 58Å using AV modelling on PDB structure 4YLN, corresponding to an estimated FRET efficiency E ~ 0.55 (Fig. 4.12A). Displacement of the σ-finger towards the upstream DNA would result in an increased FRET efficiency. Similar measurements to section 4.3 were performed on this construct to assess the conformation of the σ-finger in RPo and complexes containing 7- and 11-nt of RNA with a 5'-OH end.

Fig. 4.12B shows E* histograms for RPo and complexes containing 7- and 11-nt of RNA. The E* distribution for RPo is multimodal with three subpopulations; E* ~ 0.17 (24%), E* ~ 0.50 (55%) and E* ~ 0.77 (21%). The intermediate FRET conformation E* ~ 0.50 closely corresponds to that expected from the 'in-cleft' conformation of the σ-finger expected from structures. To assess whether the remaining low (E* ~ 0.17) or high FRET (E* ~ 0.77) subpopulations could be due to non-specific associations between RNAP holoenzyme and DNA, RPo was also formed without heparin challenge that minimises non-specifically bound RNAPs (Duchi et al., 2017). The resulting E* histogram in Fig. 4.12C shows a large increase in the abundance of the low FRET subpopulation whilst that

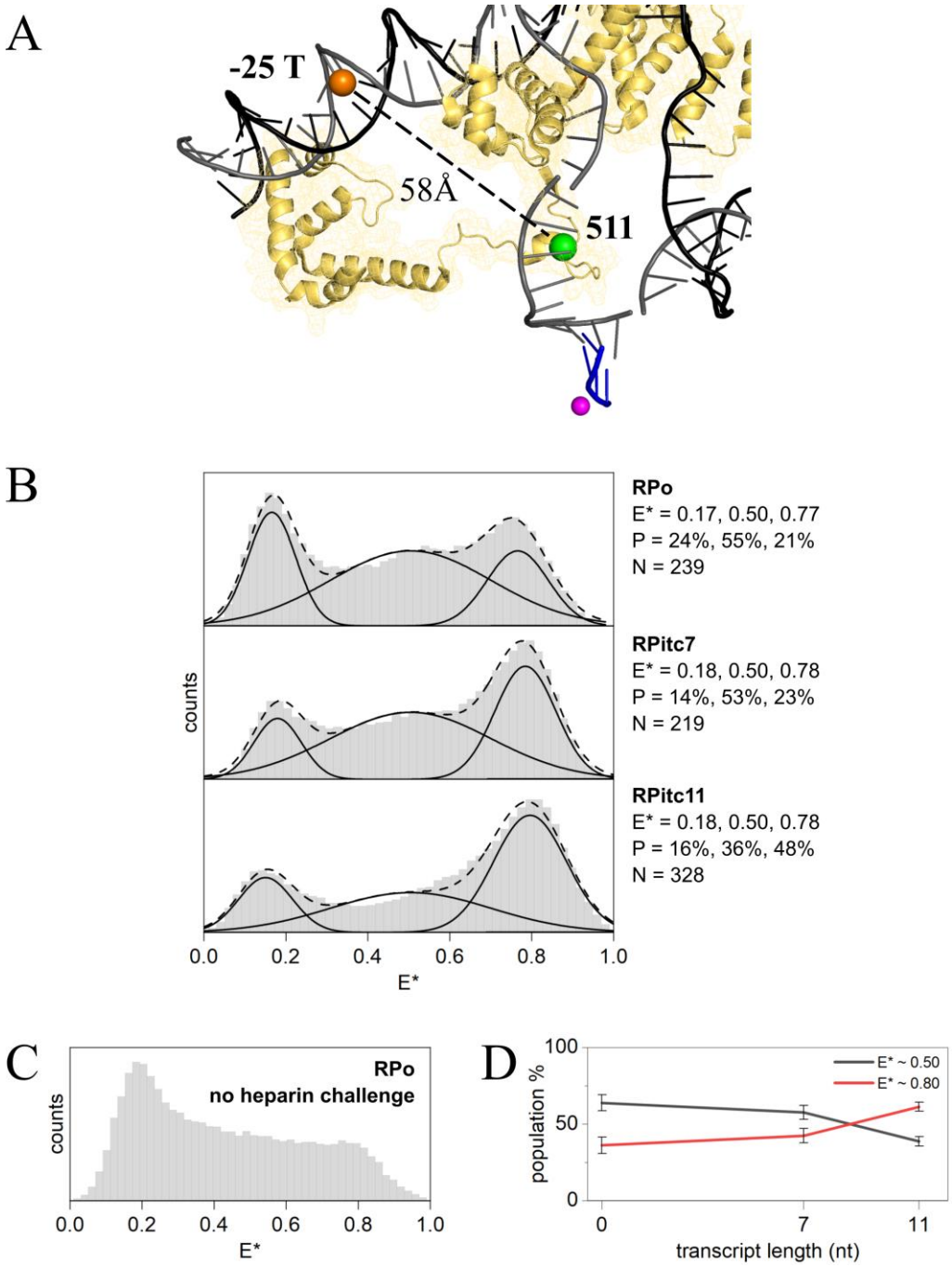

Figure 4.12: Conformation of the σ-finger relative the DNA**. (A)** smFRET dye labelling scheme with Cy3B labelled at σ^{70}-factor residue 511 (green sphere) and Atto647N labelled at position -25 of the template DNA (dark orange sphere). Colours as in Fig. 3.1. (PDB 4YLN; Zuo and Steitz, 2015). **(B)** E* histograms of the σ-finger conformation in RPo and in transcription initiation complexes involving the lacCONS promoter and 7- and 11-nt of RNA with 5'-OH end. Estimations of mean E* values and abundances are shown. **(C)** E* histogram of RPo complexes when formed without heparin challenge. **(D)** Relative abundance of intermediate (E* ~ 0.50) and high FRET (E* ~ 0.80) subpopulations and associated errors calculated from the standard error of the mean in different datasets.

of the high FRET subpopulation remained similar. This suggests that the low FRET subpopulation may be due to inherent unspecific associations that could not be removed even in an excess of heparin (250µg/ml; Fig. 4.12B). Different experimental datasets of RPo formed with heparin challenge also showed large fluctuations in the low FRET subpopulation abundance (15-35% across four datasets). Comparisons between E* histograms of RPo, RPitc7 and RPitc11 complexes were therefore made by comparing the abundance of the intermediate FRET population corresponding to the 'in-cleft' σ-finger conformation to the

high FRET population only. For comparison purposes, the mean and width of the intermediate FRET population was fixed to that found in RPo (E* ~ 0.50, w ~ 0.38) during multipeak fitting of RPitc7 and RPitc11 E* histograms.

Fig. 4.12D shows the relative abundances of the intermediate and high FRET populations with errors calculated from the standard error of the mean across different datasets. Whilst the relative abundance of the intermediate FRET subpopulation between RPo and RPitc7 are similar within error, the relative abundance of the intermediate FRET subpopulation decreases in complexes containing 11-nt of RNA (RPitc11) – with a corresponding increase in the abundance of the high FRET population. This suggests that for a significant number of molecules, synthesis of RNA between 7- and 11-nt in length results in displacement of the σ-finger away from the 'in-cleft' position to one positioned closer to position -25 on the template DNA strand – so that the σ-finger moves towards DNA upstream of the transcription bubble upon displacement.

4.8 Conclusions

This chapter shows the following results for transcription with the consensus bacterial promoter:

- In primer-dependent transcription the σ-finger is displaced upon extension to a 10-mer RNA and towards DNA upstream of the transcription bubble.

This result is consistent with recent structural work examining initial transcribing complexes showed that the 5'-end of RNA clashes with the σ-finger upon formation of a 5-mer RNA (Li et al., 2020). Here, further RNA extension caused the tip of the σ-finger to fold back over itself like a 'protein-spring' and led authors to hypothesise that σ-finger displacement is directly caused by collision of the 5'-end RNA and the σ-finger. In this scheme, the point of displacement is dictated by the extent to which the σ-finger protrudes inside the transcription bubble and strength of the contacts made with the template DNA and the RNA.

- In primer-independent transcription the σ-finger is displaced earlier, upon extension to a 5-mer RNA, due to unfavourable interactions between the 5'-ppp and σ-finger.

Transcription during the exponential growth phase occurs by primer-independent initiation with primer-dependent initiation increasing during the stationary phase (Nickels, 2012). The early σ-finger displacement observed in primer-independent transcription may be used to ease promoter escape and increase transcription during exponential growth, while later σ-finger displacement in primer-dependent initiation could lead to reduce transcription in

stationary phase. In this way, σ-finger interactions between the 5'-end of RNA could be used to regulate gene expression in different growth phases.

- For the lacCONS promoter, the σ-finger is displaced before promoter escape.

Li et al. (2020) suggested that folding back of the tip of the σ-finger togther with DNA scrunching both contribute to storage of stress energy required to disrupt RNAP-promoter contacts in promoter escape. Displacement of the σ-finger before promoter escape means that the energy stored by σ-finger folding is released before promoter escape to disrupt the stabilising contacts between the σ-finger and RNA. This release results in the full displacement of the σ-finger and clears the path for promoter escape.

- σ-finger displacement occurs on heterogeneous timescales; ~14s (88%) and ~188s (12%) for primer-dependent transcription, and ~1.4s (94%) and ~70s (6%) for primer-independent transcription.
- This heterogeneity in displacement times correlates to conformationally different RPitc2 molecules.

Measured displacement times using the lacCONS promoter (~1-190s) were of a similar range to promoter escape times (~13 s, 96% and ~200 s, 4%) observed in another single-molecule study also using the lacCONS promoter (G. Wang et al., 2016). These kinetically distinct molecules also exhibited different initial σ-finger conformations – showing correlation between the heterogeneity in displacement times and conformationally different RPitc2 molecules. This could have two causes; different σ-finger positioning creating differences in the steric clash with the 5'-end of RNA and/or different σ-finger positioning causing alterations to the pre-organisation of the template strand.

Chapter 5: Conformation of the σ-finger with natural promoters, pR and rrnB P1

5.1 Introduction

Almost no naturally occurring promoters contain the full consensus sequences described in Chapter 4. Since initial transcription and promoter escape depend heavily on the promoter sequence and initial transcribed sequence (Heyduk and Heyduk, 2018; Winkelman et al., 2020), this chapter assesses the σ-finger conformation in two natural promoters with very different properties; pR and the ribosomal RNA P1 (rrnB P1).

pR

The sequence of the pR phage promoter used in this chapter is shown in Fig. 5.1A with the -10 and -35 regions highlighted. The open complex formed with the pR promoter is very stable, with a lifetime > 10^4 s (Ruff, Drennan, et al., 2015) and, similar to the lacCONS open complex, contains a 13bp DNA transcription bubble with a start site seven bases downstream from the -10 region. Transcription initiation complexes containing the pR promoter are found to undergo abortive initiation (Johnston and McClure, 1976) with promoter escape occurring at an RNA length of 11-nt (Henderson et al., 2017). In agreement with biochemical studies performed by Plaskon et al. (2021), recent structural studies show two classes of the pR open complex; 'active' and 'inactive' with different template strand arrangements (Saecker et al., 2021). Active molecules contained no density in the entire template strand DNA (bases -11 to +2), suggestive of dynamic behaviour, whereas inactive molecules displayed a well-resolved template DNA strand that is incompatible with binding of initiating nucleotides and therefore transcription initiation. In both these classes, the σ-finger is positioned to clash with the 5'-end of RNA when the

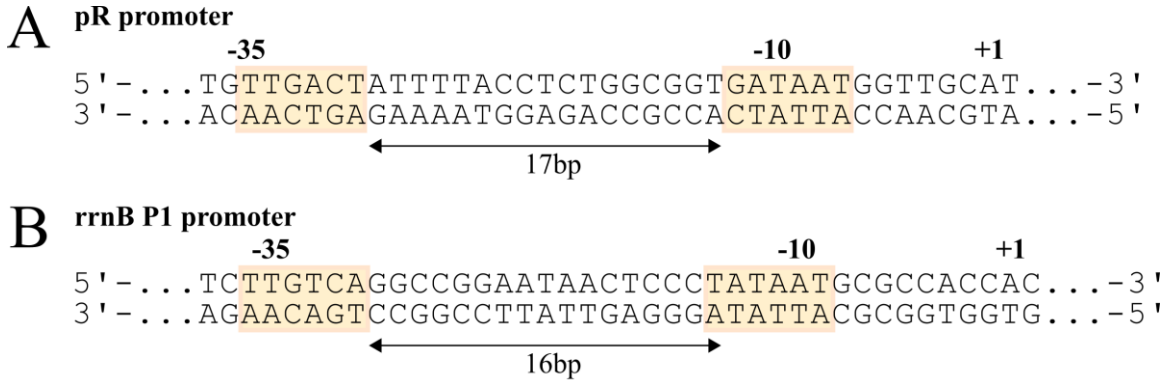

Figure 5.1: Two natural promoters with -10 and -35 regions highlighted. **(A)** pR **(B)** rrnB P1. Created with BioRender.com.

RNA is 4- to 6-nt in length (Fig. 5.2).

rrnB P1

Ribosomal RNA promoters are responsible for a large fraction of RNA synthesis (~70%) in *E. coli* during exponential phase growth (Nomura, 1999; Zengel and Lindahl, 1994). The most widely studied of these is the rrnB P1 *E. coli* promoter, the sequence of which is shown in Fig. 5.1B. Whilst this promoter contains consensus or near-consensus -10 and -35 hexamers that favour binding to the σ^{70} RNAP holoenzyme, it also contains several different features compared to the pR and lacCONS promoters such as a short (16bp) -35/-10 spacer length and a transcription start site nine base pairs downstream from the -10 region (Winkelman, Chandrangsu, et al., 2016). This different promoter architecture leads to the formation of an unstable open complex (lifetime<~1s, (Ruff, Drennan, et al., 2015) with a scrunched 15bp DNA transcription bubble that has been suggested to reduce abortive initiation and decrease the energy barrier required for promoter escape (Henderson et al., 2017). Experiments by Pupov et al. (2018) using a RNAP mutant containing a seven residue (513-519) deletion of the tip of the σ-finger increased the RNAP-rrnB P1 promoter

complex stability and shifted the transcription start site to one that is six base pairs downstream from the -10 region. These results suggest the σ-finger act as a stabilising influence on the scrunched open complex. Despite DNA scrunching in the open complex, recent structures show that similar to the pR promoter, the σ-finger is still positioned to clash with the 5'-end of RNA when the RNA is 5- to 6-nt in length (Fig. 5.2, Shin et al. 2021).

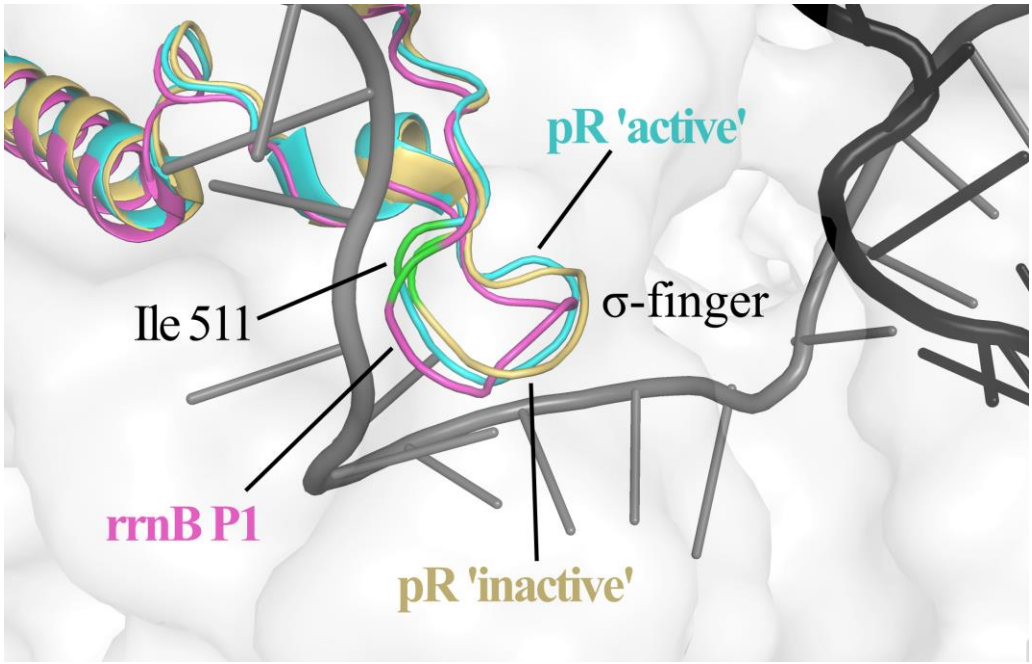

Figure 5.2: Conformation of the σ-finger in aligned structures of RPo complexes containing the pR (PDB 7MKD, 7MKE; Saecker et al., 2021) and rrnB P1 promoters (PDB 7KHB; Shin et al., 2021). The position of residue 511 where the fluorescent probe is attached in the smFRET assay is shown in green on all structures. For clarity, only the DNA template (grey) and non-template (black) strand from the PDB structure 7MKD are shown. Alignment was carried out using the 'align' function in PyMOL on the *E. coli* RNAP core subunits (α1, α2, β, β' and ω). This function performs a sequence alignment followed by a structural superposition and minimisation of the root mean square deviation between aligned residues.

Chapter summary

In this chapter, the DL σ^{70} construct described in Chapter 3 was used to assess the σ-finger conformation in primer-dependent transcription for two natural promoters; pR and rrnB P1. The results showed that the σ-finger was displaced upon the addition of the 3rd RNA nucleotide for pR promoter and by the 2nd nucleotide for rrnB P1 – both occurring before promoter escape. For both natural promoters, the σ-finger was displaced on heterogeneous timescales; ~17s (88%) and ~185s (12%) for pR; ~3.8s (84%) and ~36s (16%) for rrnB P1 – with those for the pR promoter being unexpectedly long. The origin of these long timescales was further investigated via the temperature dependence of RPo formation and it was found that RPo formation at 22ºC compared to 37ºC led to decreased displacement times for all promoters studied in primer-dependent initiation. Additionally, measurements with the rrnB P1 promoter showed σ-finger displacement to be an irreversible process, which lead to defects in promoter escape for rounds of transcription occurring after σ-finger displacement.

5.2 Experimental approach

Experiments were performed using the same approach as in Chapter 4.

The σ-finger conformation at different points of initial transcription was measured by adding the relevant dinucleotide primer with a subset of RNA nucleotides to double-labelled RPo molecules prepared with either the pR or rrnB P1 promoter (Materials and Methods 2.4). Withholding of relevant nucleotides lead to synthesis of RNA with 5'-OH group up to 4-nt for the pR promoter and 5-nt for the rrnB P1 promoter. Nucleotides were washed out of solution before imaging (Materials and Methods 2.4).

The timescale of σ-finger displacement was found by taking smFRET measurements of complexes formed with either the pR or rrnB P1 promoter during real-time RNA synthesis (Materials and Methods 2.4). Molecules exhibiting σ-finger displacement were split into two groups: 'fast-displacers' and 'slow-displacers' to analyse σ-finger conformation before nucleotide addition.

5.3 pR

σ-finger displacement point

Fig. 5.3 shows smFRET measurements of the σ-finger conformation in complexes made with the pR promoter. In RPo complexes, the E* histogram shows a major subpopulation E*~ 0.43 (86%) which corresponds to the 'in-cleft' σ-finger conformation and a minor subpopulation E* ~ 0.14 (14%) which corresponds to the 'displaced' σ-finger conformation. This is similar to results from RPo formed with the lacCONS promoter (Chapter 4.2). The E* histogram for RPitc2 also shows a small displaced σ-finger population (18%).

In the E* histogram for RPitc3, the abundance of the low FRET state E*~ 0.15 more than doubles to 40%. This shows that for a significant number of molecules, synthesis of a 3-mer RNA leads to σ-finger displacement (Fig. 5.3B). Additionally, the RPitc3 E* histogram shows a 'high FRET tail'. This contribution is from a large fraction of dynamic molecules observed in RPitc3 (36%, Fig. 5.3C) compared to RPo (15%) and RPitc2 (6%). Similarly to RPitc3, the E* histogram for RPitc4 also shows a large low FRET subpopulation corresponding to σ-finger displacement (E*~ 0.15, 39%).

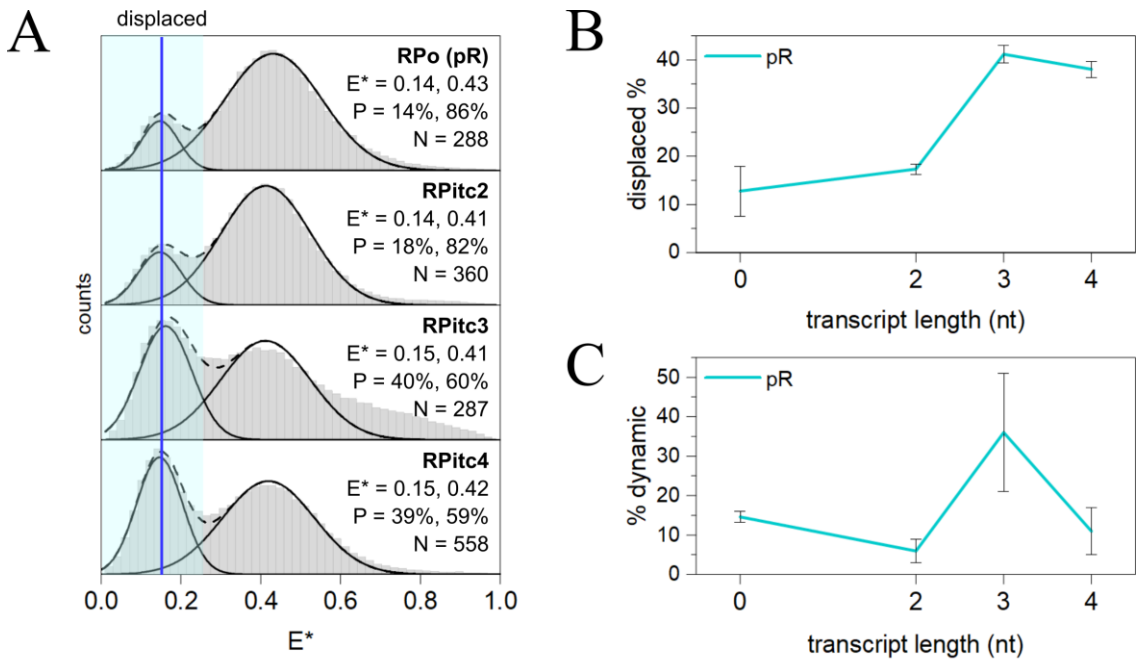

Figure 5.3: Conformation of the σ-finger in transcription initiation complexes involving the pR promoter and RNA with a 5'-OH end. **(A)** E* histograms. Estimations of mean E* values and abundances are shown for the displaced (blue line, magenta highlighting) and 'in-cleft' populations. **(B)** Occupation probabilities of the displaced σ-finger conformation and associated errors calculated from the standard error of the mean in different datasets. **(C)** Percentage of dynamic molecules and associated errors calculated from the standard error of the mean in different datasets.

σ-finger displacement at an RNA length of 3-nt is short and well before the 5- to 6-nt of RNA required to clash with the σ-finger. Previous studies show that promoter escape for the pR promoter occurs at a RNA length of 11-nt (Henderson et al., 2017) – therefore σ-finger displacement occurs before promoter escape.

σ-finger displacement in actively transcribing complexes

Fig. 5.4 shows example E*-time trajectories and the σ-finger displacement time, t_d, histogram fitted to a biexponential decay curve. The two displacement times are; $t_{d,\,short}$ ~ 17s (~88% of events) and $t_{d,\,long}$ ~ 185s (~12% of events). These timescales are almost

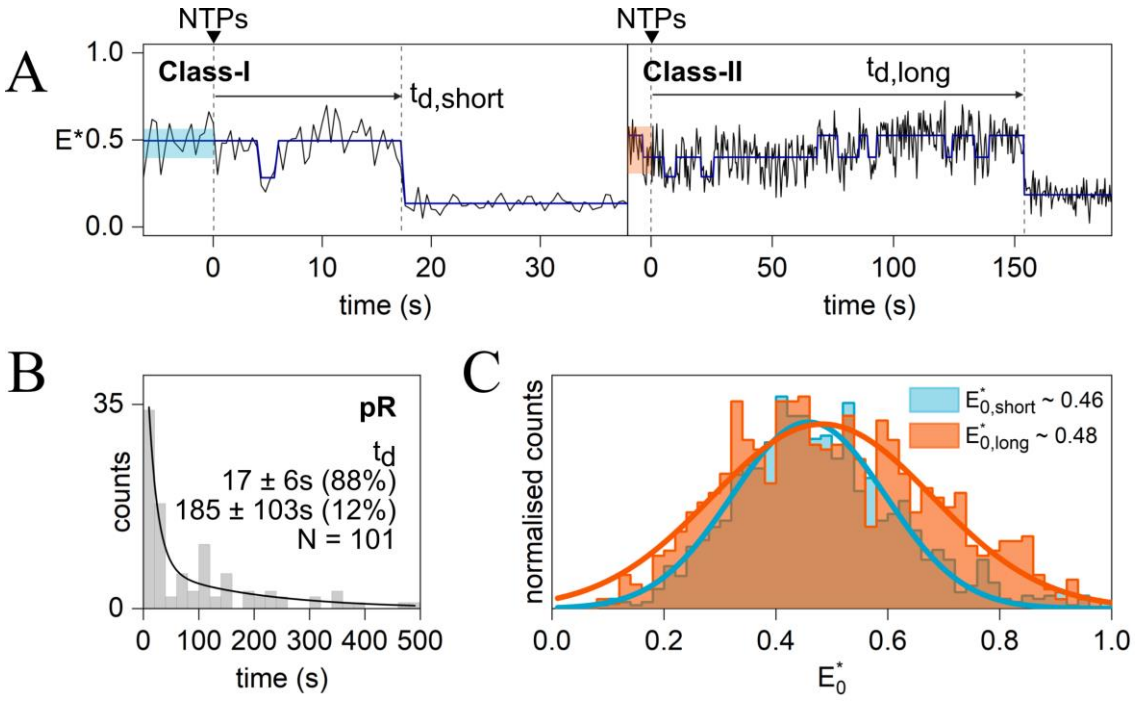

Figure 5.4: Real-time measurements of the σ-finger conformation in complexes containing the pR promoter and RNA with a 5'-OH end. **(A)** Example E*-time trajectories of Class-I and Class-II molecules showing σ-finger displacement on two timescales. E^*_0 values before NTP addition are highlighted blue (Class-I) and orange (Class-II). **(B)** Displacement time histogram (grey; bin size 20s) fitted to a double-exponential decay (black line). **(C)** E^*_0 histograms for trajectories involving a short displacement time (blue) and trajectories showing a long displacement time (orange). Created with BioRender.com.

identical to those found for the lacCONS promoter, which is surprising considering σ-finger displacement at a much earlier point in transcription for the pR (3-nt for pR cf. 10-nt for lacCONS) and *in vitro* transcription gels showing formation of even 6-mer RNA lengths within 10s (Duchi et al., 2016).

Fig. 5.4B shows E* histograms of 'fast-displacing' (t_d<36s) and 'slow-displacing' (t_d>36s) molecules before nucleotide addition. Unlike results for the lacCONS promoter, the two groups exhibit similar mean E* values (E*~0.46 for 'fast-displacers' and E*~0.48 for 'slow-displacers').

5.4 rrnB P1

σ-finger displacement point

Fig. 5.5 shows smFRET measurements of the σ-finger conformation in complexes made with the rrnB P1 promoter. The E* histogram for RPo shows a major subpopulation (E*~ 0.41, 94%) for the 'in-cleft' σ-finger conformation and a minor subpopulation for the 'displaced' σ-finger conformation (E*~ 0.14, 6%). These results are similar to RPo formed with the lacCONS (Chapter 4.2) and pR promoters. Remarkably, upon addition of the ApC dinucleotide primer to form RPitc2, the low FRET state (E*~ 0.14) abundance increased dramatically (29%). E* histograms for complexes RPitc3 onwards also show large abundances of this low FRET state (E*~ 0.14, 31% for RPitc3, 29% for RPitc4 and 44% for RPitc5; Fig. 5.5B). The large abundances of the displaced σ-finger observed in rrnB P1 complexes (29-44%) compared to lacCONS complexes (14-37%) are consistent with previous studies showing that rrnB P1 is a very strong promoter with high efficiency of promoter escape (Winkelman, Chandrangsu, et al., 2016).

σ-finger displacement and promoter escape

In order to find the point of promoter escape for the rrnB P1 promoter, a smFRET ruler similar to the one used in Chapter 4.5 was used to measure the state of the transcription bubble at various lengths of RNA. In this construct, donor and acceptor probes were placed at positions -15 on the non-template strand and +20 on the template strand of the rrnB P1 promoter (Materials and Methods Table 1). As in Chapter 4.5, the double-labelled rrnB P1 promoter was incubated with unlabelled RNAP holoenzyme and observed using TIRF-ALEX. ApC dinucleotide primer and subsets of RNA nucleotides were added to measure

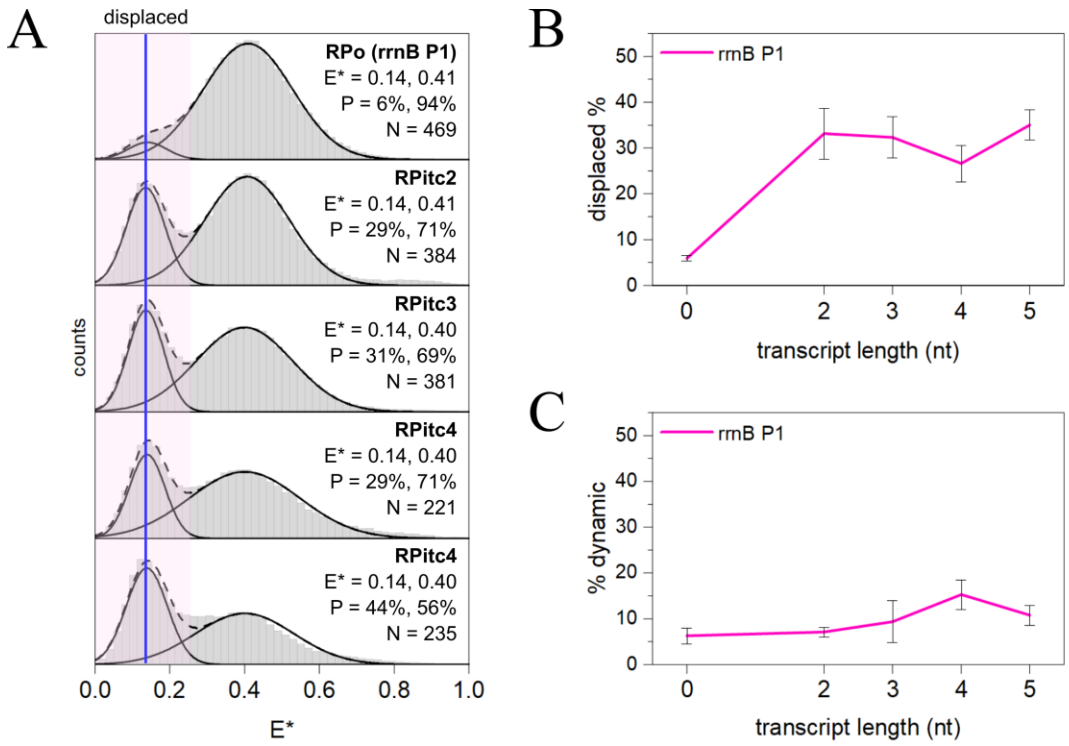

Figure 5.5: Conformation of the σ-finger in transcription initiation complexes involving the rrnB P1 promoter and RNA with a 5'-OH end. **(A)** E* histograms. Estimations of mean E* values and abundances are shown for the displaced (blue line, magenta highlighting) and 'in-cleft' populations. **(B)** Occupation probabilities of the displaced σ-finger conformation and associated errors calculated from the standard error of the mean in different datasets. **(C)** Percentage of dynamic molecules and associated errors calculated from the standard error of the mean in different datasets.

the status of the transcription bubble at RNA lengths up to 5-nt (Materials and Methods 2.4).

Fig. 5.6 shows smFRET data using this construct. As in Chapter 4.5, intermediate FRET subpopulations around E*~0.35-0.40 correspond to scrunched transcription bubble conformations. The E* histograms show a large increase in the intermediate FRET subpopulation from RPo to RPitc2. Further extensions to a 3-mer and 4-mer RNA similarly result in large intermediate FRET subpopulations, which corresponds to DNA scrunching characteristic of initial transcription. However, the E* histogram for RPitc5 shows a

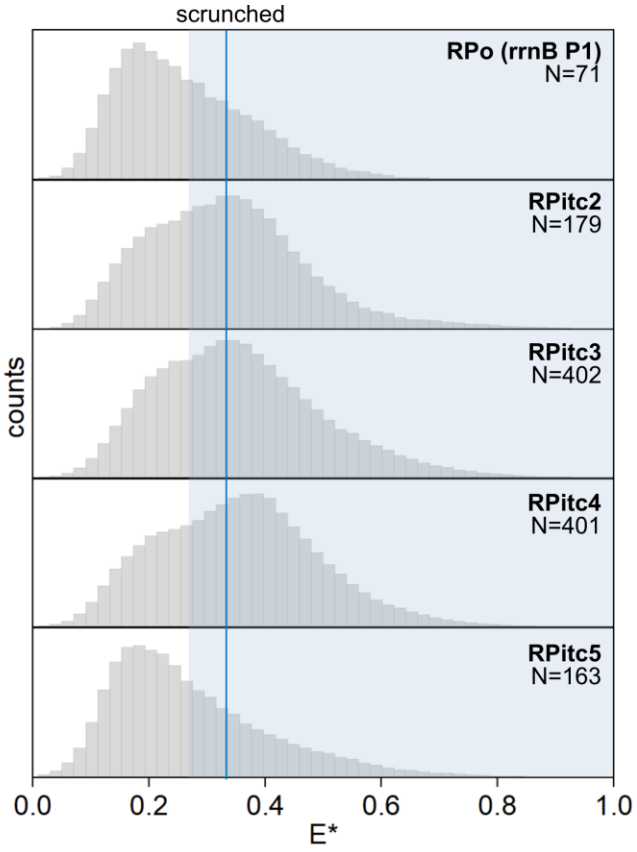

Figure 5.6: E* histograms for the conformation of the transcription bubble in transcription initiation complexes involving the rrnB P1 promoter and RNA with a 5'-OH end.

decreased abundance of this intermediate FRET subpopulation indicating the loss of the scrunched DNA transcription bubble characteristic of promoter escape. Promoter escape for the rrnB P1 promoter therefore occurs on the addition of the 5th RNA nucleotide. Combining this with σ-finger displacement at 2-nt shows that σ-finger displacement occurs before promoter escape.

Observations of σ-finger displacement in actively transcribing complexes

Fig 5.7 shows example E*-time trajectories and the σ-finger displacement time, t_d, histogram fitted to a biexponential decay curve. The two displacement times are; $t_{d,\ short}$ ~

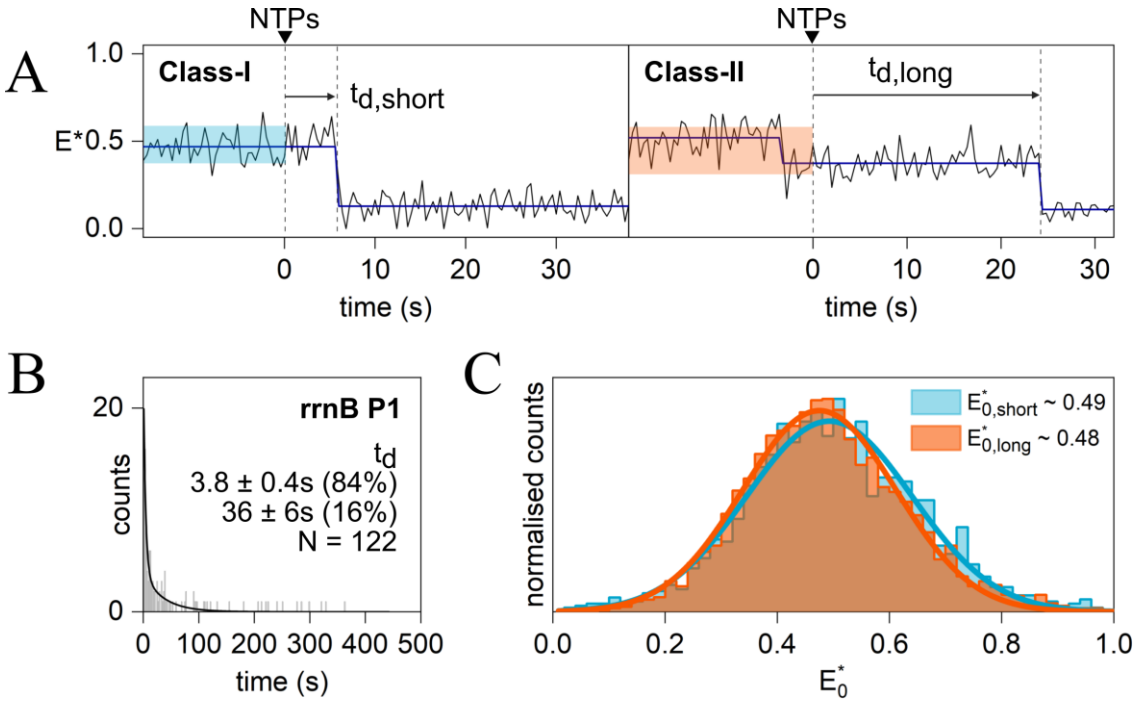

Figure 5.7: Real-time measurements of the σ-finger conformation in complexes containing the rrnB P1 promoter and RNA with a 5'-OH end. **(A)** Example E*-time trajectories of Class-I and Class-II molecules showing σ-finger displacement on two timescales. E^*_0 values before NTP addition are highlighted blue (Class-I) and orange (Class-II). **(B)** Displacement time histogram (grey; bin size 2s) fitted to a double-exponential decay (black line). **(C)** E^*_0 histograms for trajectories involving a short displacement time (blue) and trajectories showing a long displacement time (orange). Created with BioRender.com.

3.8s (~84% of events) and $t_{d,\,long}$ ~ 36s (~16% of events). These timescales are faster than that of the lacCONS and pR promoters in primer-dependent transcription, likely corresponding to the early point of displacement by 2-nt for the rrnB P1 promoter.

Fig. 5.7C shows E* histograms of 'fast-displacing' (t_d<7.3s) and 'slow-displacing' (t_d>7.3s) molecules before nucleotide addition. Unlike results for the lacCONS promoter, the two groups exhibit similar mean E* values (E*~0.48 for 'fast-displacers' and E*~0.49 for 'slow-displacers').

σ-finger displacement reversibility

In order to assess the reversibility of σ-finger displacement, RPo containing the double-labelled σ-factor was first incubated with 500μM ApC to form RPitc2 and displace the σ-finger. The dinucleotide primer was then washed out of the solution leading to dissociation of the RNA dinucleotide. This dissociation was checked by monitoring the status of the transcription bubble by taking a similar measurement but with RPo containing the double-labelled DNA construct (with Cy3B labelled at position -15 of the non-template strand DNA and Atto647N labelled at position +20 of the template strand DNA, as used to assess promoter escape; Materials and Methods Table 1).

Fig. 5.8A shows the E* histogram for 'RPitc2' complexes containing the double-labelled σ-factor after washing out of the dinucleotide primer. Remarkably, there is still a significant fraction of complexes in the displaced σ-finger conformation (E*~0.14, 24%) and it is similar to measurements shown in Fig. 5.5A taken in the presence of the dinucleotide primer (E*~ 0.14, 29%).

Fig. 5.8B shows the E* histogram for 'RPitc2' complexes containing the double-labelled DNA construct after washing out of the dinucleotide primer. Compared to RPitc2 measurements taken in the presence of ApC (Fig. 5.6), the E* histogram has a low abundance of the intermediate FRET subpopulation and closely resembles that of RPo (Fig. 5.6) – confirming dissociation of the RNA dinucleotide.

Taken together, these results show that formation of the first phosphodiester bond during synthesis of the dinucleotide permanently displaces the σ-finger.

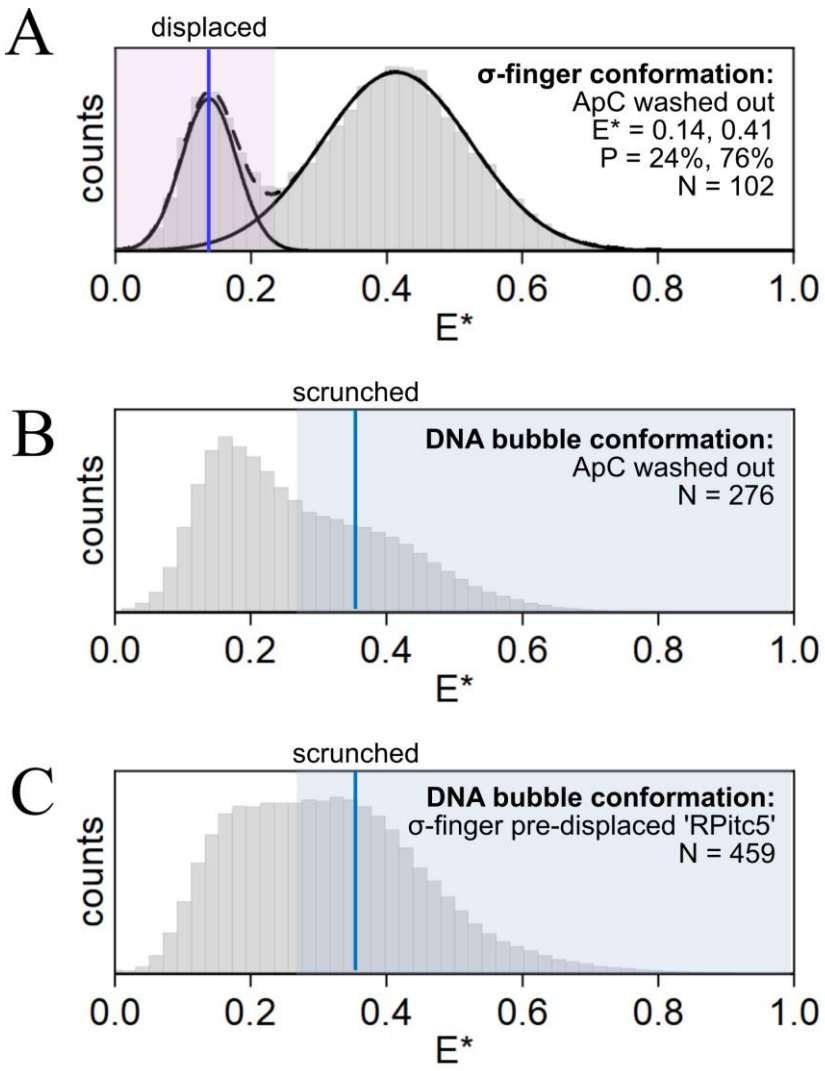

Figure 5.8: Reversibility of σ-finger displacement in complexes containing the rrnB P1 promoter and RNA with a 5'-OH end. **(A)** E* histogram showing the σ-finger conformation after forming RPitc2 and subsequently washing out the dinucleotide ApC. **(B)** E* histogram of the DNA bubble conformation after forming RPitc2 and subsequently washing out the dinucleotide ApC. **(C)** E* histogram of the DNA bubble conformation after incubating complexes with the dinucleotide ApC for 10 minutes and subsequently adding 100uM each of GTP, UTP and ATP to form'RPitc5'.

Implications for promoter escape

Results from the previous section show that σ-finger displacement is irreversible. This suggests that if the first round of transcription failed, the σ-finger would not be able to contact the DNA template strand and initiate transcription in future rounds. It is possible that this could be used as an NTP-sensing mechanism to 'turn-off' transcription when cellular nucleotide concentrations are low – in the case for the rrnB P1 promoter which has been observed to require unusually high nucleotide concentrations to initiate transcription (Gaal et al., 1997).

In order to investigate the effect on promoter escape, RPo containing the double-labelled DNA construct to monitor the transcription bubble was first incubated with ApC dinucleotide primer to form RPitc2 and displace the σ-finger. This was followed by incubation with the subset of nucleotides required for promoter escape at RPitc5. Nucleotides were washed out of solution before imaging (Materials and Methods 2.4).

Fig. 5.8C shows the result of these measurements. Remarkably, the E* histogram shows a large abundance of the intermediate FRET subpopulation corresponding to a scrunched transcription bubble – indicating that promoter escape did not occur when σ-finger contacts with the DNA template strand were removed. This is consistent with biochemical studies showing defects in promoter escape when σ-finger residues are deleted (Kulbachinskiy and Mustaev, 2006).

The abundance of the intermediate FRET subpopulation instead represents a complex involved in initial transcription that is not RPo. It is therefore possible that since the σ-

finger pre-organises the template strand DNA to initiate transcription from a particular start site, in its absence an altered start site was used and primer-independent initiation occurred.

To understand this interesting result further, future work requires an *in vitro* transcription assay.

5.5 Kinetic heterogeneity and temperature

The timescales for σ-finger displacement in transcription involving the pR promoter were surprisingly long given that displacement was found to occur upon extension to a 3-mer RNA (section 5.3). Recent studies revealed two structurally different RPo molecules with an inactive and active RPo conformation (Saecker et al., 2021). Biochemical studies, also on the pR promoter, revealed that this active RPo is the dominant subpopulation when complexes are formed at 19°C and that inactive RPo is the dominant subpopulation when complexes are formed at 37°C (Plaskon et al., 2021).

In order to test if the long σ-finger displacement times observed in transcription with the pR promoter were due to a dominant subpopulation of inactive RPo complexes, experiments in section 5.3 to observe the σ-finger conformation in actively transcribing complexes were repeated but using RPo formed by incubating DL holoenzyme with the pR promoter at 22°C instead (Materials and Methods 2.4).

Fig. 5.9A shows the σ-finger displacement time, t_d, histogram for complexes formed with the pR promoter at 22°C fitted to a biexponential decay curve. Remarkably, the two displacement times are much shorter; $t_{d,\,short}$ ~1.9s (~88% of events) and $t_{d,\,long}$ ~36s (~12% of events) compared to those found when RPo was pre-formed at 37°C (~17s and ~185s). For comparison purposes only, Fig. 5.9 also shows σ-finger displacement time data

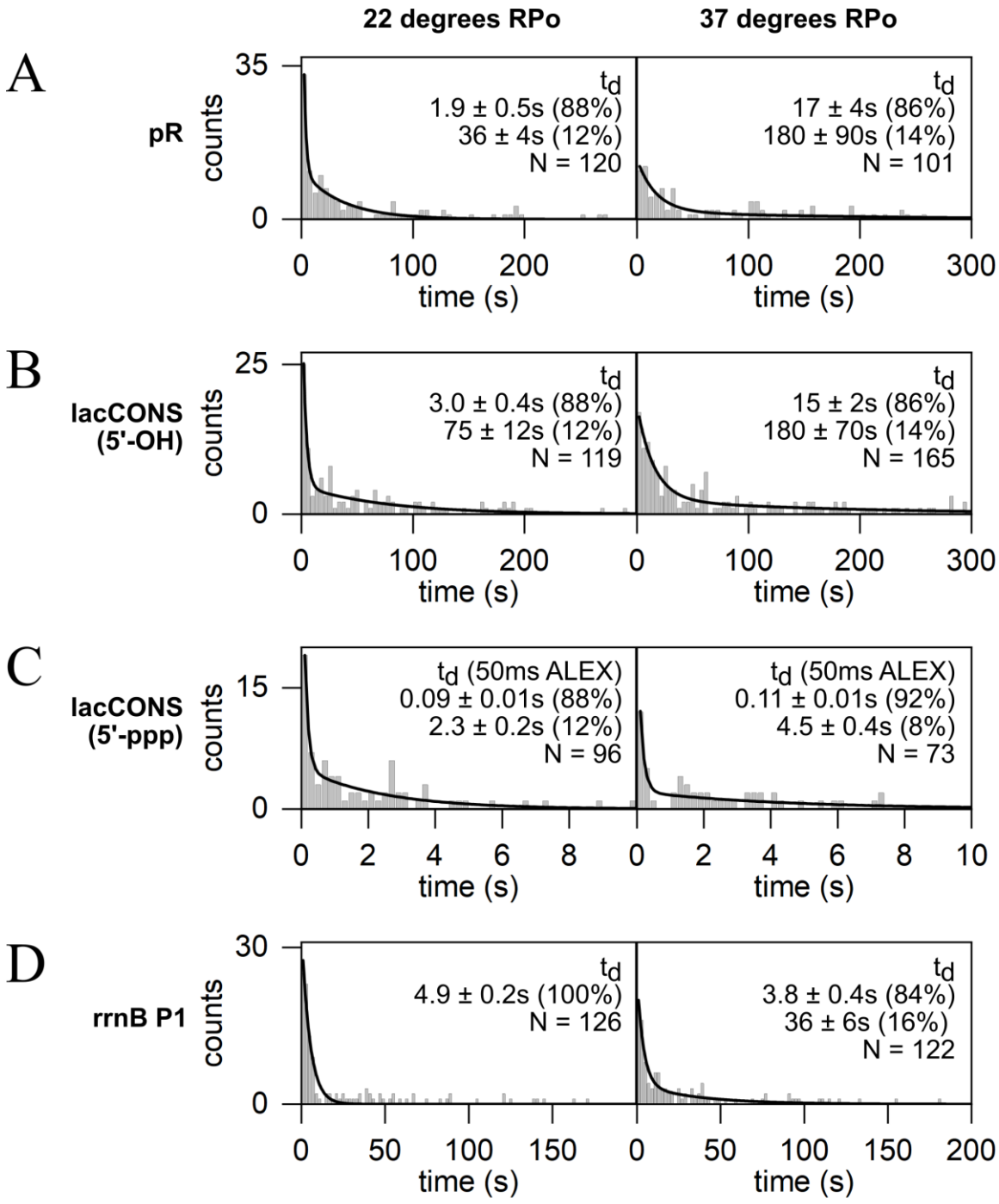

Figure 5.9: σ-finger displacement times for RPo pre-formed at different temperatures. (*left*) σ-finger displacement time histograms for RPo complexes pre-formed at 22°C for (A) pR, (B) lacCONS, (D) rrnB P1 promoters in primer-dependent transcription and (C) the lacCONS promoter in primer-independent transcription. (*right*) Panels are included for comparison purposes only and show data obtained from RPo complexes pre-formed at 37°C plotted with the same bin sizes (5s for pR, 4s for lacCONS (5'-OH), 0.2s for lacCONS (5'-ppp) and 2s for rrnB P1) as data obtained from RPo complexes pre-formed at 22°C. Histogram time axes have been truncated for clarity.

obtained from RPo pre-formed at 37°C but plotted using the same bin size as data obtained from RPo pre-formed at 22°C. The shorter displacement times found when RPo is pre-formed at 22°C suggest that the longer duration displacement events measured with RPo pre-formed at 37°C may originate from slow interconversions from an inactive to an active RPo before transcription is able to start.

In order to investigate if other promoters also exhibit this property, similar experiments were performed using RPo pre-formed at 22°C with lacCONS and rrnB P1 promoters (Materials and Methods 2.4). Experiments with the lacCONS promoter were performed for both primer-dependent and primer-independent transcription. Due to the short displacement timescales, observations involving primer-independent transcription and the lacCONS promoter used a frame time of 100ms (50ms ALEX).

Fig. 5.9B-D shows the σ-finger displacement time, t_d, histogram for complexes formed with the lacCONS and rrnB P1 promoters at 22°C fitted to decay curves. Under primer-dependent transcription, the lacCONS results show decreased displacement timescales when RPo is pre-formed at 22°C compared to 37°C (~3s and ~75s compared to ~14s and ~188s; Fig. 5.9B). This trend is also present for the rrnB P1 promoter where the σ-finger displacement time histogram obtained for RPo pre-formed at 22°C fitted equally well to a single-exponential decay (Fig. 5.9D) – showing that the abundance of the short displacement time increases when RPo is pre-formed at 22°C compared to 37°C (84% to 100%; Fig. 5.9D). However, for complexes containing the lacCONS promoter and involved in primer-independent transcription no such differences in σ-finger displacement time was observed either under 50ms ALEX (Fig. 5.9C) or 200ms ALEX (data not shown).

This may be due to different ratios of active and inactive RPo subpopulations in the presence of the ApA or pppApA dinucleotide primer.

5.6 Conclusions

This chapter shows the following results for transcription with two naturally occurring promoters, pR and rrnB P1:

- For complexes involved in primer-dependent transcription containing natural promoters pR and rrnB P1, the σ-finger is displaced upon formation of a 3- and 2-mer RNA respectively.

These RNA lengths are too short for steric clash with the σ-finger, which requires a 5- to 6-mer RNA (Murakami et al., 2002), to cause displacement. In these cases, full displacement of the σ-finger may be triggered by either a steric clash between the σ-finger and scrunched template DNA and or weakening of interactions between the σ-finger and template DNA due to scrunching instead – and happening before steric clash between the σ-finger and 5'-end of RNA can occur. Comparing these results to the late point of displacement found for the lacCONS promoter (at 10-nt) shows that promoter sequence plays a large part in determining the mechanism behind σ-finger displacement and regulation of transcription initiation kinetics

- For pR and rrnB P1 promoters, the σ-finger is displaced before promoter escape.

This is similar to results for the lacCONS promoter and therefore for all promoters studied, σ-finger displacement occurs before promoter escape. In the context of the 'protein spring' proposed by Li et al. (2020), this shows that energy stored by σ-finger folding is released before promoter escape to disrupt the stabilising contacts between the σ-finger and RNA.

As with the lacCONS promoter, this release results in the full displacement of the σ-finger and clears the path for promoter escape.

- For the rrnB P1 promoter, σ-finger displacement is irreversible.
- For the rrnB P1 promoter, complexes in which the σ-finger was first displaced did not undergo promoter escape upon RNA extension up to a 5-nt, where promoter escape had previously occurred.

This result is attributed to changes of transcription start site in the absence of the σ-finger, which would normally pre-organise and position the DNA template strand. In support of this, biochemical assays using a RNAP mutant, where σ-finger residues were deleted, showed a shifted transcription start site compared to the RNAP wildtype, where σ-finger residues were present (Pupov et al., 2018). To investigate whether first displacing the σ-finger leads to similarly shifted transcription start sites, an *in vitro* transcription assay would be required.

Gaal et al. (1997) found that transcription from the rrnB P1 promoter requires high nucleotide concentrations. Irreversible displacement of the σ-finger has interesting implications as a possible NTP-sensing mechanism since it is possible that when cellular concentrations are low and the first round of transcription fails, early irreversible displacement of the σ-finger may 'turn-off' transcription by causing defects in promoter escape.

- σ-finger displacement occurs on heterogeneous timescales of ~17s (88%) and ~185s (12%) for pR and ~3.8s (84%) and ~36s (16%) for rrnB P1 in complexes where RPo was pre-formed at 37°C.

- For the pR and rrnB P1 promoters, the heterogeneity in displacement times does not correlate to conformational differences of the σ-finger in RPo molecules.
- Pre-forming RPo at 22°C leads to decreased σ-finger displacement timecales for all promoters involved primer-dependent transcription but not necessarily primer-independent transcription.

Similar to results from the lacCONS promoter, the pR and rrnB P1 promoters also showed heterogeneous displacement times. However, unlike results from the lacCONS promoter, this kinetic heterogeneity did not correlate with the initial σ-finger in RPo. The heterogeneous displacement times are therefore not explained by either σ-finger driven alterations to template strand pre-organisation, since no difference in the initial σ-finger conformations are seen, or by any factors affecting the steric clash, since displacement occurs too early at an RNA length of 2- or 3-nt for any clash to occur. It is possible that these heterogeneous displacement timescales instead arise from conformationally different RPo molecules with similar σ-finger positions but which contain different template strand organisations, as reported in a recent structural study on the pR promoter (Saecker et al., 2021). The authors observed two types of RPo molecules; an ‘active’ one which was compatible with initial nucleotide binding and an ‘inactive’ one which was not compatible and therefore could not undergo transcription initiation. The longer displacement times may therefore originate from a rate-limiting step in the transition from an ‘inactive’ RPo to an ‘active’ RPo. This hypothesis is supported by the dramatically reduced σ-finger displacement times observed when RPo containing the pR promoter was biased towards the ‘active’ conformation by pre-forming RPo at 22°C rather than 37°C. Remarkably, all other promoters studied, lacCONS and rrnB P1, also displayed this property.

Chapter 6: Interpretation and future work

Introduction

The previous three chapters used a new smFRET construct to detect conformational changes in the σ-finger during the entirety of initial transcription with three different promoters. This chapter contains an interpretation of the results and possible future work.

6.1 Displacement of the σ-finger from the 'in-cleft' conformation during initial transcription

The smFRET construct followed the base of the σ-finger (with the fluorescent probe labelled at σ^{70}-factor residue 511) and is therefore blind to motions at the tip of the σ-finger (residues 513-519) which were observed in a recent structural study (Li et al., 2020). Following σ-finger motions through the base allows the assay to report on the full displacement of the σ-finger from the 'in-cleft' conformation in early stage initial transcription via a large change in apparent FRET efficiency, E*. For all promoters studied, displacement lead to a clear E* decrease in complexes containing the double-labelled σ-factor construct, showing σ-finger movement away from the leading edge of RNAP, since the static probe is attached at residue 366. Results using an alternative smFRET construct to follow the base of the σ-finger (with fluorescent probes labelled at σ^{70}-factor residue 511 and position -25 of the template DNA) and structural considerations suggest that the σ-finger likely moves towards the dsDNA upstream of the transcription bubble. Such a displaced conformation has been previously proposed but has never been directly observed in structural studies. This is likely due to the high mobility and disorder of the displaced structural module.

6.2 Mechanistic implications for transcription

Promoter dependent mechanism of σ-finger displacement

For all promoters studied, the σ-finger was fully displaced before promoter escape – displacement occurred upon synthesis of a 10-mer, 3-mer and 2-mer RNA for lacCONS, pR and rrnB P1 promoters respectively, whereas promoter escape occurred after synthesis of a 12-, 11- and 5-mer RNA respectively. However, observations in previous chapters also indicate distinct mechanisms through which σ-finger displacement may regulate transcription initiation kinetics for different promoters.

Recent structural work examining initial transcribing complexes showed that the 5'-end of RNA clashes with the σ-finger upon formation of a 5-mer RNA (Li et al., 2020). Further RNA extension caused the tip of the σ-finger to fold back over itself like a 'protein-spring' and led authors to hypothesise that σ-finger displacement is directly caused by collision of the 5'-end RNA and the σ-finger. In this scheme, the point of displacement is dictated by the extent to which the σ-finger protrudes inside the transcription bubble and strength of the contacts made with the template DNA and the RNA.

Consistent with this, σ-finger displacement from the 'in-cleft' position occurred late for the lacCONS promoter – upon RNA extension to 10-nt. The σ-finger was found to form contacts with the RNA from 4-nt in length (Li et al., 2020). Continued RNA extension results in further DNA scrunching and folding back of the tip of the σ-finger, both of which can contribute to storage of stress energy required to disrupt RNAP-promoter contacts. Full displacement of the σ-finger in complexes containing the lacCONS promoter in primer-dependent transcription occurs late, after synthesis of a 10-nt RNA but before promoter

escape, suggesting that the energy stored by σ-finger folding is released before promoter escape to disrupt the stabilising contacts between the σ-finger and RNA. This release results in the full displacement of the σ-finger and clears the path for promoter escape.

However, in complexes containing pR and rrnB P1 promoters, displacement occurred much earlier, upon extension to 3-nt and 2-nt respectively. These RNA lengths are too short for steric clash with the σ-finger, which requires a 5- to 6-mer RNA (Murakami et al., 2002), to cause displacement. In these cases, full displacement of the σ-finger may be triggered by either a steric clash between the σ-finger and scrunched template DNA and or weakening of interactions between the σ-finger and template DNA due to scrunching instead – and happening before steric clash between the σ-finger and 5'-end of RNA can occur. A summary of these different mechanisms is shown in Fig. 6.1.

The results show that promoter sequence plays a large part in determining the mechanism behind σ-finger displacement and regulation of transcription initiation kinetics. The smFRET construct used to detect σ-finger displacement here for lacCONS, pR and rrnB P1 promoters could be easily extended to other promoter sequences e.g. the C-7G mutation of the rrnB P1 promoter, which does not exhibit RPo scrunching and undergoes abortive initiation (Winkelman et al., 2016) or a variant of the lacCONS promoter that does not undergo pausing (Bauer et al., 2016). The development of high throughput single-molecule sequencing assays in future work would also allow a more complete investigation into specific and general features of σ-finger displacement for a large library of sequences in a systematic fashion (Andrews et al., 2022).

Furthermore, extension of the smFRET assay to more complex labelling and imaging schemes could allow simultaneous monitoring of both σ-finger displacement and

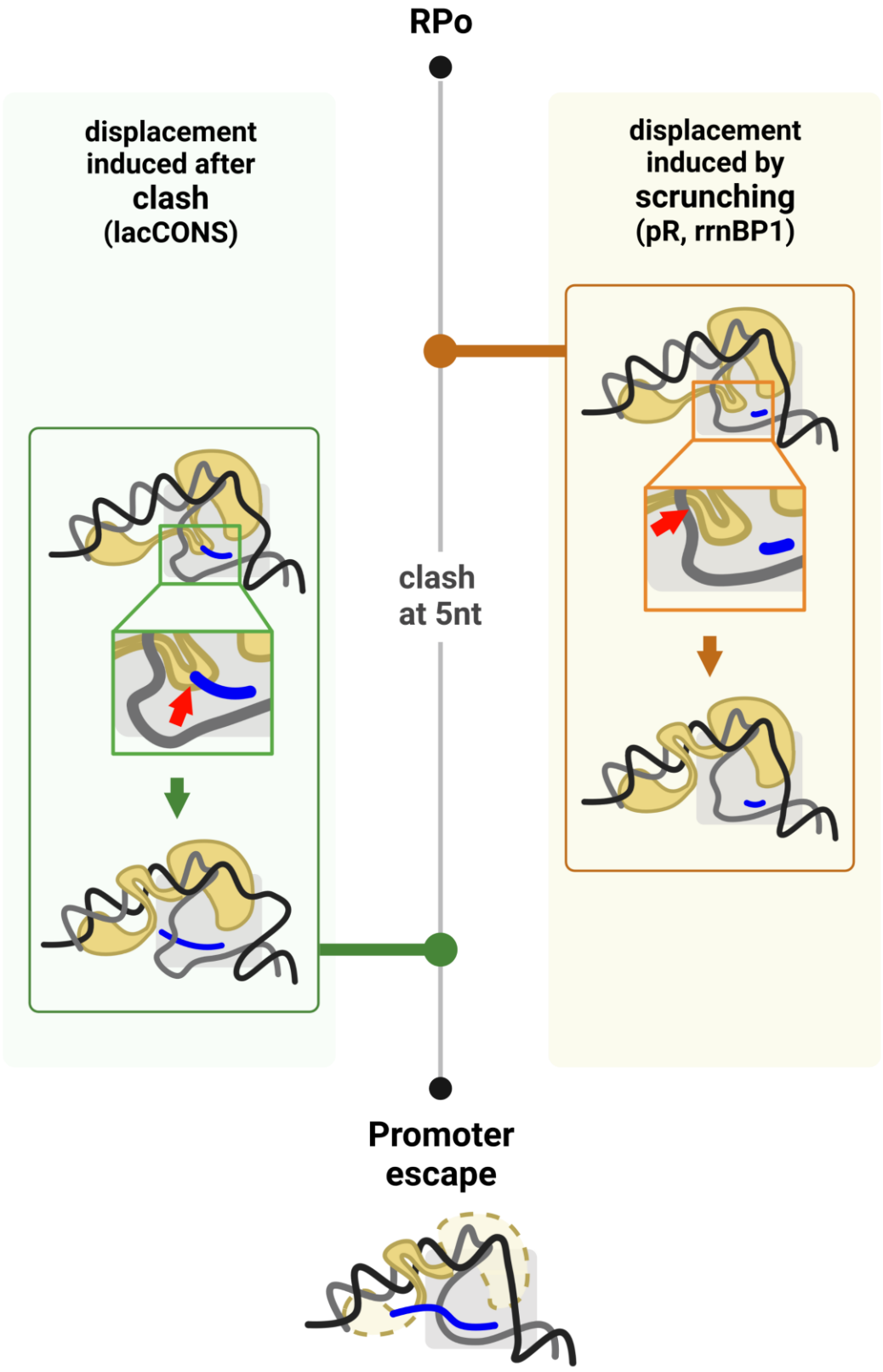

Figure 6.1: Summary of the different mechanisms for σ-finger displacement for the different promoters; lacCONS, pR and rrnB P1. Created with BioRender.com.

transcription bubble status in the same molecule – thus providing a more detailed picture of the coupling and timescales between σ-finger displacement and promoter escape.

Effect of the different mechanisms on initiation pause

Single-molecule studies of initiation complexes containing the lacCONS promoter have observed an 'initiation pause' upon extension from a 6- to 7-mer RNA, lasting around ~20s (Duchi et al., 2016; Dulin et al., 2018). Using a mutant where σ-finger resides were deleted resulted in significantly reduced pausing lifetimes, and established the σ-finger as a major pause-lifetime determinant that acts in addition to the initial transcribed sequence. In the context of the late σ-finger displacement observed for the lacCONS promoter in primer-dependent transcription (upon extension of RNA to 10-nt), the σ-finger first clashes with the RNA at a length of 5-nt and makes stabilising contacts with the 5'-end of RNA, as observed in with structural studies (Li et al., 2020). These contacts raise the translocation energy barrier, leading to pausing.

However, in the case of pR and rrnB P1 promoters where displacement of the σ-finger occurs before steric clash with the RNA chain, stabilising contacts and therefore the additional translocation energy barrier would be absent. This should result in reduced pausing lifetimes and a shorter time spent in initial transcription. Future work could test the lifetime of the pause in the pR promoter with two RNAPs – one with the σ-finger and one where σ-finger residues are deleted.

Irreversibility

Results using the rrnB P1 promoter found σ-finger displacement to be irreversible. Experiments where the σ-finger was first displaced before RNA extension up to a 5-nt,

where promoter escape had previously occurred, showed instead formation of an initial transcribing complex. This large difference is attributed to possible changes in transcription start site due to the absence of the σ-finger, which would normally pre-organise and position the DNA template strand and is supported by biochemical assays showed a shifted transcription start site and increased stability of the RNAP-rrnB P1 complex when using RNAP mutants where σ-finger residues were deleted (Pupov et al., 2018). To investigate whether first displacing the σ-finger leads to similarly shifted transcription start sites, an *in vitro* transcription assay would be required.

Gaal et al. (1997) found that transcription from the rrnB P1 promoter requires high nucleotide concentrations. The result here has interesting implications as a possible NTP-sensing mechanism since it is possible that when cellular concentrations are low and the first round of transcription fails, early irreversible displacement of the σ-finger 'turns-off' transcription by causing defects in promoter escape.

Interactions of the 5'-end of RNA on σ-finger displacement

Observations showed that σ-finger displacement during primer-independent transcription occurs much earlier and on shorter timescales than in primer-dependent transcription for the lacCONS promoter (after RNA extension to 5-nt instead of 10-nt). Primer-dependent transcription results in synthesis of RNA with a 5'-hydroxyl end whilst primer-independent transcription results in synthesis of RNA with a 5'-triphosphate end. The triphosphate group is bulky (~163Å^3 solvent accessible volume compared to ~7Å^3 solvent accessible volume of the 5'-hydroxyl end) and negatively-charged (-4 compared to 0 of the 5'-hydroxyl end) which both act to repel the negatively charged σ-finger (from four acidic residues D513-D516). In a complex containing 5-nt of RNA, the distance between the 5'-

triphosphate end and tip of the σ-finger is estimated to be ~ 3Å, which must result in electrostatic repulsion (Li et al., 2020). Furthermore, the α-phosphate of the RNA 5'-triphosphate end makes several interactions with β subunit residues not present in complexes containing RNA with a 5'-OH (Li et al., 2020). Together, these differences may contribute to earlier displacement observed for RNA with a 5'-triphosphate end compared to a 5'-hydroxyl end.

Transcription during the exponential growth phase occurs by primer-independent initiation with primer-dependent initiation increasing during the stationary phase (Nickels, 2012). The early σ-finger displacement observed in primer-independent transcription may be used to ease promoter escape and increase transcription during exponential growth, while later σ-finger displacement in primer-dependent initiation could lead to reduce transcription in stationary phase. In this way, σ-finger interactions between the 5'-end of RNA could be used to regulate gene expression in different growth phases.

One other form of primer-independent initiation not investigated here is the use of a non-canonical initiating nucleotide (NCINs; Barvík et al., 2017; Bird et al., 2016). The differing 5'-end moieties of NCINs, which is positively charged in NAD+ and has no charge in dp-CoA, may form different interactions with the negatively charged σ-finger and affect displacement kinetics. The smFRET assay used in this thesis could be easily adapted to investigate this.

Allosteric effects

Although not investigated here, allosteric effects may also influence σ-finger displacement. Aside from interacting with the σ-finger directly, cross-linking studies show that the

scrunched template DNA strand may also impinge on the β flap and β' clamp (Winkelman et al., 2015). Interactions between the β flap and β' clamp 'lid' domain form a steric barrier to displacement of the σ-finger and therefore led Winkelman et al. (2015) to suggest that disruption of these interactions by DNA template scrunching could allow for σ-finger displacement.

6.3 σ-finger displacement is heterogeneous

Different functional classes

Measurements of the σ-finger in complexes containing the lacCONS promoter during real-time primer-dependent transcription show only ~25% of molecules exhibit σ-finger displacement. Rough estimates suggest ~20% of displacement events are missed due to premature termination of FRET time-trajectories due to photobleaching and therefore that ~55% of molecules are transcriptionally inactive or stuck in early rounds of transcription initiation during which the σ-finger is not yet displaced. This picture is consistent with previous biochemical and single-molecule studies that report a subpopulation of molecules in abortive initiation not proceeding to promoter escape (Dulin et al., 2018; Henderson et al., 2017).

Different promoters

The σ-finger was displaced on two timescales for all promoters studied.

Measured displacement times using the lacCONS promoter (~1-190s) were of a similar range to promoter escape times (~13 s, 96% and ~200 s, 4%) observed in another single-molecule study also using the lacCONS promoter (G. Wang et al., 2016). These kinetically distinct molecules also exhibited different initial σ-finger conformations – showing

correlation between the heterogeneity in displacement times and conformationally different RPitc2 molecules. This could have two causes; different σ-finger positioning creating differences in the steric clash with the 5'-end of RNA and/or different σ-finger positioning causing alterations to the pre-organisation of the template strand.

Whilst the pR promoter also showed heterogeneous displacement times, the correlation with the initial σ-finger conformation was not observed. In this case, the heterogeneous displacement times are therefore not explained by either σ-finger driven alterations to template strand pre-organisation, since no difference in the initial σ-finger conformations are seen, or by any factors affecting the steric clash, since displacement occurs too early at an RNA length of 3-nt for any clash to occur. It is possible that these heterogeneous displacement timescales instead arise from conformationally different RPo molecules with similar σ-finger positions but which contain different template strand organisations, as reported in a recent structural study on the pR promoter (Saecker et al., 2021). The authors observed two types of RPo molecules; an 'active' one which was compatible with initial nucleotide binding and an 'inactive' one which was not compatible and therefore could not undergo transcription initiation. The longer displacement times may therefore originate from a rate-limiting step in the transition from an 'inactive' RPo to an 'active' RPo. This hypothesis is supported by the dramatically reduced σ-finger displacement times observed when RPo containing the pR promoter was biased towards the 'active' conformation by pre-forming RPo at 22°C rather than 37°C. Remarkably, all other promoters studied, lacCONS and rrnB P1, also displayed this property.

Displacement times for the rrnB P1 promoter were also heterogeneous, although very fast. Since displacement occurs at an RNA length of 2-nt, this displacement time reports on the

binding of the dinucleotide primer, and not on the rate of RNA synthesis. Similar to the pR promoter, the heterogeneity in σ-finger displacement times likely also indicate the presence of different RPo conformations with different template strand organisation and therefore differing binding efficiencies to the dinucleotide primer.

Chapter 7: Conclusions

This thesis has scratched the surface of how the σ-finger regulates prokaryotic transcription. *In vitro* transcription experiments using single-molecule FRET have been analysed and the first observations of full σ-finger displacement have been made. Further, σ-finger displacement has been identified as a key step that influences the kinetics of initial transcription on many levels such as different genes or physiological states.

A promoter-dependent mechanism for σ-finger displacement has been suggested. In the case of the consensus lacCONS promoter, displacement is seen to depend on a steric clash between the σ-finger and growing RNA chain, so that different σ-finger positionings directly influence the time for displacement. This interaction may be modulated by the 5'-end of RNA to regulate gene expression during different phases of prokaryotic growth. For instance, early σ-finger displacement seen in primer-independent transcription may cause higher levels of transcription in the exponential growth phase versus the stationary growth phase. However, for the natural promoters pR and rrnB P1, displacement is seen to be dependent on interactions between the σ-finger and template DNA, and not on the steric clash, so that displacement timescales are instead dictated by the different template DNA conformations in 'active' and 'inactive' RPo molecules.

This model is still far from a global understanding of the role the σ-finger plays in the regulation of gene expression. However, the assays and mechanisms introduced in this thesis are a starting point and have initiated questions that could be explored further. For example, can irreversibility of σ-finger displacement provide a possible NTP-sensing mechanism? The concept of using σ-finger conformations to control and regulate transcription may also offer a new target for antibiotics which could be in the form of small

molecules that are able to delay or block promoter escape of crucial genes, such as those of rRNA.

Finally, all other cellular polymerases in archaea and eukaryotes contain a similar structural module to the bacterial σ-finger which protrudes into RNAP active-centre cleft, interacts with the template DNA and occupies the path of the nascent RNA; TFB zinc ribbon and CSB for archaeal RNAP (Renfrow et al., 2004), Rrn7 zinc ribbon and B reader for RNAP-I (Engel et al., 2017; Han et al., 2017), TFIIB zinc ribbon and B reader for RNAP-II (He et al., 2016; Kostrewa et al., 2009; Liu et al., 2010; Plaschka et al., 2016) and Brf1 zinc ribbon for RNAP-III (Abascal-Palacios et al., 2018; Vorländer et al., 2018). The conclusions reached in this thesis provide a basis from which to form a complete model of the mechanisms by which these structural modules regulate protein synthesis in other kingdoms of life.

Future studies using single-molecule FRET have the potential to unpick mechanisms, such as the σ-finger, that regulate transcription initiation and promoter escape. In the short term, current methods should allow studies of the σR1.1 domain that, similar to the σ-finger, acts as a nucleic acid mimic but instead prevents the association of nonspecific DNA with RNAP. By combining smFRET with different single fluorescence-based methods such as smPIFE or smUIFE and recently developed systems such as spectral microscopes, it may also be possible to determine the sequence of events and timings of promoter escape. The work presented here on the σ-finger may be useful for this as it provides a potential labelling site to track σ-finger movements and relate displacement to the breakage of contacts between σR3 and the -35 hexamer, and also those of σR2 and the -10 hexamer. In the longer term, it may be important to comprehensively determine the sequence

determinants for the kinetic parameters of initial transcription and promoter escape. In order to achieve this, the combination of high-throughput transcription assays and with single-molecule sequencing assays is an exciting prospect.

‘The Moving Finger writes; and, having writ,

Moves on: nor all thy Piety nor Wit

Shall lure it back to cancel half a Line,

nor all thy tears wash out a word of it.’

~ Omar Khayyam

Bibliography

Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. *Nature*, *553*(7688), 301–306. https://doi.org/10.1038/nature25441

Andrecka, J., Lewis, R., Brückner, F., Lehmann, E., Cramer, P., and Michaelis, J. (2008). Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proceedings of the National Academy of Sciences*, *105*(1), 135–140. https://doi.org/10.1073/pnas.0703815105

Andrecka, J., Treutlein, B., Arcusa, M. A. I., Muschielok, A., Lewis, R., Cheung, A. C. M., … Michaelis, J. (2009). Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. *Nucleic Acids Research*, *37*(17), 5803–5809. https://doi.org/10.1093/nar/gkp601

Andrews, R., Steuer, H., El-Sagheer, A. H., Mazumder, A., Sayyed, H. el, Shivalingam, A., … Kapanidis, A. N. (2022, February 27). *Transient DNA binding to gapped DNA substrates links DNA sequence to the single-molecule kinetics of protein-DNA interactions*. bioRxiv. https://doi.org/10.1101/2022.02.27.482175

Axelrod, D., Thompson, N. L., and Burghardt, T. P. (1983). Total internal reflection fluorescent microscopy. *Journal of Microscopy*, *129*(1), 19–28. https://doi.org/10.1111/j.1365-2818.1983.tb04158.x

Barne, K. A., Bown, J. A., Busby, S. J. W., and Minchin, S. D. (1997). Region 2.5 of the Escherichia coli RNA polymerase σ70 subunit is responsible for the recognition of the 'extended -10' motif at promoters. *EMBO Journal*, *16*(13), 4034–4040. https://doi.org/10.1093/emboj/16.13.4034

Barvík, I., Rejman, D., Panova, N., Šanderová, H., and Krásný, L. (2017). Non-canonical transcription initiation: The expanding universe of transcription initiating substrates. *FEMS Microbiology Reviews*, *41*(2), 131–138. https://doi.org/10.1093/femsre/fuw041

Bauer, D. L. V., Duchi, D., and Kapanidis, A. N. (2016). E. Coli RNA Polymerase Pauses during Initial Transcription. *Biophysical Journal*, *110*(3, Supplement 1), 21a. https://doi.org/10.1016/j.bpj.2015.11.170

Bird, J. G., Zhang, Y., Tian, Y., Panova, N., Barvík, I., Greene, L., … Nickels, B. E. (2016). The mechanism of RNA 5′ capping with NAD+, NADH and desphospho-CoA. *Nature*, *535*(7612), 444–447. https://doi.org/10.1038/nature18622

Browning, D. F., and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology*, *14*(10), 638–650. https://doi.org/10.1038/nrmicro.2016.103

Burgess, R. R., and Travers, A. A. (1970). Escherichia coli RNA polymerase: Purification, subunit structure, and factor requirements. *Federation Proceedings*, *29*(3), 1164–1169.

Callaci, S., Heyduk, E., and Heyduk, T. (1998). Conformational Changes of Escherichia coli RNA Polymerase ς70 Factor Induced by Binding to the Core Enzyme*. *Journal of Biological Chemistry*, *273*(49), 32995–33001. https://doi.org/10.1074/jbc.273.49.32995

Callaci, S., Heyduk, E., and Heyduk, T. (1999). Core RNA Polymerase from E. coli Induces a Major Change in the Domain Arrangement of the σ70 Subunit. *Molecular Cell*, *3*(2), 229–238. https://doi.org/10.1016/S1097-2765(00)80313-5

Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., … Darst, S. A. (2002). Structure of the Bacterial RNA Polymerase Promoter Specificity σ Subunit. *Molecular Cell*, *9*(3), 527–539. https://doi.org/10.1016/S1097-2765(02)00470-7

Carpousis, A. J., and Gralla, J. D. (1980). Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry*, *19*(14), 3245–3253. https://doi.org/10.1021/bi00555a023

Cashel, M., Hsu, L. M., and Hernandez, V. J. (2003). Changes in conserved region 3 of Escherichia coli sigma 70 reduce abortive transcription and enhance promoter escape. *The Journal of Biological Chemistry*, *278*(8), 5539–5547. https://doi.org/10.1074/jbc.M211430200

Cha, J., and Lee, J. Y. (2021). A novel high-throughput single-molecule technique: DNA curtain. *Journal of the Korean Physical Society*, *78*(5), 442–448. https://doi.org/10.1007/s40042-020-00031-9

Chen, J., Chiu, C., Gopalkrishnan, S., Chen, A. Y., Olinares, P. D. B., Saecker, R. M., … Darst, S. A. (2020). Stepwise Promoter Melting by Bacterial RNA Polymerase. *Molecular Cell*, *78*(2), 275–288. https://doi.org/10.1016/j.molcel.2020.02.017

Crick, F. H. (1958). On protein synthesis. *Symposia of the Society for Experimental Biology*, *12*, 138–163.

Danson, A. E., Jovanovic, M., Buck, M., and Zhang, X. (2019). Mechanisms of σ54-Dependent Transcription Initiation and Regulation. *Journal of Molecular Biology*, *431*(20), 3960–3974. https://doi.org/10.1016/j.jmb.2019.04.022

Darst, S. A. (2001). Bacterial RNA polymerase. *Current Opinion in Structural Biology*, *11*(2), 155–162. https://doi.org/10.1016/s0959-440x(00)00185-8

Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S. (1975). Genetic Regulation: The Lac Control Region. *Science*, *187*(4171), 27–35. https://doi.org/10.1126/science.1088926

Duchi, D., Bauer, D. L. V., Fernandez, L., Evans, G., Robb, N., Hwang, L. C., … Kapanidis, A. N. (2016). RNA Polymerase Pausing during Initial Transcription. *Molecular Cell*, *63*(6), 939–950. https://doi.org/10.1016/j.molcel.2016.08.011

Duchi, D., Gryte, K., Robb, N. C., Morichaud, Z., Sheppard, C., Brodolin, K., … Kapanidis, A. N. (2017). Conformational heterogeneity and bubble dynamics in single bacterial transcription initiation complexes. *Nucleic Acids Research*, *46*(2), 677–688. https://doi.org/10.1093/nar/gkx1146

Duchi, D., Mazumder, A., Malinen, A. M., Ebright, R. H., and Kapanidis, A. N. (2018). The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. *Nucleic Acids Research*, *46*(14), 7284–7295. https://doi.org/10.1093/nar/gky482

Dulin, D., Bauer, D. L. V., Malinen, A. M., Bakermans, J. J. W., Kaller, M., Morichaud, Z., … Kapanidis, A. N. (2018). Pausing controls branching between productive and non-productive pathways during initial transcription in bacteria. *Nature Communications*, *9*(1), 1478. https://doi.org/10.1038/s41467-018-03902-9

Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., … Cramer, P. (2017). Structural Basis of RNA Polymerase I Transcription Initiation. *Cell*, *169*(1), 120–131. https://doi.org/10.1016/j.cell.2017.03.003

Feklístov, A., Bae, B., Hauver, J., Lass-Napiorkowska, A., Kalesse, M., Glaus, F., … Darst, S. A. (2017). RNA polymerase motions during promoter melting. *Science*, *356*(6340), 863–866. https://doi.org/10.1126/science.aam7858

Feklístov, A., Sharon, B. D., Darst, S. A., and Gross, C. A. (2014). Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. *Annual Review of Microbiology*, *68*(1), 357–376. https://doi.org/10.1146/annurev-micro-092412-155737

Förster, Th. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen Der Physik*, *437*, 55–75. https://doi.org/10.1002/andp.19484370105

Friedman, L. J., and Gelles, J. (2012). Mechanism of Transcription Initiation at an Activator-Dependent Promoter Defined by Single-Molecule Observation. *Cell*, *148*(4), 679–689. https://doi.org/10.1016/j.cell.2012.01.018

Friedman, L. J., Mumm, J. P., and Gelles, J. (2013). RNA polymerase approaches its promoter without long-range sliding along DNA. *Proceedings of the National Academy of Sciences*, *110*(24), 9740–9745. https://doi.org/10.1073/pnas.1300221110

Gaal, T., Bartlett, M. S., Ross, W., Turnbough, C. L., and Gourse, R. L. (1997). Transcription Regulation by Initiating NTP Concentration: RRNA Synthesis in Bacteria. *Science*, *278*(5346), 2092–2097. https://doi.org/10.1126/science.278.5346.2092

Gourse, R. L., and Landick, R. (2012). CoSMoS Unravels Mysteries of Transcription Initiation. *Cell*, *148*(4), 635–637. https://doi.org/10.1016/j.cell.2012.01.042

Gralla, J. D., Carpousis, A. J., and Stefano, J. E. (1980). Productive and abortive initiation of transcription in vitro at the lac UV5 promoter. *Biochemistry*, *19*(25), 5864–5869. https://doi.org/10.1021/bi00566a031

Grohmann, D., Nagy, J., Chakraborty, A., Klose, D., Fielden, D., Ebright, R. H., … Werner, F. (2011). The Initiation Factor TFE and the Elongation Factor Spt4/5 Compete for the

RNAP Clamp during Transcription Initiation and Elongation. *Molecular Cell*, *43*(2), 263–274. https://doi.org/10.1016/j.molcel.2011.05.030

Gueroui, Z., Place, C., Freyssingeas, E., and Berge, B. (2002). Observation by fluorescence microscopy of transcription on single combed DNA. *Proceedings of the National Academy of Sciences*, *99*(9), 6005–6010. https://doi.org/10.1073/pnas.092561399

Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996). Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences*, *93*(13), 6264–6268. https://doi.org/10.1073/pnas.93.13.6264

Han, Y., Yan, C., Nguyen, T. H. D., Jackobel, A. J., Ivanov, I., Knutson, B. A., and He, Y. (2017). Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *ELife*, *6*, e27414. https://doi.org/10.7554/eLife.27414

Harada, Y., Funatsu, T., Murakami, K., Nonoyama, Y., Ishihama, A., and Yanagida, T. (1999). Single-Molecule Imaging of RNA Polymerase-DNA Interactions in Real Time. *Biophysical Journal*, *76*(2), 709–715. https://doi.org/10.1016/S0006-3495(99)77237-1

Haugen, S. P., Ross, W., Manrique, M., and Gourse, R. L. (2008). Fine structure of the promoter–σ region 1.2 interaction. *Proceedings of the National Academy of Sciences*, *105*(9), 3292–3297. https://doi.org/10.1073/pnas.0709513105

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. *Nature*, *533*(7603), 359–365. https://doi.org/10.1038/nature17970

Heller, I., Hoekstra, T. P., King, G. A., Peterman, E. J. G., and Wuite, G. J. L. (2014). Optical Tweezers Analysis of DNA–Protein Complexes. *Chemical Reviews*, *114*(6), 3087–3119. https://doi.org/10.1021/cr4003006

Henderson, K. L., Felth, L. C., Molzahn, C. M., Shkel, I., Wang, S., Chhabra, M., … Record, M. T. (2017). Mechanism of transcription initiation and promoter escape by E. coli RNA polymerase. *Proceedings of the National Academy of Sciences*, *114*(15), 3032–3040. https://doi.org/10.1073/pnas.1618675114

Heyduk, E., and Heyduk, T. (2018). DNA template sequence control of bacterial RNA polymerase escape from the promoter. *Nucleic Acids Research*, *46*(9), 4469–4486. https://doi.org/10.1093/nar/gky172

Holden, S. J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O. J., and Kapanidis, A. N. (2010). Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophysical Journal*, *99*(9), 3102–3111. https://doi.org/10.1016/j.bpj.2010.09.005

Hsu, L. M. (2002). Promoter clearance and escape in prokaryotes. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, *1577*(2), 191–207. https://doi.org/10.1016/S0167-4781(02)00452-9

Hsu, L. M. (2009). Monitoring abortive initiation. *Methods*, *47*(1), 25–36. https://doi.org/10.1016/j.ymeth.2008.10.010

Hwang, H., Kim, H., and Myong, S. (2011). Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences*, *108*(18), 7414–7418. https://doi.org/10.1073/pnas.1017672108

Hwang, H., and Myong, S. (2014). Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chemical Society Reviews*, *43*(4), 1221–1229. https://doi.org/10.1039/C3CS60201J

Johnston, D. E., and McClure, W. R. (1976). Abortive Initiation of In Vitro RNA Synthesis on Bacteriophage λ DNA. *Cold Spring Harbor Monograph Archive*, *6*(0), 413–428. https://doi.org/10.1101/0.413-428

Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., … Seidel, C. A. M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature Methods*, *9*(12), 1218–1225. https://doi.org/10.1038/nmeth.2222

Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences*, *101*(24), 8936–8941. https://doi.org/10.1073/pnas.0401690101

Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S., and Ebright, R. H. (2006). Initial Transcription by RNA Polymerase Proceeds Through a DNA-Scrunching Mechanism. *Science*, *314*(5802), 1144–1147. https://doi.org/10.1126/science.1131399

Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., … Weiss, S. (2005). Retention of Transcription Initiation Factor σ70 in Transcription Elongation: Single-Molecule Analysis. *Molecular Cell*, *20*(3), 347–356. https://doi.org/10.1016/j.molcel.2005.10.012

Ko, J., and Heyduk, T. (2014). Kinetics of promoter escape by bacterial RNA polymerase: Effects of promoter contacts and transcription bubble collapse. *Biochemical Journal*, *463*(1), 135–144. https://doi.org/10.1042/BJ20140179

Koh, H. R., Roy, R., Sorokina, M., Tang, G.-Q., Nandakumar, D., Patel, S. S., and Ha, T. (2018). Correlating Transcription Initiation and Conformational Changes by a Single-Subunit RNA Polymerase with Near Base-Pair Resolution. *Molecular Cell*, *70*(4), 695–706. https://doi.org/10.1016/j.molcel.2018.04.018

Kostrewa, D., Zeller, M. E., Armache, K.-J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature*, *462*(7271), 323–330. https://doi.org/10.1038/nature08548

Kulbachinskiy, A., and Mustaev, A. (2006). Region 3.2 of the σ Subunit Contributes to the Binding of the 3′-Initiating Nucleotide in the RNA Polymerase Active Center and Facilitates Promoter Clearance during Initiation. *Journal of Biological Chemistry*, *281*(27), 18273–18276. https://doi.org/10.1074/jbc.C600060200

Lerner, E., Chung, S., Allen, B. L., Wang, S., Lee, J., Lu, S. W., … Weiss, S. (2016). Backtracked and paused transcription initiation intermediate of Escherichia coli RNA polymerase. *Proceedings of the National Academy of Sciences*, *113*(43), 6562–6571. https://doi.org/10.1073/pnas.1605038113

Li, L., Molodtsov, V., Lin, W., Ebright, R. H., and Zhang, Y. (2020). RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. *Proceedings of the National Academy of Sciences*, *117*(11), 5801–5809. https://doi.org/10.1073/pnas.1920747117

Lisser, S., and Margalit, H. (1993). Compilation of E.coli mRNA promoter sequences. *Nucleic Acids Research*, *21*(7), 1507–1516. https://doi.org/10.1093/nar/21.7.1507

Liu, X., Bushnell, D. A., Wang, D., Calero, G., and Kornberg, R. D. (2010). Structure of an RNA Polymerase II–TFIIB Complex and the Transcription Initiation Mechanism. *Science*, *327*(5962), 206–209. https://doi.org/10.1126/science.1182015

Ma, J., Bai, L., and Wang, M. D. (2013). Transcription Under Torsion. *Science*, *340*(6140), 1580–1583. https://doi.org/10.1126/science.1235441

Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2006). Direct Observation of Abortive Initiation and Promoter Escape within Single Immobilized Transcription Complexes. *Biophysical Journal*, *90*(4), 1419–1431. https://doi.org/10.1529/biophysj.105.069252

Mazumder, A., Ebright, R. H., and Kapanidis, A. N. (2021). Transcription initiation at a consensus bacterial promoter proceeds via a 'bind-unwind-load-and-lock' mechanism. *ELife*, *10*, e70090. https://doi.org/10.7554/eLife.70090

Mazumder, A., and Kapanidis, A. N. (2019). Recent Advances in Understanding σ70-Dependent Transcription Initiation Mechanisms. *Journal of Molecular Biology*, *431*(20), 3947–3959. https://doi.org/10.1016/j.jmb.2019.04.046

Mazumder, A., Lin, M., Kapanidis, A. N., and Ebright, R. H. (2020). Closing and opening of the RNA polymerase trigger loop. *Proceedings of the National Academy of Sciences*, *117*(27), 15642–15649. https://doi.org/10.1073/pnas.1920427117

McClure, W. R., Hawley, D. K., Youderian, P., and Susskind, M. M. (1983). DNA Determinants of Promoter Selectivity in Escherichia coli. *Cold Spring Harbor Symposia on Quantitative Biology*, *47*, 477–481. https://doi.org/10.1101/SQB.1983.047.01.057

Michaelis, J., and Treutlein, B. (2013). Single-Molecule Studies of RNA Polymerases. *Chemical Reviews*, *113*(11), 8377–8399. https://doi.org/10.1021/cr400207r

Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001). Translocation of σ70 with RNA Polymerase during Transcription: Fluorescence Resonance Energy Transfer Assay for Movement Relative to DNA. *Cell*, *106*(4), 453–463. https://doi.org/10.1016/S0092-8674(01)00464-0

Murakami, K. S., Masuda, S., and Darst, S. A. (2002). Structural Basis of Transcription Initiation: RNA Polymerase Holoenzyme at 4 Å Resolution. *Science*, *296*(5571), 1280–1284. https://doi.org/10.1126/science.1069594

Muschielok, A., Andrecka, J., Jawhari, A., Brückner, F., Cramer, P., and Michaelis, J. (2008). A nano-positioning system for macromolecular structural analysis. *Nature Methods*, *5*(11), 965–971. https://doi.org/10.1038/nmeth.1259

Muschielok, A., and Michaelis, J. (2011). Application of the Nano-Positioning System to the Analysis of Fluorescence Resonance Energy Transfer Networks. *The Journal of Physical Chemistry B*, *115*(41), 11927–11937. https://doi.org/10.1021/jp2060377

Nagy, J., Grohmann, D., Cheung, A. C. M., Schulz, S., Smollett, K., Werner, F., and Michaelis, J. (2015). Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. *Nature Communications*, *6*(1), 6161. https://doi.org/10.1038/ncomms7161

Narayanan, A., Vago, F. S., Li, K., Qayyum, M. Z., Yernool, D., Jiang, W., and Murakami, K. S. (2018). Cryo-EM structure of Escherichia coli σ70 RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation. *Journal of Biological Chemistry*, *293*(19), 7367–7375. https://doi.org/10.1074/jbc.RA118.002161

Neuman, K. C., Abbondanzieri, E. A., Landick, R., Gelles, J., and Block, S. M. (2003). Ubiquitous Transcriptional Pausing Is Independent of RNA Polymerase Backtracking. *Cell*, *115*(4), 437–447. https://doi.org/10.1016/S0092-8674(03)00845-6

Nickels, B. E. (2012). A new way to start. *Transcription*, *3*(6), 300–304. https://doi.org/10.4161/trns.21903

Nickels, B. E., Garrity, S. J., Mekler, V., Minakhin, L., Severinov, K., Ebright, R. H., and Hochschild, A. (2005). The interaction between σ70 and the β-flap of Escherichia coli RNA polymerase inhibits extension of nascent RNA during early elongation. *Proceedings of the National Academy of Sciences*, *102*(12), 4488–4493. https://doi.org/10.1073/pnas.0409850102

Nomura, M. (1999). Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae: Diversity and Common Principles. *Journal of Bacteriology*, *181*(22), 6857–6864. https://doi.org/10.1128/JB.181.22.6857-6864.1999

Oguienko, A., Petushkov, I., Pupov, D., Esyunina, D., and Kulbachinskiy, A. (2021). Universal functions of the σ finger in alternative σ factors during transcription initiation by bacterial RNA polymerase. *RNA Biology*, *18*(11), 2028–2037. https://doi.org/10.1080/15476286.2021.1889254

Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. *Nature*, *533*(7603), 353–358. https://doi.org/10.1038/nature17990

Plaskon, D. M., Henderson, K. L., Felth, L. C., Molzahn, C. M., Evensen, C., Dyke, S., … Record, M. T. (2021). Temperature effects on RNA polymerase initiation kinetics reveal

which open complex initiates and that bubble collapse is stepwise. *Proceedings of the National Academy of Sciences*, *118*(30), e2021941118. https://doi.org/10.1073/pnas.2021941118

Ploetz, E., Ambrose, B., Barth, A., Börner, R., Erichson, F., Kapanidis, A. N., … Lerner, E. (2023, February 24). *A new twist on PIFE: Photoisomerisation-related fluorescence enhancement*. arXiv. https://doi.org/10.48550/arXiv.2302.12455

Pupov, D., Kuzin, I., Bass, I., and Kulbachinskiy, A. (2014). Distinct functions of the RNA polymerase σ subunit region 3.2 in RNA priming and promoter escape. *Nucleic Acids Research*, *42*(7), 4494–4504. https://doi.org/10.1093/nar/gkt1384

Pupov, D., Petushkov, I., Esyunina, D., Murakami, K. S., and Kulbachinskiy, A. (2018). Region 3.2 of the σ factor controls the stability of rRNA promoter complexes and potentiates their repression by DksA. *Nucleic Acids Research*, *46*(21), 11477–11487. https://doi.org/10.1093/nar/gky919

Qayyum, M. Z., Molodtsov, V., Renda, A., and Murakami, K. S. (2021). Structural basis of RNA polymerase recycling by the Swi2/Snf2 family of ATPase RapA in Escherichia coli. *Journal of Biological Chemistry*, *297*(6), 101404. https://doi.org/10.1016/j.jbc.2021.101404

Rasnik, I., McKinney, S. A., and Ha, T. (2006). Nonblinking and long-lasting single-molecule fluorescence imaging. *Nature Methods*, *3*(11), 891–893. https://doi.org/10.1038/nmeth934

Renfrow, M. B., Naryshkin, N., Lewis, L. M., Chen, H. T., Ebright, R. H., and Scott, R. A. (2004). Transcription Factor B Contacts Promoter DNA Near the Transcription Start Site of the Archaeal Transcription Initiation Complex. *Journal of Biological Chemistry*, *279*(4). https://doi.org/10.1074/jbc.M311433200

Revyakin, A., Ebright, R. H., and Strick, T. R. (2004). Promoter unwinding and promoter clearance by RNA polymerase: Detection by single-molecule DNA nanomanipulation. *Proceedings of the National Academy of Sciences*, *101*(14), 4776–4780. https://doi.org/10.1073/pnas.0307241101

Revyakin, A., Liu, C., Ebright, R. H., and Strick, T. R. (2006). Abortive Initiation and Productive Initiation by RNA Polymerase Involve DNA Scrunching. *Science*, *314*(5802), 1139–1143. https://doi.org/10.1126/science.1131398

Robb, N. C., Cordes, T., Hwang, L. C., Gryte, K., Duchi, D., Craggs, T. D., … Kapanidis, A. N. (2013). The Transcription Bubble of the RNA Polymerase–Promoter Open Complex Exhibits Conformational Heterogeneity and Millisecond-Scale Dynamics: Implications for Transcription Start-Site Selection. *Journal of Molecular Biology*, *425*(5), 875–885. https://doi.org/10.1016/j.jmb.2012.12.015

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., … Gourse, R. L. (1993). A Third Recognition Element in Bacterial Promoters: DNA Binding by the α Subunit of RNA Polymerase. *Science*, *262*(5138), 1407–1413. https://doi.org/10.1126/science.8248780

Ruff, E. F., Drennan, A. C., Capp, M. W., Poulos, M. A., Artsimovitch, I., and Record, M. T. (2015). E. coli RNA Polymerase Determinants of Open Complex Lifetime and Structure. *Journal of Molecular Biology*, *427*(15), 2435–2450. https://doi.org/10.1016/j.jmb.2015.05.024

Ruff, E. F., Record, M. T., and Artsimovitch, I. (2015). Initial Events in Bacterial Transcription Initiation. *Biomolecules*, *5*(2), 1035–1062. https://doi.org/10.3390/biom5021035

Saecker, R. M., Chen, J., Chiu, C. E., Malone, B., Sotiris, J., Ebrahim, M., … Darst, S. A. (2021). Structural origins of Escherichia coli RNA polymerase open promoter complex stability. *Proceedings of the National Academy of Sciences*, *118*(40), e2112877118. https://doi.org/10.1073/pnas.2112877118

Sainsbury, S., Niesser, J., and Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature*, *493*(7432), 437–440. https://doi.org/10.1038/nature11715

Schafer, D. A., Gelles, J., Sheetz, M. P., and Landick, R. (1991). Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature*, *352*(6334), 444–448. https://doi.org/10.1038/352444a0

Shimada, T., Yamazaki, Y., Tanaka, K., and Ishihama, A. (2014). The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of Escherichia coli. *PLOS ONE*, *9*(3), e90447. https://doi.org/10.1371/journal.pone.0090447

Shin, Y., Qayyum, M. Z., Pupov, D., Esyunina, D., Kulbachinskiy, A., and Murakami, K. S. (2021). Structural basis of ribosomal RNA transcription regulation. *Nature Communications*, *12*(1), 528. https://doi.org/10.1038/s41467-020-20776-y

Skalenko, K. S., Li, L., Zhang, Y., Vvedenskaya, I. O., Winkelman, J. T., Cope, A. L., … Nickels, B. E. (2021). Promoter-sequence determinants and structural basis of primer-dependent transcription initiation in Escherichia coli. *Proceedings of the National Academy of Sciences*, *118*(27), e2106388118. https://doi.org/10.1073/pnas.2106388118

Sorokina, M., Koh, H.-R., Patel, S. S., and Ha, T. (2009). Fluorescent Lifetime Trajectories of a Single Fluorophore Reveal Reaction Intermediates During Transcription Initiation. *Journal of the American Chemical Society*, *131*(28), 9630–9631. https://doi.org/10.1021/ja902861f

Tang, G.-Q., Roy, R., Bandwar, R. P., Ha, T., and Patel, S. S. (2009). Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase. *Proceedings of the National Academy of Sciences*, *106*(52), 22175–22180. https://doi.org/10.1073/pnas.0906979106

Tang, G.-Q., Roy, R., Ha, T., and Patel, S. S. (2008). Transcription Initiation in a Single-Subunit RNA Polymerase Proceeds through DNA Scrunching and Rotation of the N-Terminal Subdomains. *Molecular Cell*, *30*(5), 567–577. https://doi.org/10.1016/j.molcel.2008.04.003

Treutlein, B., Muschielok, A., Andrecka, J., Jawhari, A., Buchen, C., Kostrewa, D., … Michaelis, J. (2012). Dynamic architecture of a minimal RNA polymerase II open promoter complex. *Molecular Cell*, *46*(2), 136–146. https://doi.org/10.1016/j.molcel.2012.02.008

van de Meent, J.-W., Bronson, J. E., Wiggins, C. H., and Gonzalez, R. L. (2014). Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophysical Journal*, *106*(6), 1327–1337. https://doi.org/10.1016/j.bpj.2013.12.055

Vo, N. V., Hsu, L. M., Kane, C. M., and Chamberlin, M. J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 2. Formation and characterization of two distinct classes of initial transcribing complexes. *Biochemistry*, *42*(13). https://doi.org/10.1021/bi0269613

Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H., and Müller, C. W. (2018). Molecular mechanism of promoter opening by RNA polymerase III. *Nature*, *553*(7688), 295–300. https://doi.org/10.1038/nature25440

Vuthoori, S., Bowers, C. W., McCracken, A., Dombroski, A. J., and Hinton, D. M. (2001). Domain 1.1 of the σ70 Subunit of Escherichia coli RNA Polymerase Modulates the Formation of Stable Polymerase/Promoter Complexes. *Journal of Molecular Biology*, *309*(3), 561–572. https://doi.org/10.1006/jmbi.2001.4690

Vvedenskaya, I. O., Sharp, J. S., Goldman, S. R., Kanabar, P. N., Livny, J., Dove, S. L., and Nickels, B. E. (2012). Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs. *Genes & Development*, *26*(13), 1498–1507. https://doi.org/10.1101/gad.192732.112

Wang, F., Redding, S., Finkelstein, I. J., Gorman, J., Reichman, D. R., and Greene, E. C. (2013). The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion. *Nature Structural & Molecular Biology*, *20*(2), 174–181. https://doi.org/10.1038/nsmb.2472

Wang, G., Hauver, J., Thomas, Z., Darst, S. A., and Pertsinidis, A. (2016). Single-Molecule Real-Time 3D Imaging of the Transcription Cycle by Modulation Interferometry. *Cell*, *167*(7), 1839–1852. https://doi.org/10.1016/j.cell.2016.11.032

Wang, M. D., Schnitzer, M. J., Yin, H., Landick, R., Gelles, J., and Block, S. M. (1998). Force and Velocity Measured for Single Molecules of RNA Polymerase. *Science*, *282*(5390), 902–907. https://doi.org/10.1126/science.282.5390.902

Wilson, C., and Dombroski, A. J. (1997). Region 1 of σ70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. *Journal of Molecular Biology*, *267*(1), 60–74. https://doi.org/10.1006/jmbi.1997.0875

Winkelman, J. T., Chandrangsu, P., Ross, W., and Gourse, R. L. (2016). Open complex scrunching before nucleotide addition accounts for the unusual transcription start site of E. coli ribosomal RNA promoters. *Proceedings of the National Academy of Sciences*, *113*(13), 1787–1795. https://doi.org/10.1073/pnas.1522159113

Winkelman, J. T., Nickels, B. E., and Ebright, R. H. (2021). The Transition from Transcription Initiation to Transcription Elongation: Start-site Selection, Initial Transcription, and Promoter Escape. In *Chemical Biology*. *RNA Polymerases as Molecular Motors: On the Road*. RSC Publishing. Retrieved from https://books.rsc.org/books/edited-volume/783/chapter/518247/The-Transition-from-Transcription-Initiation-to

Winkelman, J. T., Pukhrambam, C., Vvedenskaya, I. O., Zhang, Y., Taylor, D. M., Shah, P., … Nickels, B. E. (2020). XACT-Seq Comprehensively Defines the Promoter-Position and Promoter-Sequence Determinants for Initial-Transcription Pausing. *Molecular Cell*, *79*(5), 797–811. https://doi.org/10.1016/j.molcel.2020.07.006

Winkelman, J. T., Vvedenskaya, I. O., Zhang, Y., Zhang, Y., Bird, J. G., Taylor, D. M., … Nickels, B. E. (2016). Multiplexed protein-DNA cross-linking: Scrunching in transcription start site selection. *Science*, *351*(6277), 1090–1093. https://doi.org/10.1126/science.aad6881

Winkelman, J. T., Winkelman, B. T., Boyce, J., Maloney, M. F., Chen, A. Y., Ross, W., and Gourse, R. L. (2015). Crosslink Mapping at Amino Acid-Base Resolution Reveals the Path of Scrunched DNA in Initial Transcribing Complexes. *Molecular Cell*, *59*(5), 768–780. https://doi.org/10.1016/j.molcel.2015.06.037

Wu, C. W., Yarbrough, L. R., Hillel, Z., and Wu, F. Y. (1975). Sigma cycle during in vitro transcription: Demonstration by nanosecond fluorescence depolarization spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(8), 3019–3023.

Yin, H., Wang, M. D., Svoboda, K., Landick, R., Block, S. M., and Gelles, J. (1995). Transcription Against an Applied Force. *Science*, *270*(5242), 1653–1657. https://doi.org/10.1126/science.270.5242.1653

Zengel, J. M., and Lindahl, L. (1994). Diverse Mechanisms for Regulating Ribosomal Protein Synthesis in Escherichia coli. In *Progress in Nucleic Acid Research and Molecular Biology* (Vol. 47, pp. 331–370). Academic Press. https://doi.org/10.1016/S0079-6603(08)60256-1

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M. X., Arnold, E., and Ebright, R. H. (2012). Structural Basis of Transcription Initiation. *Science*, *338*(6110), 1076–1080. https://doi.org/10.1126/science.1227786

Zhang, Z., Revyakin, A., Grimm, J. B., Lavis, L. D., and Tjian, R. (2014). Single-molecule tracking of the transcription cycle by sub-second RNA detection. *ELife*, *3*, e01775. https://doi.org/10.7554/eLife.01775

Zuo, Y., and Steitz, T. A. (2015). Crystal Structures of the E. coli Transcription Initiation Complexes with a Complete Bubble. *Molecular Cell*, *58*(3), 534–540. https://doi.org/10.1016/j.molcel.2015.03.010