Fluorescence-Detected Intermediates In Open Complex Formation by *E.coli* RNA Polymerase: Analysis of Large-scale Conformational Changes and Effects of Lipiarmycin

By

Munish Chhabra

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biophysics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final Oral Examination: 8/16/2019

This dissertation is approved by the following members of the Final Oral Committee:

M.Thomas Record Jr., Professor, Chemistry and Biochemistry
Richard Gourse, Professor, Bacteriology
James C.Weisshaar, Professor, Chemistry
James L.Keck, Professor, Biochemistry
Katherine Henzler-Wildman, Professor, Biochemistry

Abstract

In transcription the key first steps of gene expression, the enzyme RNA polymerase (RNAP) catalyzes the synthesis of RNA (ribonucleic acid) complementary to one strand of a DNA (deoxyribonucleic acid) template. All steps of transcription, and particularly those of initiation at the start site region of promoter DNA, are highly regulated by promoter sequences, transcription factors and other environmental conditions. In initiation, RNA polymerase binds specifically to double-stranded promoter DNA and operates on it as a biophysical machine to open 13-14 bp including the transcription start site, place the template strand of the DNA in its active site, and initiate the synthesis of RNA transcripts when provided with nucleotide triphosphates (NTP). There is plethora of structural information available about RNAP and the open promoter complex (OC), but the mechanism of OC formation is not yet understood. In my thesis research, I have developed and applied several novel real-time biophysical fluorescence assays that use DNA-DNA FRET (energy transfer) between cyanine (Cy3 and Cy5) dyes located at far upstream (-100 bp from the start site) and downstream (+14 bp from the start site) positions, and RNAP-DNA PIFE (enhanced fluorescence of these individual dyes) to determine the kinetics and mechanism of OC formation by bacterial *E.coli* RNAP and λP_R promoter DNA. First, using equilibrium FRET, another lab member and I showed that the upstream region of promoter DNA bends and wraps around RNAP in the OC and in an ensemble of closed complexes (CC) stabilized by low temperature (2°C). Then I used fast kinetic (stopped-flow) studies on this system to discover the mechanism by which bending and wrapping of far-upstream promoter DNA facilitates bending the downstream promoter region into the active site cleft of RNAP (Figure) prior to DNA opening. This research can be translated to understand the action of different broad-spectrum antibiotics targeting the different intermediates formed in the process. In the third part of the thesis, I describe kinetic-mechanistic experiments with Lipiarmycin (LpM), an active ingredient of Fidaxomicin antibiotic (FDA-approved for recurrent *Clostridium difficile* associated diarrhea infection in adults), using FRET and PIFE kinetic assays. I find that LpM acts on a relatively late CC intermediate in which the clamp is open (I_{1M}) forming an off-pathway intermediate (I_{1M}-LpM) in the mechanism of OC formation. My findings regarding the mechanism of OC formation and LpM inhibition forms a basis to understand the sites of action and effects of transcriptional activators and repressors that modulate the kinetics of transcription initiation to achieve gene regulation.

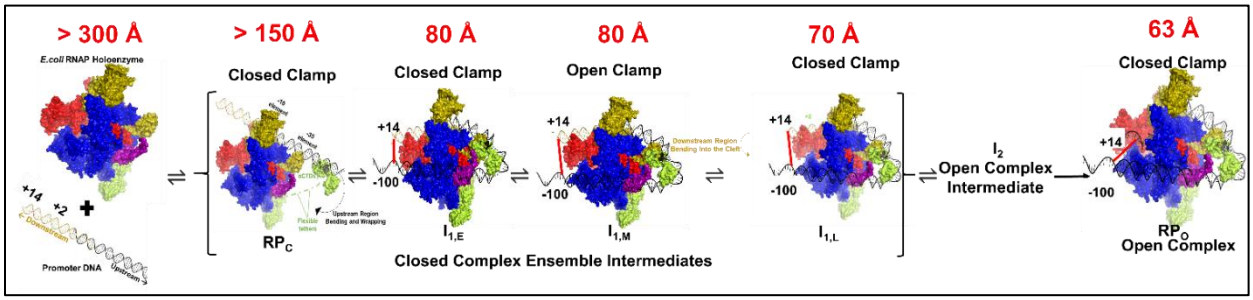


Figure: Structural Mechanism of Spontaneous Open Complex Formation from FRET and PIFE Characteristics of Closed Intermediates in Transcription Initiation. *Subunits of RNAP are β (red), β' (blue), σ^{70} (wheat), ω (purple), flexibly-tethered αCTDs (light green), and αNTDs (light green). FRET distances between promoter positions +14 (in downstream (golden) DNA) and -100 (upstream (black)) in each of the four closed intermediates deduced from the kinetic analysis. These closed intermediates, which make up the closed complex ensemble ({CC}), are the initial closed complex (RP_C) and three more advanced closed intermediates (I_{1E}, I_{1M} and I_{1L}). RP_O is the stable open complex.*

Table of Contents

Appendix-III

List of Tables

Chapter 2

Chapter 3

Appendix I

Appendix II

Appendix III

List of Figures

Chapter 3

Chapter 4

APPENDIX III

Acknowledgements

The thesis research is a collective effort of all the lab members, friends, family, reviewers and others (enemies too, if any!) that I have met in my Ph.D. training program for the past 5 years from 2014. I believe all of them have directly and indirectly helped me to think about my research project. The ones that had a direct impact whom I would like to acknowledge are:

First and foremost, Professor Tom Record who provided me an opportunity to work in his lab and introduced me to the interesting Biophysical project of Transcription Initiation which ultimately became my thesis research project. I really enjoyed working with him and learnt a lot from him on how to think critically and analytically about experimental data and research (the skill is still under construction for me!). He is one of the most influential and inspiring figures that I have met in my lifetime.

All the previous and current lab members who have helped me in collecting data, analyzing it, designing experiments, performing them, having scientific discussions and providing very useful insights. Current lab members include Scientist Irina Shkel, Dr. Kate Henderson, Hao-Che Wang, Dylan Plakson, Maxwell Rector, Emily Zytkiewicz. Previous lab members include Dr. Raashi Sreenivasan.

The team of undergraduates and summer intern students who helped me in data collection. Christina McNerney, Malavika Abhineshababu Sridevi, Clare Cimperman, Hertina Kan, Katelyn Callies, Jeremiah Dane Williams, Priya Chitturi, Feng Jiang, Andrew Xue.

Friends and family who encouraged me in both ups and downs of Ph.D. journey! They also helped me to think about my project with a broader perspective providing a holistic view.

Thesis Committee Members to annually review my research progress and provide other different scientific perspectives. NIH Funding Source.

Finally, UW Madison Biophysics Committee Members who provided me admission in the Biophysics Ph.D. program and found potential in my application to pursue Ph.D. research.

Chapter 1

Introduction

1. Introduction to The Cellular Process of Transcription

The cell is the most basic unit of a living organism. The cell stores genetic information in the form of DNA (deoxyribonucleic acid) polymer and creates copies of it in the form of RNA (ribonucleic acid) transcripts varying in amounts and length which is later utilized to form polypeptide proteins, the building blocks of the cell. The process of flow of genetic information from DNA to RNA and then to protein is the central dogma occurring inside every living biological cell. Synthesis of RNA from DNA information is called transcription and synthesis of polypeptides and protein from RNA information is called translation.

The enzyme RNA polymerase serves as the central protagonist that carries out transcription. Transcription can be divided into three stages: intiation, elongation and termination. A schematic representation of the three stages is shown in Figure 1 [1].

1.1. **Initiation Phase** includes recognition of promoter DNA by RNAP holoenzyme to form an initial closed complex (RP_C). After RPc formation, double stranded DNA is opened and separated into two-template and non-template strand in the form of transcription bubble isomerizing to form an open complex (RPo) (Figure 1). OC enters into a cycle of synthesizing short abortive RNA transcripts (3-9 mer) [2] and releasing them into the solution [3] gaining some free energy by pulling the DNA towards itself through the mechanism of scrunching to escape from the promoter before entering into the elongating phase [4 5]. Some fraction of RNAP remains in the abortive cycle and some are able to escape to form full-length RNA transcripts depending upon the strength of interaction developed between RNAP and different elements of promoter DNA [6 7]. Abortive initiation phase is analogous to revving up the engine of RNAP car/ compressing a spring gaining some potential energy before it is ready to drive on the DNA road/expanding a spring coil converting the potential energy to the kinetic energy.

1.2. **Elongation Phase** includes the transcription elongation complex (TEC) where RNAP enzyme translocates on DNA at a rate of 10-100 nucleotides/seconds in both prokaryotes and eukaryotes at 37^0C [8]. In a bacterial elongating complex (EC) approximately 12-14 bases of the DNA are in the form of transcription bubble, 15 bases ahead of the promoter present in the downstream region is protected by RNAP containing 9-10 baes-pairs (bp) of RNA-DNA hybrid [9]. EC undergoes repetitive addition of nucleotides in a cycle to the growing RNA transcript formed [10 11].

1.3. **Termination Phase** occurs when the EC enters into the termination pathway faster than the nucleotide addition cycle. In bacteria, it can occur via two types of mechanisms; intrinsic termination mediated by RNA hairpin formation or extrinsic termination in which Rho protein interacts with the RNA formed [12].

Different transcription factors serve as accessories for RNAP in the three different stages of transcription cycle to regulate gene expression at both local and global level inside the cell [13]. Mis-regulation of the gene expression programs due to errors in transcription factor binding can cause a broad ranges of diseases like cancer, diabetes, autoimmunity and inflammation, neurological and cardiovascular diseases in an eukaryotic cell [14]. Thus, it becomes highly essential to study the transcription cellular process. Transcription Initiation being the first crucial step of gene expression and regulation, it becomes important to understand the mechanism of transcription initiation first to later understand its regulatory aspects used by different transcription factors, activators, repressors and inhibitors to modulate the cellular process. The focus of this thesis research is understanding the initial interaction between promoter DNA fragment and bacterial *E.Coli* RNAP taking place in the mechanism of transcription initiation to form transcription ready open complex using biophysical fluorescence techniques.

2. Structure of *E.coli* RNAP Holoenzyme

Bacterial *E.coli* core RNAP is a multi-subunit enzyme comprising of 2 α subunits (each 40 kDa), 2 subunits of β (150 kDa) and β' (155 kDa) and 1 subunit of ω (~10 kDa). In assembly, an α_2 dimer forms first, to which β and then β' are added. Core *E.coli* RNAP interacts with a sigma specificity factor to form holoenzyme ($\alpha_2\beta\beta'\omega\sigma$). The overall dimensions of E.coli RNAP holoenzyme molecule are ~90 x 95 x 160 Å [15 16]. Sigma (σ) factors are used to recruit the core RNAP assembly to a promoter region of DNA by sequence specific interaction between the promoter elements and the different domains of σ protein. *E.coli* has 7 different types of σ factors often identified by their molecular weights ((e.g. σ^{70} is about 70 kDa). These recognize different types of promoters and direct transcription in response to different environmental conditions. Most of the housekeeping genes essentially required in growing *E.coli* bacterial cells are expressed by σ^{70} RNAP. Other sigma factors like σ^{54}, σ^{S} (~38 kDa), σ^{32}, σ^{28}, σ^{E} (~24 kDa) and σ^{FecI} (~19 kDa) are used in nitrogen assimilation, stationary phase, heat shock response, flagella synthesis or chemotaxis, periplasmic or extracellular protein and iron metabolism or transport respectively [17, 18]. Sigma factors have evolved to recognize different group of promoters in different bacterial species. Some bacterial species like *Streptomyces coelicolor* have more than 60 sigma factors [18]. The thesis research is on *E.coli* RNAP holoenzyme assembly with housekeeping σ^{70} factor ($\alpha_2\beta\beta'\omega\sigma^{70}$).

The structure of core *E.coli* RNAP ($\alpha_2\beta\beta'\omega$) resembles a right-handed crab claw where β and β' clamps act as a pincers that can open and close to accommodate duplex as well as ss DNA. The active site for phosphodiester bond synthesis where the template strand is bound is located deep in (at the base of) the clamp structure, and is identified by the catalytic Mg^{2+} ion. The 2 α subunits act as a hinge as well as an assembly platform for the β and β' clamps [19]. The functions of ω are less clear. The overall structure of the active site cleft is highly conserved across all the 3 kingdoms [20]. The amino acid sequences present in the outer surface of the enzyme are divergent

with sequence insertions (SI) present in the different subunits of RNAP in different organisms [21]. *E.coli* RNAP has lineage specific 3 major large sequence insertions in β and β' subunits: Sequence insertion 1 (SI1-β226 to β350 amino acid position in the β lobe region), Sequence Insertion 2 (SI2- β938 to β1040 amino acid position in the β flap region) and Sequence insertion 3 (SI3-β'943 to β1040 amino acid position present between the conserved β' cleft and β' jaw region) [22 23]. (Figure 2).

Sigma 70 (σ^{70}) by itself cannot recognize promoter DNA. The DNA binding domains of σ^{70} are masked by its region 1 and are exposed on interacting with core RNAP (Figure 2) [18]. σ^{70} comprises of 4 folded domains required for interaction with RNAP and promoter DNA. Domain region 1, present only in group 1 type σ^{70} factors, is bound in the active site cleft of free RNAP inhibiting DNA binding in the cleft, domain region 1.2 interacts with discriminator region present between the -10 promoter element and the transcription start site (TSS) (Figure 2) [24]. Domain region 2, a primary function of which is to interact with the -10 region of the promoter in both duplex and ss confromations, is composed of 5 conserved α helices and several other helices from the non-conserved region (NCR) [17]. $\sigma^{2.2}$ has polar residues that interfaces with RNAP, $\sigma^{2.3}$ and $\sigma^{2.4}$ forms the hydrophobic pocket of aromatic amino acid residues in which the nucleotide base of non-template strand in -10 element is buried required for DNA melting [25]. Domain 3 is a 3-helix domain interacting with the extended -10 element and domain 4 contains 4 helices interacting with the -35 consensus promoter element (Figure 2) [17]. Sigma region 3.2 is a flexible linker hairpin loop connecting domain 3 and 4 of σ^{70} protruding from towards the active center of RNAP and occupying a part of the RNA exit channel present in the protein (Figure 2) [26].

The schematic structural representation of different functional units of RNAP holoenzyme required to interact with the promoter DNA to form OC is shown in Figure 3. The roles of each of these units is discussed below in the different transcription initiation mechanisms of open complex formation.

3. Promoter Architecture

RNAP holoenzyme first binds to the region of DNA present upstream of the gene to be transcribed called as the promoter region in the transcription initiation phase. The start site where the first RNA nucleotide is incorporated and then later elongated to make RNA chain is denoted by +1 and called as the transcription start site (TSS). Sequences present ahead/downstream of TSS are denoted by '+' sign and the ones present before/upstream of the TSS are denoted by '-' sign.

E.coli RNAP σ^{70} holoenzyme system is recruited to the promoter region by sequence specific interactions with different domains of σ^{70} and the promoter elements present in the non-template DNA strand as shown in Figure 4. These interactions ensure efficient promoter binding and it's isomerization to form OC is mediated by different conformational changes taking place in the protein and DNA in the mechanism of transcription initiation. Here is a brief description of each promoter elements.

3.1. -10 element (Consensus Sequence- 5'TATAAT3'): -10 element contains AT rich sequences. Two important positions are -11 and -7 where the nucleotide base is flipped into the hydrophobic pocket of aromatic amino acid residues created by $\sigma^{2.3}$ and $\sigma^{2.4}$ [25]. Next gen sequencing analysis of 4096 variants of promoter showed an enhanced preference for base 'A' at -11 and 'T' at -7 positions in the -10 element region which could melt with the fastest kinetics by *E.coli* RNAP. Promoter sequences lacking these two bases could still melt fast if compensated by other base patterns observed from the consensus sequence 'TATAAT' present in this region [27]. Analogous to -10 element, TATA box is present in eukaryotic system utilized in forming pre-initiation complex and promoter melting. TATA box is present 25-30 and 40-100 base pairs upstream from the transcription start site in metazoans like *D. melanogaster* and yeast *S.cerevisiae*, most of them with maximal transcription activity were present on average 75bp upstream of TSS in case of yeast and human core promoters [28].

3.2. Extended -10 element (Consensus Sequence- 5'TGTG3'): Some promoters have extended -10 element 5'-'TG'-3' motif present at immediate upstream of -10 element interacting with 2 perpendicular helices spanning the σ2 ($\sigma^{2.5}$) and σ3 domains of σ^{70} causing significant bending of promoter bending towards the σ^3 domain [29 30].

3.3. Spacer Region (Consensus length- 17 bp) : -10 element is separated from the -35 element by a spacer region containing no consensus sequence but a consensus separation length of 17 base pairs on analyzing 263 E.coli promoter sequences with known TSS [31]. On changing the spacer length from 17 to 16 bp σ^4 domain rotates by 5^O which shifts the helix-turn-helix (HTH) motif of this domain to interact with -35 element, the allowance of variation in the spacer length is proposed to be dependent upon the flexibility of σ^4 domain [29]. Changes in spacer length also affect the kinetics of OC formation and transcription levels in-vitro [32 33]. The phosphate backbone present between -16/-17 DNA position is recently seen to exhibit long range electrostatic interactions with Arg-157 residue of σ^{NCR} in the Cryo-EM structure of E.coli σ^{70} RNAP holoenzyme [34]. The spacer region sequence imparts different curvature or conformations to DNA sequence aiding in downstream DNA bending and unwinding in the OC formation mechanism [35].

3.4. -35 element region (Consensus sequence- 5'TTGACA3'): It is a hexamer spanning from -30 to -36 upstream of TSS interacting with HTH motif of $\sigma^{4.2}$ which inserts in to the major groove and bends the DNA by 36^0 [36]. Flexible flap domain of β subunit interacts with the $\sigma^{4.2}$ with hydrophobic patch to position it for recognition of -35 element [37].

3.5. Discriminator region: It is present between -10 element and the TSS where the second nucleotide base downstream of the -10 element interacts with region $\sigma^{1.2}$ with the methionine amino acid at 102 position [24 38]. There is no consensus sequence or length of the discriminator region. However, its length and sequence determines the stability of the open complex lifetime

and the position of promoter escape when OC transitions from abortive initiation to the elongation phase [6 38 39].

3.6. UP element (Proximal subsite consensus sequence- 5'AAAAAARNR3', Distal subsite consensus sequence- 5'AWWWWWTTTTT3'): The UP element, present in some σ^{70} promoters, is a roughly 20 bp, AT rich region, often with A and or T tracts, located immediately upstream of -35 hexamer. The UP element interacts specifically with the 2 C-terminal domains of α subunits (αCTDs). In the absence of a near-consensus UP element, the αCTDs interact nonspecifically with the upstream region of promoter DNA. The αCTDs are linked to the corresponding αNTDs (which are the hinge of the clamp/cleft)) via flexible tethers. They play a very important role in stimulation of transcription, increasing expression of the ribosomal promoter rrnB P1 *in-vivo* and *in-vitro* by ~30-fold [40 41]. Most of this effect resulted from the proximal part of the UP element (-46 to -38) and required only one αCTD subunit. If the distal site was occupied, then the UP sequence starting at -47, called the distal subsite, had an additional substantial, if smaller (>10 fold), effect on in vivo initiation. [42]. Both αCTDs can be used interchangeably to bind proximal and distal subsites [42].

The above effects for rrnBP1 promoter are relative to nonspecific upstream DNA, and are in vivo. Upstream truncation of λP_R and lacUV5 promoters at positions between -82 (20 bp upstream of the UP element) and -42 (in the UP element) reduced the kinetics of conversion of closed to open complex by one-to-two orders of magnitude in vitro [43 44]. Based on OH footprinting kinetics, upstream bending and wrapping around RNAP is proposed to play a role in spontaneous DNA opening and formation of stable E.coli RNAP σ_{70} holoenzyme-λP_R promoter DNA open complex formation in the mechanism of transcription initiation [45] as described below in section 4 and discussed in Chapter 3.

CRE ELEMENT Another proposed promoter element, found in only some promoters, is the core recognition element (CRE) region from -4 upstream to +2 downstream in the non-template strand overlapping with the discriminator region and the transcription start site TSS. 5 of the 6 nucleotides in the CRE region interacts with the β subunit of core RNAP enzyme where nucleotide base 'G' at +2 position is flipped out of the base stacks and inserted into a deep pocket formed by the RNAP β subunit called as β pocket [46]. These interactions are a functional determinant of TSS selection**,** modulating the position of the downstream end of the transcription bubble in OC thereby modulating the TSS [47]. Gate loop region of β subunit is proposed to guide the non-template strand of the downstream DNA towards the β CRE pocket upon DNA bending into the channel in OC formation [48].

Not all promoters sequences in prokaryotic cells have all of the promoter elements described above, different promoters have different sequences close to or far from the consensus one (Figure 3) employed by the cell to regulate their gene expression with the help of other transcription factors/activators.

On observing the 3-dimenisonal parameters of ~16,500 promoter DNA across 12 prokaryotes, like the energetics of base stacking, solvation and H-bonding a common 3-D signature structure quite distinct from conventional B-DNA and coding sequences was found. The 3-D signature state mediated by the different promoter element sequences might be the hidden code recognized by RNA polymerase enzyme (RNAP) to perform transcription initiation[49].

4. Different Transcription Initiation Mechanisms of Open Complex Formation

Recent advances in Cryo-EM technology have made it possible to identify and classify the different conformers of a protein-DNA complex that are present in a frozen sample and determine their structures at high (almost crystallographic) resolution. Using this approach, in the last two years a plethora of structural information has been obtained about the different RNAP-DNA

complex intermediates formed in the mechanism of open complex formation in transcription initiation amongst eukaryotes as well as prokaryotes. Here, I have briefly described some of the mechanisms proposed for spontaneous DNA opening without ATP hydrolysis performed by the three different types of eukaryotic RNAP transcription assemblies to compare with our proposed mechanism for *E.coli* RNAP, described later.

In the first steps of transcription initiation in eukaryotes, a pre-initiation complex (PIC) is formed by a set of general transcription factors (GTFs) that bind specifically to sites in the upstream region of the promoter, forming a complex that recruits RNA polymerase. Different GTFs are involved in recruiting the three different eukaryotic RNA polymerases (pol I, II and III).

4.1. Transcription Initiation Open Complex Formation Mechanism by RNA Pol II – RNA Polymerase II (pol II) enzyme synthesizes most of mRNA. The core promoter region recognized by the pol-II transcription assembly consists of TFIIB recognition element (BRE) from ~-37 to -32 upstream (analogous extended -10 element), TATA box ~-31 to -26 upstream (analogous -10 element), Initiator region from -2 to+4 (analogous TSS) and downstream promoter element from +28 to +32 region. General transcription factors assemble on the core promoter region in a sequential order to form PIC (Figure 4B, Core PIC Assembly). TFIID is first recruited to the core promoter region followed by TFIIA and TFIIB, then RNA Pol-II-TFIIF complex and later TFIIE and TFIIH. TFIID is a multi-subunit protein comprising of TATA box binding protein (TBP) along with other 13 TBP associated transcription factors (TAFs) [50]. Binding of TBP, TFIIA and TFIIB to the TATA element introduces a ~90^0 bend in the promoter region (Figure 4B, Core PIC Assembly). Incorporation of Pol-II- TFIIF complex stabilizes the downstream DNA along the cleft of Pol-II by opening the clamps of RNAP, slightly bending the downstream DNA to Rbp1 region and rotating the upstream module towards Pol-II formed by TBP, TFIIA and TFIIB interacting with the upstream DNA. TBP–TFIIA–TFIIB–TFIIF–DNA sub-complex traps the DNA on the surface of Pol II [51 52 53]. TFIIE binds with the complex to form CC1 closed complex intermediate (Figure 5B). The core

domain of TFIIH interacts with the downstream region of the DNA ~60 bp downstream of TATA box and induces a kink imposing torsional stress on the promoter DNA to bend slightly inward over the Pol-II cleft to form CC2 closed complex intermediate (Figure 5B). DNA is subsequently melted by the translocase activity of the TFIIH subunit to form OC [52 51]. The DNA is opened ~20-30 bp downstream of the TATA box in the initially melted region (IMR) in the OC (Figure 5B).

4.2. Transcription Initiation Open Complex Formation Mechanism by RNA Pol I – RNA polymerase I (Pol I) in eukaryotes synthesizes ribosomal RNA (rRNA). The upstream region of ribosomal RNA promoters doesn't contain a TATA box (or analogous -10 element) and instead contains the core element located between -45 and +8 relative to the TSS. In yeast, formation of the PIC at a Pol-I promoter requires both the Rrn3 transcription factor and the heterotrimeric Core Factor (CF; composed of Rrn6, Rrn7 and Rrn11) to recruit Pol I to the rRNA promoter region. Initially, CF binds to the DNA region from approximately ~-40 bp to -16 bp from the TSS, immediately upstream of the region where Pol I will bind, and Rrn3 binds to Pol I to form an initiation competent complex (Figure 5A Pol-I-Rrn3 complex). The interaction of Rrn7 and Rrn11 of CF with the DNA backbone induces a ~30^{o} bend in the upstream promoter region. CF assembly on the upstream DNA then recruits the Pol I- Rrn3 complex and loads the DNA so it is threading through the wall, protrusion gate and the active site region of RNAP to form a closed complex (CC) (Figure 5A, CC). Upstream DNA bending, engagement of Rrn7 zinc ribbon domain and progressive closure of the active site cleft cause spontaneous opening of the rDNA promoter causing the transition from a CC to an OC (Figure 5A, CC to OC) [54 55 56]. The promoter sequences recognized by Pol-I transcription initiation machinery have distinct bendability around the upstream region and meltability around the TSS, this distinct property is proposed to play a role in spontaneous DNA opening [54].

4.3. Transcription Initiation Open Complex Formation Mechanism by RNA Pol III – RNA polymerase III enzyme along with transcription factor IIIB (TFIIIB) is used to transcribe the DNA

sequences specifying small structured RNAs like 5S rRNA, tRNA and other small RNAs. PIC is formed on binding of transcription factor TFIIIB to the TATA box present from -15 to -30 bases upstream to the TSS (Figure 5C, Unbound Pol III/Upstream promoter complex). TFIIIB comprises of 3 important subunits: TATA box binding protein (TBP), TFIIB related factor (Brf1/2) and B-double prime unit (Bdp1). TFIIIB on binding to TATA box recruits RNA Pol-III enzyme (Figure 5C, Recruitment). RNA Pol-III enzyme contains 17 subunits that include the conserved 10 core subunits, stalk and 2 additional sub-complexes namely the C53-C37 heterodimer and C82-C34-C31 heterotrimer. In the mechanism of promoter DNA opening, first double-stranded closed DNA is bent away by the C82 cleft loop and the clamp that blocks access of DNA into the cleft forming closed complex/closed clamp intermediate (Figure 5C). This is followed by clamp and cleft opening to slide the DNA between clamp and the lobe, subsequently C34 subunit interacting with Bdp1 tether becomes ordered enclosing double stranded DNA to form closed complex/open clamp model (Figure 5C). Then the clamp closes back again leading to downward movement of C82 cleft loop establishing stabilizing interactions with Bdp1 tether and C34 and the DNA snugged inside is melted from -11 bases to +8 base positions to form open complex/closed clamp. The non-template strand interacts with Pol-III subunits C34 and C82, the template strand interacts with Brf1 linker (Figure 5C) [57]. TFIIIB subunit interacting with the upstream promoter region fully encircles the upstream promoter DNA and restructures RNA polymerase –III enzyme by rearranging C34 subunits through Bdp1 interaction thereby promoting OC formation [58] [59] (Figure 5C, first 2 steps Unbound Pol-III upstream promoter complex and Recruitment).

4.4. Universal Mechanism Proposed for Spontaneous Promoter Opening in Eukaryotic Transcription Initiation System.

Certain promoters recognized by Pol-II (also true for pol-I and pol-III) with unstable IMRs don't require TFIIH and can undergo ATP independent spontaneous promoter opening, these promoters have upstream under-wound IMR region which along with distorted DNA into the cleft

obviates the need of ATPase activity of TFIIH in case of Pol-II for promoter opening [60] (Figure 5B –TFIIH). Recently, another distinct intermediate CC^{dist} was obtained of RNA Pol-II PIC with double stranded distorted promoter DNA in a closed Pol-II cleft with closed clamps in the presence or absence of TFIIH (Figure 4B, CC^{dist}). The DNA is distorted 7 Å deeper into the cleft at ~ +20 bp downstream of TATA box corresponding to the upstream edge position of the IMR in CC^{dist} intermediate. These findings suggest a universal model for DNA opening in transcription initiation where in the PIC, promoter DNA is first positioned above the pol-II cleft with closed clamps (CC1) followed by opening of the clamps and opening of the cleft in CC2 allowing DNA swinging into the cleft. Clamps then close back again which is coupled with DNA distortion and underwinding in the cleft in the presence or absence of TFIIH (CC^{dist}) to form OC [60]. The schematic representation of the mechanism involving these intermediates is shown in Figure 5B.

4.5. Summary of Previous Information Regarding Specific Recognition and Subsequent Large Conformational Changes in the Mechanism of Open Complex Formation by Bacterial RNAP- The most kinetic-mechanistic and structural research until recently has been done on *E.coli* RNAP (mostly σ^{70}, some also on σ^{54} and heat shock σ^{32}) and coli-phage T7 RNAP bacterial system.

4.5.1. Formation of the Initial Closed Complex Intermediate (RP_C):

In prokaryotic E.coli bacterial system, transcription initiation occurs with the assembly of initiation factor σ^{70} - core RNAP (RNAP holoenzyme) on the promoter DNA to form closed complex and later a stable open complex (OC). Initial interaction between the RNAP holoenzyme and promoter DNA forms an initial closed complex intermediate called RP_C, characterized by low temperature HO· and DNase-I footprinting experiments. These footprint results show protection and DNase-I sensitivity from -5 to -49 upstream position in the template strand and -5 to -43 upstream position in the non-template strand for lacUV5 RP_C at 0^0C with respect to TSS [61]. T7A1 and T7A3 RP_C

promoter show weak protection till -55/-56 upstream promoter DNA position at 4^0C [61 62] indicating the interaction of the proximal UP element promoter region (-35 to -56 upstream) with the αCTDs. The downstream promoter region in the footprints of lacUV5 RP_C is protected and enhanced till -5 upstream position to the TSS. This indicates the initiation of interactions of different promoter elements (-35, -10 promoter element and the spacer region) with different region of σ^{70} in RP_C formation and no interaction of the downstream duplex region from -5 to +20 with the RNAP cleft [63]. Another coat protein promoter PVIII of E.coli phage fd initially showed DNase-I enhancement/suppression at upstream region from -50 to -60 in both non-template and template strand and near -35 element on the non-template strand when forming closed complex with *E.coli* RNAP at 7^0C. On increasing the temperature from 7^0C to 17^0C and 37^0C, the DNase-I enhancement/protection begin to increase on the downstream promoter region bordering till +1 for the closed and +20 for the open complex formed respectively at these 2 temperatures [64]. These results indicate that in RP_C formation, αCTDs initially interacts with the proximal upstream promoter region followed by the other promoter elements like -10 element, -35 element and the spacer region interacting with the different σ^{70} subunit regions.

4.5.2. Formation of the Ensemble of Advanced Closed Complex Intermediates (I_1)

RP_C advances to form an advanced initial closed complex intermediate (I_1) in the mechanism of OC formation, characterized by real-time foot-printing kinetics of different promoters like λP_R, T7A1 and lacUV5 showing footprint protection from -80 upstream to +20 downstream region of promoter DNA interacting with RNAP [63]. DNase downstream footprint boundaries obtained at different promoters provide evidence for at least three ensemble of closed complexes: the initial specific complex (RP_c; boundary at -5) and two more-advanced CC (designated I_{1E} and I_{1L} for “early” and “late”) with downstream footprint boundaries at +2/+7 and at +20 [63]. In the most advanced closed complex intermediate I_{1L} the downstream DNA is bent by 90^0 into the active site cleft [65] by flipping of -11 A base nucleotide of -10 promoter element into the hydrophobic pocket

of aromatic amino acids present in the sigma region 2 to initiate bubble opening inside the cleft [25] [63]. Base 'T' at position -12 of the non-template strand interacts with the conserved tryptophan-dyad residues (W433/434) of σ^{70} to maintain the double strand/single strand (-12/-11) junction at the upstream edge of the transcription bubble [66].

Upstream truncation of λP_R and lacUV5 promoters at positions between -82 (20 bp upstream of the UP element) and -42 (in the UP element) reduced the kinetics of conversion of closed (I_{1L}) to open complex by one-to-two orders of magnitude [43] [44]. For the UT-47 λP_R truncation variant, DNase footprinting revealed downstream boundaries of protection of the template and non-template strands of the downstream duplex at +2 and +7, as compared to +20 for I_1 at FL λP_R [45]. This indicates that the closed complex population distribution is much less advanced for UT -47 than for FL λP_R behaving like I_{1E} intermediate.

There are many conformational changes taking place on the upstream and downstream promoter region on advancing from RP_C to I_{1L} intermediate in the mechanism of OC formation detected by fluorescence FRET and PIFE kinetic techniques (described below in section 5) as shown in Chapter 3 like bending and wrapping of upstream promoter region around RNAP to bend the downstream region inside the active site cleft. In this chapter, we propose that bending/wrapping of upstream promoter DNA open up the clamps of RNAP in I_{1E} to I_{1M} transition and clears the in-cleft elements like $\sigma^{1.1}$ and downstream mobile elements (DMEs) providing space for the downstream DNA to be bent inside the cleft. The DNA is then bent into the cleft in the I_{1L} intermediate followed by clamp closure of RNAP β-β' subunits. $\sigma^{1.1}$ and DMEs in free RNAP holoenzyme are positioned inside the active site cleft to prevent entry of non-promoter DNA and undergo network of interactions with the upstream and downstream promoter region serving as a gate to only allow the downstream promoter region to be bent and opened inside the active site cleft [63]. DMEs include the β lobe (β 184–224 and 344–438), β sequence insertion (SI)1 (β 225–

343), β' clamp (β' 131–347 and 1260–1368), β' upstream clamp (β' 808–912) and β' SI3 (β' 943–1130) [39].

4.5.3. Formation of the Open Complex Intermediates (I_2) and its Rapid Irreversible Transition to form stable OC (RP_O).

The DNA inside the active site cleft in I_{1L} is opened advancing to form two unstable open complex intermediates I_2 and I_3 detected by OC salt and urea upshift filter-binding and foot-printing dissociation kinetics assay [67 68 69 70 63]. Conversion of unstable open complex intermediate I_2 to a stable OC (RP_O) involves folding of β' jaw subunit and assembly of DMEs on the distal downstream promoter DNA region [39 71] (Figure 3).

4.5.4. Proposed Role of Clamp Dynamics in the Mechanism of OC Formation.

The dynamic nature of clamps of RNAP have also shown to play an important role in the mechanism of OC formation [72]. The clamps of free RNAP holoenzyme transition from partly closed/closed to open state in the order of 1.02-1.52 s^{-1} [73]. In the mechanism of OC formation by *E.coli* RNAP clamps of RNAP begin from closed state on binding with the promoter DNA and open up to accommodate the downstream DNA inside the active site cleft and close back again to form stable OC [72] (Figure 5E). Recent Cryo-EM structure of a *M.tuberculosis* RNAP-open complex intermediate formed using rrn*AP3* promoter DNA and corallopyronin A antibiotic (Cor, related to Myxopyronin in structure and function) helped to suggest a mechanism for OC formation with dynamic clamps[74]. As per the proposed mechanism, the clamps transition from closed (in RP1) to open state, close back (in RP2) and open in the late stages of promoter melting followed by one final closure to form stable RP_O [74] (Figure 5D). In chapter 4, using lipiarmycin (LpM) antibiotic and fluorescence FRET and PIFE kinetic studies we characterize an off-pathway intermediate (I_{1M}-LpM) formed when LpM acts on I_{1M} closed complex intermediate with open

clamps of RNAP and the downstream region of promoter DNA is linear and unbent, displaced from the active site cleft.

4.5.5. Our Proposed Mechanism of stable *E.coli* RNAP OC Formation using FRET and PIFE fluorescence kinetic studies.

On drawing analogy from the different OC formation mechanisms proposed in prokaryotic and eukaryotic transcription system for spontaneous promoter opening we propose the following OC formation mechanism for *E.coli* RNAP with λP_R promoter DNA (Figure 5F):

The UP element of promoter region initially interacts with the αCTDs and is bent (~30^0-$45^{0\ 75}$) to form RP_C. This is analogous to transcription factors binding to the TATA box present ~30-40 bp upstream of the TSS first and bend it by 90^0 in RNA Pol-II, 30^0 in RNA Pol-I and ~180^0 by complete encirclement in RNA Pol-III transcription system to form the PIC in eukaryotic system. In RP_C formation, different domain of σ^{70} also interacts with the -35 promoter, spacer region and -10 promoter elements. RP_C advances to I_{1E} intermediate where downstream DNA interacts with β lobe SI1 region of the clamp. I1E is similar to CC1 intermediate formed in RNA pol-II eukaryotic transcription system. The clamps of RNAP open converting from I_{1E} to I_{1M} where in-cleft elements like $\sigma^{1.1}$ and DMEs are cleared from the cleft. Similarly, in RNA pol-II eukaryotic transcription initiation mechanism CC1 (closed clamp) transitions to CC2 (open clamp). After active site cleft-clearance, the DNA is bent inside the active site cleft by 90^0 and -10 promoter element interacts $\sigma^{2.3}$ in I_{1L} intermediate to initiate DNA opening, analogous to CC2 transition to CC^{dist} intermediate where the downstream DNA is swung 7 Å deeper into the cleft in eukaryotic RNA pol-II transcription initiation. The DNA once inside the active site cleft is then opened followed by clamp closure advancing I_{1L} to I_2 intermediate. I_2 intermediate is then rapidly stabilized by folding of β' jaw and assembly of DMEs on the downstream region of promoter DNA to irreversibly form a stable open complex (RP_O).

5. Biophysical Fluorescence Techniques (FRET and PIFE) used to study transcription initiation mechanism.

In my thesis research, I have majorly employed two biophysical fluorescence spectroscopy steady-state (Chapter 2) and kinetic (Chapter 3 and 4) techniques, FRET and PIFE, to characterize the closed complex intermediates and open complex formed in bacterial *E.coli* transcription initiation mechanism.

Fluorescence is the emission of a longer wavelength radiation of lower energy by a substance after absorbing an electromagnetic radiation of higher energy. An electron in the ground state (S_0) upon absorption of high-energy photon is excited to the singlet state (S_1) that undergoes non-radiative transition to the first energy level of the excited state and finally emits fluorescence when traveling back to the ground state energy levels as shown in the energy level Jablonski diagram (Figure 6). The difference between the excitation and emission wavelength is called the Stoke's shift and the time the electron takes to get back to the ground state is the fluorescence lifetime. Fluorescence follows first order kinetic with an exponential decay rate following equation 1.

$$[S_1] = [S_1]_0 e^{-\Gamma t} \qquad \text{-Equation 1}$$

[S1] is the concentration of molecules at the excited state at time t, $[S1]_0$ is the initial concentration of excited state molecules at time t=0 and Γ is the total decay rate of radiative and non-radiative transitions.

The fraction of the photons emitted to absorbed by a substance is the fluorescence quantum yield Φ that helps to determine the fluorescence efficiency process (Equation 2).

$$\Phi = \frac{\text{Number of photons emitted}}{(Number\ of\ \text{photons}\ absorbed)} \qquad \text{-Equation 2}$$

Fluorescence phenomenon has been widely used in biological sciences to track and detect nanoseconds changes happening in biomolecules. These changes are detected by fluorescence

spectroscopy by either monitoring the differences in the intensity or the fluorescence lifetime of the molecule in response to different environmental conditions. Ensemble experiments monitors effects over a population average while single molecule experiments provide the advantage of tracking fluorescence changes occurring on a molecule under observation avoiding the average of the signal detected from ensemble experiments. The thesis research have explored ensemble fluorescence effects monitoring conformational changes in promoter DNA in the initial stages of transcription upon its interaction with *E.coli* RNA polymerase.

The fluorophores used in the study are cyanine dyes Cy3 and Cy5. The chemical structure of the dyes is shown in Figure 7A. Cy3 fluoresces greenish yellow at ~570 nm emission wavelength when excited at ~550 nm and Cy5 fluoresces in the red region at 670 emission wavelength when excited at ~650 nm. The excitation and the emission spectrum of both the dyes is shown in Figure 7B. The dyes are internally attached at the 5' phosphate end of the DNA backbone and have been observed to give more stable fluorescence by this mode of attachment [76]. Cy3 and Cy5 dyes when attached at the 5' end of the duplex with a linker stacks onto the terminal bases and predicted to improve the thermodynamically stabilize the DNA duplexes [77 78] . Cyanine dyes have been widely used in biological research due to its higher quantum yield, higher photo stability and ease of availability in different chemical amino/thiol reactive forms [79 80].

Different biophysical fluorescence techniques have been developed and used to study the transcription process. One of the most widely used biophysical fluorescence techniques in bacterial transcription research is FRET.

5.1. **FRET (Fluorescence Resonance Energy Transfer or Förster Resonance Energy Transfer)** - FRET is the non-radiative transfer of energy between the two fluorophores when brought in close proximity to each other through dipole-dipole coupling. In this process, one fluorophore molecule serves as the donor that gets excited upon absorbing the photons and

transfers its energy to the acceptor fluorophore indirectly exciting it via the FRET communication. Photo-chemically, FRET occurs when there is an overlap in energy level differences between the ground and the excited state of the acceptor and donor fluorophores. As the electrons in the donor fluorophore transition from the ground state (S_0^d) to its excited (S_1^d) state upon photon absorption they relax back from the excited state of the acceptor fluorophore (S_1^a) to its ground state (S_0^a) and exhibit fluorescence emission due to the overlap, as shown in Jablonski FRET diagram (Figure 8).

FRET can only occur if there is sufficient overlap between the acceptor excitation and donor emission spectra. For efficient FRET communication, the donor should have higher quantum yield to have more fractions of photons/energy to emit than absorb and transfer that energy to the acceptor and the acceptor on the receiving end of the donor's energy should have higher absorption coefficient. Cy3-Cy5 dye pairs qualifies to be one of the best FRET donor-acceptor combination for the FRET studies [80 81].

FRET is observed when the fluorescence emission of the acceptor increases and correspondingly the fluorescence emission of the donor decreases. It is quantitatively measured through FRET efficiency (E) which is dependent upon the distance between the acceptor and the donor fluorophore (R), quantum yield of the donor (Φ_D), dipole-dipole orientation factor (κ^2), refractive index of the medium (η) and the overlap integral (J (v)), as shown in Equation 3-4.

$$E = \frac{R^6}{R_0^6} + 1 \qquad \text{-Equation 3}$$

R_0 is the Förster distance when FRET efficiency is 50% (E= 0.5) which is dependent upon other factors as shown in Equation 4.

$$R_0 = (9.78 \times 10^3)\left(\Phi_D\, \kappa^2 \eta^{-4} J\,(v)\right)^{\frac{1}{6}} \quad Å \qquad \text{-Equation 4}$$

Each of the parameters for Cy3-Cy5 dye FRET pairs are calculated and discussed in detail in Chapter 2.

FRET efficiency is experimentally determined by measuring the energy transfer from [82 83]:

1. Donor emission-

FRET Efficiency E is measured in terms of quantum yield of the donor in the presence (Φ_{DA}) and absence of the acceptor (Φ_D). (Equation 5)

$$E = 1 - \frac{\Phi_{DA}}{\Phi_D} \quad \text{-Equation 5}$$

However, measuring the quantum yield of the dye is impractical, so it can be determined in terms of fluorescence donor intensity in the presence (I_{DA}) and absence of the acceptor (I_D). (Equation 6).

$$E = 1 - \frac{I_{DA}}{I_D} \quad \text{-Equation 6}$$

The disadvantage of the method is the inconsistency in the concentration of the donor which shouldn't be varying with different samples. The concentration difference between donor-acceptor sample and donor only sample can be accounted using the donor absorbances at the excitation in the presence (A_{DA}) and absence of the acceptor (A_D) as shown in Equation 7.

$$E = 1 - \frac{A_{DA}}{A_D}\frac{I_{DA}}{I_D} \quad \text{-Equation 7}$$

Obtaining accurate value of A_{DA} can be a difficult task due to overlap in donor and acceptor absorption spectra.

2. Acceptor emission-

To measure FRET efficiency from the acceptor emission one requires either of the two references/controls for obtaining standard acceptor intensity; one excitation wavelength which is

absorbed only by the acceptor but not by the donor or second use a sample containing only the acceptor without donor present in it.

In the first case when the standard acceptor intensity is obtained by exciting at a wavelength which excites only the acceptor but not the donor, the FRET efficiency E is calculated as shown in Equation 8.

$$E = \frac{I_{AD}A_{AA} - I_{AA}A_{AD}}{I_{AA}A_{DD}} \qquad \text{-Equation 8}$$

I_{AA} and A_{AA} are the acceptor intensity and absorbance upon acceptor excitation. I_{AD} is the acceptor intensity upon donor excitation. A_{AD} and A_{DD} are the acceptor an donor absorbances on exciting the donor.

If the donor and acceptor are in 1:1 ratio then E is calculated using their extinction coefficients at their excitation wavelength. (Equation 9)

$$E = \frac{I_{AD}\varepsilon_{AA} - I_{AA}\varepsilon_{AD}}{I_{AA}\varepsilon_{DD}} \qquad \text{-Equation 9}$$

In the second case if an additional sample without donor containing only the acceptor is used, then Equation 8 and 9 can be applied in a similar manner separately quantifying intensity and absorbance of the acceptor form the sample containing only the acceptor.

In Chapter 2 we have used the acceptor sensitized method to calculate FRET efficiency. We first obtained the ratio of the fluorescence emission of the acceptor in the presence and absence of the donor after normalizing the spectra for donor-acceptor sample to the donor-only sample in the wavelength range in which only the donor emits ($Ratio_A$) and multiplied with extinction coefficients to get FRET efficiency E using Equation 10 [84 85 86].

$$E = \mathrm{Ratio}_A * \frac{\varepsilon_A^{\lambda ex1} - \varepsilon_A^{\lambda ex2}}{\varepsilon_D^{\lambda ex2}} \qquad \text{-Equation 10}$$

$\varepsilon_A^{\lambda ex1}$ is the extinction coefficient of the acceptor upon direct excitation of the acceptor, $\varepsilon_A^{\lambda ex2}$ is the extinction coefficient of the acceptor upon excitation at donor excitation wavelength and $\varepsilon_D^{\lambda ex2}$ is the extinction coefficient of the donor at donor excitation wavelength.

3. Donor Lifetime-

FRET efficiency (E) is determined by comparing the fluorescence lifetime decay rate of the donor in the presence (τ_{DA}) and absence (τ_D) of the acceptor using Equation 11.

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad \text{-Equation 11}$$

If the life-time follows multi-exponential decay one can use the amplitude weighted life times of the donor fluorescence in the presence and absence of the acceptor.

4. Using Donor/Acceptor Intensity ratio-

The relative FRET efficiency, also known as the proximity ratio is given by,

$$E_{rel} = \frac{I_A}{I_D + I_A} \quad \text{-Equation 12}$$

I_A and I_D is the total acceptor and donor fluorescence intensities when exciting at donor excitation wavelength. The mixed spectrum obtained in the presence of both donor and acceptor must be decomposed into the isolated donor and acceptor component spectra to accurately determine I_A and I_D . One also needs to account for background concentration, leakage of donor emission into the acceptor intensity and stoichiometry gamma factor correction when determining relative FRET efficiency [87] .

FRET assay has been widely used in transcription initiation research studies to understand the promoter DNA bending, unbending, DNA bubble opening and closing, movement of RNAP on DNA, conformational changes taking place in the RNAP clamp opening/closing and DNA scrunching happening in the initiation and elongation complexes [88 7 73 89 2 90]. In chapter 3 we use

FRET signal to detect upstream and downstream promoter DNA region conformational changes taking place in the closed complex and open complex formation by E.coli RNAP and Cy3-Cy5 labeled λP_R promoter DNA fragment.

5.2. PIFE (Protein Induced Fluorescence Enhancement) - PIFE fluorescence phenomenon is an increase in the intensity of a fluorophore attached to DNA molecule when brought in close proximity to a protein. The single dye fluorescence technique has been extensively used to interpret the signal as the change in the distance between the dye and the protein as the fluorescently labeled DNA interacts with protein to form a complex or dissociate from it [91]. However, recent human replication protein A and single stranded DNA binding protein (SSB) interaction studies with Cy3 labeled DNA suggest that PIFE signal is not a through space distance-dependent phenomenon and it's magnitude is influenced by the region of protein directly interacting with the Cy3 probe [92].

Fluorophores that can isomerize between cis and trans state and are sensitive to the change in environmental conditions like Cy3 dyes are used as PIFE probes. The other dyes that can be used to get PIFE signal are Cy5, DY547 and Alexa dyes [91]. Cy5 dyes have reduced PIFE effects compared to Cy3 as it has more possibilities for bond rotations compared to Cy3 [93].

The molecular basis of PIFE is explained by the Jablonski diagram in Figure 8. PIFE probes in the ground state (S_0) can isomerize between cis (D_{cis}) and trans (D_{trans}) state. When excited to S_1 state they can form three energy sinks of excited trans (D^*_{trans}), excited cis (D^*_{cis}) and an isomerization 90^0 intermediate state (D^*_{90}). The de-excitation rates $k_{D,C}$ and $k_{D,90}$ are non-radiative and happens when electron transitions from the excited cis (D^*_{cis}) and excited isomerization intermediate (D^*_{90}) state to the ground cis (D_{cis}) state. When the electrons in the excited trans state return to the ground trans state at a rate of $k_{D,T}$ fluorescence is emitted. The longer Cy3 survives in the trans excited state isomer after excitation, the higher fluorescence is emitted upon

relaxation back to the ground state. PIFE results upon the reduction in the interconversion rate from the excited trans state (D^*_{trans}) to the excited 90^0 isomerization intermediate state (D^*_{90}) which leads to build up of more electrons in the excited trans state (D_{trans}) and increase in fluorescence when encountered with steric restriction or change in microenvironment (Figure 8) [94 95].

PIFE assay have been used in bacterial transcription initiation research studies to see change in Cy3 fluorescence tagged at +2 position of the promoter DNA when opened to form transcription bubble on interacting with RNAP. The technique has been used to determine the mechanism of transcription initiation and promoter escape in mycobacterial transcription system. In the thesis we have used +14 (downstream) Cy3/Cy5 and -100 (upstream) Cy3/Cy5 PIFE kinetics signal to understand the promoter DNA conformational changes taking place when interacting with *E.coli* RNAP to form OC in the absence (Chapter 3) and presence of Lipiarmycin drug (Chapter 4).

5.3. Instrumentation used to detect FRET, PIFE signal

In Chapter 2, we used a Quantamaster C-60/2000 spectrofluorimeter (Photon Technology Instruments) to obtain steady state FRET and PIFE fluorescence intensity and anisotropy measurements. In Chapter 3 and 4 we used SF300X Kintek Corp stopped flow fluorimeter to measure FRET and PIFE kinetics signal. The details of setting up the instrument for perfoming experiments are explained in the Method section of the corresponding Chapter.

Figures

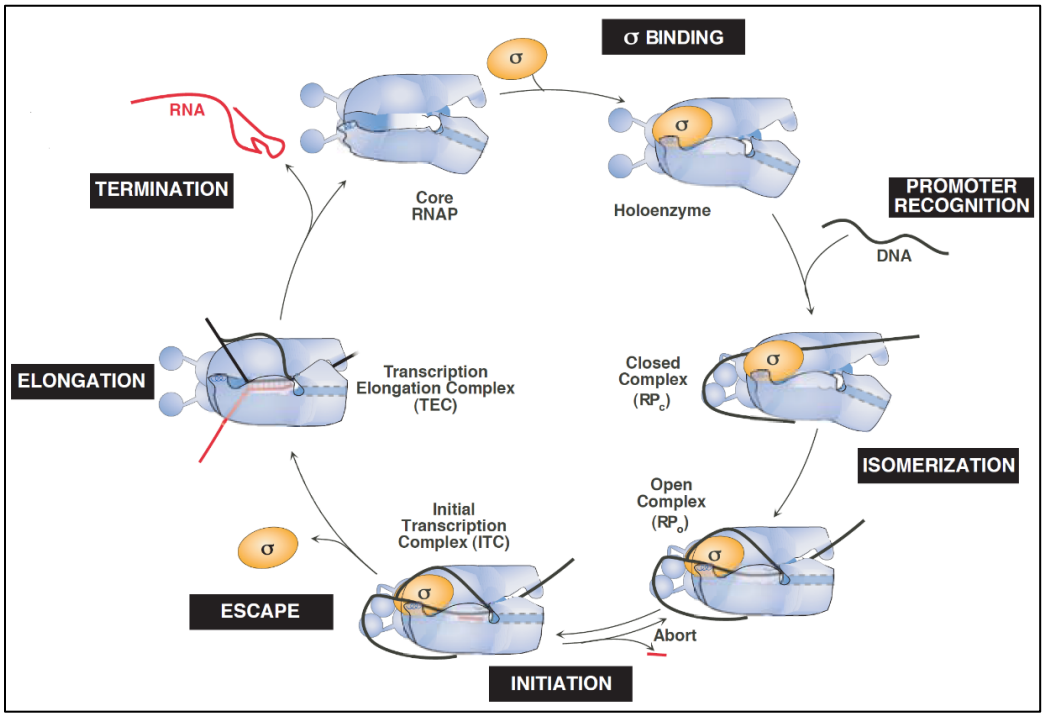


Figure 1: Schematic representation of 3 broad stages of Prokaryotic Transcription process[1].

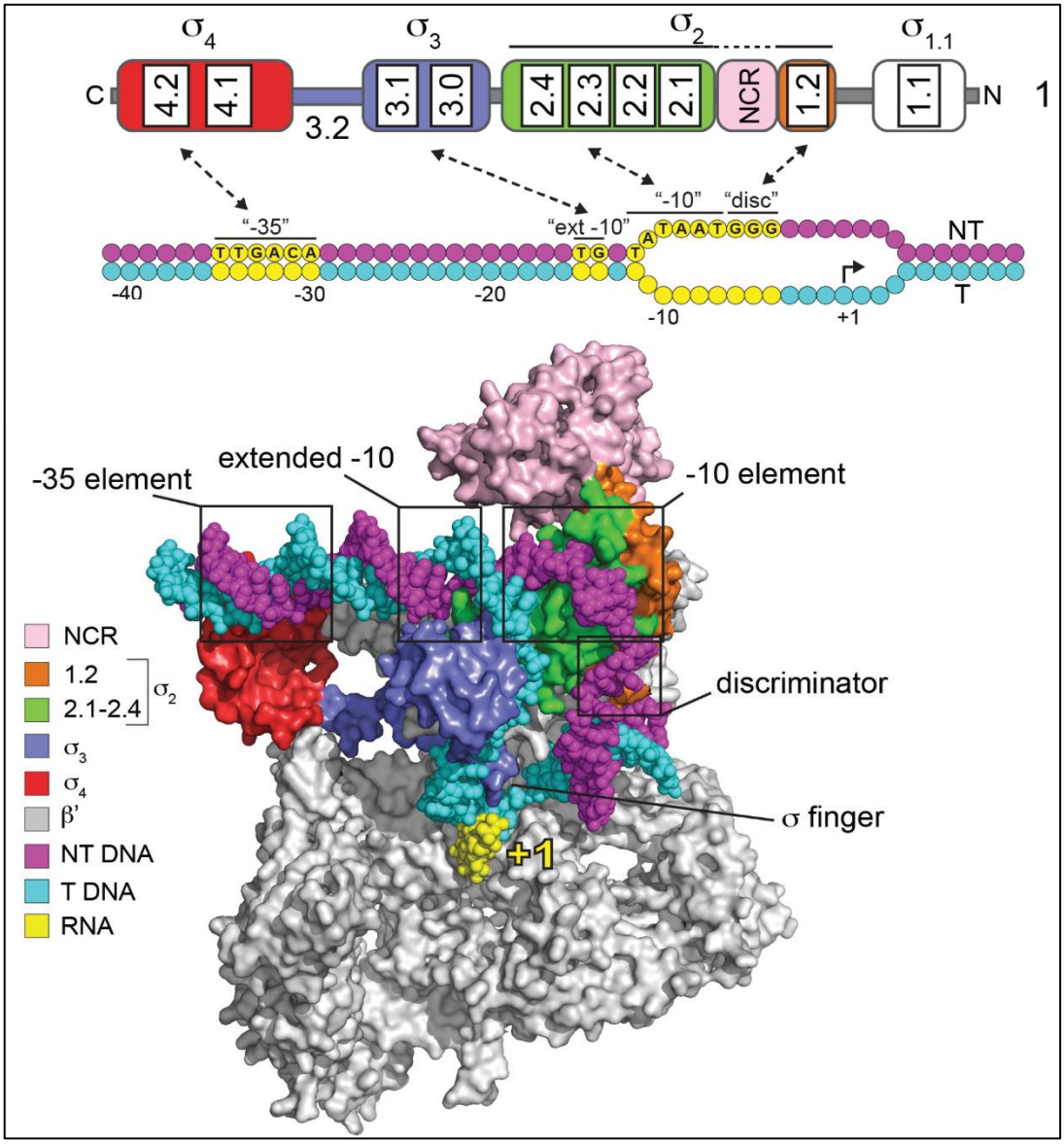


Figure 2: Schematic representation of linear and structural organization of different domains of σ^{70} required for promoter recognition and interaction with RNAP holoenzyme [96].

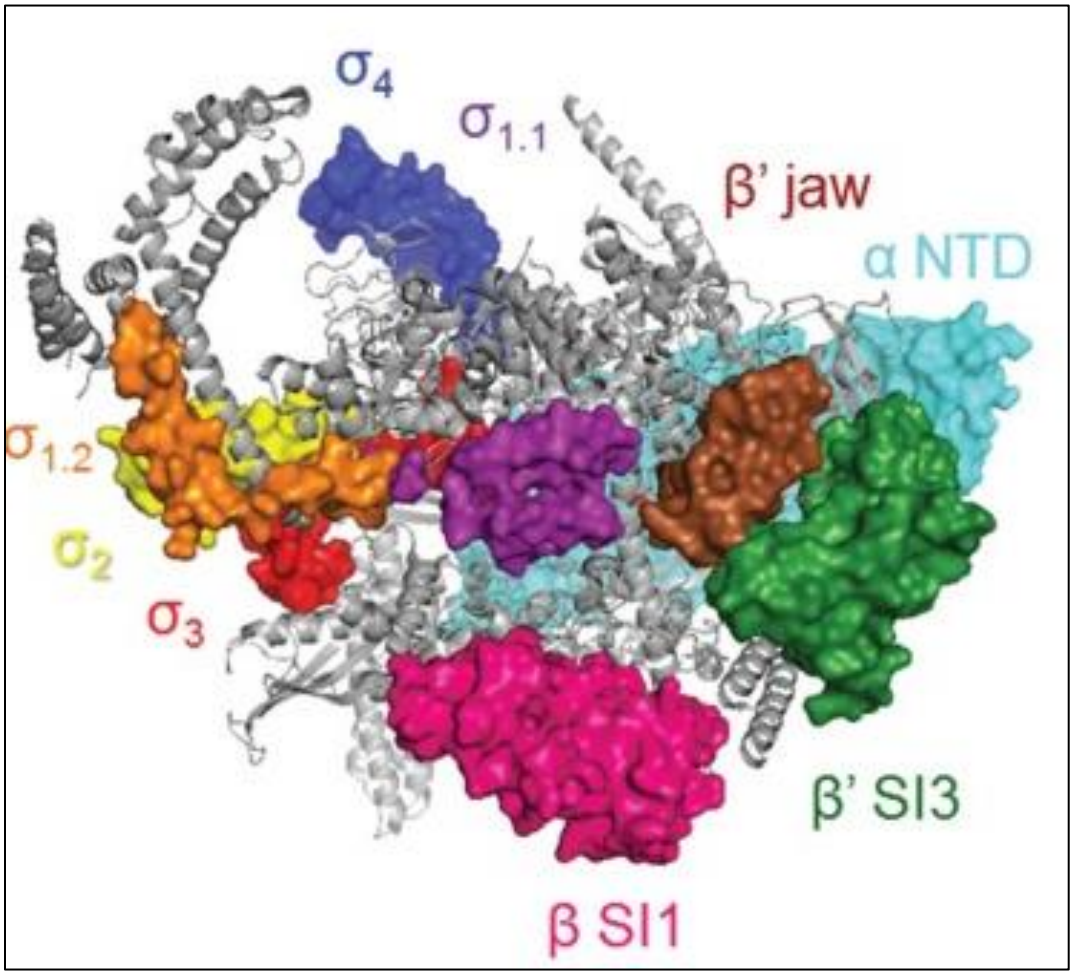


Figure 3: Structural representations of functional regions of E.coli σ^{70} holoenzyme [63].

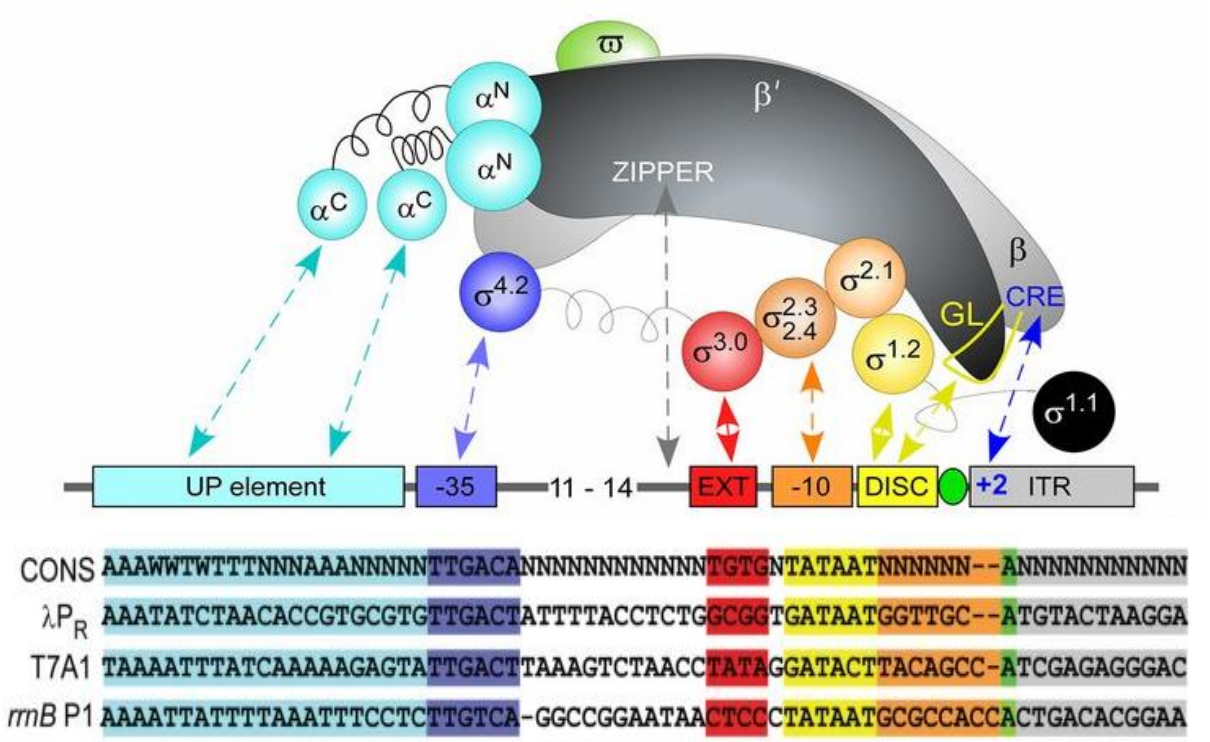


Figure 4: Schematic representation of RNAP-holoenzyme interacting with different promoter elements present on the non-template strand [48] along with consensus and different promoter sequences [63].

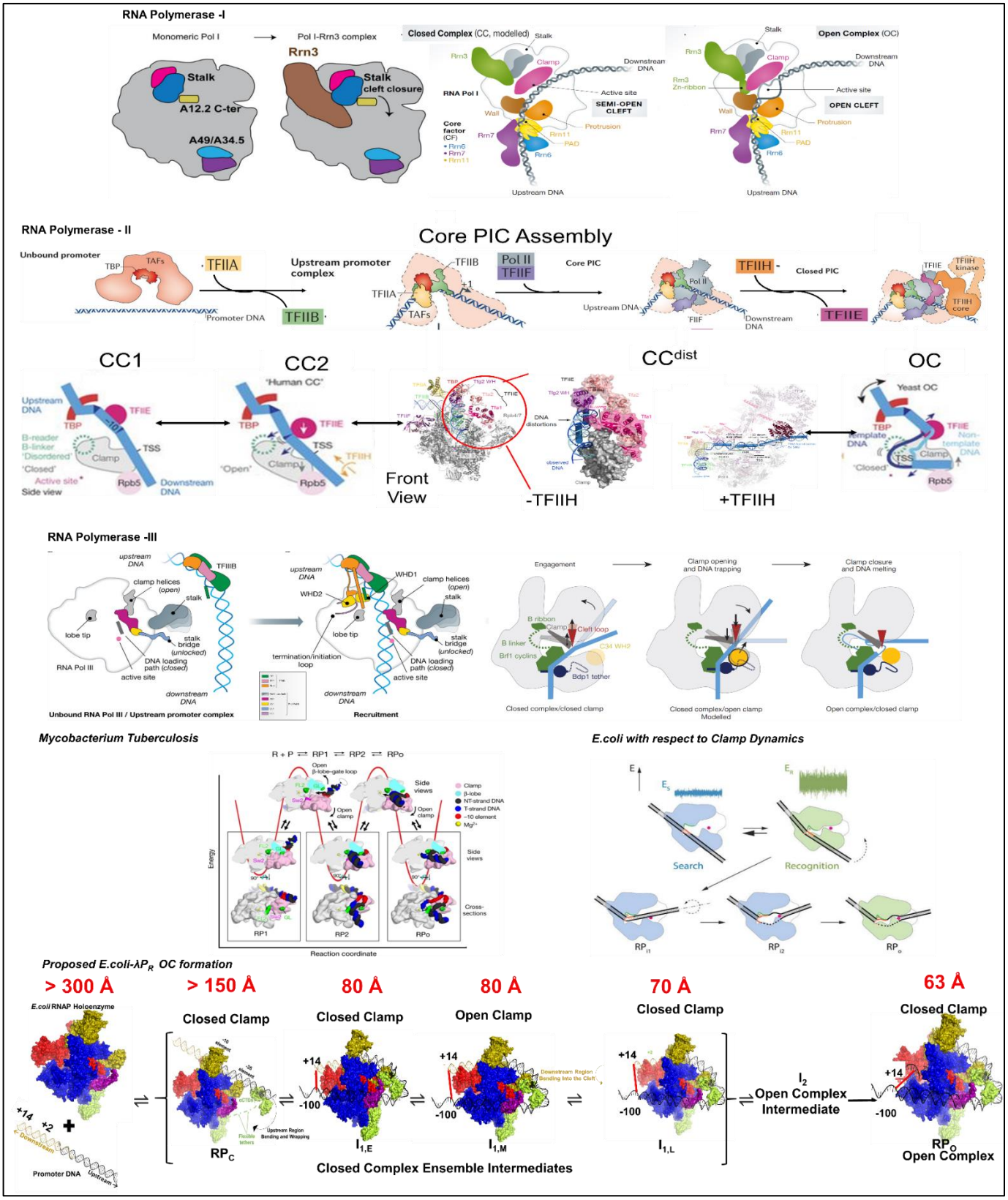

RNA Polymerase -I
Monomeric Pol I
Pol I-Rrn3 complex
Rrn3
Stalk
A12.2 C-ter
A49/A34.5
Stalk cleft closure
Closed Complex (CC, modelled)
Open Complex (OC)
Clamp
Downstream DNA
Active site
SEMI-OPEN CLEFT
OPEN CLEFT
Wall
Protrusion
Core factor (CF)
Rrn6
Rrn7
Rrn11
PAD
Upstream DNA
RNA Polymerase - II
Core PIC Assembly
Unbound promoter
TBP
TAFs
TFIIA
TFIIB
Promoter DNA
Upstream promoter complex
Pol II
TFIIF
Core PIC
Pol II
FIIF
TFIIH
Closed PIC
TFIIE
TFIIH kinase
TFIIH core
CC1
CC2
CCdist
OC
Upstream DNA
B-reader
B-linker
'Disordered'
'Closed'
Active site
Side view
TSS
'Human CC'
'Open'
Front View
-TFIIH
+TFIIH
Yeast OC
Template DNA
Non-template DNA
RNA Polymerase -III
upstream DNA
clamp helices (open)
stalk
lobe tip
stalk bridge (unlocked)
DNA loading path (closed)
active site
downstream DNA
RNA Pol III
TFIIIB
WHD1
WHD2
termination/initiation loop
Unbound RNA Pol III / Upstream promoter complex
Recruitment
Engagement
Clamp opening and DNA trapping
Clamp closure and DNA melting
B ribbon
B linker
Brf1 cyclins
Cleft loop
C34 WH2
Bdp1 tether
Closed complex/closed clamp
Closed complex/open clamp Modelled
Open complex/closed clamp
Mycobacterium Tuberculosis
E.coli with respect to Clamp Dynamics
R + P ⇄ RP1 ⇄ RP2 ⇄ RPo
Open β-lobe-gate loop
Open clamp
Side views
Cross-sections
Clamp
β-lobe
NT-strand DNA
T-strand DNA
-10 element
Energy
RP1
RP2
RPo
Reaction coordinate
Search
Recognition
Proposed E.coli-ΔP_R OC formation
> 300 Å
> 150 Å
80 Å
80 Å
70 Å
63 Å
E.coli RNAP Holoenzyme
Closed Clamp
Open Clamp
Promoter DNA
Upstream Region Bending and Wrapping
Flexible tethers
Downstream Region Bending Into the Cleft
RPC
I1,E
I1,M
I1,L
I2
Open Complex Intermediate
RPo
Open Complex
Closed Complex Ensemble Intermediates

Figure 5: Compilation of different mechanisms proposed for open complex formation in basal transcription initiation by RNA Pol-I (panel A) [55] [56], RNA Pol-II (panel B) [53] [51] [60], RNA Pol-III (panel C) [57] [58] eukaryotic transcription assembly and by prokaryotic *M.tb* [74] (panel D) and *E.coli* (panel E-F) [72] (Chapter 3-4) RNAP system for spontaneous DNA opening.

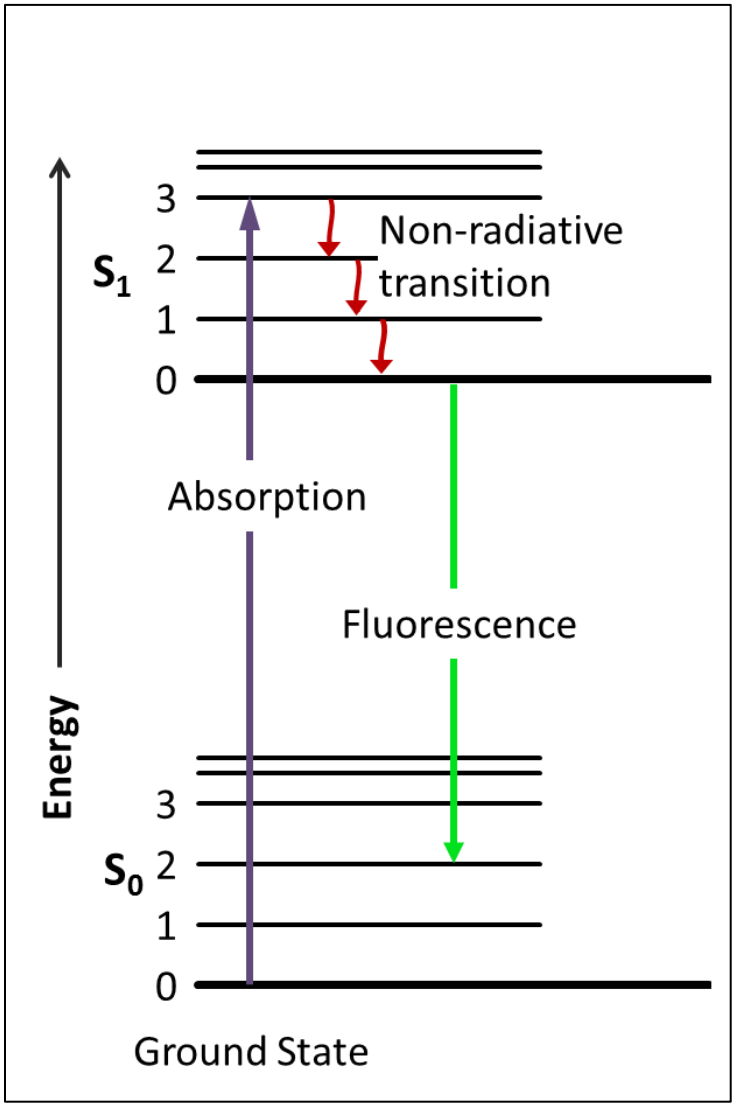


Figure 6 Jablonski Diagram to explain Fluorescence

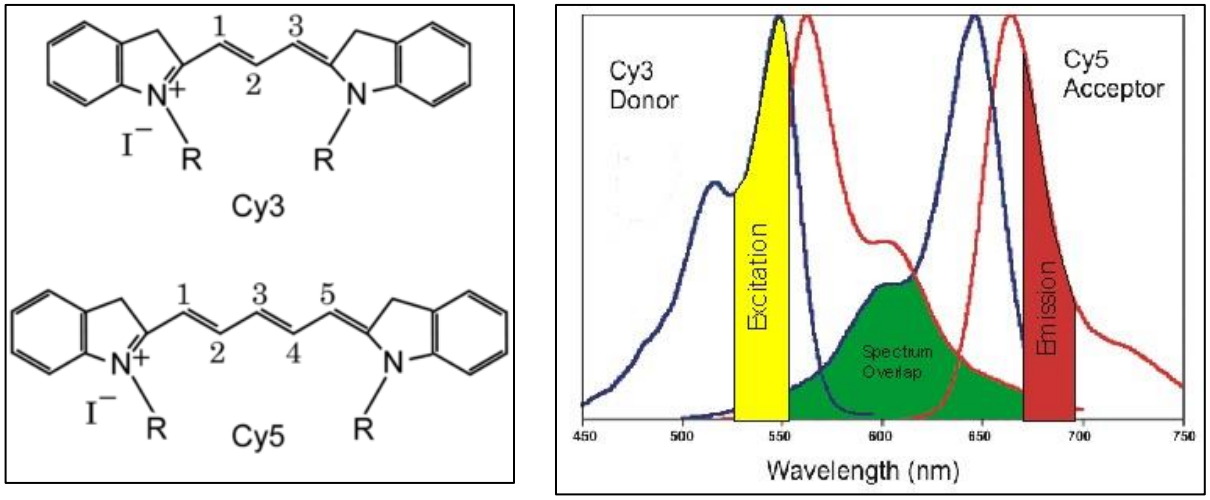


Figure 7 A. Chemical Structure of Cy3 and Cy5 cyanine fluorophores. 7 B. Absorption and Emission spectra of Cy3 and Cy5 dyes [97].

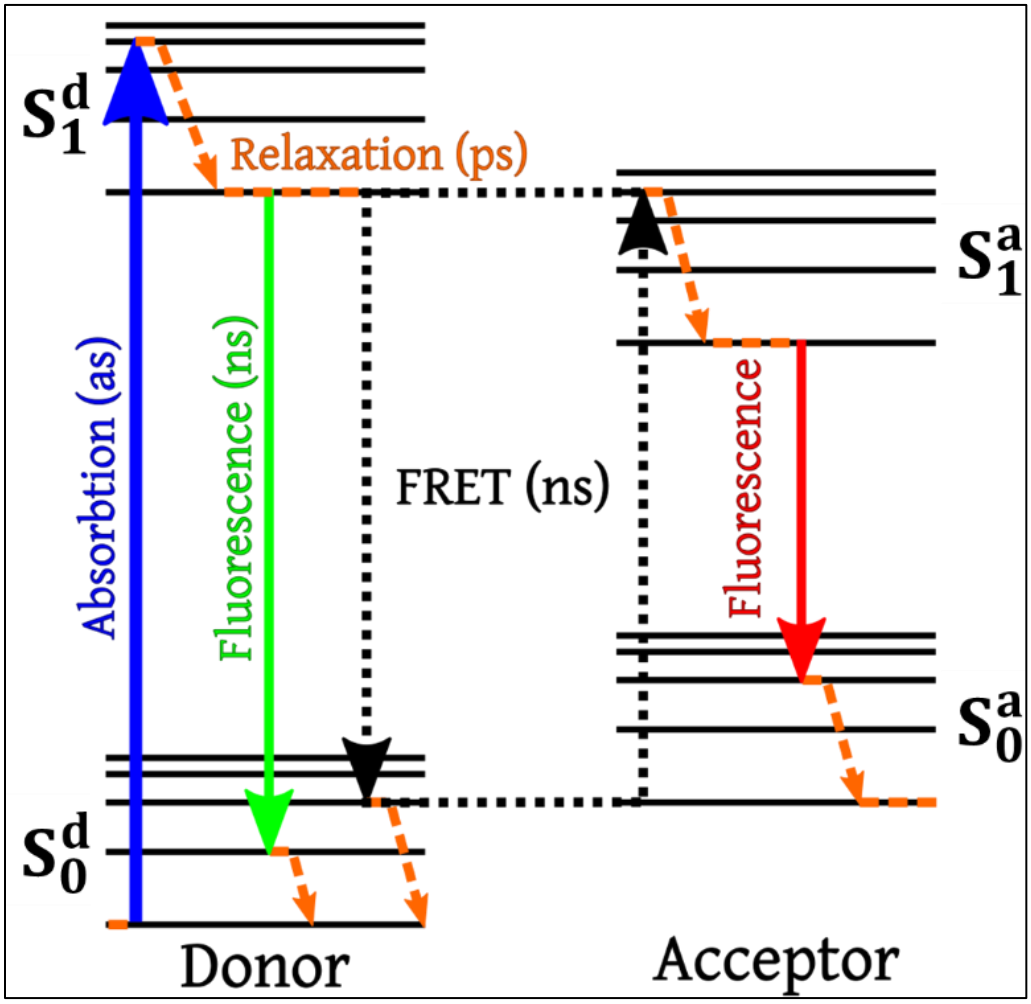


Figure 8 Jablonski Diagram of FRET

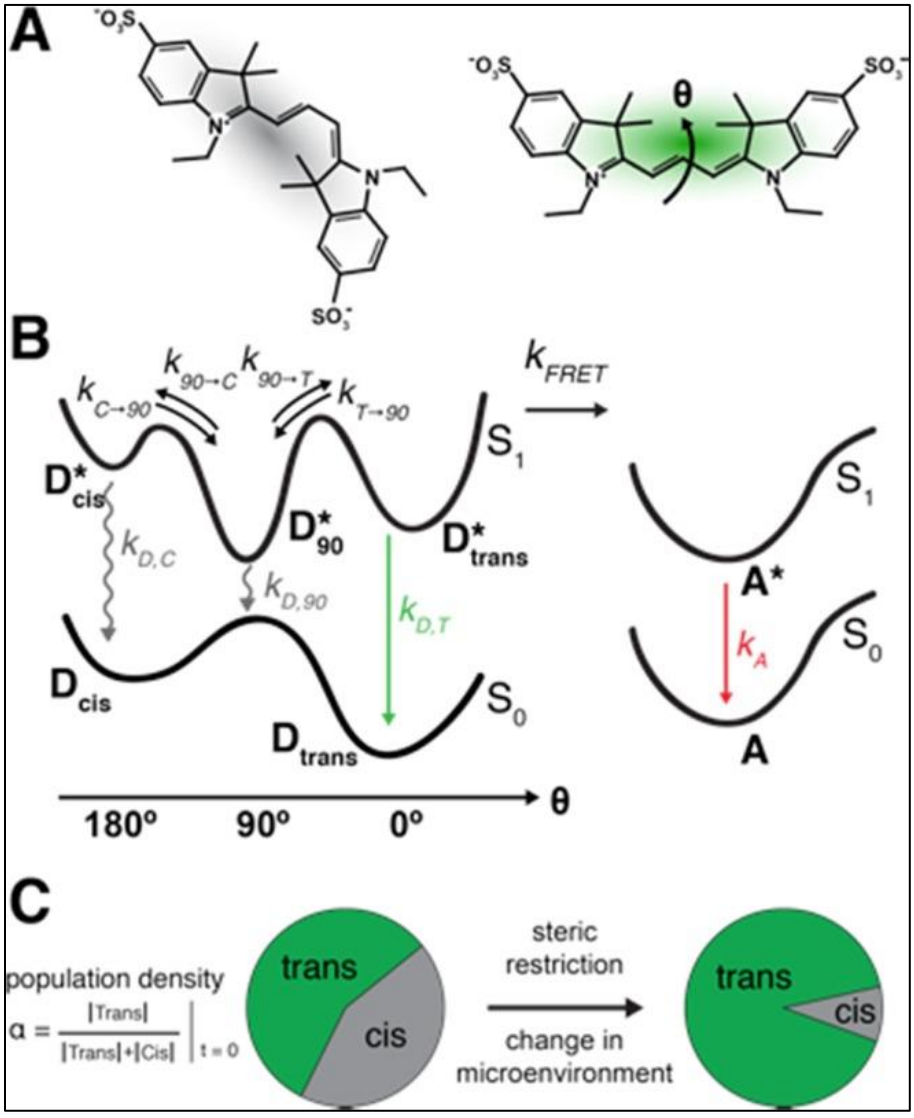


Figure 9 Jablonski Diagram of PIFE[94]

References

[1] Geszvain K.M., L. R. (2005) The structure of bacterial RNA polymerase, *University of Wisconsin Dept. of Bacteriology. Web Site Edition*, 12.

[2] 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*, 1139-1143.

[3] Goldman, S. R., Ebright, R. H., and Nickels, B. E. (2009) Direct detection of abortive RNA transcripts in vivo, *Science 324*, 927-928.

[4] 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*, 1144-1147.

[5] Winkelman, J. T., Winkelman, B. T., Boyce, J., 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*, 768-780.

[6] Henderson, K. L., Felth, L. C., Molzahn, C. M., Shkel, I. A., Wang, S., Chhabra, M., Ruff, E. F., Bieter, L., Kraft, J. E., and Record, M. T. J. (2017) Mechanism of transcription initiation and promoter escape by *E. coli* RNA polymerase, *Proceedings of the National Academy of Sciences 114*, E3032-E3040.

[7] Duchi, D., Bauer, D. L. V., Fernandez, L., Evans, G., Robb, N., Hwang, L. C., Gryte, K., Tomescu, A., Zawadzki, P., Morichaud, Z., Brodolin, K., and Kapanidis, A. N. (2016) RNA polymerase pausing during initial transcription, *Molecular Cell 63*, 939-950.

[8] Ron Milo, R. P. (2015) Cell Biology by the Numbers: What is faster, transcription or translation?, 1 ed., pp 231-236, Taylor and Francis Group.

[9] Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000) A structural model of transcription elongation, *Science 289*, 619-625.

[10] Brueckner, F., Ortiz, J., and Cramer, P. (2009) A movie of the RNA polymerase nucleotide addition cycle, *Curr Opin Struct Biol 19*, 294-299.

[11] Landick, J. Z. a. R. Substrate Loading, Nucleotide Addition, and Translocation by RNA Polymerase, In *RNA Ploymerases as Moelcular Motors* (Henri Buc, T. S., Ed.), pp 206-235.

[12] Ray-Soni, A., Bellecourt, M. J., and Landick, R. (2016) Mechanisms of Bacterial Transcription Termination: All Good Things Must End, *Annu Rev Biochem 85*, 319-347.

[13] Browning, D. F., and Busby, S. J. (2016) Local and global regulation of transcription initiation in bacteria, *Nat Rev Microbiol 14*, 638-650.

[14] Lee, T. I., and Young, R. A. (2013) Transcriptional Regulation and Its Misregulation in Disease, *Cell 152*, 1237-1251.

[15] Tichelaar, W., Schutter, W. G., Arnberg, A. C., van Bruggen, E. F., and Stender, W. (1983) The quaternary structure of Escherichia coli RNA polymerase studied with (scanning) transmission (immuno)electron microscopy, *Eur J Biochem 135*, 263-269.

[16] Darst, S. A., Ribi, H. O., Pierce, D. W., and Kornberg, R. D. (1988) Two-dimensional crystals of Escherichia coli RNA polymerase holoenzyme on positively charged lipid layers, *J Mol Biol 203*, 269-273.

[17] Gruber, T. M., and Gross, C. A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space, *Annu Rev Microbiol 57*, 441-466.

[18] Feklistov, A., Sharon, B. D., Darst, S. A., and Gross, C. A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective, *Annu Rev Microbiol 68*, 357-376.

[19] Murakami, K. S. (2015) Structural biology of bacterial RNA polymerase, *Biomolecules 5*, 848-864.

[20] Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution, *Science 292*, 1863-1876.

[21] Lane, W. J., and Darst, S. A. (2010) Molecular evolution of multisubunit RNA polymerases: sequence analysis, *J Mol Biol 395*, 671-685.

[22] Artsimovitch, I., Svetlov, V., Murakami, K. S., and Landick, R. (2003) Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions, *J Biol Chem 278*, 12344-12355.

[23] Chlenov, M., Masuda, S., Murakami, K. S., Nikiforov, V., Darst, S. A., and Mustaev, A. (2005) Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase beta' subunit, *J Mol Biol 353*, 138-154.

[24] Haugen, S. P., Ross, W., Manrique, M., and Gourse, R. L. (2008) Fine structure of the promoter-sigma region 1.2 interaction, *Proceedings of the National Academy of Sciences 105*, 3292-3297.

[25] Feklistov, A., and Darst, S. A. (2011) Structural basis for promoter -10 element recognition by the bacterial RNA polymerase σ subunit, *Cell 147*, 1257-1269.

[26] Kulbachinskiy, A., and Mustaev, A. (2006) Region 3.2 of the s 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*, 18273-18276.

[27] Heyduk, E., and Heyduk, T. (2014) Next generation sequencing-based parallel analysis of melting kinetics of 4096 variants of a bacterial promoter, *Biochemistry 53*, 282-292.

[28] Lubliner, S., Keren, L., and Segal, E. (2013) Sequence features of yeast and human core promoters that are predictive of maximal promoter activity, *Nucleic Acids Research 41*, 5569-5581.

[29] Zuo, Y., and Steitz, T. A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Molecular Cell 58*, 534-540.

[30] Barne, K. A., Bown, J. A., Busby, S. J., and Minchin, S. D. (1997) Region 2.5 of the Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters, *EMBO J 16*, 4034-4040.

[31] Harley, C. B., and Reynolds, R. P. (1987) Analysis of E. coli promoter sequences, *Nucleic Acids Res 15*, 2343-2361.

[32] Aoyama, T., Takanami, M., Ohtsuka, E., Taniyama, Y., Marumoto, R., Sato, H., and Ikehara, M. (1983) Essential structure of *E. coli* promoter: effect of spacer length between two consensus sequences on promoter function, *Nucleic Acids Research 11*, 5855-5864.

[33] Stefano, J. E., and Gralla, J. D. (1982) Spacer mutations in the *lac* ps promoter, *Proceedings of the National Academy of Sciences 79*, 1069*1072.

[34] 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 sigma(70) RNA polymerase and promoter DNA complex revealed a role of sigma non-conserved region during the open complex formation, *J Biol Chem 293*, 7367-7375.

[35] Hook-Barnard, I. G., and Hinton, D. M. (2009) The promoter spacer influences transcription initiation via sigma70 region 1.1 of Escherichia coli RNA polymerase, *Proc Natl Acad Sci U S A 106*, 737-742.

[36] Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002) Structure of the bacterial RNA polymerase promoter specificity σ subunit, *Molecular Cell 9*, 527-539.

[37] Geszvain, K., Gruber, T. M., Mooney, R. A., Gross, C. A., and Landick, R. (2004) A hydrophobic patch on the flap-tip helix of E.coli RNA polymerase mediates sigma(70) region 4 function, *J Mol Biol 343*, 569-587.

[38] Haugen, S. P., Berkmen, M. B., Ross, W., Gaal, T., Ward, C., and Gourse, R. L. (2006) rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase, *Cell 125*, 1069-1082.

[39] Drennan, A., Kraemer, M., Capp, M. W., Gries, T., Ruff, E. F., Sheppard, C., Wigneshweraraj, S., Artsimovitch, I., and Record, M. T., Jr. (2012) Key roles of the downstream mobile jaw of *Escherichia coli* RNA polymerase in transcription initiation, *Biochemistry 51*, 9447-9459.

[40] Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase, *Science 262*, 1407-1413.

[41] Rao, L., Ross, W., Appleman, J. A., Gaal, T., Leirmo, S., Schlax, P. J., Record, M. T., Jr., and Gourse, R. L. (1994) Factor independent activation of rrnB P1. An "extended" promoter with an upstream element that dramatically increases promoter strength, *J Mol Biol 235*, 1421-1435.

[42] Estrem, S. T., Ross, W., Gaal, T., Chen, Z. W., Niu, W., Ebright, R. H., and Gourse, R. L. (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit, *Genes Dev 13*, 2134-2147.

[43] Ross, W., and Gourse, R. L. (2005) Sequence-independent upstream DNA-aCTD interactions strongly stimulate *Escherichia coli* RNA polymerase-*lac*UV5 promoter association, *Proceedings of the National Academy of Sciences 102*, 291-296.

[44] Davis, C. A., Capp, M. W., Record, M. T., Jr., and Saecker, R. M. (2005) The effects of upstream DNA on open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 102*, 285-290.

[45] Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr., and Saecker, R. M. (2007) Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 104*, 7833-7838.

[46] 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*, 1076-1080.

[47] Vvedenskaya, I. O., Vahedian-Movahed, H., Zhang, Y., Taylor, D. M., Ebright, R. H., and Nickels, B. E. (2016) Interactions between RNA polymerase and the core recognition element are a determinant of transcription start site selection, *Proc Natl Acad Sci U S A 113*, E2899-2905.

[48] NandyMazumdar, M., Nedialkov, Y., Svetlov, D., Sevostyanova, A., Belogurov, G. A., and Artsimovitch, I. (2016) RNA polymerase gate loop guides the nontemplate DNA strand in transcription complexes, *Proc Natl Acad Sci U S A 113*, 14994-14999.

[49] Mishra, A., Siwach, P., Misra, P., Jayaram, B., Bansal, M., Olson, W. K., Thayer, K. M., and Beveridge, D. L. (2018) Toward a Universal Structural and Energetic Model for Prokaryotic Promoters, *Biophys J 115*, 1180-1189.

[50] Butler, J. E., and Kadonaga, J. T. (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression, *Genes Dev 16*, 2583-2592.

[51] Nogales, E., Louder, R. K., and He, Y. (2017) Structural Insights into the Eukaryotic Transcription Initiation Machinery, *Annu Rev Biophys 46*, 59-83.

[52] Han, Y., and He, Y. (2016) Eukaryotic transcription initiation machinery visualized at molecular level, *Transcr-Austin 7*, 203-208.

[53] Sainsbury, S., Bernecky, C., and Cramer, P. (2015) Structural basis of transcription initiation by RNA polymerase II, *Nat Rev Mol Cell Biol 16*, 129-143.

[54] Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C., and Cramer, P. (2017) Structural Basis of RNA Polymerase I Transcription Initiation, *Cell 169*, 120-131 e122.

[55] Sadian, Y., Tafur, L., Kosinski, J., Jakobi, A. J., Wetzel, R., Buczak, K., Hagen, W. J., Beck, M., Sachse, C., and Muller, C. W. (2017) Structural insights into transcription initiation by yeast RNA polymerase I, *EMBO J 36*, 2698-2709.

[56] Jochem, L., Ramsay, E. P., and Vannini, A. (2017) RNA polymerase I, bending the rules?, *EMBO J 36*, 2664-2666.

[57] Vorlander, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H., and Muller, C. W. (2018) Molecular mechanism of promoter opening by RNA polymerase III, *Nature 553*, 295-300.

[58] Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., and Vannini, A. (2018) Structural basis of RNA polymerase III transcription initiation, *Nature 553*, 301-306.

[59] Han, Y., Yan, C., Fishbain, S., Ivanov, I., and He, Y. (2018) Structural visualization of RNA polymerase III transcription machineries, *Cell Discov 4*, 40.

[60] Dienemann, C., Schwalb, B., Schilbach, S., and Cramer, P. (2019) Promoter Distortion and Opening in the RNA Polymerase II Cleft, *Mol Cell 73*, 97-106 e104.

[61] RT, K. (1987) The 0^0C Closed Complexes between Escherichia coli RNA Polymerase and Two Promoters, T7-A3 and ZacUV5, *Journal of Biological Chemistry 262*, 13654-13661.

[62] Schikor, P., Metzeger, W., Werel, W., Lederer, H., and Heumann, H. (1990) Topography of intermediates in transcription initiation of *E.coli*., *EMBO J 9*, 2215-2220.

[63] Ruff, E. F., Record, M. T., Jr. , and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation, *Biomolecules 5*, 1035-1062.

[64] Hofer, B., Muller, D., and Koster, H. (1985) The pathway of E. coli RNA polymerase-promoter complex formation as visualized by footprinting, *Nucleic Acids Res 13*, 5995-6013.

[65] Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., and Record, M. T. J. (2002) Kinetic studies and structural models of the association of *E. coli* sigma 70 RNA polymerase with the LPR promoter: large scale conformational changes in forming the kinetically significant intermediates, *Journal of Molecular Biology 319*, 649-671.

[66] Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S. A. (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex, *Elife 4*.

[67] Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., Saecker, R. M., and Record, M. T. J. (1998) DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA, *Journal of Molecular Biology 283*, 741-756.

[68] Craig, M. L., Suh, W. C., and Record, M. T., Jr. (1995) HO. and DNase I probing of E sigma 70 RNA polymerase--lambda PR promoter open complexes: Mg2+ binding and its structural consequences at the transcription start site, *Biochemistry 34*, 15624-15632.

[69] Gries, T. J., Kontur, W. S., Capp, M. W., Saecker, R. M., and Record, M. T. J. (2010) One-step DNA melting in the RNA polyerase cleft opens the initiation bubble to form an unstable open complex, *Proceedings of the National Academy of Sciences 107*, 10418-10423.

[70] Kontur, W. S., Saecker, R. M., Capp, M. W., and Record, M. T. J. (2008) Late steps in the formation of *E. coli* RNA polymerase-λP_R promoter open complexes: characterization of conformational changes by rapid [perturbant] upshift experiments, *Journal of Molecular Biology 376*, 1034-1047.

[71] Ruff, E. F., Drennan, A. C., Capp, M. W., Poulos, M. A., Artsimovitch, I., and Record, M. T., Jr. (2015) *E. coli* polymerase determinants of open complex lifetime and structure, *Journal of Molecular Biology 427*, 2435-2450.

[72] Andrey Feklistov, B. B., Jesse Hauver, Agnieszka Lass-Napiorkowska, Markus Kalesse, Florian Glaus, Karl-Heinz Altmann, Tomasz Heyduk, Robert Landick, and Seth A. Darst. (2017) RNA polymerase motions during promoter melting., *Science 356*, 4.

[73] 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 Res 46*, 7284-7295.

[74] Boyaci, H., Chen, J., Jansen, R., Darst, S. A., and Campbell, E. A. (2019) Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding, *Nature 565*, 382-385.

[75] Benoff, B., Yang, H., Lawson, C. L., Parkinson, G., Liu, J., Blatter, E., Ebright, Y. W., Berman, H. M., and Ebright, R. H. (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex, *Science 297*, 1562-1566.

[76] Lee, W., von Hippel, P. H., and Marcus, A. H. (2014) Internally labeled Cy3/Cy5 DNA constructs show greatly enhanced photo-stability in single-molecule FRET experiments, *Nucleic Acids Res 42*, 5967-5977.

[77] Ouellet, J., Schorr, S., Iqbal, A., Wilson, T. J., and Lilley, D. M. (2011) Orientation of cyanine fluorophores terminally attached to DNA via long, flexible tethers, *Biophys J 101*, 1148-1154.

[78] Moreira, B. G., You, Y., and Owczarzy, R. (2015) Cy3 and Cy5 dyes attached to oligonucleotide terminus stabilize DNA duplexes: predictive thermodynamic model, *Biophys Chem 198*, 36-44.

[79] Di Fiori, N., and Meller, A. (2010) The Effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids, *Biophys J 98*, 2265-2272.

[80] Roy, R., Hohng, S., and Ha, T. (2008) A practical guide to single-molecule FRET, *Nat Methods 5*, 507-516.

[81] Gust, A., Zander, A., Gietl, A., Holzmeister, P., Schulz, S., Lalkens, B., Tinnefeld, P., and Grohmann, D. (2014) A starting point for fluorescence-based single-molecule measurements in biomolecular research, *Molecules 19*, 15824-15865.

[82] Valeur, B., and Berberan-Santos, M. N. (2012) *Molecular Fluorescence*, Wiley.

[83] Lakowicz, J. R. (2006) *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer.

[84] Sreenivasan, R., Heitkamp, S., Chhabra, M., Saecker, R. M., Lingeman, E., Poulos, M. A., McCaslin, D., Capp, M. W., Artsimovitch, I., and Record, M. T. J. (2016) Fluorescence resonance energy transfer characterization of DNA wrapping in closed and open *Escherichia coli* RNA polyerase-λP_R promoter complexes, *Biochemistry 55*, 2174-2186.

[85] Clegg, R. M. (1992) Fluorescence resonance energy transfer and nucleic acids, *Methods Enzymol 211*, 353-388.

[86] Vander Meulen, K. A., Saecker, R. M., and Record, M. T., Jr. (2008) Formation of a wrapped DNA-protein interface: experimental characterization and analysis of the large contributions of ions and water to the thermodynamics of binding IHF to H' DNA, *J Mol Biol 377*, 9-27.

[87] Hohlbein, J., Craggs, T. D., and Cordes, T. (2014) Alternating-laser excitation: single-molecule FRET and beyond, *Chem Soc Rev 43*, 1156-1171.

[88] 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*, 22175-22180.

[89] 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, *Mol Cell 70*, 695-706 e695.

[90] Duchi, D., Gryte, K., Robb, N. C., Morichaud, Z., Sheppard, C., Brodolin, K., Wigneshweraraj, S., and Kapanidis, A. N. (2018) Conformational heterogeneity and bubble dynamics in single bacterial transcription initiation complexes, *Nucleic Acids Res 46*, 677-688.

[91] Hwang, H., and Myong, S. (2014) Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions, *Chem Soc Rev 43*, 1221-1229.

[92] Nguyen, B., Ciuba, M. A., Kozlov, A. G., Levitus, M., and Lohman, T. M. (2019) Protein Environment and DNA Orientation Affect Protein-Induced Cy3 Fluorescence Enhancement, *Biophys J 117*, 66-73.

[93] Ploetz, E., Lerner, E., Husada, F., Roelfs, M., Chung, S., Hohlbein, J., Weiss, S., and Cordes, T. (2016) Forster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers, *Sci Rep 6*, 33257.

[94] Lerner, E., Ploetz, E., Hohlbein, J., Cordes, T., and Weiss, S. (2016) A Quantitative Theoretical Framework For Protein-Induced Fluorescence Enhancement-Forster-Type Resonance Energy Transfer (PIFE-FRET), *J Phys Chem B 120*, 6401-6410.

[95] Stennett, E. M., Ciuba, M. A., Lin, S., and Levitus, M. (2015) Demystifying PIFE: The Photophysics Behind the Protein-Induced Fluorescence Enhancement Phenomenon in Cy3, *J Phys Chem Lett 6*, 1819-1823.

[96] Paget, M. S. (2015) Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution, *Biomolecules 5*, 1245-1265.

[97] Wade M, M. J., Coussens NP, et al. (2017) Inhibition of Protein-Protein Interactions: Cell-Based Assays. , Eli Lilly & Company and the National Center for Advancing Translational Sciences.

Chapter 2

Fluorescence Resonance Energy Transfer Characterization of DNA Wrapping in Closed and Open *Escherichia coli* RNA Polymerase-λP_R Promoter Complexes.

PREFACE

This chapter has been previously published in ACS Biochemistry Journal with myself as third author; Sreenivasan R, Heitkamp S, Chhabra M, et al. Fluorescence Resonance Energy Transfer Characterization of DNA Wrapping in Closed and Open Escherichia coli RNA Polymerase-λP(R) Promoter Complexes. Biochemistry. 2016;55(14):2174–2186. My contributions to this research included data collection for fluorescence experiments with non-promoter DNA fragment and its preparation, determining anisotropy measurements, and performing permanganate footprinting experiment with closed complex formed by Cy3/Cy5-100 λP_R promoter fragment and *E.coli* RNAP at 2^0C.

ABSTRACT (248 words)

Initial recognition of promoter DNA by RNA polymerase (RNAP) is proposed to trigger a series of conformational changes beginning with bending and wrapping of the 40-50 bp of DNA immediately upstream of the -35 region. Kinetic studies with upstream truncations demonstrated that the presence of upstream DNA facilitates bending and entry of the downstream duplex (to +20) into the active site cleft to form an advanced closed complex (CC), prior to melting of ~13 bp (-11 to +2) including the transcription start site (+1). Atomic force microscopy and footprinting revealed that the stable open complex (OC) is also highly wrapped (-60 to +20). To test the proposed bent-wrapped model of duplex DNA in advanced RNAP-λP_R CC and compare wrapping in CC and OC, we use fluorescence energy transfer (FRET) between cyanine dyes at far-upstream (-100) and downstream (+14) positions of promoter DNA. Similarly large intrinsic FRET efficiencies are observed for CC (0.30 ± 0.11) and OC (0.34 ± 0.11) for both probe orientations, while a non-promoter control exhibits no FRET. Fluorescence enhancements at +14 are observed in single-dye-labeled CC and OC. These results demonstrate that upstream DNA is extensively wrapped and the start-site region is bent into the cleft in the advanced CC, reducing the distance between -100 and +14 positions on promoter DNA from ~400 Å to <100 Å. The close proximity of upstream DNA to the downstream cleft in the advanced CC is consistent with the proposed mechanism for facilitation of OC formation by upstream DNA.

Introduction

Efficient conversion of closed promoter complexes (CC) to stable open complexes (OC) requires interactions of RNA polymerase (RNAP) with upstream and downstream DNA as well as with the central recognition region from the -35 element to the transcription start site (+1) [1 2]. RNAP interactions with the 20 bp region downstream of +1 are key players in the conversion of the initial unstable OC to more stable OC [2 3 4 5]. Interactions of RNAP with 40-50 bp of DNA upstream of the -35 element, including both the region of the UP element and far-upstream DNA, are necessary for rapid isomerization of CC to OC [1 2]. Truncations in this upstream region reduce isomerization rates of lacUV5 and λP_R promoters by up to 2 orders of magnitude [6 7, 8].

How do upstream and downstream DNA exert these effects? Formation of the initial CC (often called RP_c, in which the promoter DNA is presumed to be linear [1 2]) sets in motion a series of conformational changes in RNAP and promoter DNA [2 1 9]. To explain the large reduction of the isomerization rate in OC formation that results from truncation of upstream DNA, as well as the features observed in real-time footprints of CC at full length and truncated promoter DNA [8], we proposed that bending and wrapping of the UP element region and far-upstream DNA on RNAP are among the first of these large conformational changes [1 7]. Wrapping of upstream DNA around RNAP brings it near the αNTD (which act as hinges at the base of the active site cleft) and also near the downstream cleft, triggering conformational changes in the cleft and in downstream mobile elements (DME) of RNAP. These conformational changes facilitate bending of the downstream duplex into the cleft to form the most advanced CC prior to opening, in which more than 100 bp of promoter DNA (-82 to +20) interact with RNAP [1 2 4 7]. Subsequently these DME form downstream interactions that stabilize the initial OC [4 5]. While far-upstream (-65 to -82) interactions are eliminated in conversion of the advanced CC to the stable λP_R OC, footprinting reveals that at least 80 bp (-65 to +20) of λP_R promoter DNA remain in tight contact with RNAP [7], twice the length of the elongation complex (40 bp) [10]. AFM DNA compaction studies with

stable λP_R OC reveal that promoter DNA is wrapped around RNAP [11] [12]. Abortively initiating complexes at λP_R are also wrapped as judged by AFM [13]. Hence, evidence exists for extensive bending and wrapping of promoter DNA in both advanced λP_R CC and OC. Unwrapping must occur at or before promoter escape.

Bending and/or wrapping of DNA appear to be general properties of CC and OC at other bacterial promoters and for eukaryotic RNAP. Extensive bending and wrapping of TyrT promoter DNA in a CC with RNAP both in the presence and absence of the transcriptional activator FIS was observed by AFM [14]. Extensive wrapping of promoter DNA in CC formed by eukaryotic RNAP was observed by crosslinking and electron microscopy experiments [15] [16] [17]. To date, however, no direct spectroscopic assay to investigate structural, thermodynamic and kinetic-mechanistic aspects of promoter bending and wrapping has been reported. Here we use ensemble Förster (fluorescence) resonance energy transfer (FRET) from Cy3 to Cy5 cyanine dyes attached at the upstream (-100) and downstream (+14) ends of λP_R promoter duplexes as a quantitative assay for the large reduction in end-to-end distance of promoter DNA anticipated from bending and wrapping. Results are obtained for both an advanced λP_R CC formed at 2˚C, below the temperature range of the opening transition of this promoter (5 – 15°C) [1], and for the stable λP_R OC at 19°C [1] [2]. We also use RNAP-induced cyanine dye fluorescence enhancements [18] [19] [20] [21] [22], often called PIFE effects, at +14 to monitor the proximity of the +14 position of promoter DNA to RNAP in closed and open complexes. In ongoing research, we are applying these FRET and PIFE assays in fast kinetic studies of the roles of promoter bending and wrapping in the mechanism of formation and advancement of the initial CC, the isomerization to an OC, and the subsequent stabilization of the OC by in-cleft and downstream conformational changes.

Materials and methods

Buffers: Storage Buffer (SB) for RNAP holoenzyme is 50% glycerol (v/v), 10 mM Tris (pH 7.5 at 4 °C), 100 mM NaCl, 0.1 mM dithiothreitol (DTT), and 0.1 mM Na_2EDTA. Annealing buffer (AB) is 0.1 M potassium glutamate, 25 mM tris-acetate pH 7.6, 10 mM magnesium acetate ("KGB buffer" [23 24]). Buffer for FRET experiments (FB) is 40 mM Tris (pH 8), 10 mM $MgCl_2$, 120 mM KCl, 0.05 mg/ml BSA, 1 mM DTT, 0.02 % Tween. Buffer for transcription assays (TB) is 40 mM Tris (pH 8), 5 mM $MgCl_2$, 60 mM KCl, 0.1 mg/ml BSA, 1 mM DTT. Quench dye buffer (QB) for transcription assays is TB with 8 M urea, 15 mM Na_2EDTA, 0.05 % (w/v) xylene cyanol, 0.05 % (w/v) bromophenol blue. Stop solution for permanganate footprinting is 1 M β-mercaptoethanol and 2.6 M ammonium acetate. Unless specified, all reagents were from Sigma Aldrich (St. Louis, MO).

RNA Polymerase: Wild-type *E.coli* RNAP core (α_2 ββ'ω) was overexpressed in *E.coli* BL21(λDE3) cells from plasmid pVS10 and purified using standard protocols [25]. σ^{70} was overexpressed in *E.coli* strain M5219 containing plasmid pMRG8 and purified as previously described [26]. RNAP holoenzyme was reconstituted by incubating 5 μM core RNAP, with a two-fold molar excess of purified σ^{70} for 1 hour at 37°C. Holoenzyme was stored at -20°C for short term use. Different preparations of reconstituted RNAP were 70% to 100% active in OC formation.

Preparation of Cy3- and/or Cy5-labeled λP_R promoter and Cy3(+14)Cy5(-100) control DNA: Single dye labeled λP_R (Cy3(+14) or Cy5(+14) λP_R and Cy3(-100) or Cy5(-100) λP_R), two-dye labeled λP_R (Cy3(-100)Cy5(+14) or Cy3(+14)Cy5(-100) λP_R) and two-dye labeled control DNA for FRET (Cy3(+14)Cy5(-100) DNA control) were prepared from dye-labeled primers using the sequences and methods described in Supplemental.

Fluorescence experiments with RNAP-λP_R promoter DNA complexes: RNAP-promoter complexes were formed by incubating 100 nM Cy3Cy5 λP_R DNA and 20 -120 nM RNAP for 20-90 min in FB at 2°C (for CC) or 19°C (for OC) [9 27]. No differences were observed between

complexes incubated for these different times. FRET experiments with dye-labeled OC were performed at 19°C because the fluorescence emission intensity of cyanine dyes decreases strongly with increasing temperature (Supplemental Figure 3 [28]). Parallel series of experiments were performed with singly and doubly labeled 100 nM (Cy3, Cy5) promoter DNA. A Quantamaster C-60/2000 spectrofluorimeter (Photon Technology Instruments, New Jersey) equipped with a 75-W xenon compact arc lamp and single monochromators on the excitation and emission beam path was used for all experiments. Cuvettes (170 µl) were coated with Sigmacote before each experiment. Excitation and emission slit widths were 3 nm and 5 nm respectively. Samples were excited at 515 nm (for Cy3) or at 610 nm (for Cy5); these wavelengths were chosen to minimize direct excitation of Cy5 from excitation of Cy3 and to minimize overlap of Cy5 excitation and emission spectra. Emission spectra were collected between 550 – 720 nm for two-dye FRET experiments, single dye Cy3 experiments and between 620 -720 nm for single-dye Cy5 experiments. Samples were removed from the instrument or shuttered when not being observed to minimize photodamage. For the temperature series experiment (Supplemental Figure 3), the temperature was increased in steps and the cuvette equilibrated at each temperature before the emission scan.

Determination and analysis of FRET efficiency: FRET efficiency E was calculated by the '$Ratio_A$' method [29 30]. For given Cy3Cy5 labeled DNA complexed with RNAP, $Ratio_A$ at Cy5 $\lambda_{\max}^{\mathrm{em}}$ (660 nm) is obtained by:

$$RatioA_{\,660\,nm} = \frac{I_{660\,nm}^{Cy3\ exc}}{I_{660\,nm}^{Cy5\ exc}} \qquad \text{Eq (1)}$$

Where $\mathrm{I}_{660\ \mathrm{nm}}^{\mathrm{Cy3\ exc}}$ is the Cy5 emission intensity obtained by Cy3 excitation at 515 nm and $\mathrm{I}_{660\ \mathrm{nm}}^{\mathrm{Cy5\ exc}}$ is the Cy5 emission intensity obtained by direct Cy5 excitation at 610 nm. For this calculation,

the Cy5 portion of the emission spectrum for excitation of Cy3Cy5-labeled DNA (515 nm) was determined by subtracting a reference spectrum obtained for Cy3 DNA (Supplemental Figure 8). Prior to subtraction, this reference spectrum was normalized to match the fluorescence intensity of the Cy3Cy5-labeled DNA – RNAP spectra over the wavelength range where only Cy3 emits (555-570 nm). In Eq 1, the resulting Cy5 emission at 660nm is divided by the Cy5 emission at 660 nm for the same Cy3-Cy5 labeled DNA obtained by direct excitation at 610 nm to obtain '$Ratio_A$'. Extinction coefficients ∈ for Cy3 and Cy5 in the dye-labeled constructs at the excitation wavelengths used (515 nm for Cy3; 515 and 610 nm for Cy5) were determined using variable pathlength technology (SoloVPE system from C Technologies). Absorbance of the Cy3 or Cy5 labeled construct of known concentration was measured for at least 10 different path lengths from 0.05 mm – 0.2 mm, from which $\in_{Cy5}^{610nm}, \in_{Cy5}^{515\,nm}$ and $\in_{Cy3}^{515\,nm}$ were determined [31]. Values obtained ($\in_{Cy5}^{610nm}$= 9.3*10^4 M^{-1} cm^{-1}; $\in_{Cy5}^{515nm}$= 5.6*10^3 M^{-1} cm^{-1}; $\in_{Cy3}^{515nm}$ = 9.6*10^4 M^{-1} cm^{-1}) are consistent with previously published values at other wavelengths when converted using the spectra of these dyes [32].

The FRET efficiency E is then obtained from $Ratio_A$ determined at λ_{max}^{em}(660 nm) for Cy5, from Eq 2 [30]:

$$E = (Ratio_A * \in_{Cy5}^{610\,nm} - \in_{Cy5}^{515nm}) \,/\, \in_{Cy3}^{515nm} \qquad \text{Eq (2)}$$

Values of E reported in Table 1 and Supplemental Tables 1-3 have been corrected for the population fraction of free promoter DNA at each concentration of RNAP investigated. Binding is observed to be stoichiometric up to a 1:1 [RNAP]:[DNA] ratio at 100nM concentrations of both reagents. For CC and OC formed by Cy3(-100)Cy5(+14) λP_R DNA, Cy5(+14) PIFE makes a ~30% contribution (occupancy corrected) to both numerator and denominator of Eq 1 for $Ratio_A$ (Figure 5**,** Supplemental figure 5), and has no net effect on the calculation of the FRET E. For CC and OC formed by Cy3(+14)Cy5(-100) λP_R, Cy3(+14) PIFE makes a ~20% contribution (occupancy

corrected) to the numerator of $Ratio_A$ (Supplemental figure 4), and is accounted for in calculating $Ratio_A$.

The interdye distance R is calculated from the efficiency E:

$$R = R_o (E^{-1} - 1)^{1/6} \qquad \text{Eq (3)}$$

where the Förster distance (R_o), the distance at which E = 0.5 for the Cy3-Cy5 pair, is given by [31]:

$$R_o = (9.78 * 10^3)(\Phi_D \kappa^2 \eta^{-4} J(v))^{1/6} \text{ Å} \qquad \text{Eq (4)}$$

In Eq 4, the refractive index Ƞ = 1.33, the spectral overlap integral for Cy3 and Cy5 is $J(v) = 7.2*10^{-13} M^{-1} cm^3$ [33], Φ_D is the quantum yield of Cy3 and κ^2 is the orientation factor. Information about κ^2 is provided by fluorescence anisotropy experiments (below).

The quantum yield of donor Cy3 (Φ_D) depends on temperature and on its environment, including attachment to ss or dsDNA [28]. For cyanine dyes attached at +14, RNAP(protein)-induced fluorescence enhancement (often called PIFE; [18 19]) contributes to the quantum yield and hence the emission intensity. In our DNA constructs, Cy3 or Cy5 are inserted between the 5' position of the end deoxyribose of ds DNA and the 3' position of the first deoxyribose in a short ss region at the end of the duplex (See Supplemental methods). Occupancy corrected Cy3 PIFE effects at +14 are ~ 20 % (Supplemental Figure 5). Correcting for PIFE effects by short extrapolation of a linear fit of Cy3 - 5' dsDNA Φ_D vs. temperature [28] to 2°C gives $\Phi_D = 0.34\pm0.03$ and by interpolation $\Phi_D = 0.21\pm0.02$ at 19°C when Cy3 is at +14. Similarly $\Phi_D = 0.31\pm0.03$ at 2°C and $\Phi_D = 0.19\pm0.02$ at 19°C when Cy3 is at -100. These results are applicable at the salt concentrations of our experiments [28].

Determination and analysis of fluorescence anisotropy (FA): These experiments were performed with a Quantamaster C-60/2000 spectrofluorimeter (Photon Technology Instruments, New

Jersey) equipped with two polarizers in the paths for excitation and emission using the L-format method in the time-based mode. Experimental settings were as in the FRET experiments. Intensity values, *I*, were recorded at 565 nm for Cy3 and 660 nm for Cy5 as the λ_{max}^{em} as seen from FRET experiments.The anisotropy, r was calculated as [34]:

$$r = \frac{I_{\parallel} - G \cdot I_{\perp}}{I_{\parallel} + 2G \cdot I_{\perp}} \qquad \text{Eq (5)}$$

The subscripts on *I* indicate the relative orientations (parallel and perpendicular) of the excitation and emission polarizers. G, the instrumental correction factor for different detection sensitivities for vertically and horizontally polarized light was measured with Cy3 and Cy5-labeled primers separately. Individual values of $I_{\parallel}, I_{\perp}$ used to calculate r from Eq. 4 were determined by averaging data from two scans. Values of r at 19°C and 2°C reported here are averages of three and two independent determinations, respectively. Uncertainties are reported as deviations from the average.

Minimum and maximum values of κ^2 for our DNA- RNAP complexes were estimated from anisotropy measurements of donor and acceptor dyes (r_a, r_d) from Eqs 6,7 [31 35]:

$$\kappa^2_{min} \approx \frac{2}{3}\left(1 - \frac{\sqrt{\frac{5\,r_d}{2}} + \sqrt{\frac{5\,r_a}{2}}}{2}\right) \qquad \text{Eq (6)}$$

$$\kappa^2_{max} \approx \frac{2}{3}\left(1 + \sqrt{\frac{5\,r_d}{2}} + \sqrt{\frac{5\,r_a}{2}} + 3\sqrt{\frac{5\,r_d}{2}}\sqrt{\frac{5\,r_a}{2}}\right) \qquad \text{Eq (7)}$$

where r_d and r_a are emission anisotropy of donor and acceptor measured at λ_{max}^{em} and the factor $5/2 = 1/r_{max}$ where r_{max} is the maximum anisotropy obtained for an immobilized dye [31]. Together, Eqs. 6 and 7 determine a range of κ^2 values bracketing $\kappa^2 = 2/3$. Both κ^2_{min} and κ^2_{max} approach 2/3 as r approaches zero (the freely rotating limit) and diverge to define a wider range of possible κ^2 values, approaching $0 \leq \kappa^2 \leq 4$ as r_a and r_d increase to 0.4 (the immobilized dye anisotropy).

Permanganate footprinting assay: Permanganate footprinting was performed to characterize the position and extent of opening of promoter DNA in a dye-labeled construct (Supplemental Figure 9). ^{32}P end-labeling of dye-labeled primers was unsuccessful and permanganate footprinting was therefore performed only for Cy5 (-100) λP_R where the upstream primer had an internal Cy5 while downstream primer with no dye ('downstream 1-5') was end-labeled with [γ ^{32}P]-ATP with T4 PNK (NEB, Ipswich, MA). The MnO_4^- reference was ^{32}P-labeled λP_R without any fluorescent probe. DNA was purified as indicated above. DNA- RNAP complexes were formed in FB using ~1 nM promoter DNA and ~80 nM RNAP on ice (0°C; CC) and at 19°C (OC) for ~ 60 minutes. $KMnO_4$ was added to a final concentration of 10 mM and reaction carried out for 10 s at 19°C and 60 s on ice (0°C). Subsequent workup was performed as described previously [9]. Footprinting gels were imaged using Typhoon FLA-9000 (GE Healthcare Bio-sciences, Pittsburgh, PA), and analyzed using Image Quant TL. Band intensity analyses were performed with ImageJ software.

Transcription assay: To form OC, solutions containing 100 nM λP_R promoter DNA and 80 nM RNAP were incubated as above. Transcription was initiated with 50 μM ATP, CTP, and GTP and 170 nM/10 μCi [α ^{32}P]-UTP. Samples were incubated for 10 min at 19°C after the addition of [α ^{32}P]-UTP. QB was then added to the samples which were heated to 90°C for 2 min. Samples were centrifuged and loaded on a 8 M urea, 20% denaturing acrylamide gel along with a 10-100 nt low molecular weight RNA ladder from USB-Affymetrix (Santa Clara, CA) end-labeled with [γ ^{32}P]-ATP by T4 polynucleotide kinase (PNK; New England Biolabs (NEB), Ipswich, MA) Transcription gels were imaged using Typhoon FLA-9000 (GE Healthcare), and line scans analyzed using Image Quant TL or ImageJ [37].

Construction of models of RNAP-promoter complexes: Models of RNAP-promoter CC and OC were made and manipulated using PyMOL Molecular Graphics System, Version 1.7.0.3 (Schrödinger LLC, New york) and Coot [38]. For CC, our starting point was the published model of a wrapped CC [7], which predated the *E. coli* RNAP structures and was therefore based on the

T. thermophilus RNAP structure. The modeling procedure for promoter DNA and *E. coli* RNAP was done as described in [1]. The distance R between -100 to +14 in this model is ~ 85 Å.

To update this model, we replaced *T. thermophilus* RNAP by *E. coli* RNAP (PDB ID: 4LJZ) [36]. αCTD and β' sequence insertion 3 (β'SI3) is absent from this *E. coli* RNAP structure. We modeled the αCTD by analogy to the published wrapped CC model [7] using PDB structure -1LB2 for the αCTD and centering them at UP element positions -42 and -52. Key residue R265 is placed in the minor groove and the acidic patch on proximal αCTD (D259, E261) is aligned with R603 on σ region 4 [37].

No molecular model for bent, wrapped upstream DNA in an OC is available. To construct an OC model we used the recent structure of the *E. coli* RNAP transcription initiation complex (PDB ID: 4YLN [38]). In this structure, the promoter DNA is open from -11 to +2 and surrounding duplex DNA extends from -37 to +13. We modeled αCTD (1LB2) as in the CC model. Upstream DNA (-38 to approximately -112) was bent and wrapped on this structure as in the CC model above, to be consistent with OC $HO^{\cdot}$ footprinting data [7].

Results

Large Cy5 FRET Acceptor Signals from Exciting Cy3 Demonstrate DNA Wrapping in Advanced Closed and Open RNAP-λP_R Promoter Complexes

As a direct spectroscopic test of the proposed model of DNA bending and wrapping in the advanced CC [7] and to compare extents of DNA wrapping in the advanced CC and OC, far upstream (-100) position of the non-template strand and downstream (+14) position of the template strand backbone of a λP_R promoter DNA fragment were internally labeled with the Cy3/Cy5 donor/acceptor pair of cyanine dyes (*iCy3, iCy5*). Previous DNaseI footprinting [39] revealed that the equilibrium 2°C RNAP - λP_R complex, like the intermediate characterized in real time at higher temperature [7], is an advanced CC with the downstream duplex protected to +20 by

interactions with RNAP. Far-upstream contacts in the 2°C complex could not be characterized by footprinting due to low DNaseI reactivity of this region of this promoter DNA. Here we investigate FRET between cyanine dyes at -100 and +14 in this advanced λP_R CC and compare with results for the stable λP_R OC (RP_o) and with the fluorescence behavior of a nonpromoter Cy3-Cy5 labeled DNA fragment.

The signature of FRET is an increase in emission of the acceptor dye (Cy5) in a two-dye construct upon excitation of the donor (Cy3), relative to the single (Cy5)-dye control. Figure 1 shows Cy5 emission spectra as a function of the [RNAP] : [λP_R promoter DNA] ratio obtained using 100 nM promoter DNA for both orientations of the FRET pair of Cy3 and Cy5 dyes at -100 and +14 in the advanced λP_R CC. Figure 2 shows the analogous series of Cy5 emission spectra for the stable λP_R OC. A very significant FRET effect is demonstrated in both Figure 1 and 2 by the increases in Cy5 emission (λ_{em}^{max} = 660 nm) with increasing [RNAP]:[λP_R promoter DNA] ratio up to approximately 0.6-0.8:1. Additional series of titrations showing FRET for both CC and OC for both orientations of Cy3 and Cy5 are shown in Supplemental figures 1,2. In all cases the excitation wavelength is 515 nm, near the Cy3 excitation maximum and a wavelength at which direct excitation of Cy5 is minimized. Observation of FRET unambiguously demonstrates extensive bending and wrapping of promoter DNA on RNAP in both CC and OC, reducing the separation distance between the -100 and +14 positions of the DNA backbone from ~ 400 Å (linear DNA) to <100 Å (the upper bound distance for detecting a FRET effect [31]) in these complexes. These results provide the first spectroscopic demonstration of large-scale bending and wrapping in any closed or open RNAP–promoter DNA complex. No FRET is observed at either 2°C or 19°C for a 1:1 mixture of RNAP and Cy3Cy5-labeled non-promoter DNA (each 100 nM) (Figure 3), demonstrating that the FRET effects and DNA bending and wrapping by RNAP are properties of promoter CC and OC.

At [RNAP] : [λP_R promoter DNA] ratios of 0.8 -1.0, the FRET effect plateaus, indicating complete binding of the limiting (RNAP) reagent at these high nM concentrations to form wrapped CC and OC. In some cases the FRET effect is reduced upon further titration with RNAP, perhaps indicating competition from weaker end-binding modes of RNAP. Fluorescence intensities are significantly higher for both dyes at 2°C than at 19°C (Figures 1,2), as expected because of the reduction in quantum yield with increasing temperature for cyanine dyes (Supplemental figure 3) [28]. Because of this, OC FRET measurements were made at 19°C, a temperature well above the opening transition range (5-15°C) for the λP_R promoter [9 40 41].

Figure 4 plots the acceptor:donor λ_{max}^{em} emission intensity ratio I_{Cy5}/I_{Cy3} as a function of [RNAP]:[λP_R promoter DNA] ratio for both locations of the probes and both temperatures investigated. This analysis shows that I_{Cy5}/I_{Cy3} increases monotonically and linearly with increasing [RNAP]:[DNA] ratio to a plateau value attained at a ratio of 0.8:1, completely consistent with the behavior of the Cy5 acceptor signal in Figures 1 and 2. The data in these figures and Supplemental Figures 1,2 provide the information needed to calculate FRET efficiencies and estimate distances between far-upstream (-100) and downstream (+14) positions of promoter DNA in wrapped CC and OC.

FRET Efficiencies and Estimates of the Extent of Promoter DNA Wrapping in Closed and Open λP_R Complexes

The FRET efficiency E (the fraction of Cy3 excitation energy that is transferred to Cy5) of a Cy3Cy5-labeled λP_R promoter-RNAP complex is calculated from the ratio of the Cy5 emission spectrum for excitation at the Cy3 wavelength (515 nm) to that for excitation at the Cy5 wavelength (610 nm), using the "$Ratio_A$" method [30] (see Eq 2 in Methods). Table 1 and Supplemental Tables 1-3 list occupancy-corrected efficiencies E calculated for each [RNAP]:[λP_R

promoter DNA] ratio investigated, for 1:1 stoichiometric binding as demonstrated by the linearity of Figure 4.

For CC and OC, the two different arrangements of donor and acceptor dyes at -100 and +14 yield similar FRET efficiencies (Table 1), averaging to E_{CC} = 0.30 ± 0.11 and E_{OC} = 0.34 ± 0.11. FRET efficiencies of this magnitude indicate that the through-space distance between positions -100 and +14 is significantly less than 100 Å; model-building (see Discussion) indicates that for this to occur both upstream and downstream regions of the promoter DNA must be highly bent and wrapped around RNAP in both CC and OC.

Interpretation of these FRET efficiencies using Eq 3 shows that dye-dye distances in the CC and OC are approximately 15% and 10% larger than the relevant Förster distance R_o (Eq 4) for these dyes. For the Cy3-Cy5 pair, if the population-average orientation of one dye with respect to the other in CC and OC were relatively random so that $\kappa^2 = 2/3$, then R_o would be ~ 60 Å for CC and ~ 56 Å for OC. From this R_o and measured E values (**table 1**), the through-space distances R between -100 and +14 positions on promoter DNA would be ~ 69 Å for CC and ~ 62 Å for OC. These are also similar to the mid-range values of κ^2 (between κ^2_{min} and κ^2_{max}) and R_o, from Eqs. 6 and 7, and so are relevant even if the random orientation limit is inapplicable. Determination of the anisotropies in donor dye and acceptor dye samples also allows us to estimate lower and upper bound on κ^2 [31 35].

Fluorescence Anisotropies of Cyanine Dyes on DNA and in RNAP- λP_R complexes

To obtain maximum and minimum estimates of κ^2 and test the relevance of $\kappa^2 = 2/3$ for analysis of FRET data for CC and OC, as well as to determine relative mobilities of Cy3 and Cy5 dyes in CC and OC as compared to promoter DNA, primer DNA and the free dyes, we measured steady state fluorescence anisotropies r. Average values obtained from multiple independent determinations on these different dye-labeled DNAs and RNAP-DNA complexes are listed in

Table 2. Also listed are literature values of the anisotropies of the free dyes (r ≈ 0.24 for Cy3 [28 42 43], independent of temperature in the range examined (5 – 25 °C [28]); r ≈ 0.17 for Cy5 [42 43 44]), and the upper limit for anisotropies (r = 0.4 [31]), applicable to situations in which the dye is immobile on the time scale of its fluorescence lifetime (~2 ns [28 31]). If these dyes were completely rigidly attached to DNA, and only rotated on the time scale of the DNA or RNAP-DNA complex, their anisotropy is predicted to be 0.4 [31]. Anisotropies in between those of the free dyes and the immobilized limit indicate the extent to which the dye can rotate independently of the DNA to which it is attached.

Table 2 shows that, as expected, covalent attachment to the DNA backbone increases the anisotropy of both Cy3 and Cy5 as compared to the free dyes. Anisotropies of both Cy3 and Cy5 increase in progressing from the free dye to the ss dye-labeled primer (26 or 40 bases) to the 114 bp promoter duplex with the dye at an end. Anisotropies of both dyes on both DNAs are somewhat larger at 2°C than at 19°C. Since anisotropies of the free Cy3 have been shown to be independent of temperature [28], these small differences may result from increased stacking of these dyes on neighboring nucleobases at 2°C than that at 19°C. The larger anisotropies of Cy3 and Cy5 on ss primers and ds promoter DNA than that for the free dyes result from restrictions on rotation of these dyes by incorporation into the backbone of a ss region of DNA and from stacking interactions especially with the end of the duplex in promoter DNA.

For Cy3 located at +14, the anisotropy determined for OC (19°C; 0.36±0.01) probably exceeds that of CC (2°C; 0.34 ± 0.02), and both definitely exceed the anisotropy of Cy3 +14 on unbound duplex promoter DNA (0.31 ± 0.01 at 19 °C; 0.32 ± 0.01 at 2 °C). These anisotropy increases are consistent with footprinting results that indicate that +14 position of promoter DNA is contacted by RNAP in CC and more strongly contacted in OC. We conclude that the Cy3+14 is less mobile in OC and CC than in free promoter DNA, but nonetheless is capable of some

rotational reorientation independent of the complex, since these anisotropies are significantly less than 0.4.

For Cy5 located at -100, the anisotropy determined for OC (19˚C; 0.31 ± 0.03) is the same or slightly less than that of CC (2˚C; 0.33 ± 0.03), and neither is significantly different from the anisotropy of Cy5 -100 on unbound duplex promoter DNA (0.31±0.03 at 19˚C; 0.32±0.02 at 2˚C). These similar anisotropies at -100 for DNA in CC and OC consistent with footprinting results that indicate that -100 position of promoter DNA is not in contact with RNAP in either CC or OC. We conclude that Cy5-100 is no less mobile in OC and CC than in free promoter DNA.

Enhanced fluorescence (PIFE) is Observed at +14 in the Advanced CC and Stable OC.

Equilibrium studies with a single dye (Cy3 or Cy5) at position +14 of λP_R promoter DNA were performed to study the interactions of this region of the downstream duplex with RNAP in advanced CC and OC. Cy3 samples were excited at 515 nm and Cy5 samples at 610 nm. Representative emission spectra for 100 nM Cy5(+14) λP_R in complex with RNAP are shown in Figure 5. With Cy5 at +14, fluorescence emission intensity at λ_{max}^{em} increases with increasing [RNAP]: [DNA] ratio up to a 1:1 ratio in both CC (2°C, Figure 5A) and OC (19°C, Figure 5B). The fluorescence increase upon RNAP binding observed here (often called PIFE; protein induced fluorescence enhancement) is from proximity of the dye to the protein [20 21 45 46], and is consistent with DNA footprinting evidence [7] that the +14 region of DNA interacts with RNAP in both the advanced CC and OC [7]. PIFE effects at +14 (~30% increase in Cy5 fluorescence intensity and ~20% increase in Cy3 fluorescence intensity for complexes as compared to free DNA) are similar in CC and OC (Supplemental table 4). At -100, no systematic PIFE effects are observed for either Cy3(-100) or Cy5(-100) at either 2°C or 19°C, consistent with the observed upstream boundaries of HO· footprints that are 20 (CC) to 40 (OC) bp removed from the -100 dye position (Supplemental figures 6,7).

Strong Cy3(+14) PIFE effects mask its FRET donor effects in experiments with Cy3(+14)Cy5(-100) λP_R, as shown in the Cy3 donor emission spectra centered between 560-580 nm in Supplemental figure 2. Enhanced fluorescence of the Cy3 donor is observed at all [RNAP] : [promoter DNA] ratios ≤1:1, relative to a single-dye Cy3 control in the absence of RNAP. This indicates dominance of the PIFE effect over the donor quenching FRET effect when Cy3 is at +14.

Dye-labeled RNAP-λP_R promoter OC are transcriptionally active and permanganate reactive in the open region

Transcription assays were performed to determine whether OC formed by RNAP on the cyanine dye-labeled promoter DNAs are capable of specific initiation. Figure 6 demonstrates that all dye-labeled λP_R promoter fragments initiate transcription and yield significant amounts of the expected-length run-off transcript at 19°C. The observed differences in transcript length result from differences in the length of duplex DNA downstream of the transcription start site at +1. For promoter constructs with a cyanine dye at +14, the major transcript is 13-14 bases in length, indicated by the black boxes in Figure 6. For promoters with no dye at +14 (unlabeled λP_R and λP_R with a cyanine dye only at -100), the downstream duplex extends to +70, and the longest transcript is observed to be 70 bases long (also indicated by black boxes in Figure 6).

We also compared MnO_4^- footprints of dye-labeled and unlabeled (control) RNAP-promoter OC at 19°C to test whether these are equally open and have the same pattern of reactive thymine bases in the open region. Non-template strands of these promoters were dye labeled at the upstream end (-100) and template strands of both promoters were ^{32}P-labeled at the downstream end (see Methods). Transcription experiments (Figure 6) reveal no systematic differences between unlabeled and Cy3(-100) or Cy5(-100) labeled constructs. Supplemental Figure 9A compares template strand MnO_4^- footprints of RNAP complexes with unlabeled and

Cy5(-100)-labeled λP_R promoter DNA. Thymine bases at the expected positions (-11, -9, -8, +1) were detected in both the Cy5(-100) and control λP_R – RNAP complexes, demonstrating that the dye does not perturb the open region. No such bands are seen in the –RNAP control lanes for both constructs. Intensities of the thymine bands at -11, -9, -8 and +1 were quantified and normalized by the intensity of the full-length (unreacted) DNA band to correct for the loading difference between lanes. After correction, no significant difference in permanganate reactivity between dye-labeled and control λP_R OC is observed for any of these four thymines. By contrast, MnO_4^- experiments performed at 0°C on CC formed with dye-free λP_R DNA and Cy3(-100) λP_R DNA exhibit no reactivity above background (Supplemental Figure 9B).

Discussion

FRET between Far-Upstream and Downstream Dyes Indicates Extensive Bending/ Wrapping of Promoter DNA on RNAP in Both Closed and Open Promoter Complexes

The observation of FRET between dyes at -100 and +14 demonstrates that the through-space distance between these dyes is less than 100 Å in both CC and OC. Since the dye-dye distance (-100 to +14) in linear DNA is almost 400 Å, both upstream and downstream regions of promoter DNA must be highly bent and wrapped by *E.coli* RNAP in both complexes. For the λP_R OC, this conclusion is consistent with that derived from AFM compaction measurements, DNaseI and $HO^{\cdot}$ footprinting, and other assays at both λP_R and other promoters, summarized below. For the advanced λP_R CC, this conclusion is consistent with HO footprinting data and the model for this advanced CC proposed by [7].

Based on the kinetic and footprinting results summarized in Introduction, we [2 7] proposed that upstream wrapping occurs as the RNAP molecular machine converts the earliest CC to an advanced CC, and that this upstream wrapping causes conformational changes in RNAP that make it possible to bend the downstream duplex at -11 into the active site cleft, nucleating opening

of -11 to +2 bp by RNAP. The significance of wrapping for the properties of OC, including its stability and ability to initiate efficiently upon addition of NTP, is not yet known.

+14 PIFE Effect at 2 ˚C *Indicates CC is Advanced, with Downstream DNA Bent into RNAP Cleft*

Enhanced fluorescence of DNA-attached cyanine dyes was originally observed upon binding of SSB, BamHI, RecA, RadH1 to neighboring sites [18 19 22]. Myong *et al* designated this effect PIFE and showed that it is detectable at a dye-protein distance of 34 Å and increases with decreasing dye-protein distance [19]. PIFE between RNAP and a cyanine dye at +2 on promoter DNA has recently been used to study initiation and backtracking events upon NTP addition to OC [20 21], OC dissociation kinetics [46], and effects of the CarD regulatory protein on OC formation [45]. PIFE (and also FRET) have also been used to investigate histone-transcription factor binding [47]. Here the observation of +14 PIFE in the 2 ˚C CC corroborates information from footprinting [39] and FRET that this CC is advanced, with the downstream duplex bent into the active site cleft as proposed originally by [9]. These +14 PIFE effects in CC may result from interactions of this region of the downstream duplex with downstream mobile elements (DME) of RNAP including the lineage specific insert β i4 (also called Sequence Insertion 1: SI1), the β' clamp, and/or the β' jaw (See Figure 7A and B below) [2 4 38]. In early CC at other promoters (RP_c), where the promoter DNA is linear (not yet bent), the downstream boundary of the $HO^{\cdot}$ or DNaseI footprint is at -5. Hence in RP_c the distance between RNAP and the +14 position of promoter DNA greatly exceeds 34 Å and no PIFE would be expected. The +14 PIFE effect observed for the stable OC is expected based on footprinting and RNAP deletion variant data that indicate that this position is contacted by the β' jaw and possibly other DME [2 4 39].

Through-Space Distances between Far-Upstream and Downstream Regions of Promoter DNA Bent and Wrapped on RNAP in Closed and Open Complexes

Here we interpret FRET efficiencies and anisotropies for CC and OC to estimate upper and lower bounds on interdye distances in bent/wrapped CC and OC and compare with predictions from AFM compaction data and, in a subsequent section, from structural models built with input from DNA footprinting data.

OC and CC anisotropy data (Table 2) demonstrate that Cy3 at +14 is relatively immobile while Cy5 at -100 is more mobile. For the situation where one dye is immobile and the other is free to rotate independently, a fourfold range of values of κ^2 is possible, centered at κ^2 =2/3, with κ^2_{min} = 1/3 and κ^2_{max} = 4/3. This four-fold range of possible values of κ^2 translates into only an additional ± 26% uncertainty in the distance R between -100 and +14 positions because R varies with $(\kappa^2)^{1/6}$. Combined with the experimental uncertainty in the FRET efficiency, this leads to an uncertainty of ± 36% in dye-dye distances for both CC and OC. However, while the anisotropy of Cy5(-100) is about the same within uncertainty in the RNAP-promoter CC and OC as in free DNA, it exceeds that of the free DNA and the ss primer DNA (Table 2), presumably because cyanine dyes at the end of a duplex (like Cy5(-100)) are partially stacked on the duplex and not completely free to reorient [28,51]. Application of Eqs 6 and 7 to obtain minimum and maximum values of κ^2, Ro and R from the anisotropy results increases the combined uncertainty in determination of R for CC and OC from the measured E and r values to ± 60%, in a range centered on the R values for these complexes calculated for κ^2 = 2/3.

Structural Models For CC and OC Consistent with FRET and Footprinting Data

To interpret the strong FRET efficiencies of closed and open RNAP-promoter complexes and obtain possible structures for bent and wrapped DNA on RNAP, we model upstream (and, for CC, downstream) DNA onto recent crystal structures of *E. coli* RNAP holoenzyme [36] (for CC) and a ternary initiation complex with an open region from -11 to +3 and a downstream duplex

extending to +13 [38] (for OC), using HO· footprinting data for promoter DNA in advanced CC and stable OC as a guide.

Advanced Closed Complex: Our modeling of DNA bending and wrapping in the advanced CC updates and extends the previously-published model, which was based on the structure of *T. thermophilus* RNAP [7]. In that CC model, the promoter DNA is wrapped in an approximate U-turn around RNAP, with an overall bend angle of > 200°, resulting in upstream contacts between RNAP and promoter DNA to -81/-82 as in the HO· footprint. The distance R between upstream (-100) and downstream (+14) positions in this model is ~85 Å, near the upper end of the range consistent with the FRET efficiency reported here (Table 1). Replacement of *T.th* RNAP by *E.coli* RNAP has little effect on upstream contacts or on the distance R between -100 and +14. Slight bending of the upstream DNA at -65 in this CC model creates a series of wrapped complexes with smaller distances R but similar upstream boundaries near -82 (see Supplemental Figure 10). For example an ~8° bend at -65 reduces R to ~ 50 Å, the minimum distance obtainable by this approach without introducing steric clash. A ~5° bend at -65 yields the CC structure of Figure 7A with R = 68 Å, the mid-range distance deduced from the FRET efficiency and isotropy of CC.

Stable Open Complex: Though the FRET efficiencies for energy transfer between -100 and +14 are similar for the advanced CC and OC λP_R -RNAP complexes, the structures of these complexes must differ significantly in both upstream and downstream regions. Upstream DNA in both the advanced CC and the stable OC is predicted to be strongly bent between -35 and -55 ([7], [27]), indicating the trajectory of the upstream wrap begins similarly in both complexes, but they exhibit very different upstream boundaries of OH protection (-81/-82 in the advanced CC; -60/-65 in the stable OC) [7]. Hence the trajectory of the far-upstream DNA (between -60/-65 and -81/-82), and/or the positions of elements of RNAP including β' residues 808-912 (previously designated the UP clamp [1,4]) that appear to make contact with this far-upstream DNA, must shift in the conversion of the advanced CC to the stable OC.

Likewise the trajectories of the downstream DNA in the advanced CC and stable OC appear different, although both have a ~90° bend at the upstream end of the -10 region and the downstream boundaries of HO protection for both the advanced CC and the stable OC at LPR are the same (+20 [7]). The DNA downstream of the bend at -11 in the advanced CC is thought to be linear, while additional strong bending in the open region observed in the recent structure of an initiation complex [41] makes the trajectory of the downstream duplex quite different from that of linear downstream DNA in the model of the advanced CC. The +14 region of downstream DNA is in close proximity to (and possibly interacting with) β-SI1 (residues 227-308) in advanced CC (Figure 7A), but is further away from it in OC, indicating rearrangement of either this region of RNAP or of the downstream DNA in the conversion of CC to OC. The β' jaw (residues 1150-1208) is also in closer proximity to DNA downstream of +10 in OC (Figure 7B), supporting earlier evidence of stabilizing interactions between the β' jaw and this region of downstream DNA in OC [4].

Starting from the structure of the initiation complex [23], we modeled in the upstream DNA as described in Methods to create upstream-wrapped structures with distances R between -100 and +14 in the range 40 Å to 80 Å as specified by the FRET efficiency, in which the upstream DNA contacts RNAP only to the upstream end of the UP element at -60/-65 to match the upstream HO footprint boundary, and with the extreme wrapping deduced from DNA compaction by AFM [11]. If upstream DNA (from -37 to -99) from the CC model (Fig 7A) is modeled into the structure of the initiation complex without any additional bending, we obtain an OC model with R > 100 Å and in which the tethers of the αCTD must be somewhat extended to make contacts with the UP-element DNA.

This OC model does not have sufficient upstream bending and wrapping to explain the OC FRET efficiency. To reduce R, bends in upstream duplex DNA at positions of DNaseI or HO enhancements [7 39] were increased by small amounts (~ 3°) at -38 and -44 and substantially (~

25^{o}) at -54, relative to the path of the DNA in the original CC model. This OC structure (with R= 60 Å) is shown in Figure 7B. More extensive bending at these or other positions would force contacts between upstream DNA and RNAP beyond -60/-65, not observed in the OC HO footprint [7]. Insertion of a ~ 25^{o} bend at -54 changes the trajectory of upstream DNA from that of the CC model (in which it is near the α-NTD and the β' groove) shifting the upstream DNA toward the ω subunit. Upstream DNA beyond -65 passes near the β' subunit without a direct interaction, consistent with the end point of the HO footprint at -65 on the non-template strand [7]. The upstream(-100) – downstream (+14) distance R in this structure is 60 Å, which is also the mid-range distance from FRET/anisotropy data. Smaller values of R are not possible in this model because of steric interactions, unless there are large conformational differences in the region of RNAP contacting the upstream DNA between the stable OC and the initiation complex. While phasing and positioning of upstream DNA with respect to RNAP in this model are approximate and traced by eye, we think it provides a good estimate of the most probable distance in OC between -100 and +14.

Possible Roles of Promoter DNA Wrapping and Unwrapping in OC Formation and Transcription Initiation

Truncations of far-upstream promoter DNA (between -65 and -100 [6]), and within the UP element (-38 to -65 [6 8]) both dramatically reduce the rate of isomerization of closed to open complexes at λP_R and lacUV5 promoters. While the real-time footprint of the ensemble of advanced closed complexes at λP_R extends from -82 to +20, truncation at -47 results in a less advanced closed complex, with a downstream boundary in real-time footprinting of only +2/+7 that indicates that the downstream duplex is only partially bent into the active site cleft. From this information, we proposed that upstream bending and wrapping, set in motion by initial binding to the -35 region and UP element to form a linear complex (RP_c), is necessary to bend the downstream duplex fully into the cleft in the advanced CC prior to opening. Presumably this is

because upstream bending and wrapping plays key roles in opening the cleft and moving obstacles like sigma 1.1 and DME of core RNAP to allow bending of the downstream duplex at -11 and its entry into the active site cleft before opening [4 5 9 48].

The significance of wrapping for the properties of OC is unknown. AFM studies with FL λP_R indicate similar degrees of wrapping in OC and in an initiation complex paused at +3 [13]. Since elongation complexes (after escape from the promoter) exhibit a short (40 bp) footprint [10] and are therefore unwrapped, unwrapping of the upstream DNA may therefore be a significant element of the escape transition from initiation to elongation that occurs after formation of a moderately long transcript (~10 bp RNA-DNA hybrid) which necessitates and may be driven by scrunching of promoter DNA in the cleft [49 50]. Release of upstream wrapping presumably occurs together with release of σ-promoter DNA contacts.

Comparison of Promoter DNA Wrapping in CC Formed by Prokaryotic and Eukaryotic RNAP

Wrapping of upstream DNA and bending of downstream DNA appear to be general features of preinitiation complexes for eukaryotic as well as bacterial RNAP. After binding the TATA box, human and yeast TATA binding proteins (TBP) and transcription factor TFIIB together bend this sequence of duplex DNA by 80 – 100° [51 52] in a manner reminiscent of the bending of the upstream region from around -35 to -100 by *E.coli* RNAP to form an advanced CC [1 39]. Assisted by the essential factor TFIIF (required for RNA pol II binding to promoter DNA) and TFIIE, another ~90° bend is formed near the transcription initiation site (+1), which may be analogous to the downstream bending of DNA at -11 into active site cleft in *E.coli* RNAP in advanced CC [1] that also has been implicated in nucleation of DNA opening [46]. Extensive protein-DNA crosslinking is observed over a 90 bp region of promoter DNA (-56 to +34 [17]) in an advanced closed complex formed with the proteins TBP, TFIIB, RNA pol II, TFIIF, and TFIIE [17], indicating extensive DNA wrapping analogous to the ~100 bp of DNA wrapped by *E.coli* RNAP in the

advanced CC (from -80/-81 to +20 [7]). We hypothesize that transcriptional factors such as TFIIE, TFIIF may play similar roles in wrapping promoter DNA in the advanced CC to those of the β lobe (including SI1) and the β' upclamp *for E. coli* RNAP. Kinetic studies are underway in our laboratory to determine the progression of upstream DNA wrapping and downstream DNA bending as the initial closed complex (RP_c) advances to bend the downstream duplex into the active site cleft and open it.

Acknowledgements

This research was supported by NIH GM103061 (MTR) and GM 67153 (IA). Fluorescence data were obtained at the University of Wisconsin - Madison Biophysics Instrumentation Facility, which was established with support from the University of Wisconsin - Madison and grants BIR-9512577 (NSF) and S10RR13790 (NIH).

Table 1: FRET Efficiencies and Distances R between Promoter Positions -100 and +14 in Advanced Closed and Stable Open Complexes

Property	Advanced Closed Complex (CC; 2°C)	Stable Open Complex (OC; 19°C)
FRET Efficiency E [a]		
Cy3(-100) Cy5(+14) λP_R	0.36 ± 0.12	0.37 ± 0.13
Cy3(+14) Cy5(-100) λP_R	0.23 ± 0.06	0.32 ± 0.11
Average	0.30 ± 0.11	0.34 ± 0.11
Estimates of Interdye Distance R from FRET Efficiency E		
Random orientation limit ($\kappa^2 = 2/3$)	$R_{2/3}$ = 69 ± 8 Å	$R_{2/3}$ = 62 ± 6 Å
Anisotropy maximum ($\kappa^2_{ub} = 3.59$)	R_{max} = 90 ± 9 Å	R_{max} = 82 ± 8 Å
Anisotropy minimum ($\kappa^2_{lb} = 0.052$)	R_{min} = 47 ± 6 Å	R_{min} = 41 ± 4 Å
Mid-range Value from Anisotropy	R_{mid} = 68 ± 22 Å	R_{mid} = 60 ± 20 Å
Structural minimum value (from models in Figure 7)	$R_{min\text{-}str}$ ~ 50 Å	$R_{min\text{-}str}$ ~ 60 Å

[a]FRET efficiencies E are averages of results in **Figure 1, 2** and **Supplemental Figure 1, 2**, corrected for occupancy and excluding results for [RNAP]:[DNA] ≥ 1:1. E value uncertainties are ±1 SD. R value uncertainties are deviations from the average.

Table 2: Fluorescence Anisotropies (r) of Cyanine Dyes, Dye-Labeled DNAs and RNAP-Promoter Complexes

Sample	2°C	19°C
Dye immobilized on DNA [a]	0.40	0.40
Cy3(+14) λP_R DNA-RNAP Complex	(CC) 0.34 ± 0.02	(OC) 0.36±0.01
Cy3(+14) Duplex λP_R DNA	0.32 ± 0.01	0.31±0.01
Single-stranded Cy3 (+14) Primer	0.28 ± 0.01	0.27±0.01
Cy3 free dye	0.24±0.02 [b,c]	0.24±0.02 [b,c]
Cy5(-100) λP_R DNA-RNAP Complex	(CC) 0.33 ± 0.03	(OC) 0.31±0.03
Cy5(-100) Duplex λP_R DNA	0.32 ± 0.02	0.31±0.03
Single-stranded Cy5(-100) Primer	0.29 ± 0.02	0.25±0.03
Cy5 free dye	0.17±0.04[b,c]	0.17±0.04[b]

[a] Calculated from estimates of Cy3, Cy5 fluorescence lifetimes ($\tau \sim 2$ ns) and rotational diffusion coefficient of RNAP-DNA CC, OC ($D_r \sim 1.4\ \mu s^{-1}$) from the relationship $r = r_o/(1 + D_r\tau) \approx r_o = 0.4$ [33]

[b] Average of values reported at 20-25 °C for Cy3 [28,45,46] and Cy5 [45–47].

[c] Fluorescence of free Cy3 dye independent of T in range investigated (5-25°C) [28]; temperature-independent fluorescence assumed for Cy5.

Figure Captions

Figure 1. **Cy3 to Cy5 Fluorescence energy transfer (FRET) demonstrating wrapping in CC.** Cy5 fluorescence emission spectra (620-720 nm; counts/s) obtained by excitation of Cy3 at 515 nm as a function of [RNAP]:[DNA] ratio (0 – 1.2:1; 100 nM DNA) for closed (CC; 2°C) RNAP - λP_R promoter DNA complexes. (A) Cy3(-100)-Cy5(+14) λP_R (B) Cy3(+14)-Cy5(-100) λP_R.
DNA spectra (-RNAP) were obtained with 100 nM Cy3-labeled DNA.

Figure 2. Cy3 to Cy5 Fluorescence energy transfer (FRET) demonstrating wrapping in OC. Cy5 fluorescence emission spectra (620-720 nm; counts/s) obtained by excitation of Cy3 at 515 nm as a function of [RNAP]:[DNA] ratio (0 – 1.2:1; 100 nM DNA) for closed (OC; 19°C) RNAP - λP_R promoter DNA complexes. (A) Cy3(-100)-Cy5(+14) λP_R DNA (B) Cy3(+14)-Cy5(-100) λP_R.
DNA spectra (-RNAP) were obtained with 100 nM Cy3-labeled DNA.

Figure 3. **Non – Promoter DNA Exhibits no FRET.**

Comparison of Cy5 FRET acceptor emission spectra (620-720 nm; counts/s) of 1:1 mixtures of RNAP (100 nM) and either Cy3(+14)Cy5(-100) λP_R DNA (100 nM) or Cy3(+14)Cy5(-100) nonpromoter DNA control (100 nM) at (A) 2°C and (B) 19°C obtained by excitation of Cy3 at 515 nm. Spectra of nonpromoter DNA control (black) were normalized to those of the λP_R promoter DNA (red) over the Cy3 emission range 555-570 nm to observe and compare emission at Cy5 λ_{max}^{em} (660 nm).

Figure 4. Intensity ratios (I_{cy5}/I_{cy3}) for Cy5 and Cy3 emission at their wavelength maxima (λ_{max}^{em}) upon excitation of Cy3 at 515 nm, demonstrating FRET in RNAP titrations of 100 nM dye labeled λP_R DNA. Data for both CC (2°C; red) and OC (19°C; black) formed with (A) Cy3(-100)-Cy5(+14) and (B) Cy3(+14)-Cy5(-100) are shown. λ_{max}^{em} = 660nm (Cy5); λ_{max}^{em} = 565nm (Cy3). Data are from **Figures** 1,2. Lines interpret the FRET effect as a linear increase up to [RNAP]:[promoter DNA] = 0.8:1 and a plateau at higher [RNAP]:[promoter DNA]

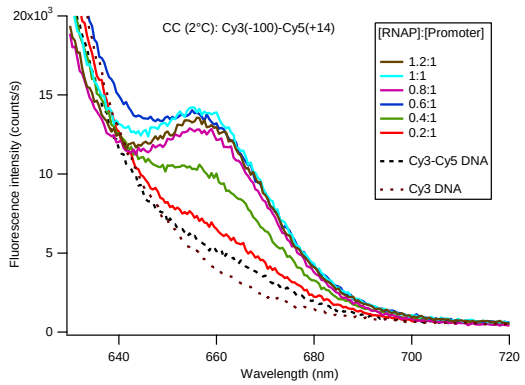
Figure 5. Observed PIFE in emission spectra of 100 nM Cy5(+14) λP_R promoter DNA titrated with RNAP (0, 50, 75, 100 nM) for CC (2°C; panel A) and OC (19°C; panel B). Samples were excited at 610 nm.

Figure 6. Gel analysis of transcription products comparing the various double-dye (left center) and single-dye (center, right) constructs with unlabeled λP_R promoter (right lane). A RNA ladder is shown at left. Boxed areas show full-length transcripts for each fragment.

Figure 7. Models of bent and wrapped promoter DNA in (A) CC and (B) OC with *E. coli* RNAP, built for midrange values of the dye-dye distances R = 68 Å for CC and R = 60 Å for OC as determined from FRET and anisotropy data (**Table 1**). The CC model is built from the structure of *E. coli* RNAP [39]; the OC model is built from the structure of the *E. coli* transcription initiation complex [41]. Cyanine dyes are shown at -100 (pink) and +14 (blue) ellipsoids. RNAP subunits αNTD (cyan), αCTD (yellow; modeled from PDB structure - 1LB2), β (pale pink), β' (green), ω (blue), σ (orange) are shown. β Sequence Insertion-1 (dark pink; residues 225-343), β' clamp (dark green; residues 8-347 and 1260-1368); β' upstream (UP) clamp (bright green; residues 808-912) and β' jaw (brown; residues 1150-1208) are highlighted. Promoter DNA is shown in light grey (NT) and dark grey (T). A "backside" view of OC is also shown in (B).

Figure 1

A.


20x10^3
CC (2°C): Cy3(-100)-Cy5(+14)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
Fluorescence intensity (counts/s)
Wavelength (nm)
0
5
10
15
640
660
680
700
720

B.

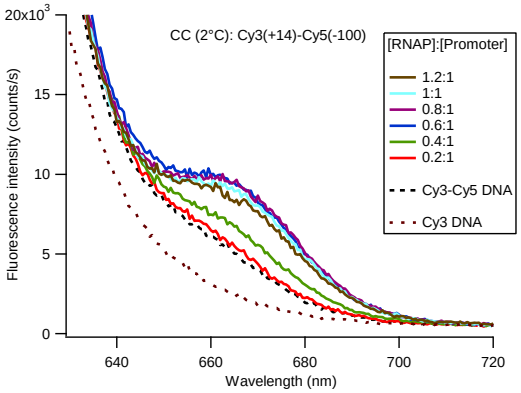
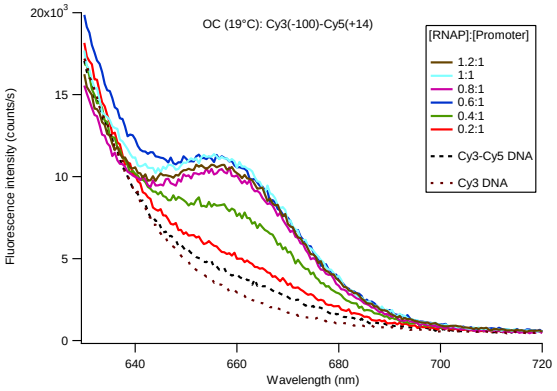

20x10^3
CC (2°C): Cy3(+14)-Cy5(-100)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
Fluorescence intensity (counts/s)
Wavelength (nm)
0
5
10
15
640
660
680
700
720

Figure 2

A.


OC (19°C): Cy3(-100)-Cy5(+14)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
20x10^3
15
10
5
0
Fluorescence intensity (counts/s)
640
660
680
700
720
Wavelength (nm)

B.

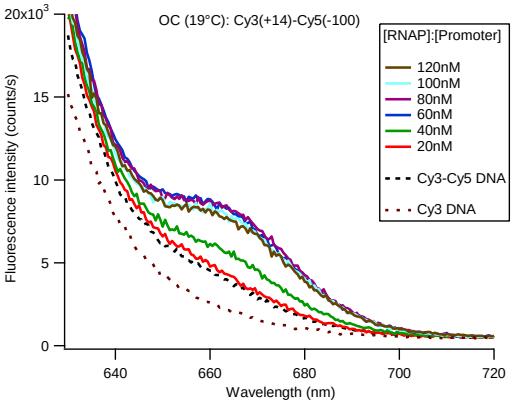
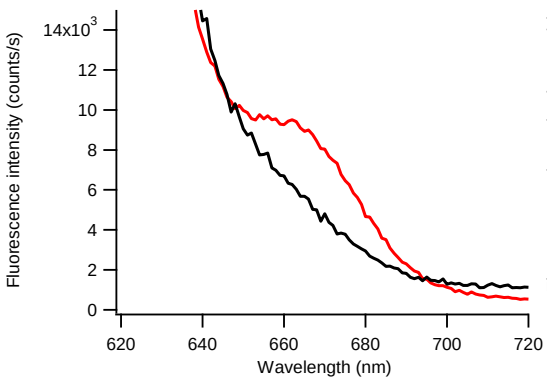

OC (19°C): Cy3(+14)-Cy5(-100)
[RNAP]:[Promoter]
120nM
100nM
80nM
60nM
40nM
20nM
Cy3-Cy5 DNA
Cy3 DNA
20x10^3
15
10
5
0
Fluorescence intensity (counts/s)
640
660
680
700
720
Wavelength (nm)

Figure 3

A.


Fluorescence intensity (counts/s)
14x10^3
12
10
8
6
4
2
0
620
640
660
680
700
720
Wavelength (nm)

B.

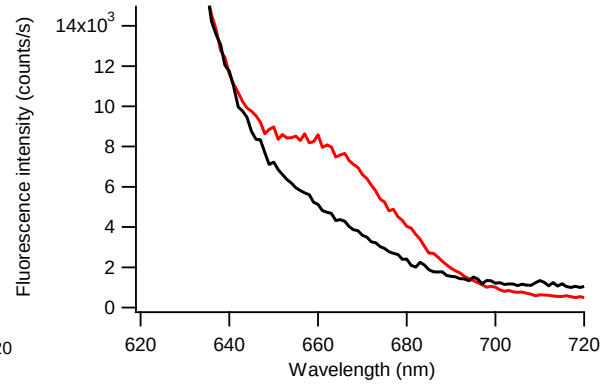
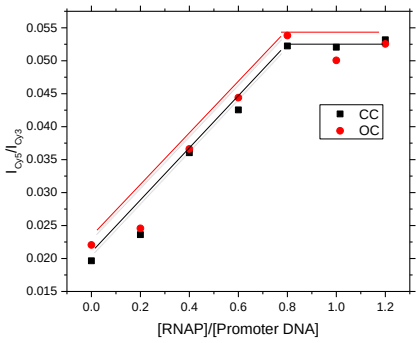

Fluorescence intensity (counts/s)
14x10^3
12
10
8
6
4
2
0
620
640
660
680
700
720
Wavelength (nm)

Figure 4

A.


0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.0
0.2
0.4
0.6
0.8
1.0
1.2
[RNAP]/[Promoter DNA]
CC
OC

B.

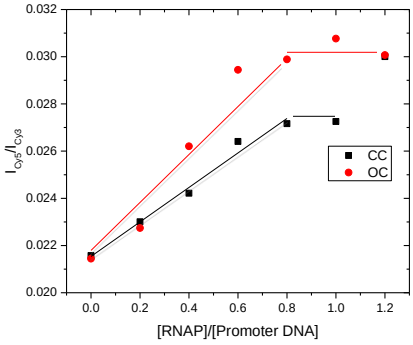
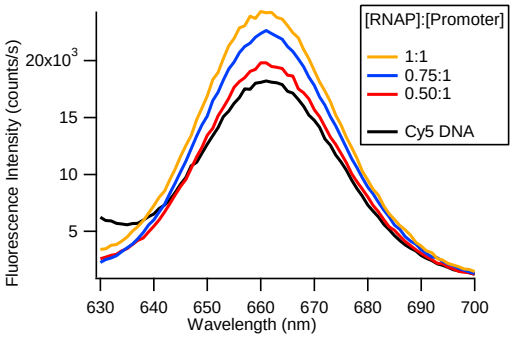

0.032
0.030
0.028
0.026
0.024
0.022
0.020
0.0
0.2
0.4
0.6
0.8
1.0
1.2
[RNAP]/[Promoter DNA]
CC
OC

Figure 5

A. CC (2°C): Cy5 (+14)



B. OC (19°C): Cy5 (+14)

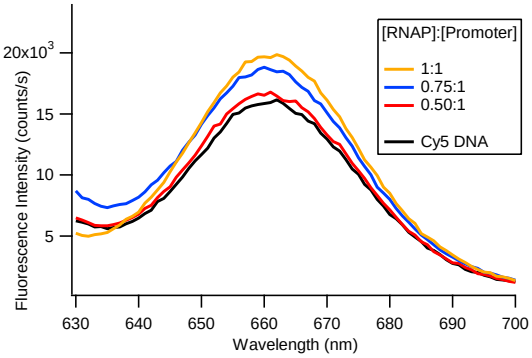


Figure 6

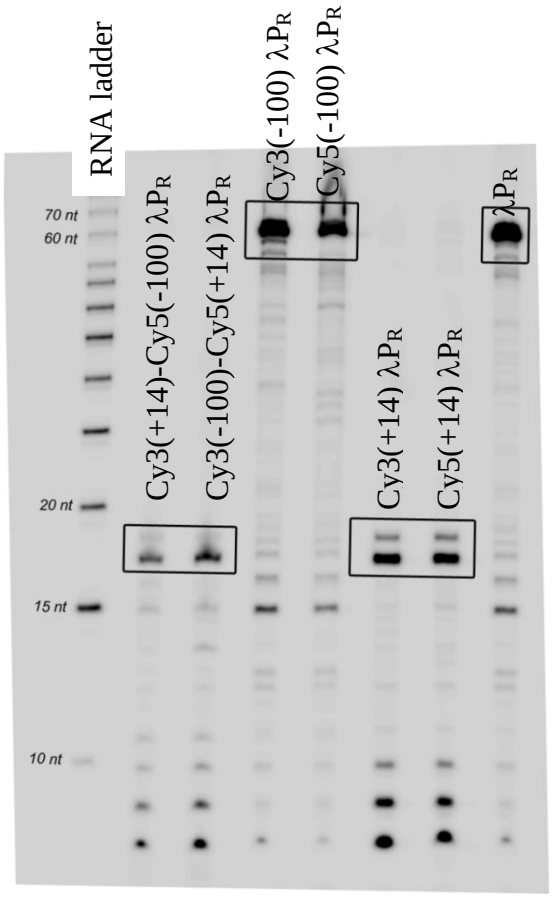
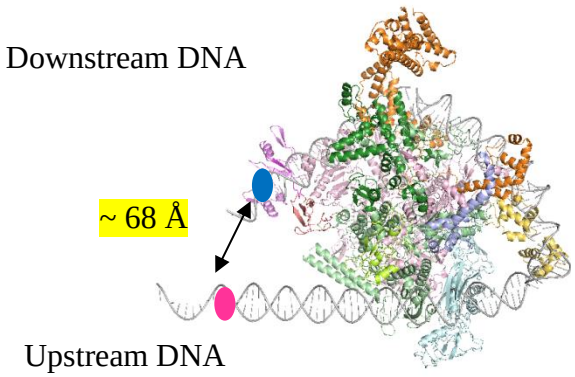
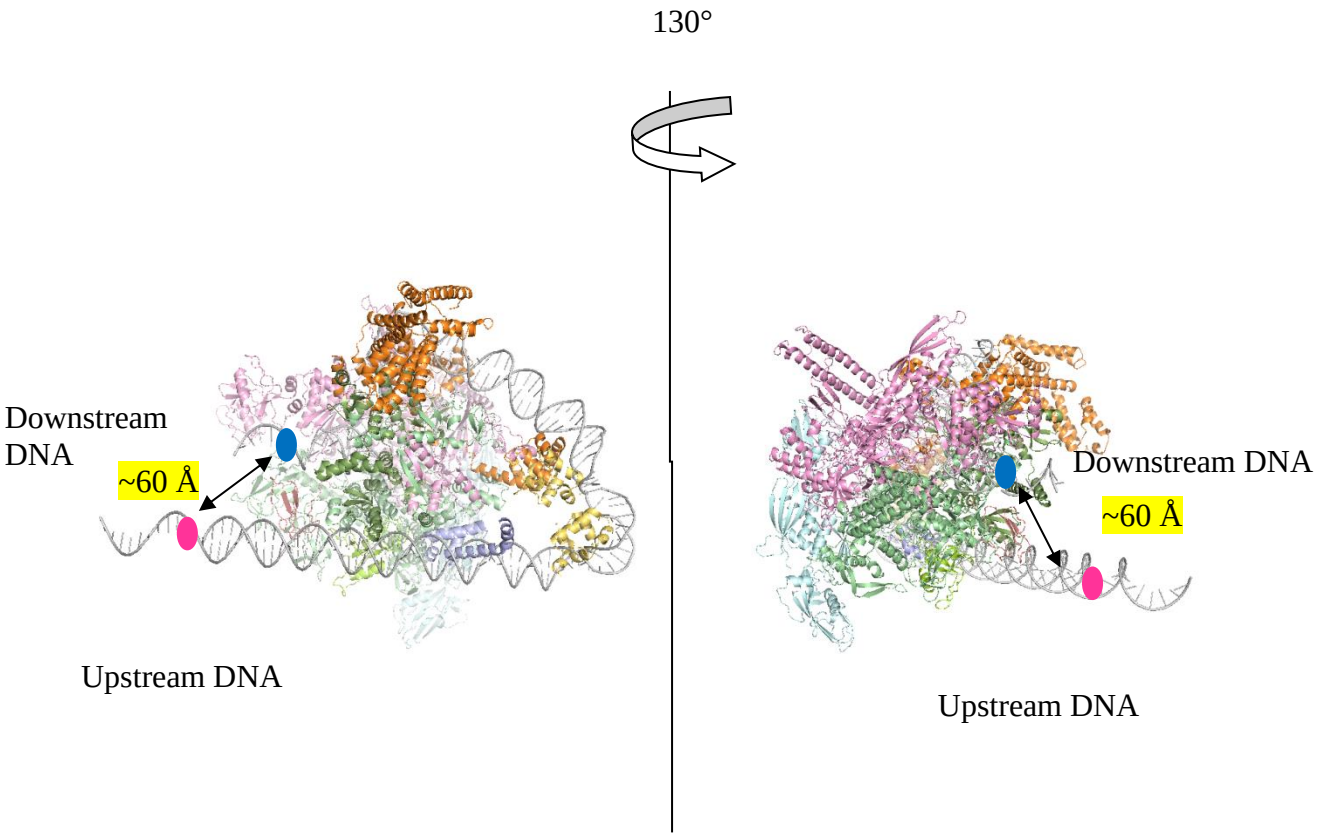

RNA ladder
Cy3(+14)-Cy5(-100) λP$_R$
Cy3(-100)-Cy5(+14) λP$_R$
Cy3(-100) λP$_R$
Cy5(-100) λP$_R$
Cy3(+14) λP$_R$
Cy5(+14) λP$_R$
λP$_R$
70 nt
60 nt
20 nt
15 nt
10 nt

Figure 7

A. Model of CC



B. Model of OC



Supplemental materials and methods

1. Table S1. DNA sequences

Promoter	Template	Upstream primer	Downstream primer	Duplex (bp) and positions relative to start site (+1)*
Unlabeled λP_R	λP_R(-59 to +34)[53]	Upstream2-4	Downstream1-5	190 (-126 to +64)
Cy3(+14) λP_R	Unlabeled λP_R or λP_R (-59 to +34)	Upstream2-4	Cy3(+14) down	140 (-126 to +14)
Cy5(+14) λP_R	Unlabeled λP_R or λP_R (-59 to +34)	Upstream2-4	Cy5(+14) down	140 (-126 to +14)
Cy3(-100) λP_R	Unlabeled λP_R or λP_R (-59 to +34)	Downstream1-5	Cy3(-100) up	164 (-100 to +64)
Cy5(-100) λP_R	Unlabeled λP_R or λP_R (-59 to +34)	Downstream1-5	Cy5(-100) up	164 (-100 to +64)
Cy3(-100)-Cy5(+14) λP_R	Unlabeled λP_R or λP_R (-59 to +34)	Cy3(-100) up	Cy5(+14) down	114 (-100 to +14)
Cy3(+14)-Cy5(-100) λP_R	Unlabeled λP_R or λP_R (-59 to +34)	Cy3(+14) down	Cy5(-100) up	114 (-100 to +14)
Cy3(+14)-Cy5(-100) DNA control	λP_R DNA control with multiple changes in UP ele-ment, -35, -10 regions	Cy3(+14) control - down	Cy5(-100) up	114 (-100 to +14)

* Dye-labeled promoter DNA constructs also have 12 base (upstream) and/or 8 base (downstream) ssDNA termini.

Sequences

Upstream2-4 primer:

5’-GTACGAATTCGATATCCAGCTATGACCATGATTACGCCAAGC-3’

Downstream1-5 primer:

5’-CAGGACCCGGGAAGCTTTTAATTAACACTCTTATACATTATTCC-3’

Cy3(+14) or Cy5(+14) down primer:

5’-CCATACAA/iCy3/CCTCCTTACTACATGCAACCATTATCACCGCC-3’

5’-CCATACAA/iCy5/CCTCCTTACTACATGCAACCATTATCACCGCC-3’

Cy3(-100) or Cy5(-100) up primer:

5’-CAGCTATGACCA/iCy3/TGATTACGCCAAGC-3’

5’-CAGCTATGACCA/iCy5/TGATTACGCCAAGC-3’

Sequence of Oligos for non-promoter DNA control

Fragment 1:

CAGCTATGACCATGATTACGCCAAGCTCGCCTGCCTACAGCTTTGCTCAGTCGATCTGTCT
CTTGTGTCTAGCCTTCATGC

Fragment 2:

CATTATTCCATACAACCTCCTTACTACATCAGGAACCGCACACCGACAGAGGTCGGATCGC
ATGAAGGCTAGACACAAGAGACAGA

Cy3 (+14) non-promoter Down primer:

5’ - CCATACAA/iCy3/CCTCCTTACTACATCAGGAACCGCACACCGAC -3’

Supplemental methods

Preparation of Cy3- and/or Cy5-labeled λP_R promoter and Cy3(+14)Cy5(-100) control DNA

Dye-labeled promoters for both labeled λP_R promoter and control constructs were prepared using internally-labeled PCR primers purchased from Integrated DNA Technologies (IDT, Coralville, IA) synthesized with Cy3 or Cy5 dye (‘iCy3’, ‘iCy5’) on the DNA backbone either 8 bases (for +14) or 12 bases (for -100) from the end. Unlabeled primers were also purchased from IDT. Both two-dye labeled promoters (Cy3(-100)Cy5(+14) λP_R, Cy3(+14)Cy5(-100) λP_R), all

four single-dye labeled promoters (Cy3(+14) λP_R, Cy5(+14) λP_R , Cy3(-100) λP_R and Cy5(-100) λP_R) and a Cy3(+14)Cy5(-100) control construct were prepared and investigated.

For template DNA for use in PCR, unlabeled λP_R promoter DNA fragments were first synthesized by PCR using Vent DNA polymerase from New England Biolabs (NEB, Ipswich, MA) from plasmid pPR59 (λP_R from -59 to +20 ; gift from Dr. Wilma Ross [53]). Unlabeled λP_R was used as template with respective dye-labeled primer oligos to synthesize dye-labeled λP_R promoter constructs labeled with cyanine dyes (Cy3,Cy5) at upstream (-100) and/or downstream (+14) positions.

DNA fragments, unlabeled or labeled with Cy3 at +14 and Cy5 at -100 were also synthesized in which -10, -35 and UP element (-38 to -65) regions of λP_R sequence were replaced with heterologous sequences as a control for steady-state FRET experiments using a similar PCR protocol. Two long oligomers incorporating these changes designated fragment 1 and 2 were purchased from IDT (Coralville, IA) that were annealed in AB and extended by Taq DNA Polymerase (NEB, Ipswich, MA). The annealed and extended construct was gel-purified. This duplex oligomer was used as template with dye-labeled primers to prepare Cy3(+14)Cy5(-100) DNA lacking λP_R core-promoter sequence, as a control. For the control Cy3(+14)Cy5(-100) construct, the upstream Cy5(-100) primer was the same as for the λP_R DNA preparation with a different downstream Cy3(+14) primer used (see supplemental) without the -10 region of λP_R. Lengths and sequences of all constructs are given in supplemental.

Sequencing showed that Vent polymerase stops synthesis at the position of the dye. Hence both two-dye labeled λP_R and control constructs are double-stranded over the entire promoter region from -100 to +14, with short single-stranded flanking regions corresponding to the 8 bases of the primer beyond the dye at +14 (t strand) and the 12 bases beyond they dye at -100 (nt strand). Single-dye-labeled λP_R promoter fragments are fully duplex at the unlabeled end.

PCR products were purified using Qiaquick PCR purification kit (Qiagen, Valencia, CA), and concentrations of labeled DNA were determined by UV spectrophotometry (Nanodrop 2000c, Thermo Scientific). Using the microarray mode in the Nanodrop that corrects for absorbance of the dye at 260nm, the efficiency of labeling was invariably found to be approximately 100%.

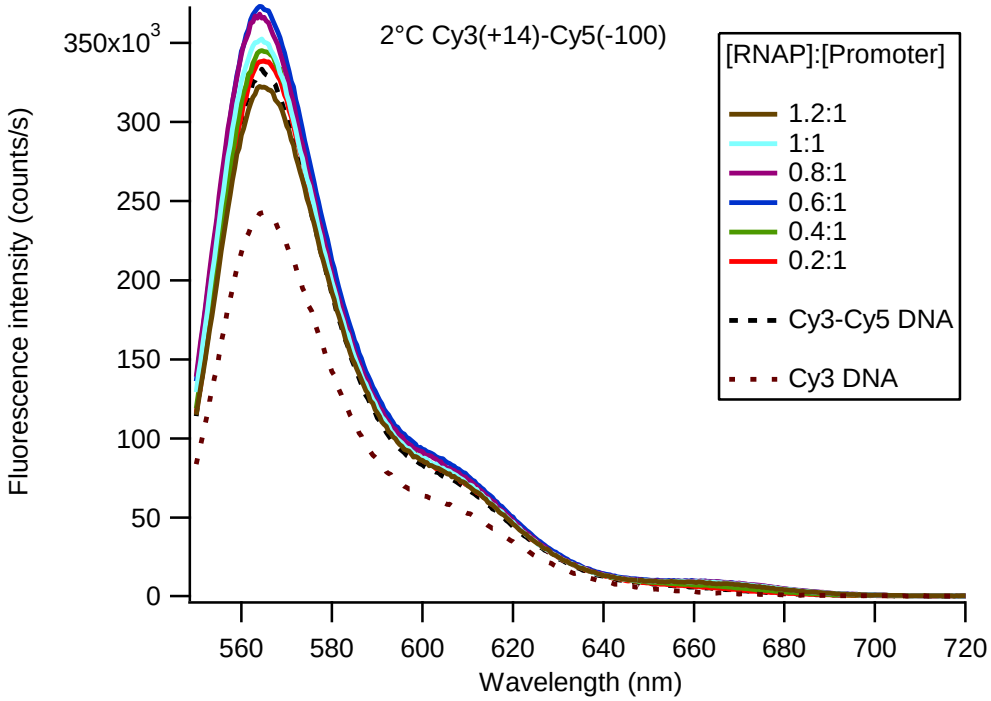
λP_R promoter and control DNA sequences were run through a bacteriophage promoter recognition software (phiSITE[54]). Putative -10, -35 elements were identified with default parameters. To compare the raw scores obtained, P_R sequence with consensus -10, -35 elements was run and a score of 10.4 was predicted. The P_R sequence with wt P_R -10 and -35 regions (above) scored 7.3; other sequences with some similarity to promoter -10, -35 regions within the P_R DNA scored < 3.3. The control DNA sequence also scored < 3.3. The score of the control DNA sequence could not be reduced without introducing potential secondary structures that would affect PCR.

Supplemental Figures

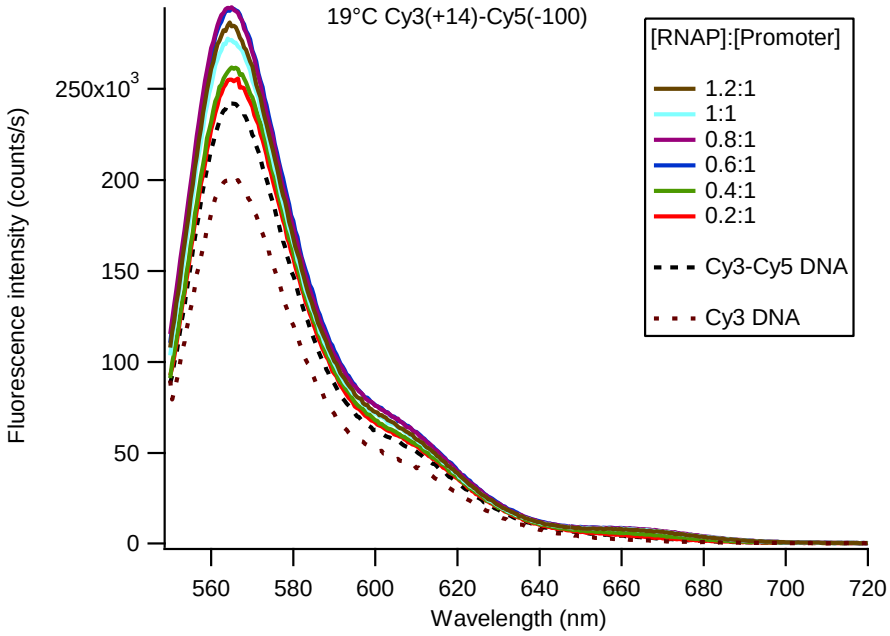
Figure S1: Fluorescence energy transfer (FRET) from Cy3 to Cy5.

Cy3 (donor) fluorescence emission spectra (550-720nm; counts/s) corresponding to the Cy5 (acceptor) spectra shown in figures 1,2, obtained by excitation of Cy3 at 515 nm as a function of [RNAP]:[DNA] ratio (0 – 1.2:1; 100 nM DNA) for both closed (CC) and open (OC) RNAP - λP_R promoter PCR-DNA complexes. Cy5 emission (λ_{max}^{em} = 660nm) is observed upon excitation of Cy3. Panels: Cy3(+14)-Cy5(-100) λP_R DNA in A) CC (2°C) and (B) OC (19°C); Cy3(-100)-Cy5(+14) λP_R DNA in (C) CC (2°C) and (D) OC (19°C). λP_R DNA spectra (-RNAP) were obtained with 100nM Cy3-labeled DNA.

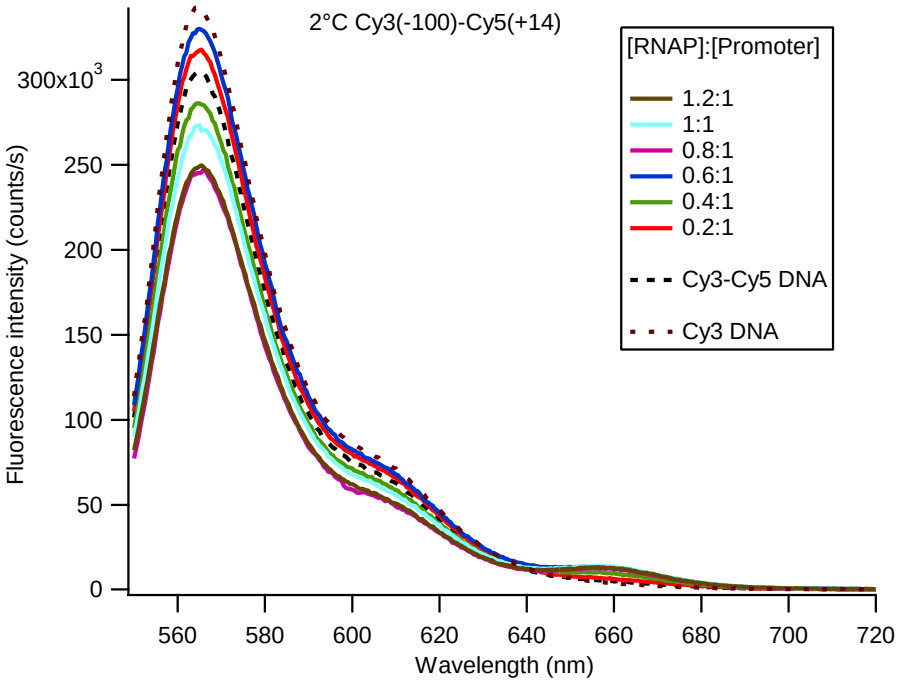
(A)


2°C Cy3(+14)-Cy5(-100)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
Fluorescence intensity (counts/s)
350x10^3
300
250
200
150
100
50
0
560
580
600
620
640
660
680
700
720
Wavelength (nm)

(B)


19°C Cy3(+14)-Cy5(-100)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
Fluorescence intensity (counts/s)
250x10^3
200
150
100
50
0
560
580
600
620
640
660
680
700
720
Wavelength (nm)

(C)


2°C Cy3(-100)-Cy5(+14)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
300x10^3
250
200
150
100
50
0
Fluorescence intensity (counts/s)
560
580
600
620
640
660
680
700
720
Wavelength (nm)

(D)

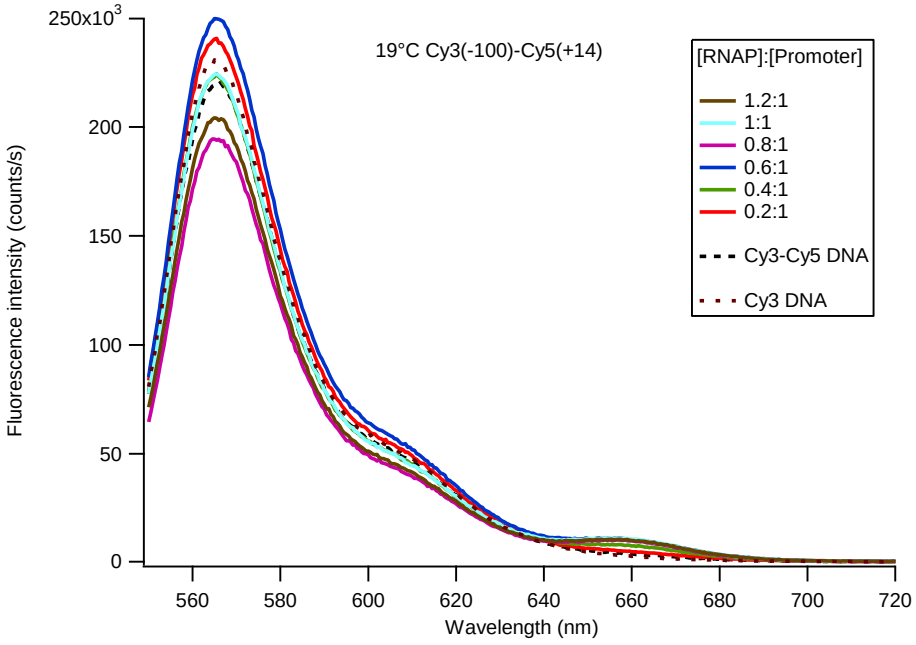
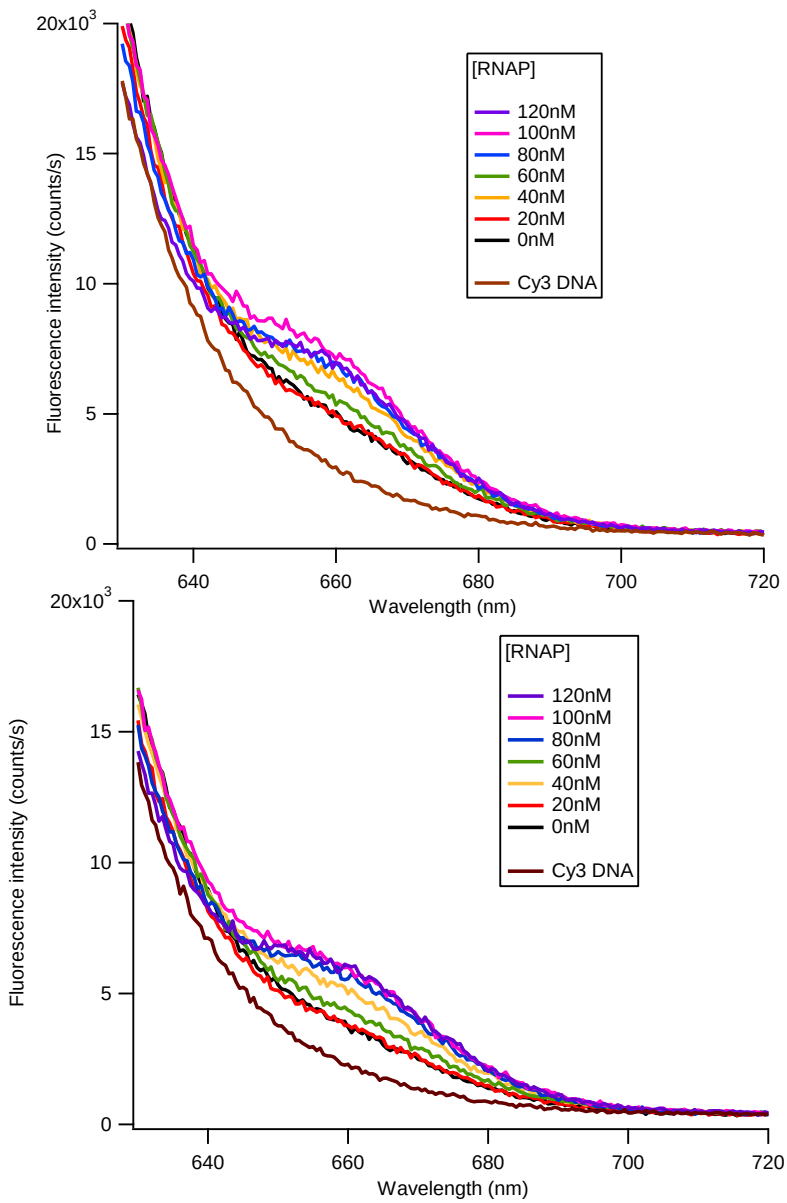

19°C Cy3(-100)-Cy5(+14)
[RNAP]:[Promoter]
1.2:1
1:1
0.8:1
0.6:1
0.4:1
0.2:1
Cy3-Cy5 DNA
Cy3 DNA
250x10^3
200
150
100
50
0
Fluorescence intensity (counts/s)
560
580
600
620
640
660
680
700
720
Wavelength (nm)

Figure S2. Fluorescence energy transfer (FRET) experiments: Cy5 emission spectra

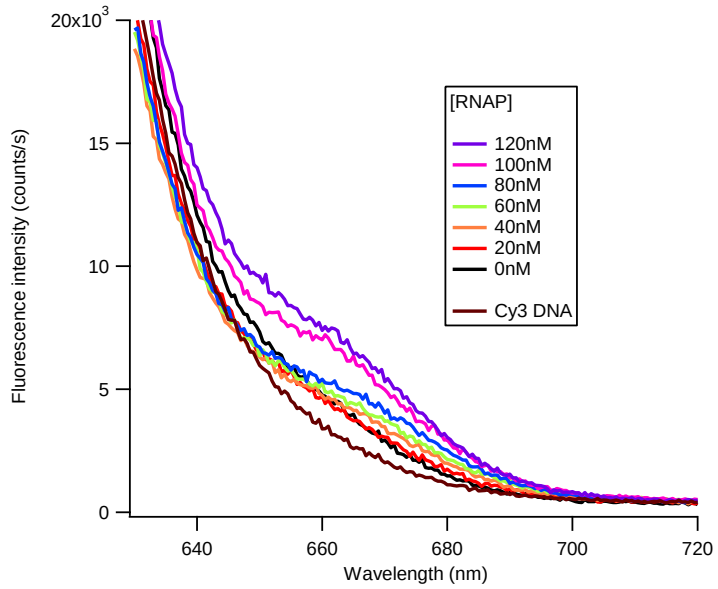
Independent reproductions of Cy5 emission spectra for Cy3-Cy5 FRET, similar to data shown in Figures 1,2. Cy5 emission spectra for Cy3-Cy5 FRET, similar to data shown in figure 2. Cy5 fluorescence emission spectra (620-720nm; counts/s; panels on the right) are obtained by excitation of Cy3 at 515 nm as a function of [RNAP]:[DNA] ratio (0 – 1.2:1; 100 nM DNA) for both closed (CC) and open (OC) RNAP - λP_R promoter-DNA complexes. Cy3(-100)-Cy5(+14) λP_R DNA in (A) CC (2°C) and (B) OC (19°C); Cy3(+14)-Cy5(-100) λP_R DNA in (C) CC (2°C) and (D) OC (19°C). λP_R DNA spectra (-RNAP) were obtained with 100nM Cy3-labeled DNA.

(A)

(B)



(C)


20x10^3
15
10
5
0
Fluorescence intensity (counts/s)
640
660
680
700
720
Wavelength (nm)
[RNAP]
120nM
100nM
80nM
60nM
40nM
20nM
0nM
Cy3 DNA

(D)

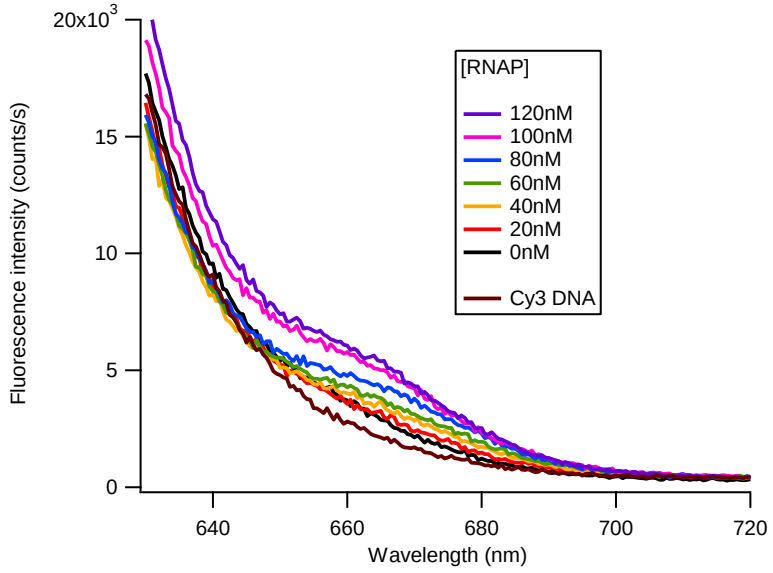

20x10^3
15
10
5
0
Fluorescence intensity (counts/s)
640
660
680
700
720
Wavelength (nm)
[RNAP]
120nM
100nM
80nM
60nM
40nM
20nM
0nM
Cy3 DNA

Figure S3: Temperature dependence of Cy3,Cy5 fluorescence emission intensity.
Temperature series plot for 100nM Cy3(-100)Cy5(+14) λP_R promoter DNA construct with 100nM RNAP shows a decrease in fluorescence emission intensity of Cy3 and Cy5 (inset plot) with increasing temperature, in parallel to the intensities observed for CCs and OCs in fig. 1,S1,S2. Samples were excited at 515nm and Cy3 emission was monitored between 550-720nm; Cy5 emission was monitored between 620nm-720nm.

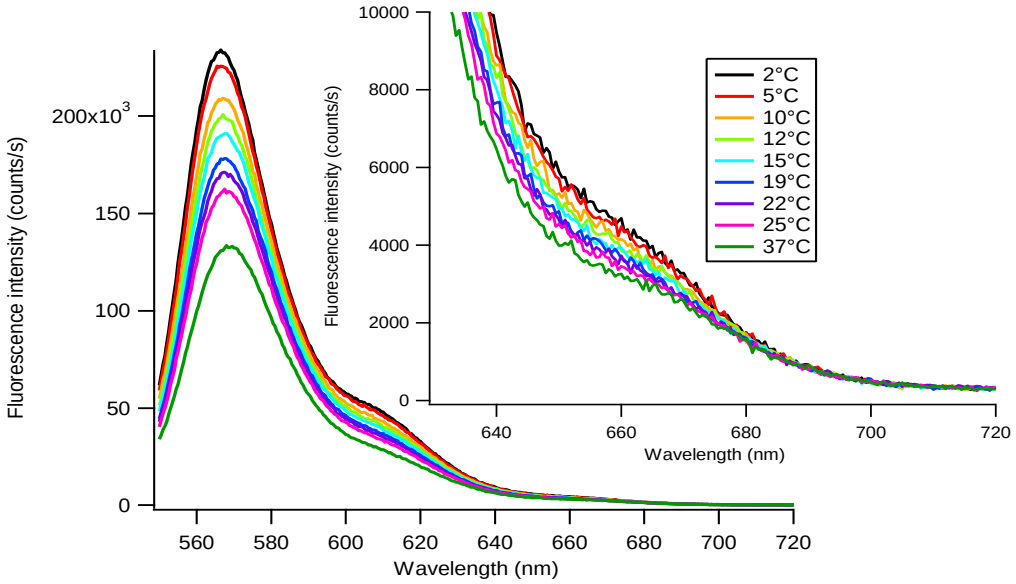


Figure S4: Demonstration of RNAP (Protein) Induced Fluorescence Enhancement (PIFE).

Emission spectra of 100nM Cy3(+14) λP_R promoter DNA from independent experiment, reproducing figure 5 with Cy3.

Emission spectra of Cy3 (550-720 nm;counts/s) titrated with RNAP (0, 50, 75, 100, 125, 150 nM) for CC (2°C; left panel) and OC (19°C; right panel). Samples were excited at 515nm.

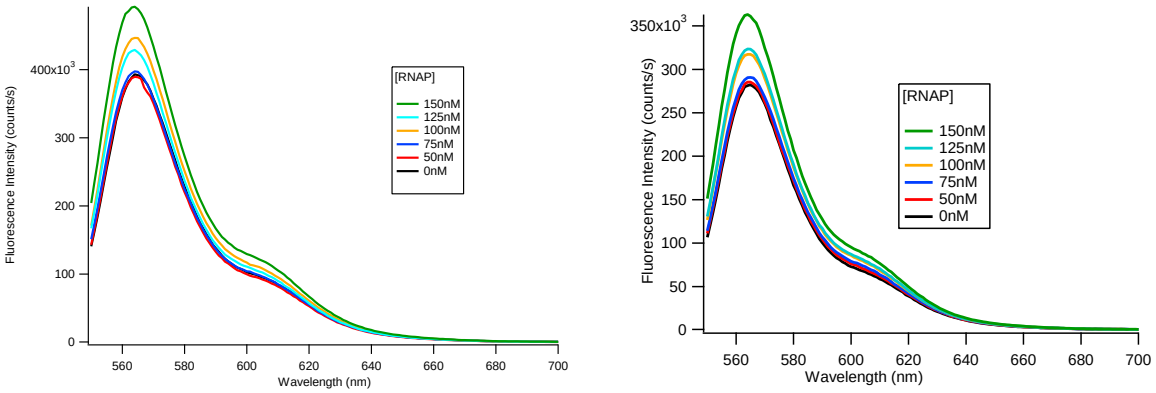


Figure S5: Demonstration of RNAP (Protein) Induced Fluorescence Enhancement (PIFE). Emission spectra of 100nM Cy5(+14) λP_R promoter DNA from independent experiment, reproducing figure 5. Emission spectra of Cy5 (620-700 nm or 620-720 nm;counts/s) titrated with RNAP (0, 50, 75, 100, 125, 150 nM) for CC (2°C; left panel) and OC (19°C; right panel). Samples were excited at 610nm.

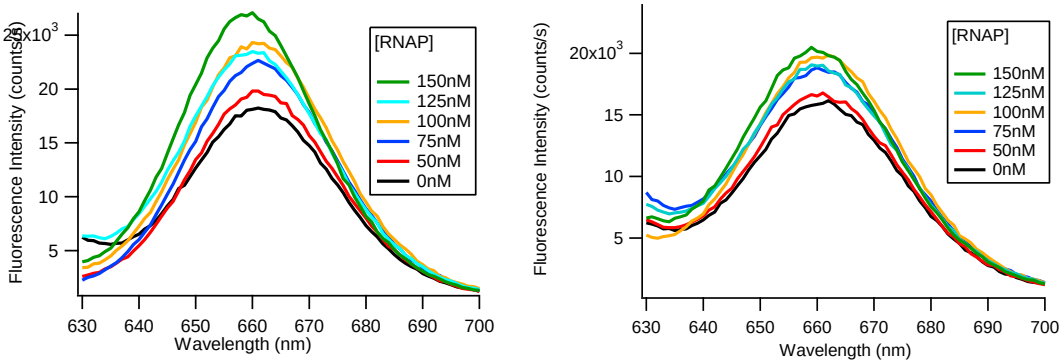


Figure S6: Emission spectra of 100nM Cy3(-100) λP_R promoter DNA.

Spectra of Cy3 (550-720 nm;counts/s) titrated with RNAP (0, 50, 75, 100, 125, 150 nM) for CC (2°C; left panel) and OC (19°C; right panel) showing absence of significant PIFE at -100. Samples were excited at 515nm.

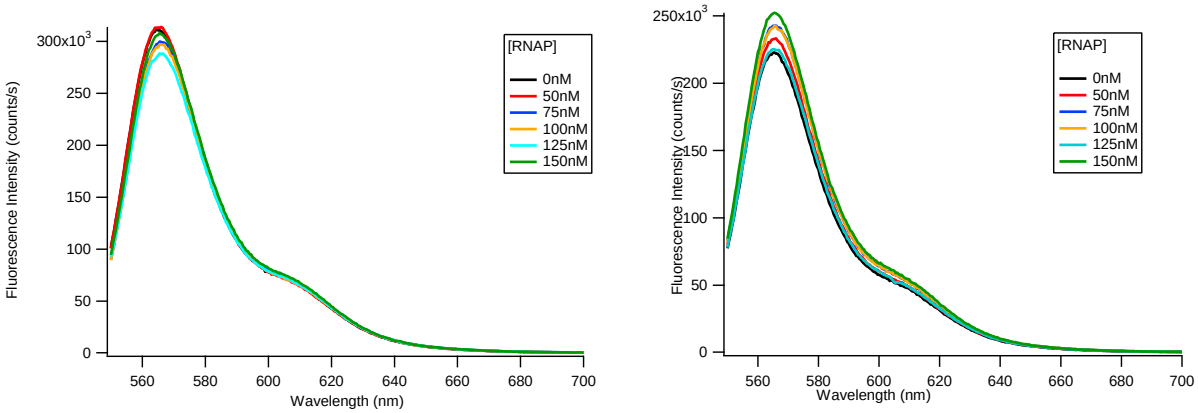


Figure S7: Emission spectra of 100nM Cy5(-100) λP_R promoter DNA.

Spectra of Cy5 (550-720 nm;counts/s) titrated with RNAP (0, 50, 75, 100, 125, 150 nM) for CC (2°C; left panel) and OC (19°C; right panel) showing absence of significant PIFE at -100. Samples were excited at 610 nm.

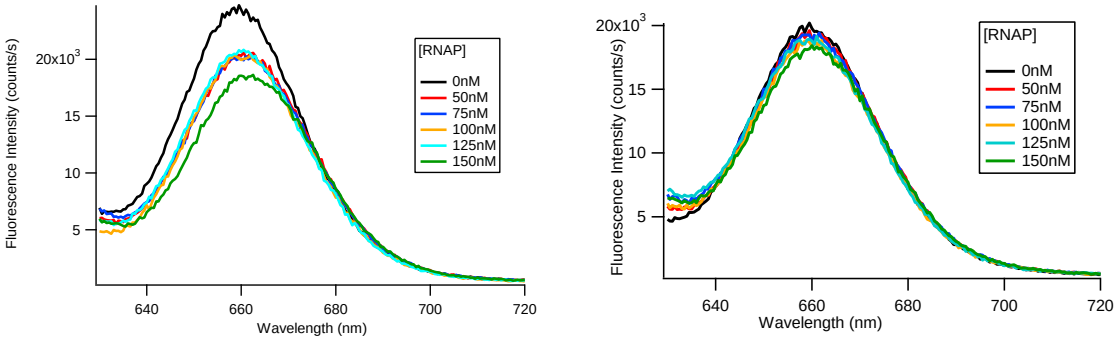
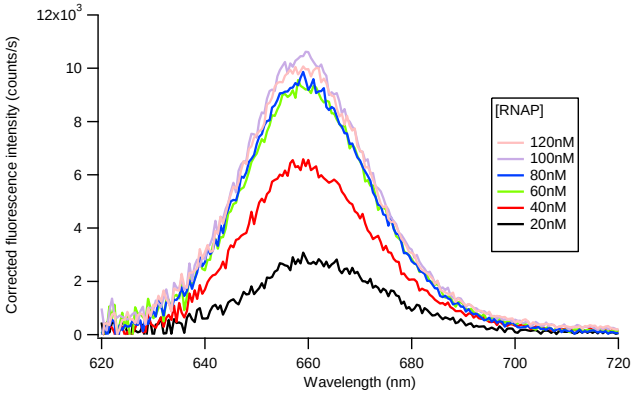


Figure S8: FRET efficiency determination: $Ratio_A$ analysis

Representative figures for (A) CC (B) OC showing increased Cy5 emission from 100nM Cy3(-100)Cy5(+14) LPR promoter DNA obtained by subtracting intensity-normalized signal of 100nM Cy3(-100) only DNA from the different spectra shown in Figure 1A (CC; 2°C data) and Figure 2A (OC; 19°C data), for $Ratio_A$ calculation to determine FRET efficiency (See Materials and Methods).

(A)



(B)

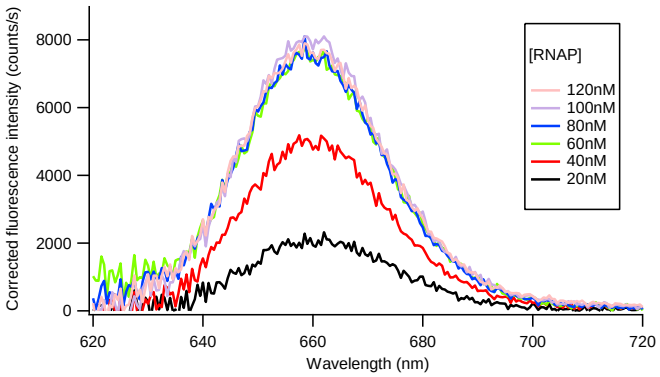
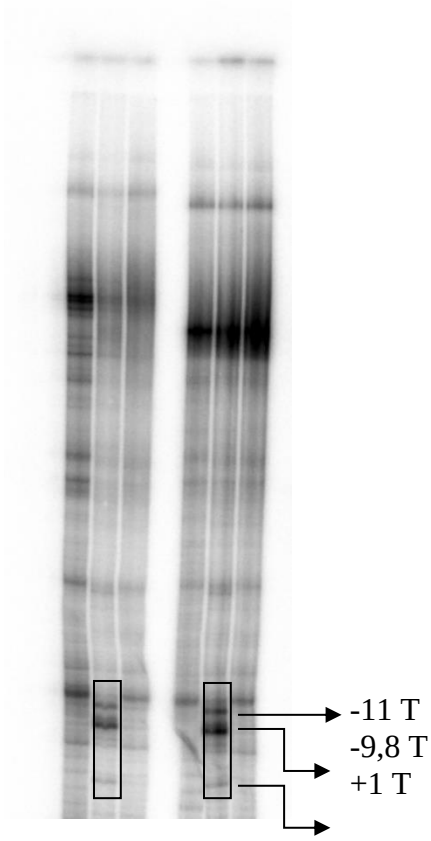


Figure S9: Comparison of permanganate footprints of T strand of Cy5 or Cy3(-100) λP_R at (A) 19°C and (B) 0°C. Boxed regions include positions of MnO_4-reactive thymines at -11, -9 & -8, and +1 in the open complex and neighboring background bands.

(A)

19 C Cy5 (-100) λP_R			**19 C λP_R**		
	RNAP			RNAP	
A+G	+	-	A+G	+	-



(B)

0 C λP_R			0 C Cy3 (-100) λP_R		
	RNAP			RNAP	
A+G	-	+	A+G	-	+

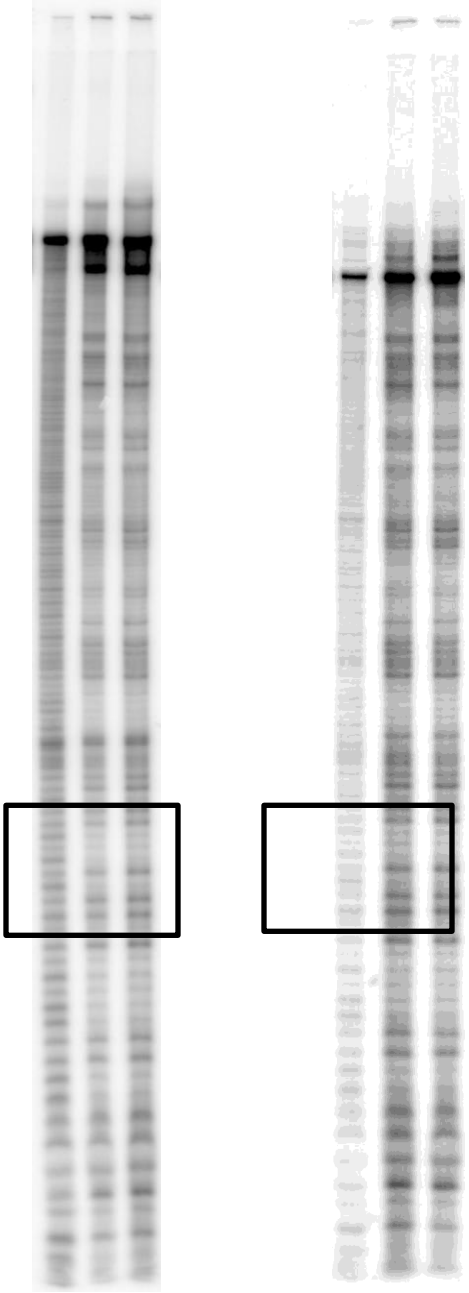


Figure S10: Models of bent and wrapped promoter DNA in CC with *E.coli* RNAP, built to estimate minimum possible dye-dye distances R consistent with structural and footprinting data. The CC model is built on the structure of *E.coli* RNAP [36]. Cyanine dyes are positioned as ellipsoids at -100 (hot pink) and +14 (blue). RNAP subunits αNTD (cyan), αCTD (yellow), β (pink), β' (green), ω (blue), σ^{70} (orange) are shown. The promoter DNA shown in different colors is represented to have different angles of bending at -65 that result in different distances between dyes at -100 and +14: Dark grey (70 Å), blue (60 Å), yellow (45 Å) and purple (30 Å). For interdye distances < 50 Å, the upstream footprint extends to or beyond -85, inconsistent with the boundary known from HO· footprinting (-82/-81) for CC [36] (See Figure 7A).

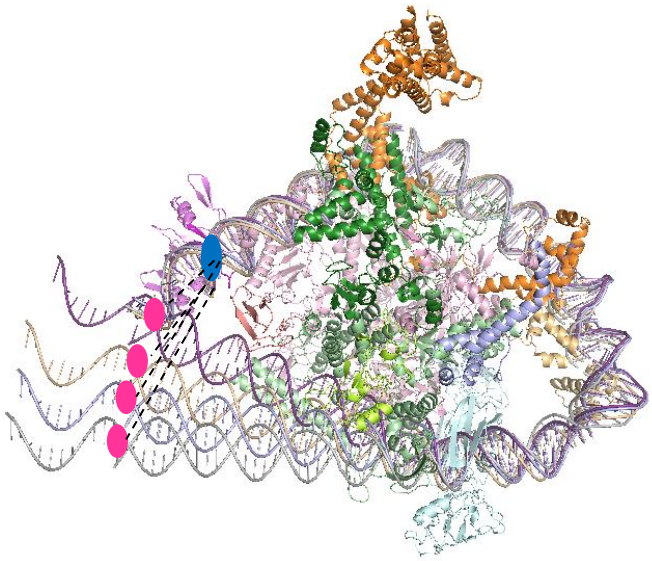
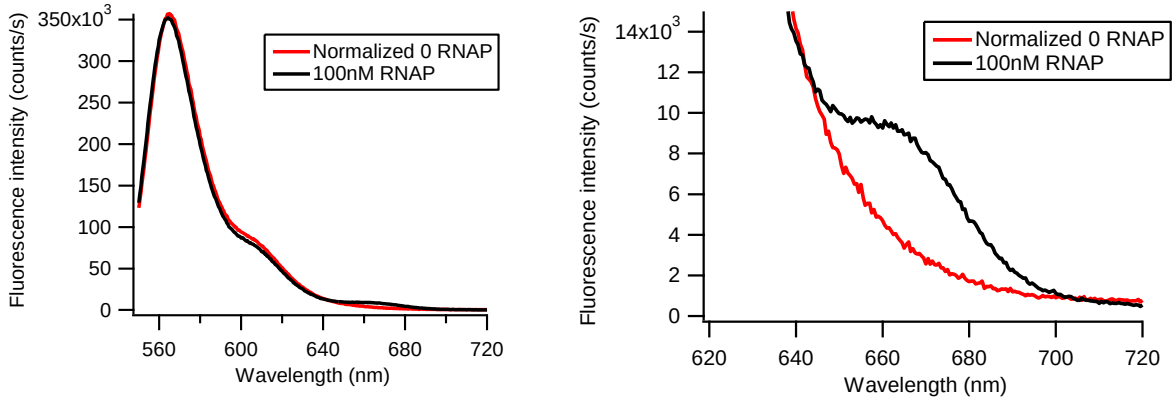
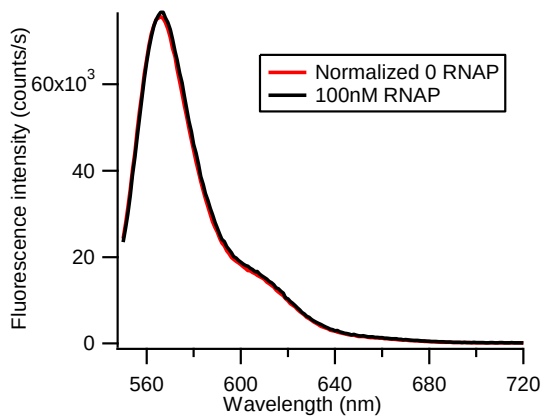


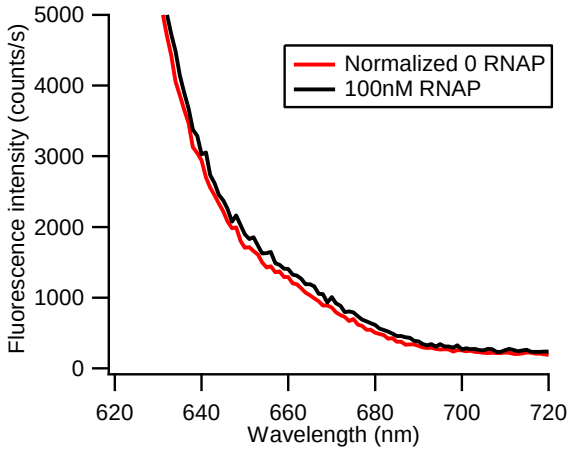
Figure S11: Side-by-side comparison of Cy3 emission - normalized curves at (A) 2°C for (A) Cy3(+14)Cy5(-100) λP_R DNA- RNAP (B) Cy3(+14)Cy5(-100) control DNA – RNAP complexes and at 19°C for (C) Cy3(+14)Cy5(-100) λP_R DNA- RNAP and (D) Cy3(+14)Cy5(-100) control DNA – RNAP complexes summarized in Figure 3. Experiments with both constructs were performed under identical conditions as described in methods. Left/Right panels shows emission over 550-720 nm/620-720 nm upon excitation of samples at Cy3 excitation wavelength of 515 nm. As described in FRET efficiency estimation (see methods), the -RNAP curves were normalized to 100 nM RNAP curves over the Cy3 emission range 555-570 nm allowing clear comparison of Cy5 emission signal between 620-720 nm in 0 nM vs 100 nM RNAP samples. λP_R DNA- RNAP complexes in (C) and (D) show increased Cy5 emission upon Cy3 excitation while the non - P_R DNA- RNAP complexes in (A) and (B) do not, proving absence of FRET in these constructs.

(A)

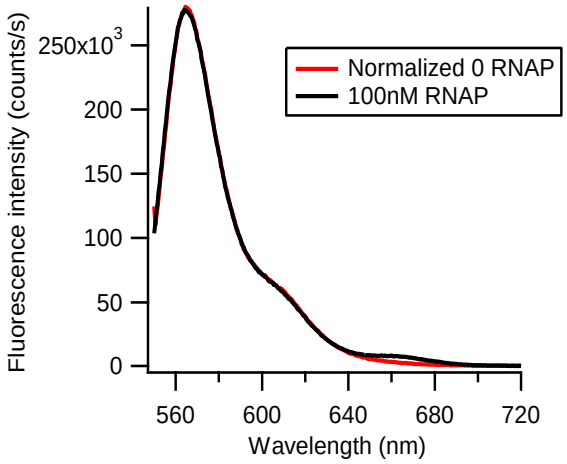


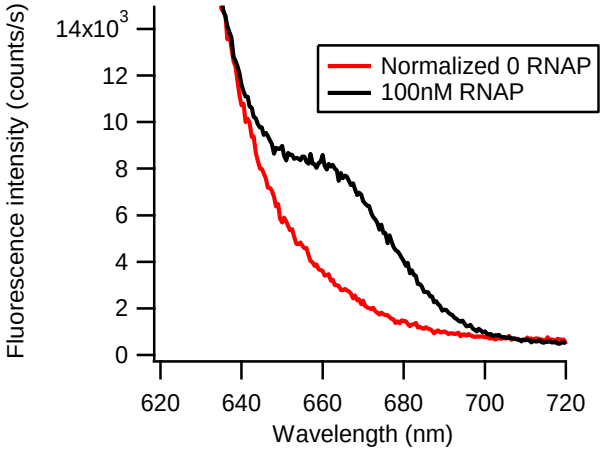
(B)


Fluorescence intensity (counts/s)
60x10^3
40
20
0
560
600
640
680
720
Wavelength (nm)
Normalized 0 RNAP
100nM RNAP

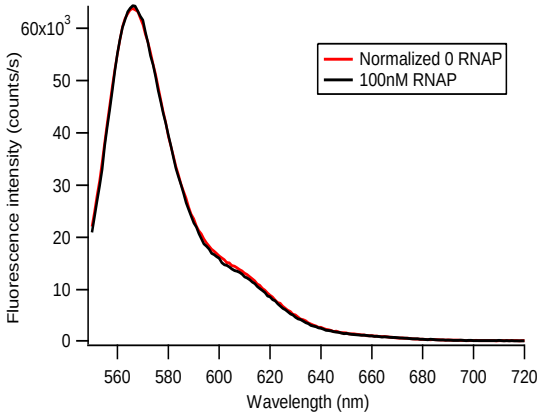

Fluorescence intensity (counts/s)
5000
4000
3000
2000
1000
0
620
640
660
680
700
720
Wavelength (nm)
Normalized 0 RNAP
100nM RNAP

(C)


Fluorescence intensity (counts/s)
250x10^3
200
150
100
50
0
560
600
640
680
720
Wavelength (nm)
Normalized 0 RNAP
100nM RNAP


Fluorescence intensity (counts/s)
14x10^3
12
10
8
6
4
2
0
620
640
660
680
700
720
Wavelength (nm)
Normalized 0 RNAP
100nM RNAP

(D)


Fluorescence intensity (counts/s)
60x10^3
50
40
30
20
10
0
560
580
600
620
640
660
680
700
720
Wavelength (nm)
Normalized 0 RNAP
100nM RNAP

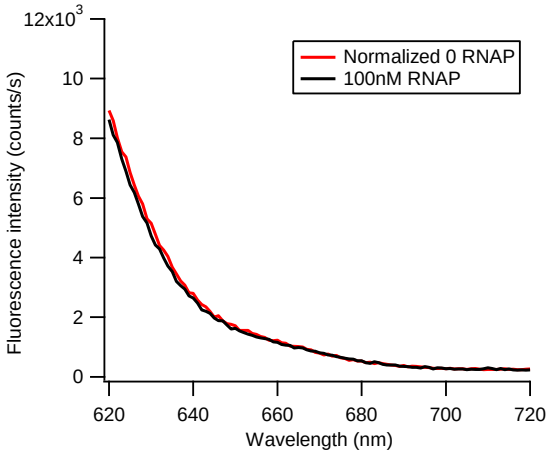

Fluorescence intensity (counts/s)
12x10^3
10
8
6
4
2
0
620
640
660
680
700
720
Wavelength (nm)
Normalized 0 RNAP
100nM RNAP

Table S1: Occupancy corrected FRET efficiency calculated from $Ratio_A$ at 660nm for each orientation of the dyes at (A) 2°C and (B) 19°C. Distances calculated from these efficiencies are given in table 1. Values in parantheses were not included in calculating average of FRET efficiencies.

(A)

[RNAP] nM	20	40	60	80	100	120
Cy3(-100)-Cy5(+14) λP_R	0.35	0.56	0.45	0.41	(0.36)	(0.28)
Cy3(+14)-Cy5(-100) λP_R	0.18	0.23	0.36	0.32	(0.21)	(0.18)

(B)

[RNAP] nM	20	40	60	80	100	120
Cy3(-100)-Cy5(+14) λP_R	0.24	0.47	0.40	0.39	(0.31)	(0.25)
Cy3(+14)-Cy5(-100) λP_R	0.10	0.25	0.41	0.36	(0.27)	(0.25)

Table S2: Occupancy corrected FRET efficiency calculated from $Ratio_A$ at 660nm for each orientation of the dyes at (A) 2°C and (B) 19°C. Distances calculated from these efficiencies are given in table 1. Values in parantheses were not included in calculating average of FRET efficiencies.

(A)

[RNAP] nM	50	75	100	125	150
Cy3(-100)-Cy5(+14) λP_R	0.28	0.20	(0.13)	(0.12)	(0.11)
Cy3(+14)-Cy5(-100) λP_R	0.19	0.16	(0.07)	(0.05)	(0.11)

(B)

[RNAP] nM	50	75	100	125	150
Cy3(-100)-Cy5(+14) λP_R	0.52	0.10	(0.08)	(0.07)	(0.14)
Cy3(+14)-Cy5(-100) λP_R	0.37	0.35	0.09	(0.06)	(0.10)

Table S3: Occupancy corrected FRET efficiency calculated from $Ratio_A$ at 660nm for each orientation of the dyes at (A) 2°C and (B) 19°C. Distances calculated from these efficiencies are given in table 1. Values in parantheses were not included in calculating average of FRET efficiencies.

(A)

[RNAP] nM	50	75	100	125	150
Cy3(-100)-Cy5(+14) λP_R	0.34	0.36	(0.23)	(0.23)	(0.20)
Cy3(+14)-Cy5(-100) λP_R	0.24	0.14	(0.24)	(0.20)	(0.18)

(B)

[RNAP] nM	50	75	100	125	150
Cy3(-100)-Cy5(+14) λP_R	0.39	0.44	(0.25)	(0.27)	(0.20)
Cy3(+14)-Cy5(-100) λP_R	0.40	0.26	(0.37)	(0.28)	(0.24)

References:

[1] Saecker, R. M., Record, M. T. J., and deHaseth, P. L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase-promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis, *Journal of Molecular Biology 412*.

[2] Ruff, E. F., Record, M. T., Jr. , and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation, *Biomolecules 5*, 1035-1062.

[3] Mekler, V., Minakhin, L., Borukhov, S., Mustaev, A., and Severinov, K. (2014) Coupling of downstream RNA polymerase-promoter interactions with formation of catalytically competent transcription initiation complex, *J Mol Biol 426*, 3973-3984.

[4] Drennan, A., Kraemer, M., Capp, M. W., Gries, T., Ruff, E. F., Sheppard, C., Wigneshweraraj, S., Artsimovitch, I., and Record, M. T., Jr. (2012) Key roles of the downstream mobile jaw of *Escherichia coli* RNA polymerase in transcription initiation, *Biochemistry 51*, 9447-9459.

[5] Ruff, E. F., Drennan, A. C., Capp, M. W., Poulos, M. A., Artsimovitch, I., and Record, M. T., Jr. (2015) *E. coli* polymerase determinants of open complex lifetime and structure, *Journal of Molecular Biology 427*, 2435-2450.

[6] Ross, W., and Gourse, R. L. (2005) Sequence-independent upstream DNA-aCTD interactions strongly stimulate *Escherichia coli* RNA polymerase-*lac*UV5 promoter association, *Proceedings of the National Academy of Sciences 102*, 291-296.

[7] Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr., and Saecker, R. M. (2007) Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 104*, 7833-7838.

[8] Davis, C. A., Capp, M. W., Record, M. T., Jr., and Saecker, R. M. (2005) The effects of upstream DNA on open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 102*, 285-290.

[9] Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., and Record, M. T. J. (2002) Kinetic studies and structural models of the association of *E. coli* sigma 70 RNA polymerase with the LPR promoter: large scale conformational changes in forming the kinetically significant intermediates, *Journal of Molecular Biology 319*, 649-671.

[10] Carpousis, A. J., and Gralla, J. D. (1985) Interaction of RNA polymerase with *lac*UV5 promoter DNAduring mRNA initiation and elongation, *Journal of Molecular Biology 183*, 165-177.

[11] Rivetti, C., Guthold, M., and Bustamante, C. (1999) Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex, *EMBO 18*, 4464-4475.

[12] Doniselli, N., Rodriguez-Aliaga, P., Amidani, D., Bardales, J. A., Bustamante, C., Guerra, D. G., and Rivetti, C. (2015) New insights into the regulatory mechanisms of ppGpp and DksA on Escherichia coli RNA polymerase-promoter complex, *Nucleic Acids Res 43*, 5249-5262.

[13] Rivetti, C., Codeluppi, S., Dieci, G., and Bustamante, C. (2003) Visualizing RNA Extrusion and DNA Wrapping in Transcription Elongation Complexes of Bacterial and Eukaryotic RNA Polymerases, *Journal of Molecular Biology 326*, 1413-1426.

[14] Maurer, S., Fritz, J., Muskhelishvili, G., and Travers, A. (2006) RNA polymerase and an activator form discrete subcomplexes in a transcription initiation complex, *EMBO J 25*, 3784-3790.

[15] Forget, D., Robert, F., Grondin, G., Burton, Z. F., Greenblatt, J., and Coulombe, B. (1997) RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box, *Proc Natl Acad Sci U S A 94*, 7150-7155.

[16] Kim, T. K., Lagrange, T., Wang, Y. H., Griffith, J. D., Reinberg, D., and Ebright, R. H. (1997) Trajectory of DNA in the RNA polymerase II transcription preinitiation complex, *Proc Natl Acad Sci U S A 94*, 12268-12273.

[17] Robert, F., Douziech, M., Forget, D., Egly, J. M., Greenblatt, J., Burton, Z. F., and Coulombe, B. (1998) Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF, *Mol Cell 2*, 341-351.

[18] Kozlov, A. G., and Lohman, T. M. (2002) Stopped-flow studies of the kinetics of single-stranded DNA binding and wrapping around the Escherichia coli SSB tetramer, *Biochemistry 41*, 6032-6044.

[19] Hwang, H., and Myong, S. (2014) Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions, *Chem Soc Rev 43*, 1221-1229.

[20] Ko, J., and Heyduk, T. (2014) Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse, *Biochemistry Journal 463*, 135-144.

[21] 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*, 22175-22180.

[22] Hwang, H., Kim, H., and Myong, S. (2011) Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity, *Proc Natl Acad Sci U S A 108*, 7414-7418.

[23] Hanish, J., and McClelland, M. (1988) Activity of DNA modification and restriction enzymes in KGB, a potassium glutamate buffer, *Gene Anal Tech 5*, 105-107.

[24] McClelland, M., Hanish, J., Nelson, M., and Patel, Y. (1988) KGB: a single buffer for all restriction endonucleases, *Nucleic Acids Res 16*, 364.

[25] Svetlov, V., and Artsimovitch, I. (2015) Purification of bacterial RNA polymerase: tools and protocols, *Methods in Molecular Biology 1276*, 13-29.

[26] Borukhov, S., and Goldfarb, A. (1993) Recombinant Escherichia coli RNA polymerase: purification of individually overexpressed subunits and in vitro assembly, *Protein Expr Purif 4*, 503-511.

[27] Craig, M. L., Suh, W. C., and Record, M. T., Jr. (1995) HO. and DNase I probing of E sigma 70 RNA polymerase--lambda PR promoter open complexes: Mg2+ binding and its structural consequences at the transcription start site, *Biochemistry 34*, 15624-15632.

[28] Sanborn, M. E., Connolly, B. K., Gurunathan, K., and Levitus, M. (2007) Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA, *J Phys Chem B 111*, 11064-11074.

[29] Vander Meulen, K. A., Saecker, R. M., and Record, M. T., Jr. (2008) Formation of a wrapped DNA-protein interface: experimental characterization and analysis of the large contributions of ions and water to the thermodynamics of binding IHF to H' DNA, *J Mol Biol 377*, 9-27.

[30] Clegg, R. M. (1992) Fluorescence resonance energy transfer and nucleic acids, *Methods Enzymol 211*, 353-388.

[31] Lakowicz, J. R. (2006) *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer.

[32] Malicka, J., Gryczynski, I., Kusba, J., and Lakowicz, J. R. (2003) Effects of metallic silver island films on resonance energy transfer between N,N'-(dipropyl)-tetramethyl-indocarbocyanine (Cy3)- and N,N'-(dipropyl)-tetramethyl- indodicarbocyanine (Cy5)-labeled DNA, *Biopolymers 70*, 595-603.

[33] Ouellet, J., Schorr, S., Iqbal, A., Wilson, T. J., and Lilley, D. M. (2011) Orientation of cyanine fluorophores terminally attached to DNA via long, flexible tethers, *Biophys J 101*, 1148-1154.

[34] Chen, R. F., and Bowman, R. L. (1965) Fluorescence Polarization: Measurement with Ultraviolet-Polarizing Filters in a Spectrophotofluorometer, *Science 147*, 729-732.

[35] Ivanov, V., Li, M., and Mizuuchi, K. (2009) Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements, *Biophys J 97*, 922-929.

[36] Bae, B., Davis, E., Brown, D., Campbell, E. A., Wigneshweraraj, S., and Darst, S. A. (2013) Phage T7 gp2 inhibition of *E. coli* RNA polymerase involves misappropriation of sigma 70 domain 1.1, *Proceedings of the National Academy of Sciences 110*, 19772-19777.

[37] Ross, W., Schneider, D. A., Paul, B. J., Mertens, A., and Gourse, R. L. (2003) An intersubunit contact stimulating transcription initiation by E coli RNA polymerase: interaction of the alpha C-terminal domain and sigma region 4, *Genes Dev 17*, 1293-1307.

[38] Zuo, Y., and Steitz, T. A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Molecular Cell 58*, 534-540.

[39] Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., Saecker, R. M., and Record, M. T. J. (1998) DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA, *Journal of Molecular Biology 283*, 741-756.

[40] Roe, J.-H., Burgess, R. R., and Record, M. T. J. (1984) Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λPR promoter, *Journal of Molecular Biology 176*, 495-521.

[41] Roe, J. H., and Record, M. T., Jr. (1985) Regulation of the kinetics of the interaction of Escherichia coli RNA polymerase with the lambda PR promoter by salt concentration, *Biochemistry 24*, 4721-4726.

[42] Kobitski, A. Y., Nierth, A., Helm, M., Jaschke, A., and Nienhaus, G. U. (2007) Mg2+-dependent folding of a Diels-Alderase ribozyme probed by single-molecule FRET analysis, *Nucleic Acids Res 35*, 2047-2059.

[43] Yasuda, R., Masaike, T., Adachi, K., Noji, H., Itoh, H., and Kinosita, K., Jr. (2003) The ATP-waiting conformation of rotating F1-ATPase revealed by single-pair fluorescence resonance energy transfer, *Proc Natl Acad Sci U S A 100*, 9314-9318.

[44] Lee, J. Y., Wei, S., and Lee, T. H. (2011) Effects of histone acetylation by Piccolo NuA4 on the structure of a nucleosome and the interactions between two nucleosomes, *J Biol Chem 286*, 11099-11109.

[45] Rammohan, J., Ruiz Manzano, A., Garner, A. L., Stallings, C. L., and Galburt, E. A. (2015) CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism, *Nucleic Acids Res 43*, 3272-3285.

[46] Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S. A. (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex, *Elife 4*.

[47] Bernier, M., Luo, Y., Nwokelo, K. C., Goodwin, M., Dreher, S. J., Zhang, P., Parthun, M. R., Fondufe-Mittendorf, Y., Ottesen, J. J., and Poirier, M. G. (2015) Linker histone H1 and

H3K56 acetylation are antagonistic regulators of nucleosome dynamics, *Nat Commun 6*, 10152.

[48] Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution, *Nature 417*, 712-719.

[49] 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*, 1139-1143.

[50] 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*, 1144-1147.

[51] Sainsbury, S., Bernecky, C., and Cramer, P. (2015) Structural basis of transcription initiation by RNA polymerase II, *Nat Rev Mol Cell Biol 16*, 129-143.

[52] Nogales, E., Louder, R. K., and He, Y. (2017) Structural Insights into the Eukaryotic Transcription Initiation Machinery, *Annu Rev Biophys 46*, 59-83.

[53] Mangiarotti, L., Cellai, S., Ross, W., Bustamante, C., and Rivetti, C. (2009) Sequence-dependent upstream DNA-RNA polymerase interactions in the open complex with lambdaPR and lambdaPRM promoters and implications for the mechanism of promoter interference, *J Mol Biol 385*, 748-760.

[54] Klucar, L., Stano, M., and Hajduk, M. (2010) phiSITE: database of gene regulation in bacteriophages, *Nucleic Acids Res 38*, D366-370.

Chapter 3

Fluorescence-Monitored Conformational Changes in Open Complex Formation by E.coli RNA Polymerase: Upstream Wrapping, Clamp Opening, and Bending of Downstream Duplex DNA into the Active Site Cleft.

PREFACE

This chapter is in preparation for publication with myself as a co-first author (Munish Chhabra and Raashi Sreenivasan). My contributions in this research included: data collection from the stopped-flow fluorimeter, its analysis, fitting the data to the mechanism and preparation of structural models of closed complex intermediates to support the FRET, PIFE data. Michael W.Capp and Rezwana Karim helped me in preparing RNAP. Christina McNerney, Clare Cimperman, Katelyn Callies, Hertina Kan and Raashi Sreenivsan helped me in preparation of samples for experiments and data collection. Dr. Irina Shkel helped me in data fitting to a multi 5 step mechanism.

Abstract (250 words)

Equilibrium FRET (fluorescence energy transfer) between far-upstream (-100) and downstream (+14) cyanine dyes showed that λP_R promoter DNA is highly bent and wrapped on *E. coli* RNA polymerase (RNAP) in both closed and open complexes (CC,OC). Here, using FRET and single-dye fluorescence enhancements (PIFE), we determine the kinetics and mechanism of DNA bending and wrapping by FRET and of contact-formation between RNAP and -100 and +14 promoter positions by PIFE. The kinetics show two phases: relatively rapid-step-by-step reversible formation of a CC ensemble ({CC}) of at-least four closed intermediates (initial (RP_c), early (I_{1E}), mid- (I_{1m}), and most advanced (late; I_{1L})), and slower conversion of {CC} to OC via I_{1L}. Initial FRET and PIFE signals are found to be from I_{1E}, not RP_C. Formation of I_{1E} involves extensive wrapping of upstream DNA, contact of -100 with RNAP, and bending of downstream DNA to contact the β clamp at +14. Conversion of I_{1E} to I_{1M} occurs on a 1 s time scale without large changes in FRET and PIFE, consistent with the large scale opening of β'-clamp. A subsequent increase in FRET in conversion of I_{1M} to I_{1L} indicates insertion of the downstream duplex deep into the clamp, and is followed by the slower kinetic phase of conversion of I_{1L} to OC by clamp closing and DNA opening. The {CC} equilibrium constant the {CC}-to-OC isomerization rate constant are consistent with previous filter-binding and MnO_4^- footprinting results. Implications of these results for facilitation of CC-to-OC isomerization by upstream DNA and by upstream-binding transcription factors are discussed.

Introduction

The rate of open complex (OC) formation by *E. coli* RNA polymerase σ^{70} holoenzyme (RNAP; R) at promoter DNA (P) is an important determinant of the rate of transcription initiation. Hence the kinetics and equilibria of the steps of OC formation are highly regulated by promoter sequence, transcription factors, ligands, and solution variables. In abortive initiation and filter binding assays that detect only the formation of long-lived OC, the kinetics of OC formation per promoter at a specified RNAP concentration typically are well-described by the two-step minimal Mechanism 1.

$$R + P \overset{K_{CC}}{\rightleftharpoons} \{CC\} \overset{k_{isom}}{\rightarrow} OC \qquad \text{Mechanism 1}$$

In Mechanism 1, K_{CC} is the composite equilibrium constant for reversible formation of an ensemble of early and advanced closed complexes (called the CC ensemble, symbolized as {CC}) from promoter DNA. Evidence for atleast three closed complexes (CC) that may be in general be part of {CC} is provided by different downstream boundaries of chemical and/or enzymatic footprints obtained at different promoters. These CC include the initial specific complex (RP_C; boundary at approximately -5) and two more–advanced CC (designated I_{1E} and I_{1L} for “early” and “late”) with downstream footprint boundaries at approximately +2/+7 and +20 [1]. The isomerization rate constant k_{isom} is the composite rate constant for conversion of {CC} and free P to OC via the most advanced member of {CC} (I_{1L}) and the subsequent DNA opening step (with rate constant k_{open}).

This interpretation of the parameters of Mechanism 1 differs subtly but significantly from those used previously. In particular, the observed rate constant k_{isom} is the product of k_{open} and the fraction of the fully-established {CC} that is I_{1L}, while the observed equilibrium constant K_{CC} for forming {CC} is the quotient of K_{RPc} and the fraction of the fully-established {CC} that is RP_c [1](

see Discussion and SI). Previous interpretations of effects of changes in promoter sequence or length, transcription factors and solution variables on K_{CC} and k_{isom} have generally not considered shifts in the composition of {CC} with these variables. In general, any shift in the proportions of RP_c, I_{1E}, and I_{1L} in {CC} will change both K_{CC} and k_{isom}. The large effects of upstream truncation of promoter DNA on k isom are best explained in this way [1](see below).

In general, the initial closed complex RP_c is stabilized by some combination of specific interactions of the flexibly-tethered αCTD with the promoter UP element (a 20 bp region containing A and T tracts immediately upstream of the -35 region [2]) and of regions 2 and 4 of the σ^{70} subunit with the -10 and -35 elements of the promoter. Downstream protection of promoter DNA in RP_c extends to approximately -5 [3 4 1]. Phased bending of UP element DNA by interaction with the αCTD of RNAP (including in one case the -35 element) is indicated in several but not all RP_c complexes investigated by periodic (10 bp) sites of enhanced DNase reactivity in both strands, offset by approximately 3 bp [3 4]. After RP_c formation, the downstream duplex must bend by at least 90^o into the β-β' clamp above the active site cleft [5 6], in order for 13-14 bp of the -10 and discriminator regions and the transcription start site to be opened by RNAP using binding free energy [1] (4YLN.pdb) [7].

CryoEM studies indicate that ATP-independent promoter opening by eukaryotic RNA polymerases may involve a similar set of conformational changes. In the assembly of the Pol II pre-initiation complex (Pol II PIC), the upstream DNA is bent by 90^o by binding to the TATA element of transcription factors TBP (TATA-binding protein) together with TFIIA, TFIIB, and TAFs (TBP-associated factors). After Pol II PIC assembly, cryoEM reveals three closed complexes, designated CC1, CC2, and CC^{dist}, which plausibly are on-pathway intermediates to form OC [8]. In CC1, the downstream promoter is positioned above the closed clamp of RNAP. In CC2 the clamp

is open and in CC^{dist} the downstream duplex is bent into the open clamp. Conversion of CC^{dist} to the final OC involves closing the clamp and opening the DNA [8].

For *E. coli* RNAP, the isomerization of {CC} to OC is greatly facilitated by the presence of far-upstream DNA. For full-length (FL) λP_R promoter DNA, real-time HO footprinting of {CC} revealed interactions with RNAP up to position -82 or beyond [9]. Upstream truncation of λP_R and lacUV5 promoters at positions between -82 (20 bp upstream of the UP element) and -42 (in the UP element) reduced the isomerization rate constant k_{isom} (Mechanism 1) for conversion of {CC} to OC by one-to-two orders of magnitude [10 11]. The effect on k_{isom} increases with increasing upstream truncation and is accompanied by a modest increase in K_{CC}.

For the UT-47 λP_R truncation variant, DNase footprinting of {CC} early in the time course of OC formation revealed downstream boundaries of protection of the template and nontemplate strands of the downstream duplex at +2 and +7, as compared to +20 for {CC} at FL λP_R [10 9]. This indicates that the {CC} population distribution is much less advanced for UT -47 than for FL λP_R. To explain these profound differences in OC-formation kinetics and {CC} footprints for FL and UT-47 λP_R, we previously proposed that upstream DNA to -82 is bent and wrapped on RNAP and that the downstream duplex is bent into the active site cleft in the most advanced member of {CC} (I_{1L}), poised to be opened by RNAP. Tests of these proposals are provided by the FRET and PIFE kinetics studies reported here.

Equilibrium FRET studies on low temperature (2 ^oC; closed) and higher temperature (19 ^oC; open) complexes revealed that upstream DNA in both complexes is highly bent and wrapped and yielded structural models for the wrapping track in both complexes [12]. Substantial structural and footprinting evidence existed previously for upstream bending and wrapping (from approximately -38 to -65), and downstream bending of the individual strands in the open complex [10 13 9 14 15 16 7]. Upstream modulators of initiation like CAP and NAPs (IHF, HU, FIS) that also

bend and/or wrap DNA may therefore affect k_{isom} by affecting the bending and wrapping of upstream DNA on RNAP in members of the CC ensemble.

In addition to FRET kinetic studies of changes of DNA end-to-end distance from bending and wrapping on RNAP, we report the kinetics of RNAP-induced fluorescence enhancements (called PIFE effects [17]) of Cy3 and Cy5 at -100 and +14 position. These PIFE effects report on the development of contacts between these DNA positions and RNAP in the kinetics of OC formation, and hence complement the FRET kinetic studies. PIFE effects using Cy3 at positions near the TSS have been used to study aspects of the kinetics of OC formation, initiation, and escape [18] [19] [20]. Cy3 PIFE kinetic measurements were also used to study the cooperative interaction between CarD and RbpA transcription factor in OC formation for *Mycobacterium tuberculosis* σ^A RNAP holoenzyme [21, 22].

Here we use FRET and PIFE in kinetic assays of OC formation, dissociation and of transcription initiation to understand the key changes in interactions of upstream and downstream promoter DNA with RNAP as the initial CC advances and converts to OC. Analysis of the FRET and PIFE kinetic data provides compelling evidence for a sequential, five step mechanism with four closed complexes (the initial closed complex (RP_c) and three more advanced members of {CC} designated I_{1E}, I_{1M}, I_{1L}) on the pathway to open complex formation.

$$R + P \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} RP_c \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} I_{1E} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} I_{1M} \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} I_{1L} \overset{k_5}{\rightarrow} RP_o \qquad \text{Mechanism 2}$$

Rate and equilibrium constants for reversible formation of the members of {CC} (RP_c, I_{1E}, I_{1M}, I_{1L}) are obtained, as well as the DNA opening rate constant k_{open} and intrinsic FRET and PIFE signatures of these closed intermediates to compare with those of OC and free P. FRET signatures of the intermediates provide information about the progression of upstream wrapping

and downstream bending as {CC} advances. Positioning of the downstream duplex in these more advanced CC intermediates deduced from FRET studies reported here appears to correspond closely to what is observed for pol II by Cryo-EM [8]. PIFE signatures of the intermediates provide information about the strength of contacts between RNAP and far-upstream (-100) and downstream (+14) regions of duplex DNA in early vs late CC and OC. We also predict the time-evolution of the populations of individual early and late CC and the final OC. Rate and equilibrium constants obtained from FRET and PIFE analyses of OC formation at the λP_R promoter are consistent with previous determinations of the composite quantities K_{CC} and k_{isom} for λP_R by filter binding assays.

These studies provide novel insights into the time-evolution of the {CC} ensemble and how wrapping of upstream DNA facilitates isomerization in OC formation. They provide strong support for the proposal that promoter sequences and transcription factors that increase k_{isom} do so by favoring the formation of advanced closed complexes in which promoter DNA is wrapped upstream, bent downstream, and therefore poised to be opened by RNAP, and not by increasing the rate of the actual opening step (k_{open}) [23 1].

Materials and Methods

Buffers:

Storage Buffer (SB) for RNA polymerase holoenzyme is 50% glycerol (v/v), 10 mM Tris base (pH 7.5 at 4 ^oC), 0.1 M NaCl, 0.1 mM DTT, and 0.1 mM Na_2EDTA. Fluorescence buffer (FB) for FRET and PIFE kinetics experiments is 40.2 mM Tris, pH 8, 2 mM NaCl, 0.12 M KCl, 10mM $MgCl_2$, 2 μM DTT, 2 mM Na_2EDTA, 0.05 mg/ml BSA,1-2% glycerol and 0.02% Tween. Permanganate footprinting buffer (PFB) is 40 mM Tris (pH 8 at 19 ^oC), 10 mM $MgCl_2$, 0.12 M NaCl and 100 μg/ml BSA. Urea loading buffer (ULB), used to resuspend footprinting samples, is 8 M urea, 0.5 X TBE (45mM Tris-borate and 1 mM Na_2EDTA), 0.05% xylene cyanol (w/v) and 0.05%

bromophenol blue (w/v). Transcription buffer (TB), initiation (IS) and quench (QS) solutions were prepared as described previously [24]. The electrophoresis buffer for transcription experiment was 1X TBE (90 mM Tris-borate and 2 mM Na_2EDTA).

Preparation of E. coli RNA Polymerase (RNAP) Holoenzyme and Labeled λP_R Promoter DNA

RNAP core enzyme was overexpressed and purified as described previously, using *E. coli* BL21(λDE3) transformed with pVS10 [12]. σ^{70} was overexpressed and purified using *E. coli* M5219 transformed with plasmid pMRG8, as described previously [25]. RNAP core enzyme and σ^{70} were incubated at a 1:2 molar ratio for 1 hour at 37°C in SB to reconstitute RNAP holoenzyme, then stored at -20°C until use. Fractions of RNAP molecules able to form open complexes ranged from 0.50 to 0.90 for different preparations, as determined by a filter binding activity assay.[26] RNAP concentrations reported are active concentrations.

Single-dye-labeled λP_R fragments [Cy3 (-100), Cy5 (-100), Cy3 (+14) and Cy5 (+14)] for PIFE experiments and two-dye-labeled λP_R promoter DNA fragments [Cy3 (+14) Cy5 (-100) and Cy3 (-100) Cy5 (+14)] for FRET experiments were prepared by PCR as described previously using dye-labeled primers [12]. Unlabeled λP_R DNA fragments for single round transcription kinetics experiment were prepared by PCR as described previously [12, 24]. ^{32}P-labeled λP_R promoter DNA fragments for MnO_4^- footprinting experiments was prepared by PCR as described previously[23, 27]. Sequences of the different primers, DNA constructs generated and their lengths are described in the supporting information.

Fast MnO_4^- Footprinting:

RNAP and λP_R promoter DNA (-59 to +34) in PFB were independently loaded into a KinTek Corporation RQF-3 Rapid Chemical Quench-Flow instrument at 19°C. The solutions of RNAP and promoter DNA were mixed and held for varying times before mixing with 66.7 mM

$NaMnO_4$ for 50 ms, previously determined to be a concentration and time adequate for accurate detection of the rate of OC formation [28]. Samples were expelled, quenched with 500 µl ethanol, and immediately precipitated with ethanol and washed. Modified fragments were cleaved by 1M piperidine at 90°C. Reactions were evaporated and resuspended in ULB and resolved on 8% acrylamide sequencing gel.

Fast MnO_4^- Footprinting Analysis

Each lane and reactive band of a given MnO_4^- footprint was boxed and the total intensity was quantified using ImaqeQuant TL. The fraction of promoter DNA modified at a given position was determined by dividing the intensity of each uncut and reactive band by the total intensities of all uncut and reactive bands within a lane, and background was subtracted to determine corrected intensity. These corrected intensities were plotted vs time and fit to a single exponential time course. Corrected intensities were normalized by the appropriate fitted plateau intensity and plotted as normalized reactivity vs time.

Kinetics of Open Complex Formation from Free RNAP and Promoter DNA by Stopped-Flow Fluorescence (FRET, PIFE):

Fluorescence kinetic experiments were performed at 19°C in a Kintek SF-300X stopped flow fluorimeter (Kintek Corporation, PA) equipped with a 150 Watt Hg-Xe lamp (Hamamatsu, Japan) by rapid-mixing of equal volumes (20 µL) of dye-labeled promoter DNA and RNAP stock solutions. Each mixing of aliquots of the same stock solutions is called a “shot”. At 19°C the equilibrium constant for forming the CC ensemble from promoter DNA is near-maximal, the rate of isomerization of the CC ensemble to the OC is sufficiently slow to separate this kinetic phase from the earlier CC phase, and OC is sufficiently stable at this temperature that its formation is irreversible [5, 29]. Previous equilibrium FRET results on OC [12] were obtained at 19°C because the fluorescence of cyanine dyes decreases strongly with increasing temperature [12, 30-32].

Stocks of DNA and RNAP were prepared at twice the desired final concentration in FB. Most experiments were performed at a 1:1 mole ratio of active RNAP to promoter DNA (50 nM final concentrations of each reactant, 1% glycerol). FRET experiments at different salt or glycine betaine concentrations in FB were also performed at 19˚C at final concentrations of 50 nM DNA and 50 nM RNAP. A limited number of experiments were performed at 1.5:1 and 2:1 mole ratios of [RNAP]:[DNA] (final concentrations 75 or 100 nM RNAP, 50 nM DNA, and 1.5 or 2% glycerol). Control experiments showed no significant effects of glycerol concentration in this range (Supplemental Fig. S3).

Control experiments (5-6 shots) were performed by mixing a DNA solution in FB with no RNAP to verify that the fluorescence signal was time-independent, without photobleaching or large instrumental drift. DNA and RNAP solutions were loaded in the stopped flow syringes and incubated for at least 10 minutes at 19°C before being mixed in a series of shots. The dye-labeled λP_R promoter constructs were excited in the observation cell at wavelengths of 515nm (Cy3) or 610nm (Cy5) for single dye Cy3/Cy5 PIFE experiments. In FRET experiments, Cy3 dye (FRET donor) was excited at 515 nm wavelength and fluorescence emission of Cy3 and Cy5 (FRET acceptor) were monitored with different PMT as a function of time. Fluorescence emission signals were monitored using a 565-625 nm band pass filter for Cy3[36] and a 660 nm long-pass filter for Cy5[33] (Omega Optical, Brattleboro, VT). Monochromator excitation slit width of 1.56 mm or 3.14 mm were used. Fluorescence intensity was monitored from 10ms to 400s with 600 data points uniformly distributed on a log time scale.

In RNAP- DNA mixing experiments the initial 4 shots were typically discarded and the next 5 - 15 shots were collected and analyzed. Data collected from each shot was normalized, as described in the supporting information and then averaged to reduce noise at earlier time points for further analysis. For each of the six orientations or positions of the FRET and PIFE probes, at least six independent experiments were performed at 1:1 RNAP : promoter ratio, using

independent dilutions (and in some cases independent preparations) of RNAP and DNA solutions. Numbers of independent experiments and the average numbers of shots for each experiment are tabulated in Supplemental Table S3.

Kinetics of Transcription Initiation and Salt Upshift from Pre-formed Open Complexes Monitored by FRET and PIFE:

Cy3(-100)-Cy5(+14) λP_R or Cy5(-100)-Cy3(+14) λP_R promoter DNA was combined with RNAP in a 1:1 mole ratio in FB and incubated at 19°C for >30 min to form OC (final concentration 100nm). This dye-labeled OC solution in FB was rapidly mixed with varying NTP concentrations in FB (final OC concentration 50 nM) in transcription initiation and KCl (0.4 -1.1 M final KCl concentration) in FB for OC dissociation by salt upshift assay. Instrument conditions and excitation/ emission wavelengths are as described above. Data collected from each shot was first normalized as described in the supporting information and then averaged. Corrected FRET acceptor and PIFE emission signal vs time was fit to a sum of two exponentials using Origin 2017 software.

Results

Fast Permanganate Footprinting Kinetics of OC Formation

To test whether opening of the initiation bubble occurs as a single kinetic step and to validate the interpretation of the kinetics of open complex formation obtained previously from filter-binding data using Mechanism 1, fast permanganate (MnO_4^-) footprinting studies were used to determine rates of opening individual thymines in the λP_R open region (-11 to +2) in open complex formation at 19°C. Fast MnO_4^- footprinting of thymines in the -10 and start site regions of the promoter was previously performed to characterize the kinetics of open complex formation at the T7A1 promoter[14] and to quantify the size of the bubble and the extent of reactivity of individual thymines in the unstable intermediate open complex (I_2) formed in the DNA-opening step at the

λP_R promoter [23]. Although MnO_4^- reaction kinetics are moderately slow, requiring a reaction time of 150 ms even at high MnO_4^- concentration, this is fast relative to the time scale of the CC-to-OC isomerization at 19°C ($1/k_{isom} \sim 70$ s).[5] Since only stable OC are detected, the MnO_4^- kinetic assay is expected to be equivalent to the filter binding kinetic assay in which brief exposure of each sample to a competitor before assaying for binding eliminates contributions to the binding assay from any short-lived (closed or non-promoter) complexes that were present.

Gel lanes in the insets in Figure 1 show the kinetics of development of MnO_4^- reactivity in the region of the initiation bubble on non-template (nt, panel A) and template (t, panel B) strands. In the first 10 s, where CC are the predominant promoter complexes, these MnO_4^- snapshots detect no reactive (open) thymines. MnO_4^- reactivity of all observable thymines (-4/-3 and +2 on the nt strand; -11 and -9/-8 on the t strand) develops on a much slower timescale, becoming visible in the gel lanes only after 15 s and increasing to a plateau at times greater than 200 s. Single exponential global fits to the data plotted on both linear (Figure 1) and logarithmic (SI Figure S1) time scales for each strand yield values of k_{obs}, the observed first order rate constant for the formation of MnO_4^--detected complexes at 55 nM excess RNAP: $k_{obs} = 0.010 \pm 0.002$ s^{-1} for the nt strand and $k_{obs} = 0.011 \pm 0.002$ s^{-1} for the t strand. No significant kinetic differences are observed between the different thymines of the two strands in the open region.

Rate constants k_{obs} obtained from MnO4- footprinting assays (Fig. 1) at 19 °C are compared with those obtained previously by filter binding[5] at 20 °C on a plot vs [RNAP] in Figure 2A, and were fit to the expected hyperbolic functional form [5, 34] :

$$k_{obs} = \frac{K_{CC} k_{isom} [RNAP]}{(1 + K_{CC}[RNAP])} \qquad \text{(Eq. 1)}$$

where K_{CC} is the equilibrium constant for forming the CC ensemble from RNAP and promoter DNA, and k_{isom} is the CC-to-OC isomerization rate constant. Values of K_{CC} ((5 ± 1) x 10^7 M^{-1}) and k_{isom} (0.014 ± 0.003 s^{-1}) agree within the uncertainty with those reported previously.[5]

Predicted Time Evolution and [RNAP] Dependence of CC and OC Populations

The kinetic results (K_{CC}, k_{isom}) from the analysis in Figure 2A, obtained in RNAP excess at a low promoter concentration (0.3 nM) allow prediction of two key aspects of the time course of open complex formation for the reactant concentrations in the fluorescence assays reported here (50 nM of each reactant). K_{CC} predicts the extent of conversion of P to {CC} in the rapidly reversible first step of Mechanism 1 and k_{isom} predicts the kinetics of conversion of this rapidly-reversible mixture of CC and P to OC. These predictions, given in Figure 2B on a logarithmic time scale, show that significant OC formation does not occur until approximately 10 s after mixing.

Hence the kinetics of OC formation are predicted to exhibit *two phases*. The first phase (0 – 10 s after mixing for the conditions simulated) is the step-wise, reversible formation of the ensemble of closed complex intermediates ({CC}) from free promoter DNA. All steps of {CC} formation must be readily reversible, because the kinetics of OC formation in assays like filter binding or MnO_4^- reactivity that only detect OC are single-exponential in excess RNAP [35]. A quasi-equilibrium between {CC} and P is therefore established in the first 10 s of the reaction, quantified by the equilibrium constant K_{CC}. Figure 2B shows that for the conditions investigated here (19 ^oC; 50 nM initial concentrations of R and P) approximately half of total promoter DNA exists as {CC} at 10 s, and the other half as un-complexed P. This {CC} ensemble (and previously un-complexed P) convert to OC in the second kinetic phase. The kinetics of the steps of evolution of {CC} in the transient first phase cannot be obtained from studies of OC formation like those in Fig 1, but are determined by fitting of FRET and PIFE kinetic data for these dye positions to the

five-step Mechanism 2 (see below). Figure 2B shows the predicted behavior of P and {CC} vs time in the transient kinetic phase obtained from these fits.

FRET Kinetics of DNA Bending and Wrapping in Closed and Open Complex Formation

Previous equilibrium FRET studies with closed (2°C) and open (19°C) complexes of RNAP and doubly-labeled (Cy3 and Cy5) λP_R promoter DNA revealed that upstream and downstream DNA regions are highly bent and upstream DNA is wrapped on RNAP in both closed and open complexes, reducing the distance between -100 and +14 positions (numbered by convention relative to the +1 start site) from >300 Å before binding RNAP to 60-70 Å in these complexes. To determine the time course and mechanism of these DNA deformations, FRET kinetics experiments were performed at a 1:1 ratio of RNAP: promoter DNA at 19° for each orientation of the donor and acceptor dyes. Representative FRET acceptor time-courses for each dye orientation, plotted on a logarithmic time scale, are shown in Figure 3. The FRET kinetics span a time range of more than four orders of magnitude, from ~20 ms to at least 400 s. As predicted in Figure 2B, two kinetic phases are observed. The first phase, in which approximately 15% of the FRET increase occurs, ends about 10 s after mixing, and the slower second phase develops from 10 s to 400 s. By comparison with the simulations of Figure 2B, these two kinetic phases clearly correspond to reversible formation of the {CC} ensemble from reactants, followed by conversion of this {CC} ensemble and free promoter DNA to OC. The FRET inflection point in Fig. 3 in the time range 1-10 s, when ~50% of promoter DNA has formed a closed complex with RNAP (Fig 2B), is only 15% of the long-time (400 s) value, when ~80% of promoter DNA is in an OC (Fig 2B). This indicates that the average closed complex in {CC} is much less bent and wrapped than the final OC.

Increases in RNAP Concentration and Addition of Glycine Betaine Increase the Transient Population of Closed Complexes Exhibiting FRET in the First Kinetic Phase

Rapid equilibrium between CC and free promoter DNA was previously deduced from the single-exponential kinetics of OC formation in filter binding kinetic assays.[5] To verify that the first FRET kinetic phase in Figure 3 corresponds to formation of an ensemble of bent and wrapped CC in rapid equilibrium with free promoter DNA as in Fig. 2B, higher RNAP concentrations and a CC-stabilizing solute (glycine betaine, GB) were used to increase the equilibrium extent of CC formation.

Figure 4A show representative FRET experiments where the concentration of RNAP was increased by 1.5- and 2-fold. The FRET signal late in the first phase (10 s), where an equilibrium mixture of CC and unbound promoter DNA is predicted (Figure 2B), increases by 35% upon increasing the stoichiometric ratio of active RNAP:DNA from 1:1 to 2:1. This is similar to the 40% increase in the fraction of CC predicted from the simulation in SI Figure S2. Likewise, Figure 4A shows that OC formation is complete at 400 s for experiments at 100 nM RNAP, 50 nM DNA but not complete at 400s for experiments at 50 nM RNAP, 50 nM DNA, also consistent with the simulation.

Figure 4B shows the effects of increasing the GB concentration to 1.5 M. The FRET signal late in the first phase (10 s), where an equilibrium mixture of CC and unbound promoter DNA is expected, increases by ~50%, while the time course of conversion of CC to OC (t > 10 s) is similar in the presence and absence of GB. These results are qualitatively consistent with previous filter-binding kinetic studies which found that glycine betaine (GB) at molar concentrations increases K_{cc} and reduces k_{isom}, stabilizing {CC} in comparison to reactants and reducing the rate of isomerization of {CC} to OC [36]. Simulations of the time-evolution of CC and OC at the RNAP and promoter concentrations of Fig 4, based on the published rate and equilibrium constants as functions of GB concentration[46], are shown in Fig S4. Figure S4A predicts a ~40% increase in the rapidly-equilibrated fraction of CC in the time range 0.3 – 10 s upon addition of GB to 1.5 M,

consistent with the observation in Fig 4B. Likewise, Figure S4B predicts very similar time courses for conversion of the CC ensemble and free promoter DNA to OC in the slower kinetic phase (10-400 s) in the presence and absence of GB, consistent with observation in Fig 4B.

RNAP-Induced Fluorescence Enhancements (PIFE) at Promoter Positions -100 and +14 Both Appear Early and Decrease Later in {CC} Phase and In OC Formation

Single-dye kinetics experiments (Figure 5) show large increases in Cy3 and Cy5 fluorescence (PIFE effects) at both -100 and +14 positions of λP_R promoter DNA. -100 and +14 PIFE both develop on the same time scale as observed for FRET (t > 10 ms). Hence, RNAP must contact these dye positions early in the process of forming the {CC} ensemble, as part of the process of bending and wrapping upstream and downstream promoter DNA on RNAP. Position +14 is 25 bp downstream of the footprint boundary for RP_c typically -5 [3 4 1] but within the +20 downstream boundary of the footprint observed for more advanced CC [9 13 5 37] and OC [9 13 37]. Position -100 is 20-40 bp upstream of footprint boundaries of all RNAP-λP_R promoter complexes. The Discussion considers the implications of these contacts for the operation of the RNAP-promoter biophysical machinery that opens the β-β' clamp and cleft, bends the downstream duplex into the clamp, and opens 13 base pairs using binding free energy.

Kinetic information regarding involvement of the downstream DNA in steps of OC formation is provided by +14 PIFE, as shown in Figure 5A. Starting at 20 ms after mixing, PIFE at +14 increases to a broad plateau near 10 s and then decreases modestly at times >20 s as the CC ensemble converts to OC, indicating that intrinsic +14 PIFE signal intensities of I_{1E} and I_{1M} somewhat exceed those of I_{1L} and RP_o. . No significant differences between +14 PIFE effects for Cy3 and Cy5 are observed at short times (<10s), while small but reproducible differences are observed at longer times (>10s). In FRET experiments with Cy3 at +14 or -100 positions, the

PIFE effect is dominant in the first kinetic phase (0.01 to 10 seconds) and donor quenching is not observed (Supplemental Figure 4).

Kinetic information provided by -100 PIFE regarding involvement of the far-upstream DNA in steps of {CC} and OC formation is shown in Figure 5B. Two kinetic phases are observed, as in FRET assays (Figs. 3, 4). As RP_c converts to more advanced closed complexes (first I_{1E}, then I_{1M}; 10 ms – 3 s), -100 PIFE increases, showing that this far-upstream position is contacted by RNAP in these intermediates. The initial strong increase is followed by a strong decrease in -100 PIFE, starting at 3 s. This decrease begins late in the first kinetic phase ({CC} formation) defined by the FRET data and FB predictions (Fig. 2), and continues in the second kinetic phase of conversion of {CC} to OC. This indicates that contacts between -100 and RNAP are less strong in the most advanced CC (I_{1L}) and in OC than in earlier CC (I_{1E}, I_{1M}). The time evolution of these -100 PIFE effects is interpreted together with FRET and +14 PIFE effects in Discussion. Late in the reaction, -100 PIFE is less reproducible than +14 PIFE, decaying in different experiments to between 0-50% of the maximum. In most cases, Cy3 -100 PIFE signal approaches zero more closely than for Cy5. These slow, less-reproducible effects may result from slow binding of inactive RNAP to the upstream end of promoter DNA, which is made accessible in the conversion of {CC} to OC when the upstream boundary of the HO footprint shrinks from -82 (nontemplate strand)/-81 (template strand) to approximately -64 (nontemplate strand)/-59 (template strand) [9].

Fluorescence Characterization of Unwrapping of Upstream Promoter DNA from RNAP Early in Response to a Salt Upshift and in Transcription Initiation

If the upstream promoter DNA is bent and wrapped around RNAP in forming a stable OC unwrapping should occur by destabilizing it or on its escape from the promoter. Unwrapping of promoter DNA from the stable OC and concomitant loss of FRET was investigated in two ways. The stable OC was destabilized by a large upshift in KCl concentration to the range 0.4 -

1.1 M. As Supplemental Figure S7 show, loss of FRET by unwrapping following a large salt upshift occurs rapidly, faster than the salt-independent rate of DNA closing and indicating that salt-induced unwrapping occurs from an open complex, probably by disrupting the far-upstream wrap. In these experiments the loss of +14 PIFE is slower than the loss of FRET though faster than the rate of DNA closing (SI Figure S7). This confirms that the rapid loss of FRET after a salt upshift is from upstream unwrapping.

Unwrapping was also induced by addition of NTP to initiate transcription of a 14-mer RNA; Supplemental Figure S8-S10 shows that the fractional loss of FRET in initiation corresponds to the fraction of the OC population that are capable of synthesis of a 14-mer full length RNA transcript.

Discussion

A Five-Step Mechanism of OC Formation with a Key Intermediate CC Not Previously Observed by DNA Footprinting

A basic principle in deducing mechanism from kinetics is to start with the simplest mechanism (e.g. two step Mechanism 2) and to introduce more intermediates and steps only as warranted. The kinetics of OC formation, determined by MnO_4^- footprinting (in Fig 1) and by filter binding, are well-fit by a single exponential and well described by a two-step mechanism. On the other hand, four exponentials with decay rate constants of spanning four orders of magnitude (20 s^{-1} to 0.02 s^{-1}) are needed to fit the FRET and PIFE kinetic data, as described in SI. It is therefore not surprising that a five-step mechanism (Mechanism 2) is needed to fit all the FRET acceptor and PIFE kinetic data using a common set of rate constants. We could not get a common set of rate constants that could globally fit well to all types of FRET, PIFE+14 and -100 data for a four-step mechanism.

Fits of the five step mechanism to the FRET and PIFE kinetic data of Figs. 3 and 5 using sets of rate constants that are constrained to be the same within the uncertainty for all FRET and PIFE data sets are shown in Figure 7 panels A-C. Fits to other data sets are shown in SI. Rate constants obtained from these fits are given in Table 1, and a simulation of the time-evolution of {CC} and of OC for these reactant concentrations and conditions is given in Figure 6. Intrinsic FRET acceptor and PIFE signal intensities for intermediates, relative to a reference state (RP_o for FRET, P for PIFE), are given in Table 2. FRET acceptor signals of the intermediates in {CC} are interpreted in terms of distances between dyes at -100 and +14 in Table 3.

Consistent with the 4-exponential fits, the sums of rate constants for each step in Mechanism 2 are approximately 20 s^{-1}, 4 s^{-1}, 0.2 s^{-1} and 0.04 s^{-1}. Consistent with the filter binding and MnO_4^- results of Figure 2, equilibrium constants for the steps of closed complex formation combine (see below) to predict a composite K_{cc} = 5.2 x 10^7 M^{-1} and k_{isom} = 0.014 s^{-1}.

Fitted rate constants in Table 1 predict that the initial specific complex (RP_c, with a typical footprint from -60 to -5 (refs)) forms in the time range <10 ms - 40 ms at the concentrations of R and P investigated (see Figure 6). the concentration of RP_c is a maximum at 40 ms). The second order rate constant k_1 ((3.1 $\pm$ 0.2) x 10^8 M^{-1} s^{-1}) is about 5% of that predicted for the diffusion-collision limit, presumably because only one in about twenty collisions leads to complex formation. This k_1 is consistent with published rate and equilibrium constants for the first detected complex at other promoters [38 22 21]. No significant FRET or PIFE signal is observed in this time range, and indeed none is expected for RP_c because the downstream promoter DNA is unbent and the probe positions (-100, +14) are not contacted. The modest RP_c binding constant (7 x 10^6 M^{-1}) means that only 10-15% of promoter DNA forms RP_c at these early times in the reaction (Figure 6). The RP_c population remains in this range for the first 1 s of the reaction and then slowly declines as the advanced CC and OC form.

Analysis of the FRET and PIFE kinetic data, summarized in Tables 1 and 2) provides evidence for at least three kinetically significant more advanced closed complexes than RP_c (designated I_{1E}, I_{1M}, I_{1L}) before the DNA-opening step. The rate constant for dissociation of RP_c (~ 44 s^{-1}) is much larger than that for formation of I_{1E} (~ 13 s^{-1}), indicating that formation of RP_c from R and P is rapidly reversible on the time scale of its conversion to I_{1E}. I_{1E} forms in the time range 0.03 – 0.3 s (Figure 6), and is somewhat favored thermodynamically over RP_c (K_2 = 1.8 (Table 1) and ΔG^o_2 = - 0.3 kcal/mol). The maximum I_{1E} population (about 20% of the total promoter population) occurs at 0.3 s and I_{1E} decreases gradually at longer times as more advanced CC and OC form. I_{1E} exhibits large -100 and +14 PIFE intensities as well as significant FRET between these positions (Table 2). The FRET signal of I_{1E} corresponds to a distance of ~80 A between these far-upstream and downstream positions (Table 3), demonstrating that both upstream and downstream promoter DNA must be bent and wrapped on RNAP (see following section).

The rate constant for dissociation of I_{1E} (~ 7 s^{-1}) is significantly larger than that for formation of I_{1M} (~ 2.5 s^{-1}), indicating that the RP_c - I_{1E} step is rapidly reversible on the time scale of conversion of I_{1E} to I_{1M}. I_{1M} forms on a time scale of 0.2 – 2 s (Figure 6), and is somewhat favored thermodynamically over I_{1E} (K_3 = 1.5 (Table 1) and ΔG^o_3 = - 0.2 kcal/mol). The maximum I_{1M} population (also about 20% of the total promoter population) occurs at 2 s and I_{1M} decreases gradually at longer times as more advanced CC and OC form. I_{1M} exhibits -100 and +14 PIFE intensities and a FRET signal between these positions which are the same as I_{1E} (Table 2), indicating no change in the bending, wrapping, and far-upstream (-100) and downstream (+14) contacts of promoter DNA.

The rate constant for dissociation of I_{1M} (~ 2 s^{-1}) is significantly larger than that for formation of I_{1L} (~ 0.1 s^{-1}), indicating that the I_{1E} - I_{1M} step is rapidly reversible on the time scale of conversion of I_{1M} to I_{1L}, on a time scale of 3 – 30 s (Figure 6), and is slightly favored thermodynamically over

I_{1M} (K_3 = 1.1 (Table 1) and ΔG^o_3 = - 0.1 kcal/mol). With the formation of I_{1L} the {CC} is complete, with all intermediates in quasi-equilibrium on the time scale of the subsequent DNA opening step ($1/k_5$ = 25 s). The maximum I_{1L} population (about 15% of the total promoter population) occurs at about 20 s and I_{1L} decreases gradually at longer times as it slowly converts to OC in the DNA-opening step (rate constant k_5 = 0.04 s^{-1}, about three times larger than the isomerization rate constant k_{isom} of Mechanism 1). PIFE intensities at -100 and +14 are smaller for I_{1L} than for I_{1M} or I_{1E} -100 and +14 PIFE, but the FRET efficiency between these positions in I_{1L} is significantly greater than for I_{1M} or I_{1E}. The increased FRET efficiency of I_{1L} shows that the distance between +14 and -100 probes has decreased from 80 Å in I_{1M} and I_{1E} to 70 Å in I_{1L} (Table 3).

Mechanism of Operation of the RNAP-Promoter Biophysical Machine: How Bending and Wrapping Upstream and Downstream DNA Facilitate Opening of the RNAP Clamp and Entry of the Downstream Duplex

The most striking finding from analysis of the FRET and PIFE kinetic data in terms of the five-step mechanism is that both FRET and strong upstream and downstream PIFE develop in the conversion of RP_c to I_{1E}. Both upstream and downstream duplex DNA must therefore be extensively bent and wrapped on RNAP in this step, which occurs reversibly on a 0.1 s time scale, which is fast by comparison to the subsequent step and is deduced to be slightly favorable ($K_{RPc \rightarrow I1E} \sim 2$) To interpret these concerted upstream and downstream effects, we propose that the contacts made in RP_c between RNAP and -35 and/or UP elements of promoter DNA initiate strong upstream bending, perhaps initiated or accompanied by contraction/folding of the flexible tethers linking α-CTD to α-NTD. Upstream bending allows far-upstream wrapping, which allows the -100 position of the promoter to be contacted by a mobile downstream element (DME) of RNAP, as proposed previously [9 6 1]. If only the upstream DNA but not the downstream DNA were bent and wrapped, the distance between -100 and +14 probe positions would be about 135 Å

(see Fig S11), much too large to exhibit FRET. Therefore bending and wrapping of upstream DNA must trigger bending of the downstream duplex, presumably at the junction between the -10 and spacer regions, reducing the distance between +14 and -100 positions on promoter DNA to 80 A to give FRET.

A possible structural model for I_{1E} is provided in Fig. 8, based on previous models proposed for wrapped CC and OC and PDB structure 4YLN [7]. The requirements are that the -100 to +14 dye-dye distance be 80 Å and that the +14 position of downstream duplex DNA be in contact with RNAP. These requirements are met if the downstream duplex is bent at the upstream end of the -10 region to bring +14 in contact with the top of the β lobe. This model utilizes the OC conformation of these regions of RNAP and may not be accurate for I_{1E}, but is similar in its placement of the downstream duplex to the CC1 pol II intermediate observed by cryoEM. Moreover, this proposal is not inconsistent with the observation of a downstream boundary of +2/+7 in DNase footprinting of {CC} formed with UT-47 LPR promoter DNA, where I_{1E} may be the dominant species as discussed in the next section, because the +14 contact appears to be an isolated contact which might be displaced by DNase, as happens in the far-upstream wrapped region. Movement of RNAP downstream mobile elements (DME) to contact far upstream region of the promoter (near -100) as part of the process of making the clamp and cleft accessible to the downstream duplex was proposed previously [9].

Interconversion of I_{1E} and I_{1M} is significantly slower (occurring on a 1s time scale) and exhibits little change in FRET or -100 PIFE, while +14 PIFE we propose it is a movement in the β' portion of clamp perhaps accompanied by movement of $\sigma_{1.1}$ that must precede entry of the downstream duplex in the subsequent step. Rate constants for opening and closing the RNAP clamp are in the range of 1.0 to 1.7 s^{-1} at 22^0C [39]. We propose that sigma region 1.1 is cleared

from the cleft in conversion from I_{1E} to I_{1M} on clamp opening to accommodate the downstream DNA positioned outside the active site cleft.

On clearance of the in-cleft elements the downstream DNA is then bent into the active site cleft forming closed complex I_{1M} intermediate. This is similar to the DNA distorted 7 Å deeper into the cleft in CC^{dist} intermediate formed preceding DNA opening in pol-II eukaryotic system [8]. The distance between the downstream +14 and upstream -100 position is reduced from 80 Å to 70 Å on conversion of I_{1M} to I_{1L} achieved by swinging the downstream DNA into the active site cleft and bringing the+14 position closer to upstream -100 position by 10 Å (Figure 8).

Once the DNA is placed in the active site cleft in I_{1L} the DNA is opened with simultaneous clamp closing. The DNA opening step (and/or subsequent rapid steps that stabilize the initial open complex) involves an additional FRET increase, with a small accompanying reduction in +14 PIFE and the elimination of any remaining -100 PIFE. This FRET increase in OC formation is anticipated based on the previous comparison of equilibrium FRET signals for {CC} and OC, and indicates a 7 Å reduction in distance between -100 and +14 to 63 Å in the stable OC (RP_o).

Insights from the Five-Step Mechanism into Origins of Large Effects of Upstream DNA Truncation, Promoter Sequence and Upstream-binding Factors on Kinetics of OC Formation

Previously, effects of upstream DNA truncation, promoter sequence and upstream-binding factors on the kinetics of OC formation were interpreted using 2-step Mechanism 1 as effects on the closed complex binding constant K_{CC} and/or the isomerization rate constant k_{isom}. Interpretation of these effects using the relationship between the parameters of the two and five step mechanisms provides insight into how promoter sequence, upstream truncation and factors regulate the rate of OC formation and transcription initiation.

Comparison of Mechanism 2 with Mechanism 1 shows that

$$k_{isom} = f_{I1L} k_5 = f_{I1L} k_{open}$$

where f_{I1L} is the fraction of the fully-formed, rapidly-reversible {CC} (for t > 10 s in our experiments; see Fig 6) that is I_{1L}:

$$f_{I1L} = ([I_{1,L}]/\{[I_{1,L}] + [I_{1,M}] + [I_{1,E}] + [RP_c]\} = K_{I1E} K_{I1M} K_{I1L}/(1 + K_{I1E}(1 + K_{I1M}(1 + K_{I1L}))$$

and

$$K_{CC} = \{([I_{1,L}] + ([I_{1,E}] + ([RP_c]\}/[R][P] = K_{RPc}/f_{RPc}$$

where f_{RPc} is the fraction of the fully-formed {CC} that is RPc:

$$f_{RPc} = 1/(1 + K_{I1E}(1 + K_{I1M}(1 + K_{I1L})).$$

The overall second order rate constant for the association direction of OC formation (k_a) is interpreted as

$$k_a = K_{CC} k_{isom} = K_{RPc} K_{I1E} K_{I1M} K_{I1L} k_{open}$$

Effects of promoter sequence, upstream DNA length and transcription factors on k_a, K_{CC} and/or k_{isom} therefore can arise from effects of these variables on the distribution of complexes in {CC} (i.e. on f_{RPc} and/or $f_{I1L,}$ which of course are not independent) as well as on the probably-more independent quantities K_{RPc} (as a determinant of K_{CC}) or k_{open} (as a determinant of k_{isom}).

The isomerization rate constant k_{isom} (Mechanism 1) is expected to be smaller and more promoter-specific than k_{open} because of promoter-specific differences in the distribution between early (e.g. RP_c) and advanced (e.g. I_{1L}) members of {CC}. Indeed k_{open} may be insensitive to promoter identity and primarily a function of temperature, as judged by the behavior of k_{close}, the DNA closing rate constant for the reverse direction of this step, which is insensitive to promoter identity and to salt or solute concentration.

Regulation of the velocity of product formation in enzyme-catalyzed reactions is typically achieved using inhibitors or activators that primarily affect the equilibrium constant(s) for reversible binding of the substrate(s), and not the catalytic rate constant. Likewise, effects of promoter sequence and transcription factors on the rate of OC formation appear to be largely on equilibrium constants for the steps that form RP_c and/or convert RP_c to more advanced members of the CC ensemble (designated as I_{1E}, I_{1L}). These reversible steps are analogous to reversible substrate binding steps. In a small number of comparisons to date, promoter sequence and solute or salt concentration have only small effects on the kinetics of DNA closing and opening steps, which are the analogs of the reverse and forward catalytic steps of this process.

Activation by Upstream DNA. Upstream truncation of λP_R promoter DNA at position -47 (UT-47) reduces k_{isom} to 1-2% of its value on a full-length (FL; to -100) promoter fragment. Similarly large effects were observed for other upstream truncations of the lacUV5 promoter. K_{CC} is either unaffected or increased somewhat (~2-fold) by upstream truncation. Hence the presence of upstream DNA (-40 to -80) affects the kinetics of OC formation similarly to a strong Type II transcription factor. The 100-fold range of effects of upstream DNA length on k_{isom} of these promoters is comparable to that observed between fast- and slow-isomerizing promoter sequences and exceeds the effects of many upstream modulators. The proposed molecular explanation of these upstream-truncation effects is that the UP element and far-upstream DNA bend and wrap on RNAP as the CC population advances and that this facilitates formation of the most advanced member of the CC population, in which the downstream duplex is bent in the cleft, poised to be opened by RNAP. Precedent for this proposal was provided by previous footprinting and AFM compaction studies revealing that the upstream region of promoter DNA (from approximately -38 to -65) is bent and wrapped on RNAP in the stable OC [40 41] and also in an advanced CC investigated at equilibrium by low-temperature footprinting.

The above analysis demonstrates that upstream DNA by itself behaves as a Class II transcription activator protein, acting upstream to increase k_{isom} greatly without much effect on K_{CC}. Addition of upstream DNA from -47 to -100 increases k_{isom} by facilitating the steps that convert I_{1E} to I_{1L}, as previously proposed, and thereby increasing the fraction of {CC} that is I_{1L} by 10-100 fold. Transcription activators/repressors that affect k_{isom} achieve by affecting the fI_{1L} and not the k_{open} rate constant, the ones that affect K_{CC} can do so by affecting K_{RPc} and/or f_{RPc}.

Conclusion

The series of steps by which DNA information directs the operation of the RNAP biophysical machinery to bend and wrap upstream promoter DNA, thereby facilitating bending of the downstream duplex into the clamp and cleft of RNAP and triggering of DNA opening, are determined here by stopped-flow fluorescence kinetic studies with the λP_R promoter. Kinetic and thermodynamic information about each intermediate is obtained. Formation of initial specific closed complex (RP_C) should be rapid (10 ms time scale) and reversible for the concentrations investigated but is not detected, indicating the DNA is not sufficiently bent to give FRET or PIFE for probes at -100 and +14. Large FRET and PIFE increases accompany reversible formation of the key closed intermediate ($I_{1,E}$) in the 0.1 s time range. These fluorescence effects demonstrate that the upstream DNA must fully bent and wrapped and the downstream DNA partially bent to bring the two probe positions to an 80 Å separation distance, in which the downstream duplex contacts the top of the RNAP β-lobe, and DME move to contact -100 as previously proposed. On a 1 s time scale, $I_{1,E}$ converts to a more advanced closed complex ($I_{1,M}$) with similar PIFE and FRET signals, indicating that this slow step is a large conformational change in RNAP, which we propose must be the movement of the β' clamp to open the cleft, perhaps accompanied by clearance of $\sigma_{1.1}$ from the cleft. Finally, in the 10 s time range, the downstream DNA (+14) moves 10 A closer to the upstream (-100) DNA as $I_{1,M}$ converts reversibly to the most advanced closed

complex ($I_{1,L}$). Clearly this represents the entry of the downstream duplex into the open clamp and cleft. The clamp and cleft must subsequently close as the DNA is opened by binding free energy in the subsequent slow (100 s time scale at 19 ^{o}C)) reversible conversion of I_{1L} to the first unstable open intermediate (I_2). Stabilizing downstream interactions involving the DME then rapidly and irreversibly convert I_2 to the stable OC (RP_o). This mechanism of OC formation provides insights into why the mere presence of upstream DNA is such a transcription activator, and into how upstream-binding transcription activators and repressors work by affecting the amount of CC formation and/or composition of the {CC} ensemble.

Figures

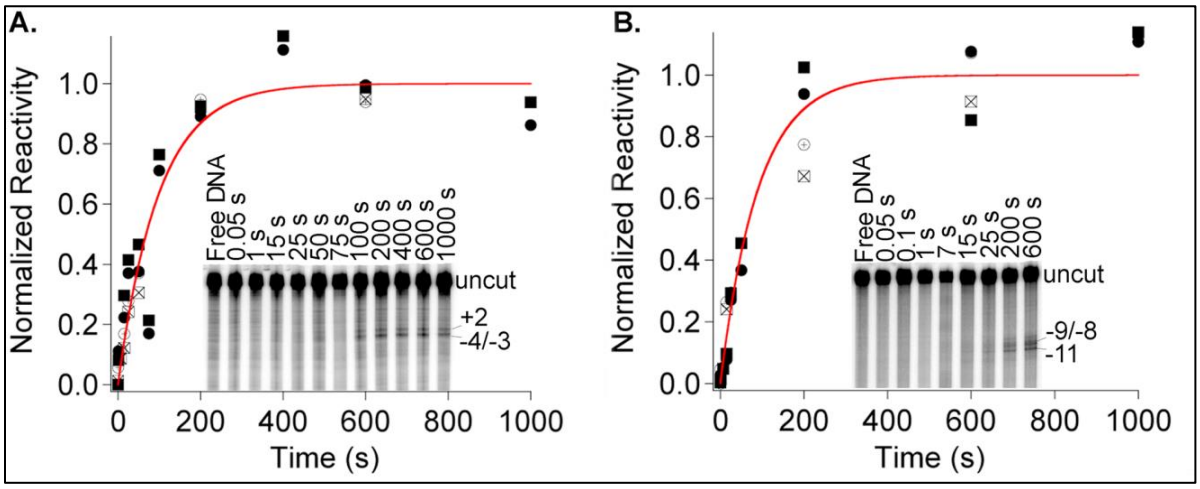


Figure 1 Kinetics of OC Formation Monitored by MnO_4^- Reactivity. Fast (0.15 s) MnO_4^- snapshots monitor the time course of opening individual thymines in OC formation after mixing excess RNAP (55 nM) with λP_R promoter DNA (0.3 nM) at 19 ºC. Representative gels are provided as insets. Kinetics of development of MnO_4^- reactivity are plotted for thymines on the nt strand (+2 (⊕, λ) and -4/-3 (⊠, ν) in panel **A** and for the t strand (-9/-8 (⊕, λ) and -11 (⊠, ν) in panel **B**. Rate constants k_{obs} for OC formation from these fits are the same within uncertainty (template strand k_{obs} = 0.010 ± 0.002 s^{-1} ; non-template strand k_{obs} = 0.011 ± 0.002 s^{-1}). Times indicated are times after mixing RNAP with DNA at which the 0.15 s MnO_4^- snapshot was initiated. Gels were quantified and results normalized as described in Materials and Methods.

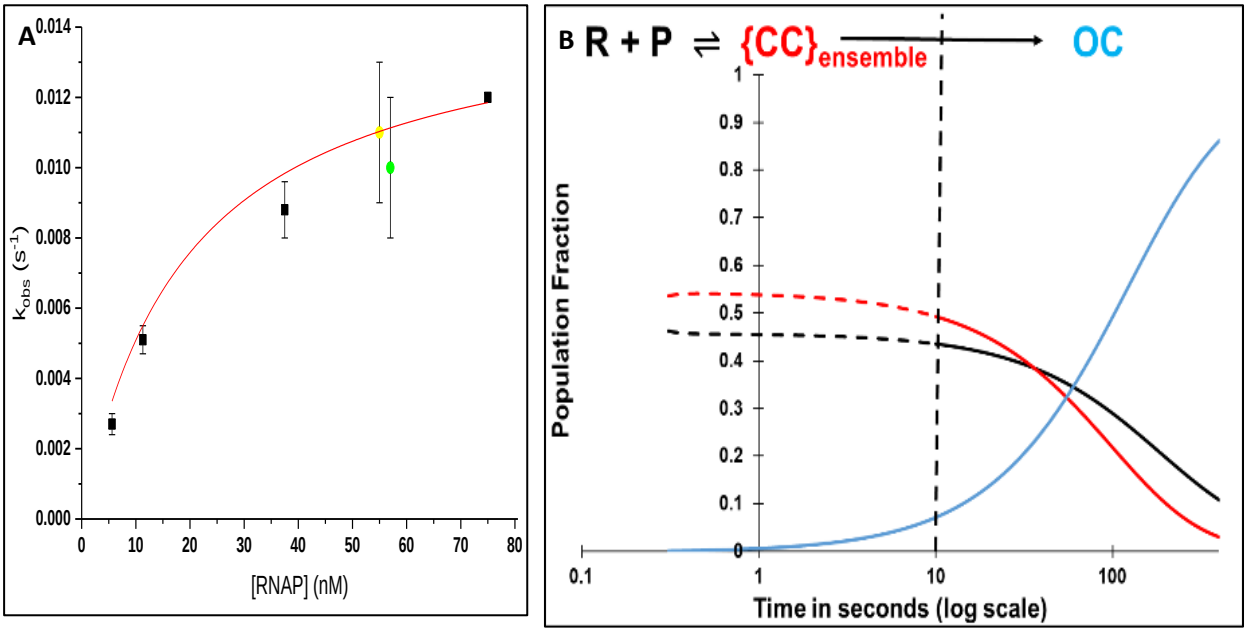


Figure 2. Panel A: Dependence of Rate Constant k_{obs} for Open Complex Formation at λP_R Promoter on RNAP Concentration. First order rate constants k_{obs} (Equation 2) for OC formation in excess RNAP plotted as a function of active [RNAP]**.** ■ : k_{obs} values from filter binding assays at 20ºC (Saecker et al JMB '02). ● , ● : k_{obs} values from MnO_4^- footprinting of template and non-template strands, respectively at 19 ºC (Fig 1). **Panel B: Predicted Time Evolution of {CC} and OC Formation for 1:1 ratio of RNAP to λP_R Promoter (50 nM each).** Predicted fractional populations of open complexes (OC) (blue line), closed complexes {CC} (red line), and free promoter DNA (black line) as a function of time (log scale) covering the range from 0.3s to 400 s. The dashed vertical line at 10 s marks the onset of OC formation from the equilibrium mixture of {CC} and free promoter DNA (see text).

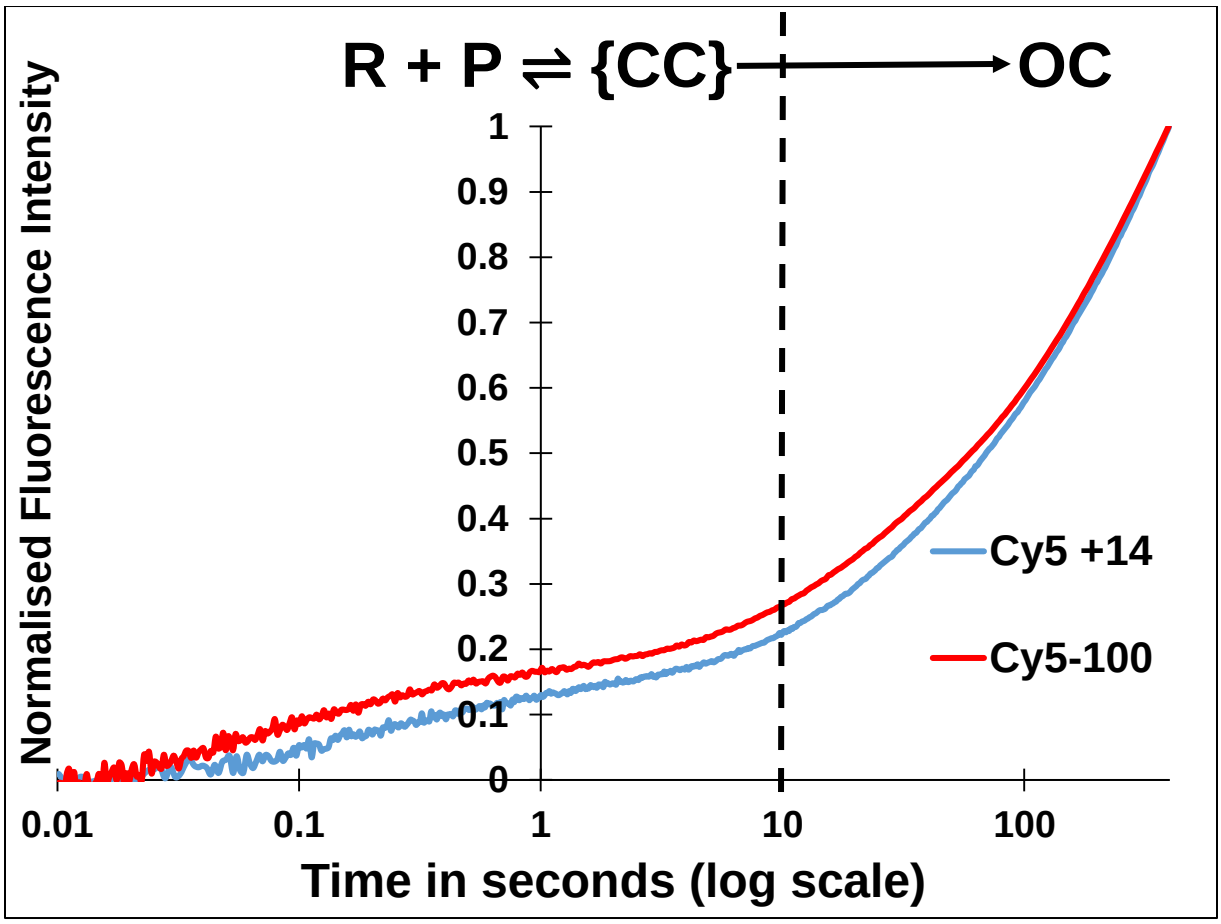


Figure 3 FRET- detected promoter DNA Bending and Wrapping by RNAP in formation of the CC ensemble and the stable OC Formation. Representative time course (log scale) of normalized FRET acceptor (Cy5) emission intensity after mixing Cy3-Cy5 dye-labeled λP_R DNA (50 nM final) with *E.coli* RNAP (50 nM final) at 19^{o}C and exciting FRET donor (Cy3) at 515nm. The vertical dashed line at 10 s corresponds to the predicted onset of OC formation at these concentrations (see Figure 3 below). Curves shown in the figure are the average of a series of normalized shots in one experiment, representative of 14-15 replicates done on different dates at same RNAP and promoter solutions (see Methods and Supplemental Table 3). Results of experiments with both variations of dye positions are shown, with Cy5(+14) Cy3(-100) represented in blue and Cy3(+14) Cy5(-100) represented in red. In most cases fluorescence intensities in the early time phase (t < 10 s) are somewhat larger when Cy5 acceptor is at -100 than at +14.

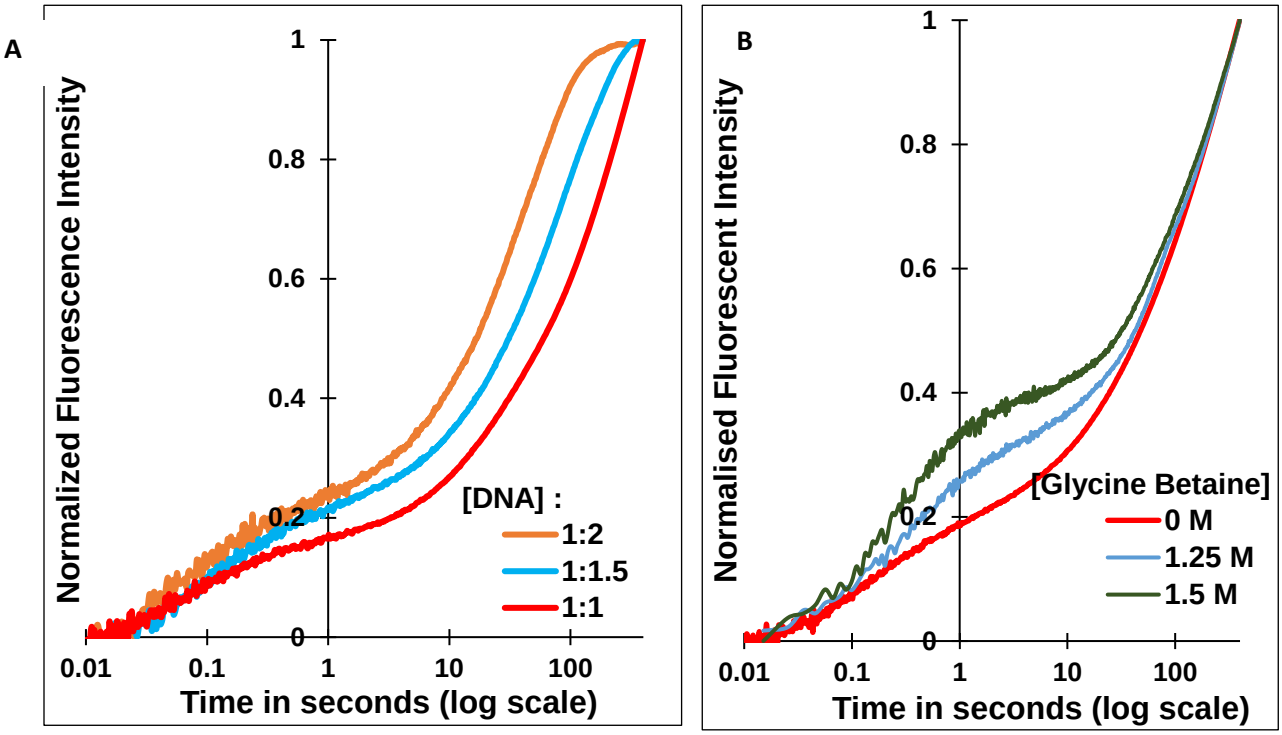


Figure 4 Effects of A) RNAP and B) Glycine Betaine Concentration on FRET-observed kinetics of OC Formation. Representative time courses of normalized FRET acceptor (Cy5) emission intensity after mixing Cy3(+14) Cy5(-100) dye-labeled λP_R DNA (50 nM final) with *E.coli* RNAP and excitation of FRET donor (Cy3) at 515nm at 19°C. Curves shown in the figure are the average of a series of normalized shots in one experiment, that is representative of various replicates done on different dates at same RNAP and promoter solutions (see Methods and Supplemental Table 3). **A)** Final concentrations of 50 nM DNA mixed with final concentrations of 50 nM (red line), 75 nM (light blue line) and 100 nM (orange line) of RNAP. **B)** FRET time courses after mixing Cy3(+14) Cy5(-100) λP_R DNA (50 nM final) with RNAP (50 nM final) in the presence of 0 -1.5 M glycine betaine.

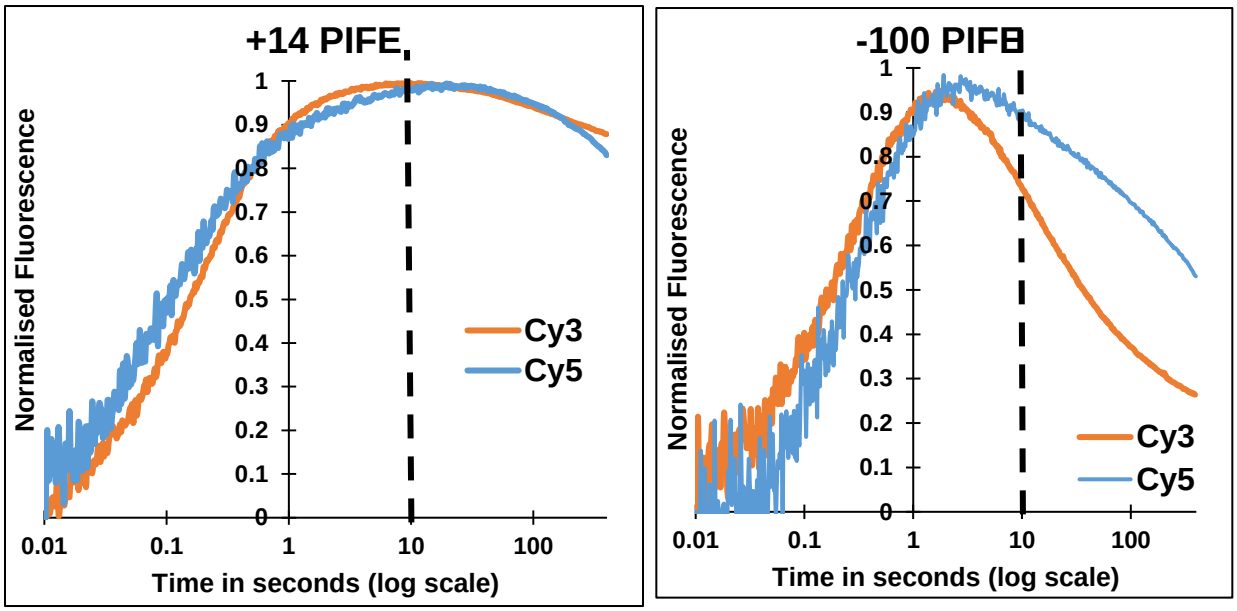


Figure 5 Time Courses of Fluorescence Enhancement (PIFE) from RNAP Interactions with -100 and +14 positions of λP_R promoter DNA. Representative time courses (log scale) of normalized single-dye-labeled λP_R DNA (50 nM final) PIFE-induced emission at +14 position **(Panel A, left)** and -100 position **(Panel B, right)** with *E.coli* RNAP (50 nM final) at 19 ºC, exciting Cy3 at 515 nm and Cy5 at 610 nm wavelength. The vertical dashed line at 10 s is the predicted onset of OC formation as described (see Figure 3B). Curves shown in the figure are the average of a series of normalized shots in one experiment, representative of 3-5 replicates done on different dates at same RNAP and promoter solutions (see Methods and Supplemental Table 3).

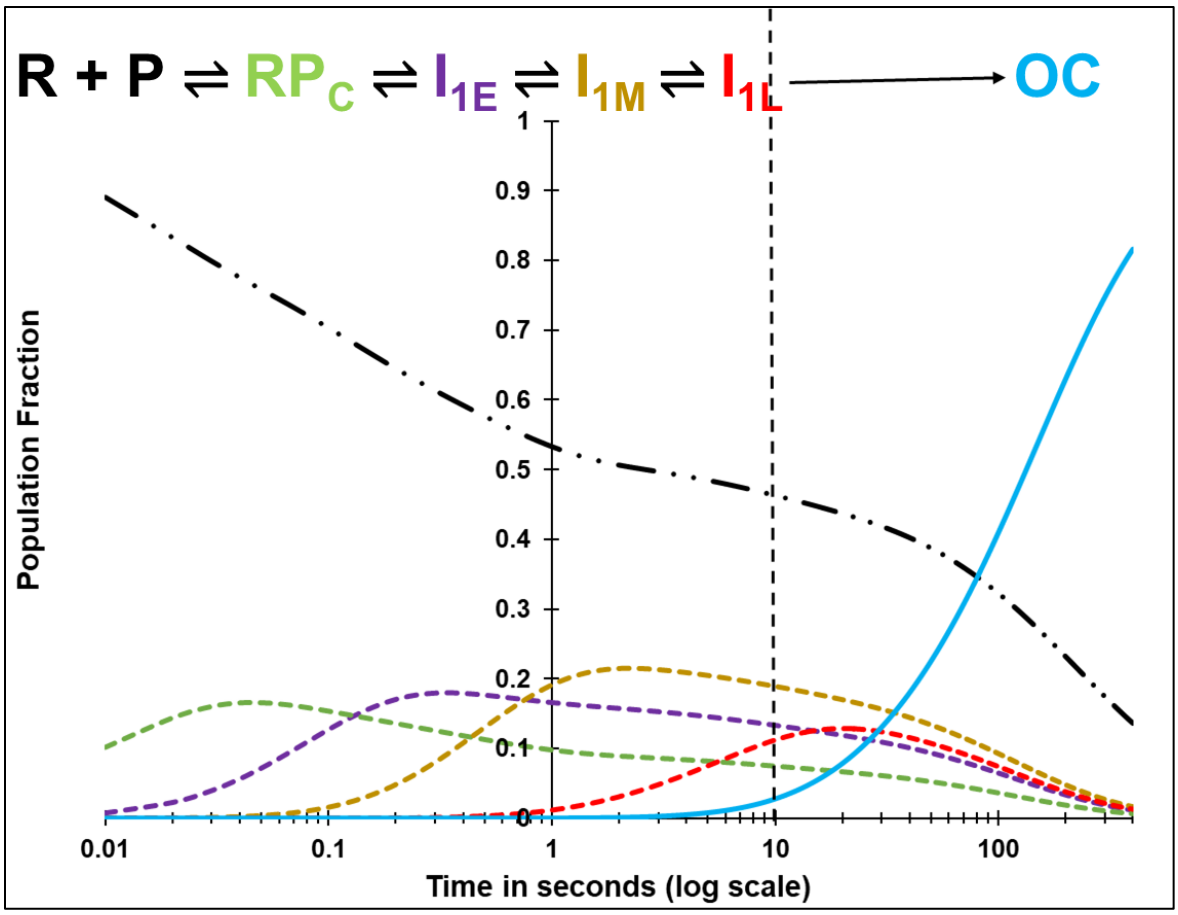


Figure 6 Time evolution of Closed Complex Intermediates, free promoter DNA and OC in a multi 5 step mechanism of OC formation. Simulation curves of population fraction of free promoter DNA (black dashed), closed complex intermediates: RP_C (green dashed), $I_{1,E}$ (purple dashed), $I_{1,M}$ (light brown), $I_{1,L}$ (red dashed) and OC (blue solid colored line) concentration from 0.01 to 400 seconds following a 5 step mechanism. The kinetic rates of the mechanism are chosen from Table S4 following the conditions described in SI Method section to be consistent the kinetics results obtained for a 2 step mechanism of OC formation by filter binding assay. The dashed vertical line at 10 s marks the onset of OC formation from the equilibrium mixture of {CC} and free promoter DNA.

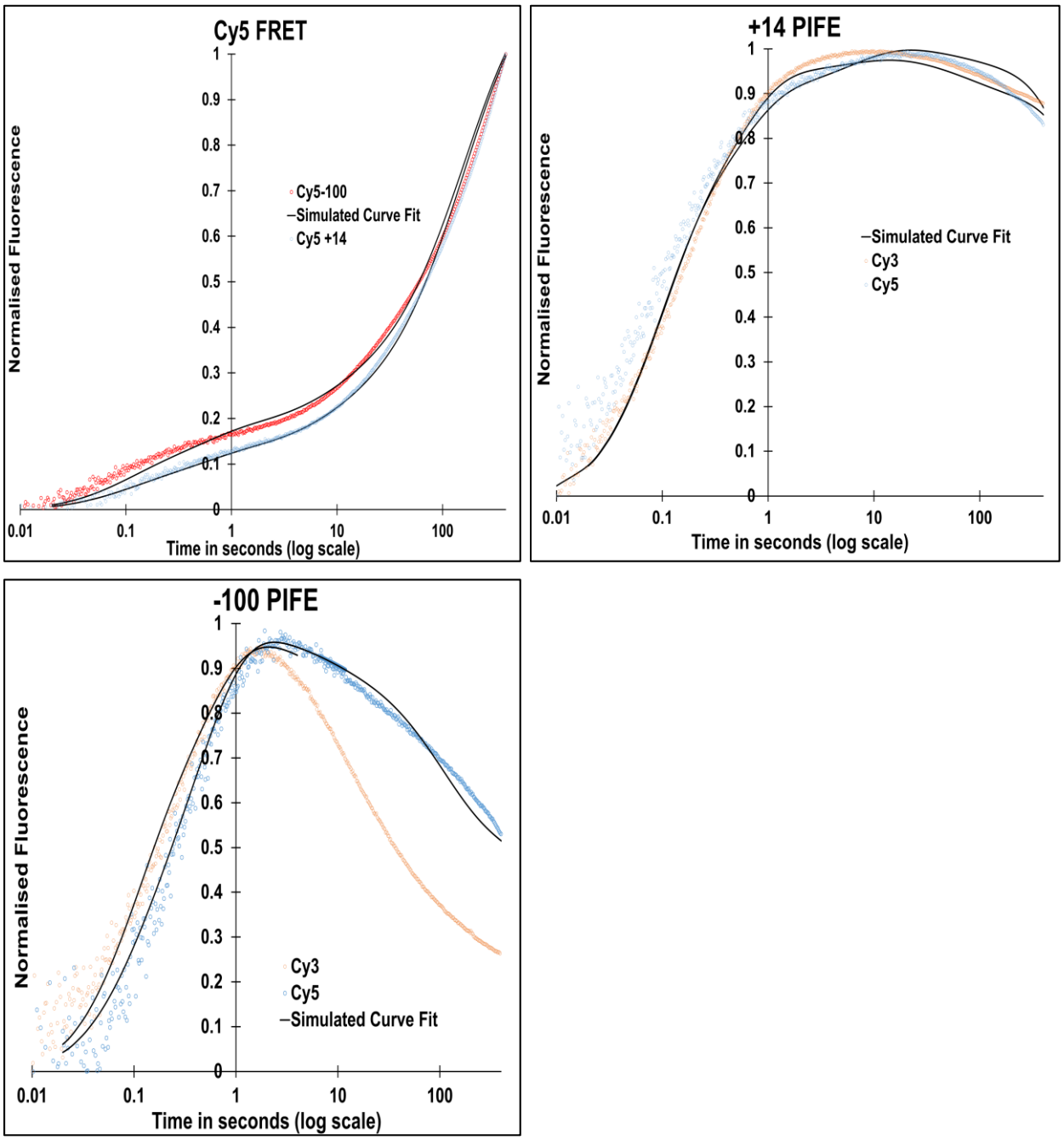


Figure 7 Fitting OC Fluorescence Kinetics FRET and PIFE representative experiments to a 5 step OC formation mechanism shown in Figure 6. Simulated curve (black line) is obtained from the rate constants and signal intensities given in Table 1 and 2 following a 5-step OC formation mechanism shown in Figure 6. The curve is fit to FRET acceptor Cy5+14 and Cy5-100(**Panel A, top left**), PIFE Cy3/Cy5 +14 (**panel B, top right**) and PIFE Cy3/Cy5 -100 (**panel C, bottom left**) representative individual experiments shown in Figure 3 and 5. For more details on fitting the data to a mechanism, see SI methods section on data fitting and analysis.

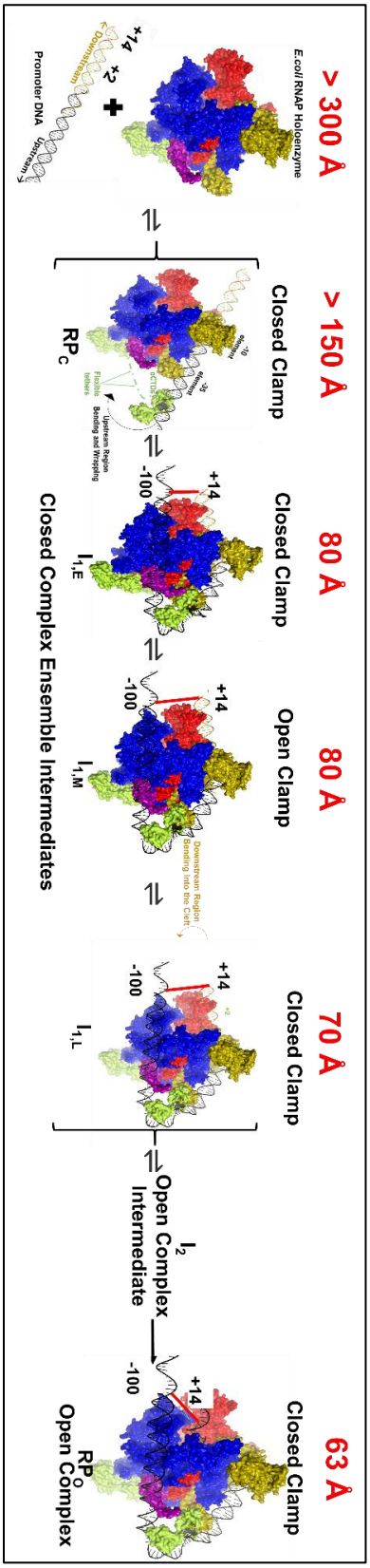


Figure 8: Structural Mechanism of Spontaneous Open Complex Formation from FRET and PIFE Characteristics of Closed Intermediates in Transcription Initiation. Subunits of RNAP are β (red), β' (blue), σ^{70} (wheat), ω (purple), flexibly-tethered αCTDs (light green), and αNTDs (light green). FRET distances between promoter positions +14 (in downstream (golden) DNA) and -100 (upstream (black)) in each of the four closed intermediates deduced from the kinetic analysis are shown. These closed intermediates, which make up the closed complex ensemble ({CC}), are the initial closed complex (RP_c) and three more advanced closed intermediates (I_{1E}, I_{1M} and I_{1L}). RP_o is the stable open complex.

Tables:

Table 1- Kinetics Rates and Equilibrium constants for ensemble of Closed Complex Intermediates formation and DNA opening steps in the mechanism of OC formation.

	Kinetic Step	Rate Constants (± Standard Deviation)[a]	Equilibrium Constants
1	$R + P \rightleftharpoons RP_C$	$k_1 = (3.07 \pm 0.11) \times 10^8\ M^{-1}\ s^{-1}$	$7 \times 10^6\ M^{-1}$
		$k_{-1} = 44 \pm 2.3\ s^{-1}$	
2	$RP_C \rightleftharpoons I_{1,E}$	$k_2 = 13.2 \pm 0.94\ s^{-1}$	1.8
		$k_{-2} = 7.3 \pm 0.33\ s^{-1}$	
3	$I_{1,E} \rightleftharpoons I_{1,M}$	$k_3 = 2.5 \pm 0.2\ s^{-1}$	1.5
		$k_{-3} = 1.7 \pm 0.075\ s^{-1}$	
4	$I_{1,M} \rightleftharpoons I_{1,L}$	$k_4 = 0.10 \pm 0.005\ s^{-1}$	1
		$k_{-4} = 0.1 \pm 0.006\ s^{-1}$	yields $K_{\{CC\}} = 5.2 \times 10^7\ M^{-1}$
5	$I_{1,M} \longrightarrow OC$	$k_5 = 0.04 \pm 0.006\ s^{-1}$	

a Standard Deviation is calculated for dataset that didn't require linear term drift correction. For more details refer to SI Methods.

Table 2-Relative Fluorescence Intensities obtained from Fits of FRET and PIFE Kinetic Data to a 5 step Mechanism.

Fluorescence of Intermediates and RP_O						
	FRET Acceptor Intensity (Relative to RP_O)		+14 Fluorescence Intensity (from PIFE) (Relative to P)		-100 Fluorescence Intensity (from PIFE) (Relative to P)	
	Cy5+14	Cy5-100	Cy3	Cy5	Cy3	Cy5
RP_C	0	0	1	1	1	1
I_{1E}	0.26 ± 0.13	0.38 ± 0.05	1.75 ± 0.05	1.22 ± 0.07	1.26 ± 0.03	1.08 ± 0.04
I_{1M}	0.28 ± 0.05	0.38 ± 0.14	1.52 ± 0.04	1.14 ± 0.05	1.30 ± 0.03	1.1 ± 0.03
I_{1L}	0.65 ± 0.2	0.69 ± 0.15	1.39 ± 0.05	1.15 ± 0.02	1.11	1.01
RP_O	1.0 ± 0.08	1.0 ± 0.07	1.31 ± 0.03	1.1 ± 0.02	1.05	1.03

Table 3- FRET Determinations of Distances between -100 and +14 Positions on Promoter DNA in Intermediates (Closed Complexes) and in RP_o.

Distance between -100 Upstream and +14 Downstream end of promoter DNA in the closed complex intermediates involved in OC formation mechanism.			
Species	**Relative average FRET Intensities**	**Scaled FRET Efficiency relative to OC**	**Distance (in Å)**
I_{1E}	0.32 ± 0.06	0.10 ± 0.02	80 ± 2.8
I_{1M}	0.33 ± 0.05	0.11 ± 0.02	80 ± 2.3
I_{1L}	0.67 ± 0.02	0.21 ± 0.006	70 ± 0.4
RP_O	1	0.32	63.5

Supporting information and Supplemental Figures

Table for Supporting Information:

Supplemental Table 1. DNA sequences-

Promoter	Upstream primer	Downstream primer	Duplex (bp) and position relative to start site (+1)[a]	Experiment used in
λP_R (-110 to +22)	Up1	Down1	132 (-110 to +22)	Single round transcription kinetics
λP_R (-110 to +14)	Up	Down2	124 (-110 to +14)	
λP_R (-59 to +34)[c]	Up	Down	192 (-128 to +64)	Permanganate footprinting kinetics
Cy3(+14) λP_R [b]	Up	Cy3+14 (Down)	142 (-128 to +14)	Fluorescent kinetics, single round transcription kinetics
Cy5(+14) λP_R [b]	Up	Cy5+14 (Down)	142 (-128 to +14)	
Cy3(-100) λP_R [b]	Cy3 (-100) Up	Down	164 (-100 to +64)	
Cy5(-100) λP_R [b]	Cy5 (-100) Up	Down	164 (-100 to +64)	
Cy3(+14)-Cy5 (-100) λP_R [b]	Cy5 (-100) Up	Cy3+14 (Down)	114 (-100 to +14)	
Cy5 (+14)-Cy3 (-100) λP_R [b]	Cy3 (-100) Up	Cy5+14 (Down)	114 (-100 to +14)	

[a] Dye-labeled DNA have 12 base (upstream) and/or 8 base (downstream) ssDNA termini.

[b] Sreenivasan et al'16, [c] Drennan et al'12.

Supplemental Table 2. Primer Sequences-

Sno	Primer	Sequence
1	Up	5’-GTACGAATTCGATATCCAGCTATGACCATGATTACGCCAAGC-3’
2	Up1	5'-CAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGT-3'
3	Up2	5’-TGCCTGCAGGTTTAAACAGTCGATAAATATCTAACACCGT-3’
4	Up3	5’ – ACACCGTGCGTGTTGACTATTTTACCTCTGGCG -3’
5	Down	5’-CAGGACCCGGGAAGCTTTTAATTAACACTCTTATACATTATTCC-3’
6	Down1	5’-CCATACAACCTCCTTACTACATGCAACCATTATCACCGCC-3’
7	Down2	5’-CCTCCTTACTACATGCAACCATTATCACCGCC-3’
8	Cy3 (+14)	5’-CCATACAA/iCy3/CCTCCTTACTACATGCAACCATTATCACCGCC-3’
9	Cy5 (+14)	5’-CCATACAA/iCy5/CCTCCTTACTACATGCAACCATTATCACCGCC-3’
10	Cy3(-100)	5’-CAGCTATGACCA/iCy3/TGATTACGCCAAGC-3’
11	Cy5 (-100)	5’-CAGCTATGACCA/iCy5/TGATTACGCCAAGC-3’

Experiment Type		[DNA]: [RNAP]	DNA Construct	Number of	
Fluorescence Kinetics				Average shots Obtained In an Individual Experiment	Independent Replicates
	Buffer conditions				
FRET Kinetics of OC formation at different DNA:RNAP	Fluorescence Buffer (FB)[c]	1:1	Cy3+14 Cy5-100	10	15
			Cy5+14 Cy3-100	10	14
	FB[b]	1:1.5	Cy3+14 Cy5-100	9	2
	FB[a]	1:2	Cy3+14 Cy5-100	8	3
FRET Kinetics of OC formation as function of [glycine betaine] 1.25 M	FB[c]	1:1	Cy3+14 Cy5-100	10	3
				11	3
PIFE kinetics of OC formation			Cy3+14	8	12
			Cy5+14	6	9
			Cy3-100	7	6
			Cy5-100	6	8
FRET OC transcription initiation in: 100 uM all NTPs			Cy3+14 Cy5-100	6	2
			Cy5+14 Cy3-100	7	3
10 uM UTP/100 uM each ATP,GTP and CTP			Cy3+14 Cy5-100	18	3

10 uM UTP/200 uM each ATP,GTP and CTP				12	3
PIFE OC salt upshift dissociation in: 1.1 M KCl			Cy3+14	17	8
0.4 M KCl				10	5
FRET OC salt upshift dissociation in: 1.1 M KCl			Cy5+14 Cy3-100	11	7
0.4 M KCl				12	3
Single ^{32}P detected round transcription kinetics		[DNA]: [RNAP]	DNA Construct	Number of Independent Experiments	
		1:1/1:2	λP_R (-110 to +22)	7	
			λP_R (-110 to +14)	5	
			λP_R (Cy5+14 Cy3-100)	5	
			λP_R (Cy3+14 Cy5-100)	5	
			λP_R (Cy5+14)	3	
			λP_R (Cy3+14)	3	

Supplemental Table 3. Summary of different fluorescence and transcription initiation kinetic experiments.

[a] FB with 2% (v/v) glycerol, [b]FB with 1.5% (v/v) glycerol, [c] FB with 1% (v/v) glycerol.

Supplemental Table 4. Kinetic Parameters and Fluorescence Amplitude values with their range floated to fit OC formation fluorescence kinetics FRET, PIFE signal shown in Figure 3 and 5 to a multi-step mechanism.

Kinetic Rates and Bounds			
		Lower Bound	**Upper Bound**
$\mathbf{k_1}$	$3.07 \times 10^8\ M^{-1}\ s^{-1}$	$2.76 \times 10^8\ M^{-1}\ s^{-1}$	$3.22 \times 10^8\ M^{-1}\ s^{-1}$
$\mathbf{k_{-1}}$	$44\ s^{-1}$	$41.8\ s^{-1}$	$50.6\ s^{-1}$
$\mathbf{k_2}$	$13.1\ s^{-1}$	$11.8\ s^{-1}$	$14.4\ s^{-1}$
$\mathbf{k_{-2}}$	$7.3\ s^{-1}$	$6.9\ s^{-1}$	$8.4\ s^{-1}$
$\mathbf{k_3}$	$2.5\ s^{-1}$	$2.2\ s^{-1}$	$2.7\ s^{-1}$
$\mathbf{k_{-3}}$	$1.7\ s^{-1}$	$1.6\ s^{-1}$	$1.9\ s^{-1}$
$\mathbf{k_4}$	$0.11\ s^{-1}$	$0.09\ s^{-1}$	$0.12\ s^{-1}$
$\mathbf{k_{-4}}$	$0.1\ s^{-1}$	$0.095\ s^{-1}$	$0.115\ s^{-1}$
$\mathbf{k_5}$	$0.041\ s^{-1}$	$0.02\ s^{-1}$	$0.045\ s^{-1}$

Amplitudes [a]	**FRET** [b]	
	Lower Bound	Upper Bound
$A_{I1,E}$	0	0.5
$A_{I1,M}$	0.3	0.9
$A_{I1,L}$	0.5	1.1
A_{RPo}	0.5	1.4

a. *Results of fit as average and standard deviation are given in Table S5.*
b. *for FRET signal at 400 s for a population of 80% RPo, 10% {CC} and 10% P,* $A_{RPo} \approx 1.25$. .

In case of PIFE +14 and -100 the amplitudes were varied with wider positive range of real numbers, (0, ∞).

Supplemental Table 5. Fitting amplitudes, A_i, and Relative Fluorescence Intensities, A_i^{rel}, obtained by fitting FRET and PIFE fluorescence kinetics data to a 5 step mechanism.

	FRET, Cy5+14		**FRET, Cy5-100**	
	A_i [a]	**A_i^{rel}** [b]	**A_i**	**A_i^{rel}**
I_{1E}	0.31 ± 0.15	0.26 ± 0.13	0.46 ± 0.05	0.38 ± 0.04
I_{1M}	0.33 ± 0.06	0.28 ± 0.05	0.45 ± 0.16	0.38 ± 0.13
I_{1L}	0.78 ± 0.2	0.65 ± 0.2	0.83 ± 0.2	0.69 ± 0.15
RP_O	1.2 ± 0.07	1.0 ± 0.06	1.2 ± 0.06	1.0 ± 0.05
	PIFE +14, Cy3 [c]		**PIFE +14, Cy5** [c]	
	A_i	**A_i^{rel}**	**A_i**	**A_i^{rel}**
I_{1E}	2.9 ± 0.2	1.75 ± 0.05	2.7 ± 0.8	1.21 ± 0.07
I_{1M}	2.0 ± 0.2	1.52 ± 0.04	1.6 ± 0.6	1.13 ± 0.05
I_{1L}	1.5 ± 0.2	1.39 ± 0.05	1.8 ± 0.2	1.15 ± 0.02
RP_O	1.2 ± 0.1	1.31 ± 0.03	1.2 ± 0.2	1.1 ± 0.02
	PIFE-100, Cy3 [d]		**PIFE-100, Cy5** [d]	
	A_i	**A_i^{rel}**	**A_i**	**A_i^{rel}**
I_{1E}	2.4 ± 0.3	1.26 ± 0.03	2.0 ± 1.2	1.08 ± 0.04
I_{1M}	2.7 ± 0.2	1.30 ± 0.03	2.7 ± 0.9	1.1 ± 0.03
I_{1L}	1.0 ±	1.11	0.2 ±	1.01
RP_O	0.5 ±	1.05	0.7 ±	1.03

[a] Average Fitting amplitudes and standard deviation as uncertainties, obtained from fitting FRET and PIFE data to 5 step mechanism with rates, amplitudes and bounds from Table S4.

[b] Relative fluorescence intensity calculated by eqs XX.

[c] PIFE+14 data was fit with linear correction term in the range of -(4.7 ± 0.7) x 10-4 for Cy3 and -(7.7 ± 1.4) x10-4 for Cy5 to account for drift observed at later time-points

[d] PIFE-100 data was fit within the time range 0.01-4 seconds due to inconsistencies in the extent of decrease observed in the fluorescence signal after 4 seconds of each individual experiments

Supplemental Table 6. Comparison of two set of rates for fit of an individual experiment of acceptor data.

Rates and Amplitudes			**Derived Quantities**		
	Panel A	**Panel B**		**Panel A**	**Panel B**
k_1	1.31 x 10^8 M^{-1} s^{-1}	3.07 x 10^8 M^{-1} s^{-1}	K_1	1.31 x 10^6 M^{-1}	6.98 x 10^6 M^{-1}
k_{-1}	100 s^{-1}	44 s^{-1}	K_2	5.08	1.79
k_2	49.7 s^{-1}	13.1 s^{-1}	K_3	1.47	1.44
k_{-2}	9.78 s^{-1}	7.3 s^{-1}	K_4	12.3	1.07
k_3	1.42 s^{-1}	2.5 s^{-1}	K_{CC}	1.30 x 10^8 M^{-1}	5.7 x 10^7 M^{-1}
k_{-3}	0.96 s^{-1}	1.7 s^{-1}	k_{isom}	0.01996	0.01398
k_4	0.49 s^{-1}	0.11 s^{-1}	$f_{P\ t=1s}$	0.7	0.5
k_{-4}	0.04 s^{-1}	0.1 s^{-1}	$f_{RPot=400s}$	0.9	0.8
k_5	0.0229 s^{-1}	0.041 s^{-1}	$f_{I_{1E}}^{\{CC\}}$	0.05	0.22
A_{I1E}	0.733	0.279	$f_{I_{1M}}^{\{CC\}}$	0.07	0.32
A_{I1M}	0.382	0.361	$f_{I_{1L}}^{\{CC\}}$	0.87	0.34
A_{I1L}	0.439	0.876			
A_{RPo}	0.679	1.08			

RMS	0.0064	0.0091			

Methods

1. Normalization of Fluorescence Kinetic Data

For both FRET and PIFE OC formation assays, time-averaged initial minimum (F_{min}) and maximum (F_{max}) fluorescence in the interval 10 ms to 400s were obtained for each shot and used to calculate the normalized fluorescence (F_{norm}) as a function of time using Equation S1.

$$F_{norm} = \frac{F - F_{min}}{F_{max} - F_{min}} \qquad \text{- Equation S1}$$

Visual inspection of the set of 10-15 normalized shots in one series was used to eliminate any systematic outliers. Then these normalized data were averaged for subsequent fitting.

In previous FRET-detected equilibrium titrations of these Cy3Cy5 labeled promoter DNAs (at 50 nM) with RNAP, evidence for a competitive nonpromoter binding mode was obtained at RNAP:promoter DNA ratios greater than 1:1 (Sreenivasan et al ’15). To avoid this, most kinetics experiments were performed at a 1:1 ratio of RNAP to promoter DNA. FRET-detected forward kinetics experiments performed in excess RNAP (1.5:1 and 2:1 RNAP:promoter ratios) typically showed a maximum in the FRET acceptor signal at about 100 s and a subsequent reduction in FRET (Supplemental Figure 4) .This quenching was empirically corrected by fitting the fluorescence data to a multi 4 exponential fit with a linear time dependent term (Equation 2). The linear term was subtracted from the raw data, and the corrected fluorescence data was then re-normalized using Equation S1.

To normalize data obtained in individual transcription initiation and salt-upshift dissociation assays, F_{max} (Equation S1) for each shot was obtained by averaging the first 5-7 fluorescence readings (10 to 20 ms). F_{min} was obtained as the fluorescence at completion of the process.

2. Fitting of Fluorescence Kinetic Data for OC Formation to a Multi-step Mechanism

PIFE and FRET Fluorescence kinetic data (10 ms - 400 s) were first fit to a sum of exponentials with a linear drift-correction term using OriginPro software (Equation S2).

$$\mathrm{F} = (\sum_{i=1}^{4} A_i e^{-k_{i,obs}t}) + k_{corr}t \qquad \text{-Equation S2}$$

Multi-exponential fit gave belter results by having the linear-term to account for the long-term drift observed in the instrument which otherwise would give rise to slower kinetics that those predicted form the filter binding analysis (Figure 2).

Four but not three exponentials were required to obtain a high quality fit to FRET and PIFE data. Rate constants $k_{i,obs}$ spanned ~4 orders of magnitude with typical values.

$$k_{1,obs} \sim 15\ s^{-1},\ \ k_{2,obs} \sim 2\ s^{-1}, k_{3,obs} \sim 0.2\ s^{-1}, k_{4,obs} \sim 0.02\ s^{-1}$$

Previous investigations of the kinetics of OC formation for λP_R promoter at 19^0C revealed that the overall process of forming {CC} was rapidly reversible on the time scale of converting {CC} to OC, and yielded an isomerization rate constant for the {CC}-OC conversion of 0.014 ± 0.02 s^{-1}. We therefore assumed that each of the above exponential-decay rate constants could be interpreted as a relaxation-to-equilibrium rate constant for a step of the mechanism that is detected by our assay (Saecker, R.M. et al, 2002), and corresponding mechanism is

$$R + P \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} RP_c \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} I_{1E} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} I_{1M} \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} I_{1L} \overset{k_5}{\rightarrow} RP_o$$

Based on previously published second order rate constants for the initial step of forming RP_C (1-3 x 10^8 $M^{-1}s^{-1}$) for different promoters and 1:1 equilibrium distribution between RP_C and free P, we estimated that RP_C equilibrates with free P in the first 20-30 ms after mixing. This step is not detectable by the instrument due to lack of fluorescence signal change observed in earlier time points when collected in logarithmic time scale from 0.01 to 400 seconds (see supplemental method section 5).

From the observation that four exponentials are required to fit FRET and PIFE fluorescence kinetic data, we deduced that at least four closed intermediate complexes are part of {CC} ensemble in the mechanism of OC formation, and that the sums of forward and reverse rate constants for forming these species are given by the exponential decay rate constants as follows:

For conversion of RP_C to I_{1E}, $k_{1,obs} \sim k_2 + k_{-2} \sim 15\ s^{-1}$.

For conversion of I_{1E} to I_{1M}, $k_{2,obs} \sim k_3 + k_{-3} \sim 2\ s^{-1}$.

For conversion of I_{1M} to I_{1L}, $k_{3,obs} \sim k_4 + k_{-4} \sim 0.2\ s^{-1}$.

And for conversion of I_{1L} to OC, $k_{4,obs} \sim k_5 = k_{open} \sim 0.2\ s^{-1}$.

Next, normalized FRET and PIFE kinetic data was fit to a 5-step mechanism (Mechanism 2 of main text) using Berkeley Madonna Simulation software. (A four-step mechanism was also tested; see supplemental method section 5). Initial guesses of rate constants were obtained from the above relaxation relationships and the assumption that the equilibrium constants for each step of advancing the CC population from RPc to I_{1L} are in the range 1-3 (justified by relationship between $K_{\{CC\}}$ and K_{RPc}).

Good fits could be obtained to 4-step mechanisms if the rate constants were allowed to be different for FRET and PIFE experiments, but to obtain a good fit to both FRET and PIFE data using the same rate constants and only allowing the intrinsic fluorescence properties of the

intermediates to vary required a five step mechanism. Constraints were applied in the fitting to ensure that the rate constants are consistent with filter binding kinetics results, see supplemental method section 3 and 5.

Normalized fluorescence kinetics signal was fit to the sum of the products of the fraction of the concentration of each intermediate species, f_i, and their fluorescence amplitudes, A_i, as shown in equation S3 and S4:

$$F_{norm} = A_{I_{1E}} f_{I_{1E}} + A_{I_{1M}} f_{I_{1M}} + A_{I_{1L}} f_{I_{1L}} + A_{RP_o} f_{RP_o} + k_{corr} t \qquad \text{-Equation S3}$$

$$f_i = \frac{[i]}{total\ conc} = \frac{[i]}{[P]+[I_{1E}]+[I_{1M}]+[I_{1L}]+[RP_o]+[RPc]},\ i = I_{1E}, I_{1M}, I_{1L}, RP_o \qquad \text{-Equation S4}$$

Based on the previous estimates of kinetics of the first step to form the initial closed complex RP_C intermediate, which is diffusion limited and in the order of 10^8 M^{-1} s^{-1} we assumed zero fluorescence contribution from RP_c species in the fluorescence kinetics signal collected from 0.01 to 400 seconds $A_{RP_c} = 0$. We tested floating the fluorescence amplitudes of A_{RPc} along with other amplitudes but got very low values, supplemental method section 5.

Other additional rapid equilibrium conditions and different tests were applied to fit the OC formation fluorescence kinetics signal to the mechanism as described below in section 3-5 which yielded following set of kinetic parameters and fluorescent amplitude values (Supplemental Table 4 and 5). The kinetic rates and amplitude values were floated within the range mentioned to fit the individual FRET and PIFE OC formation experiments using Berkeley Madonna simulation software. The obtained rates and amplitudes were averaged to determine the standard deviation, which is typically less than ± 10% of the values mentioned in the upper and lower bounds of supplemental Table 4.

3. Additional Conditions Applied to Fit the Mechanism

4 additional conditions were applied to obtain the kinetic parameters of a multi- 4 to 5 -step mechanism to satisfy the previous results of OC formation kinetics obtained using filter binding assay fit to an overall two-step mechanism (equation 1) (Tsoldikov and Ruth, 2005).

$$R + P \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \{CC\} \overset{k_{isom}}{\rightarrow} RP_o \qquad \text{-Equation S5}$$

$$K\{cc\} = \frac{k_f}{k_r} \qquad \text{-Equation S6}$$

i) **Applying Rapid Equilibrium condition -** Previous Filter Binding OC formation kinetics results when fit to a minimal 2 step mechanism (Equation 1) gave single exponential binding isotherm following the rapid equilibrium condition where $k_r >> k_{isom}$. On fitting the fluorescence signal to a multi-step mechanism we applied similar rapid equilibrium approach, that is the reverse kinetic rate of (i-1)th elementary step should be greater than the forward kinetic rate of the (i)th one (Condition I).

$$k_{(i-1)} >> k_i \quad i = \{2,3,4\} \qquad \text{-Condition I}$$

ii) **Fixing Isomerization rate constant** – Previous kinetic analysis of OC formation for full length λP_R promoter DNA estimated the isomerization rate constant k_{isom} to be in the range of 0.014 s^{-1} at 17-19^{0}C (Ruth et al, Davis et al, 2005). k_{isom} is related to the DNA opening step k_{open} by the fraction of the late advanced closed complex intermediate as shown below in Condition II.

$$k_{isom} = k_{open}\, f_{I_{1L}}^{\{CC\}} = 0.014 \pm 0.003\ s^{-1}$$

$$f_{I_{1L}}^{\{CC\}} = \frac{[I_{1L}]}{([I_{1L}] + [I_{1M}] + [I_{1E}] + [RP_C])} = \frac{K_2K_3K_4}{\left(1+K_2\left(1+K_3(1+K_4)\right)\right)} \qquad \text{-Condition II}$$

(Other fractions of Closed Complex $f_{I_{1E}}^{\{CC\}} = \frac{[I_{1E}]}{([I_{1L}] + [I_{1M}] + [I_{1E}] + [RP_C])} = \frac{K_2}{\left(1+K_2\left(1+K_3(1+K_4)\right)\right)}$,

and $f_{I_{1M}}^{\{CC\}} = \frac{[I_{1M}]}{([I_{1L}] + [I_{1M}] + [I_{1E}] + [RP_C])} = \frac{K_2K_3}{\left(1+K_2\left(1+K_3(1+K_4)\right)\right)}$)

iii) Correlating equilibrium rate constant of each step to the overall $K\{cc\}$ - The overall equilibrium constant $K\{cc\}$ of a 2 step mechanism is related to the equilibrium constant of individual steps of 4 or 5 steps pathway using Condition III. The overall equilibrium constant was kept in the range of 2 – 10 x 10^7 based on the previous estimates by filter binding assays (Ruth et al, Davis et al, 2005).

$$K_{cc} = K_1\left(1 + K_2\left(1 + K_3(1 + K_4)\right)\right) = (2-10)\ x\ 10^7\ M^{-1} \qquad \text{-Condition III}$$

$$K_{cc}[P]_{tot} = K_1\left(1 + K_2\left(1 + K_3(1 + K_4)\right)\right)[P]_{tot}$$

$$= (2-10)\ x\ 10^7\ M^{-1} x\ 5\ x\ 10^{-8} M \sim (1-5)$$

iv) Keeping consistent fraction of free Promoter DNA and OC at 1 and 400 seconds respectively – On simulating the time evolution of free promoter DNA and OC concentration using the kinetics values of a 2 tier OC formation mechanism (Equation 1) fraction of free promoter DNA is 0.5 at 1 sec and OC is 0.8 at 400 seconds in 1:1 DNA:RNAP mixing at 19^0C (Condition IV).

$$f_P \sim 0.5\ at\ t = 1\ sec\ \ and\ f_{RP_o} = 0.8\ at\ t = 400\ sec \qquad \text{-Condition IV}$$

4. Determination of kinetic rates –

We chose the set of kinetic parameters and amplitude values that follows conditions I - IV. Product $K_{\{CC\}}[P]_{tot}$=(2-10) x10^7 (M^{-1}) x 5 x 10^{-8} (M) = (1-5) is on order of 1 or slightly larger than 1 (Condition III). From RE equation $K_{\{CC\}}[P]_{tot} = K_1[P]_{tot}(1+K_2(1+K_3(1+K_4)))$~(1-5). Using for k_1 estimation k_1~3 10^8 (M^{-1} s^{-1}) and for $k_{-1} >> k_2 \sim k_{1,obs}$=15 s^{-1} (Condition I) we obtain inequality $K_1[P]_{tot} = k_1[P]_{tot}/k_{-1}$ <<3 10^8 M^{-1} s^{-1} x 5 x 10^{-8} (M) /15 s^{-1} =1 . This inequality means that $K_1[P]_{tot}$ << 1.Consider first $K_1[P]_{tot}$ ~0.3, then (1+ K_2 (1+ K_3 (1+ K_4)))~3 x (1-5) and each K_i cannot be much larger than 1, so they all ~1, then $K_2 \sim K_3 \sim K_4$ ~0.8 - 2. This also gives $k_2 \sim k_{-2}$ ~ 10 s^{-1} , $k_3 \sim k_{-3}$ ~ 1 s^{-1} , and $k_4 \sim k_{-4}$ ~ 0.1 s^{-1} . If $K_1[P]_{tot}$ ~0.1 <<1, then (1+ K_2 (1+ K_3 (1+ K_4)))~10 x (1-5), this changes inequalities for $K_2 \sim K_3 \sim K_4$ ~2 - 3 only slightly.

These conclusions are true for 4-step mechanism as well, If K_1 $[P]_{tot}$ ~0.3 , K_2 ~ K_3 ~1 -3 If K_1 $[P]_{tot}$ ~0.1 , K_2 ~ K_3 ~2.5 -6.5.

Initial fit of all data with some choice of kinetic rates gives good fit (Table S5 and Fig S10), an example of Cy5-100 FRET representative dataset is shown below (Figure S10). However, the best fit set of rates and amplitudes does not satisfy most conditions mentioned above like:

i) It overestimates $K\{cc\}$ by almost an order of magnitude: 1.4 x 10^8 M^{-1}.

ii) Not all steps are in rapid equilibrium (RE) condition, k_{-1}, . $k_{-3,\ and}$ k_{-4} only twice larger then forward rate in following steps.

iii) Fraction of free Promoter DNA plateaus at ~0.6 -0.7 instead of 0.5, and final fraction of RP_o is 0.95 instead of 0.8.

To obtain more plausible set of kinetic rates that have 0.5 plateau for free promoter DNA fraction at 1 second and follow RE conditions, linked fit and continuous simulation option of KinTek FitSpace explorer was used.

- We choose low noise FRET acceptor data, and started from values of amplitudes A_{I1E} ~ 0.3, A_{I1M} ~ 0.4, A_{I1L} ~ 0.8, and A_{RPo} ~ 1.2.
- Forward rate constant of the second step k_2 was linked to be 2.5 - 3 times smaller than the reverse rate constant of the first step, $k_{-1.}$ Similarly, k_3 was kept 3 times smaller than k_{-2}, and k_{-2} 2 times smaller than k_2.
- k_{-1} was reduced by 2.5 times and k_1 increased by 2.5 times. k_{-3} was also increased and adjusted to get a good concentration distribution with f_P~0.5 $at\ t = 1\ sec$ and visually good fit of data.
- Then the data were fitted to get estimates for the rest of the parameters: k_4 = k_{-4} ~0.01 s^{-1} and k_5 = 0.04-0.05 s^{-1} keeping kinetic rates of steps 1 to 3 fixed. The

corresponding rates and amplitudes are given in Table S4, and fit of experimental data together with concentration distribution of all species are shown on Fig S9.

The set of kinetic rates determined to follow the multi-step mechanism and all the conditions described above in section 3 are in Supplemental Table 4. These kinetic parameters and amplitude values were varied within the mentioned lower and upper bound to fit all types of FRET and PIFE OC formation fluorescence kinetics data obtained using Berkeley Madonna simulation software to get good fits.

5. Tests and Comparisons-

1. **Fitting to a 4 vs 5 step mechanism** – OC formation fluorescence kinetics data were fitting well to a 5 step mechanism with better RMS value for both FRET and PIFE +14 and -100 dataset. PIFE -100 signal (t< 4 seconds) couldn't fit to a 4 step mechanism. We couldn't get a set of rate constants that followed all the conditions described above and fit to both FRET and PIFE experiments. Moreover, the fluorescence kinetics signal analytically fits well to a sum of 4 but not 3 exponential functions with a linear drift correction term indicating that it needs minimally 4 or more steps with less contribution to the fluorescence signal from the early first intermediate species in the mechanism of OC formation.
2. **Validating the assumption that fluorescence contribution of RP_C species is negligible** – We varied the fluorescent amplitudes of RP_c while floating other variables to fit the fluorescence OC formation data to a multi-step mechanism. We got values ranging from 1×10^{-7} to 0.01 depending upon the amount of fluctuations/signal noise observed at early time points for an individual experiment. The values were at least an order of magnitude smaller on the scale of other advanced intermediates present in the mechanism. This validates our assumption that RP_c species have very little or none

fluorescence contribution to the observed fluorescence signal as compared to the other advanced closed complex intermediates.

6. Conversion of Fitted Fluorescence Amplitudes to Relative Fluorescence Intensities-

Fitting amplitudes A_i contribute to normalizes fluorescence signal, F_{norm}, while the raw signal have contribution from intrinsic intensity of species, a_i.

$$F_{norm} = A_{I_{1E}} f_{I_{1E}} + A_{I_{1M}} f_{I_{1M}} + A_{I_{1L}} f_{I_{1L}} + A_{RP_o} f_{RP_o} + k_{corr} t \quad \text{Eq. S3}$$

$$F = a_{I_{1E}} f_{I_{1E}} + a_{I_{1M}} f_{I_{1M}} + a_{I_{1L}} f_{I_{1L}} + a_{RP_o} f_{RP_o} + other\ contributions$$

In case of FRET, we convert fitting amplitudes to the relative intrinsic intensities . Tthey caharcterize intensity\ies relative to intensity of the open complex RP_o.

$$A_i^{rel} = \frac{a_i}{a_{RPo}} = \frac{A_i}{A_{RPo}} \qquad i = \{I_{1E}, I_{1M}, I_{1L}\} \qquad \text{- Equation S7a}$$

For PIFE, the fitting amplitudes are converted to relative intensities by multiplying the maximum change in absolute fluorescence signal observed from initial time-point 0.01 to final time-point 400 seconds with the fluorescence amplitude (A_i) and added with 1 for each individual experiments as shown in Equation S7. the relative intensities are with respect to the intensity of RP_c.

$$A_i^{rel} = \frac{a_i}{a_{RPc}} = 1 + A_i \left(\frac{F_{max} - F_{min}}{F_{min}}\right)_{av} \qquad i = \{I_{1E}, I_{1M}, I_{1L}, RP_o\} \qquad \text{- Equation S7}$$

Values $\left(\frac{F_{max} - F_{min}}{F_{min}}\right)_{av}$ are calculated as average of each shot that was acoounted for in normalized signal, and then as average of all experiments of same dye type and position.

6.1. Derivation of equation 11 to relative fluorescence amplitudes

The derivation presented here for equation 11 is for a 2 step mechanism for simplicity purpose which can be generalized for any multi step mechanism.

$$P + R \rightarrow I_1 \rightarrow RP_o$$

Here reactants, R and P (including RP_c) are species whose PIFE amplitudes are unaffected or have 0 FRET signal but the later participants are affected.

For FRET signal

Fluorescence signal F is proportional to the sum of the products of each species concentration and their corresponding amplitudes $a_{i\ ,}$, species before I_1 have zero amplitude, F_{init} accounts for possible instrument offset at the beginning of analysis interval.

$$F = F_{init} + a_1[I_1] + a_{RPo}[RP_o]$$

$$F_{min} = F_{init} + a_1[I_1](t_0) + a_{RPo}[RP_o](t_0) \approx F_{init}$$

Assuming $[I_1](t_0) \approx [RP_o](t_0) \approx 0$

$$F_{max} = F_{init} + a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max})$$

Then

$$\frac{F - F_{min}}{F_{max} - F_{min}} = \frac{F_{init} + a_1[I_1] + a_{RPo}[RP_o] - F_{init}}{F_{init} + a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max}) - F_{init}} = \frac{a_1[I_1] + a_{RPo}[RP_o]}{a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max})}$$

$$= \frac{a_1[P_{tot}]}{a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max})}\frac{[I_1]}{[P_{tot}]} + \frac{a_{RPo}[P_{tot}]}{a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max})}\frac{[RP_O]}{[P_{tot}]}$$

Comparing with fitting expression,

$$\frac{F - F_{min}}{F_{max} - F_{min}} = A_1\frac{[I_1]}{[P_{tot}]} + A_{RP_O}\frac{[RP_o]}{[P_{tot}]}$$

Recall that $\frac{[RP_o](t_{max})}{[P_{tot}]} = 0.8$, then RP_o should have amplitude close to 1.25

$$A_1 = \frac{a_1[P_{tot}]}{a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max})}, \quad A_{RP_O} = \frac{a_{RPo}[P_{tot}]}{a_1[I_1](t_{max}) + a_{RPo}[RP_o](t_{max})}$$
$$= 1.25 * \frac{1}{1 + \frac{a_1[I_1](t_{max})}{a_{RPo}[RP_o](t_{max})}} \leq 1.25$$

We find relative (to RP_o) intrinsic intensity as

$$A_1^{rel} = \frac{a_1}{a_{RPo}} = A_1 \frac{[RP_o](t_{max})}{[P_{tot}]} = \frac{A_1}{A_{RP_O}}$$

For other species, their amplitudes relative to RP_o are simply ratio of fitting amplitudes

$$A_i^{rel} = \frac{a_i}{a_{RPo}} = \frac{A_i}{A_{RP_O}}$$

For PIFE signal

Fluorescence signal F is proportional to the sum of the products of each species concentration and their corresponding amplitudes a_i (Equation S8), F_{init} accounts for possible instrument offset at the beginning of analysis interval.

$$F = F_{init} + a_{RPc}[P] + a_1[I_1] + a_2[RP_o] \qquad \text{-Equation S8}$$

Value F_{min} is at some initial time moment t_0

$$F_{min} = F_{init} + a_{RPc}[P](t_0) + a_1[I_1](t_0) + a_2[RP_o](t_0) \qquad \text{-Equation S9}$$

Value F_{max} is at some time t_{max}

$$F_{max} = F_{init} + a_{RPc}[P](t_{max}) + a_1[I_1](t_{max}) + a_2[RP_o](t_{max}) \qquad \text{-Equation S10}$$

Applying conservation of mass equation, one can present the signal as:

$F = F_{init} + a_{RPc}[P] + a_1[I_1] + a_2[RP_o] = F_{init} + a_{RPc}([P] + [I_1] + [RP_o]) + \Delta a_1[I_1] + \Delta a_2[RP_o] =$

$F_{init} + a_{RPc}[P]_{tot} + \Delta a_1[I_1] + \Delta a_2[RP_o]$

Where $[P_{tot}] = [P] + [I_1] + [RP_o]$, and $\Delta a_1 = a_1 - a_{RPc} > 0$ -Equation S11

Similarly for F_{min} and F_{max}:

$$F_{min} = F_{init} + a_{RPc}[P_{tot}] + \Delta a_1[I_1](t_0) + \Delta a_2[RP_o](t_0)$$ -Equation S12

Assuming affected complexes are not present in significant amount, at time t_0, then:

$[I_1](t_0) \approx [RP_o](t_0) \approx 0$, and $F_{min} \approx F_{init} + a_{RPc}[P_{tot}]$

So , $F_{max} = F_{init} + a_{RPc}[P_{tot}] + \Delta a_1[I]_1(t_{max}) + \Delta a_2[RP_o](t_{max})$ -Equation S13

For fraction $(F_{max}-F_{min})/F_{min}$:

$$\frac{F_{max}-F_{min}}{F_{min}} = \frac{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})}{F_{init}+a_{RPc}[P_{tot}]}$$ -Equation S14

Normalized signal:

$$\frac{F-F_{min}}{F_{max}-F_{min}} = \frac{a_{RPc}[P_{tot}]+\Delta a_1[I_1]+\Delta a_2[RP_o]-a_{RPc}[P_{tot}]}{a_{RPc}[P_{tot}]+\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})-a_{RPc}[P_{tot}]} = \frac{\Delta a_1[I_1]+\Delta a_2[RP_o]}{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})} =$$

$$\frac{\Delta a_1}{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})}[I_1] + \frac{\Delta a_2}{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})}[RP_o]$$

-Equation S15

Equaling Equation S15 to the fitting expression through concentration fraction (Equation S16),

$$\frac{F-F_{min}}{F_{max}-F_{min}} = A_1\frac{[I_1]}{[P_{tot}]} + A_2\frac{[RP_o]}{[P_{tot}]}$$ - Equation S16

We get for amplitudes A_1 and A_2

$$A_1 = \frac{\Delta a_1}{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})}[P_{tot}], \quad A_2 = \frac{\Delta a_2}{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})}[P_{tot}]$$

-Equation S17-S18

Note that from Equation S18:

$$\frac{1}{\Delta a_1[I_1](t_{max})+\Delta a_2[RP_o](t_{max})}[P_{tot}] = \left(\frac{F_{min}}{F_{max}-F_{min}}\right)_{av} \frac{1}{a_{RPc}}$$

Fitting amplitude A_1 (Equation S17) can be expressed through $\left(\frac{F_{min}}{F_{max}-F_{min}}\right)_{av} as$:

$$A_1 = \frac{\Delta a_1}{a_{RPc}} \left(\frac{F_{min}}{F_{max}-F_{min}}\right)_{av}, \qquad \text{-Equation S19}$$

Relative intrinsic intensity is

$$A_1^{rel} = \frac{a_1}{a_{RPc}} = 1 + \frac{\Delta a_1}{a_{RPc}} = 1 + A_1 \left(\frac{F_{min}}{F_{max} - F_{min}}\right)_{av}$$

And similarly for all other species, where i is

$$A_i^{rel} = \frac{a_i}{a_{RPc}} = 1 + \frac{\Delta a_i}{a_{RPc}} = 1 + A_i \left(\frac{F_{min}}{F_{max}-F_{min}}\right)_{av}, \; i = \{I_{1E}, I_{1M}, I_{1L}, RP_o\} \qquad \text{-Equation S20}$$

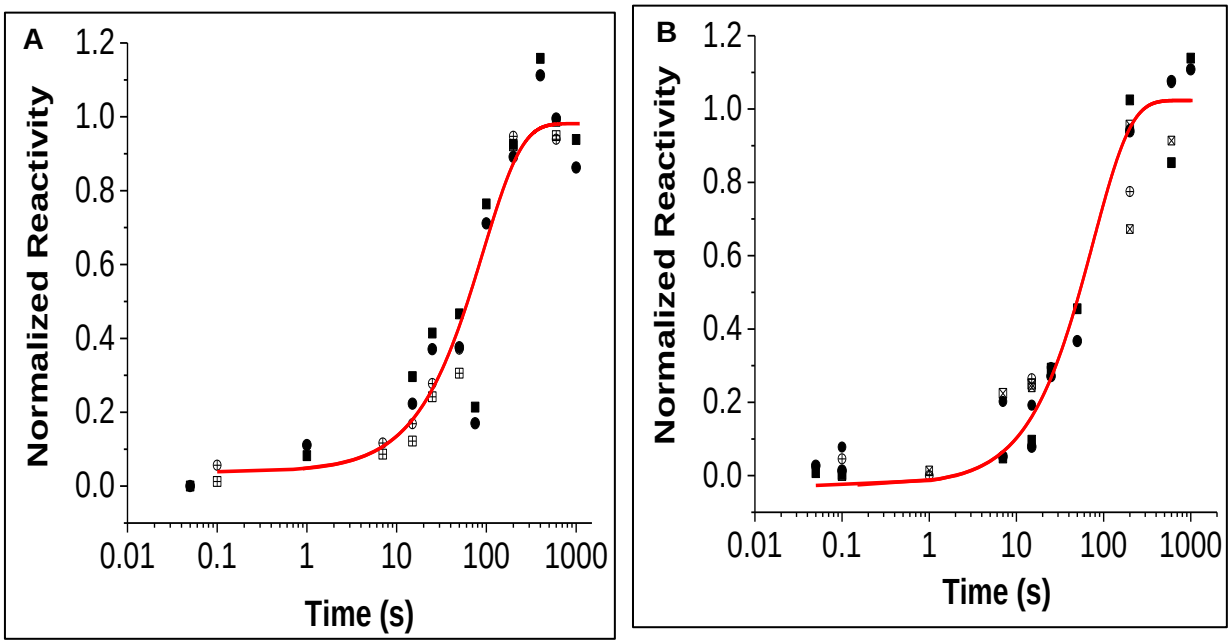
Relative fluorescence intensities of each species in a 5 step mechanism of OC formation for PIFE data is shown in supplemental Table 5.

7. **Distance estimation between upstream -100 and downstream +14 end using normalized fluorescence FRET amplitudes** – The distance between the upstream -100 end and downstream +14 position of the promoter DNA in each closed complex intermediate species is obtained using the relative FRET amplitudes shown in supplemental table 5. The amplitudes of both Cy5+14 and Cy5-100 FRET are averaged and then scaled to determine the FRET efficiency values of $\mathrm{I_{1E}}, \mathrm{I_{1M}}$ and $\mathrm{I_{1L}}$ species using the previously obtained FRET efficiency value of the open complex RP_O, 0.32 ± 0.11 (Sreenivasan et al'16) as shown in Supplemental

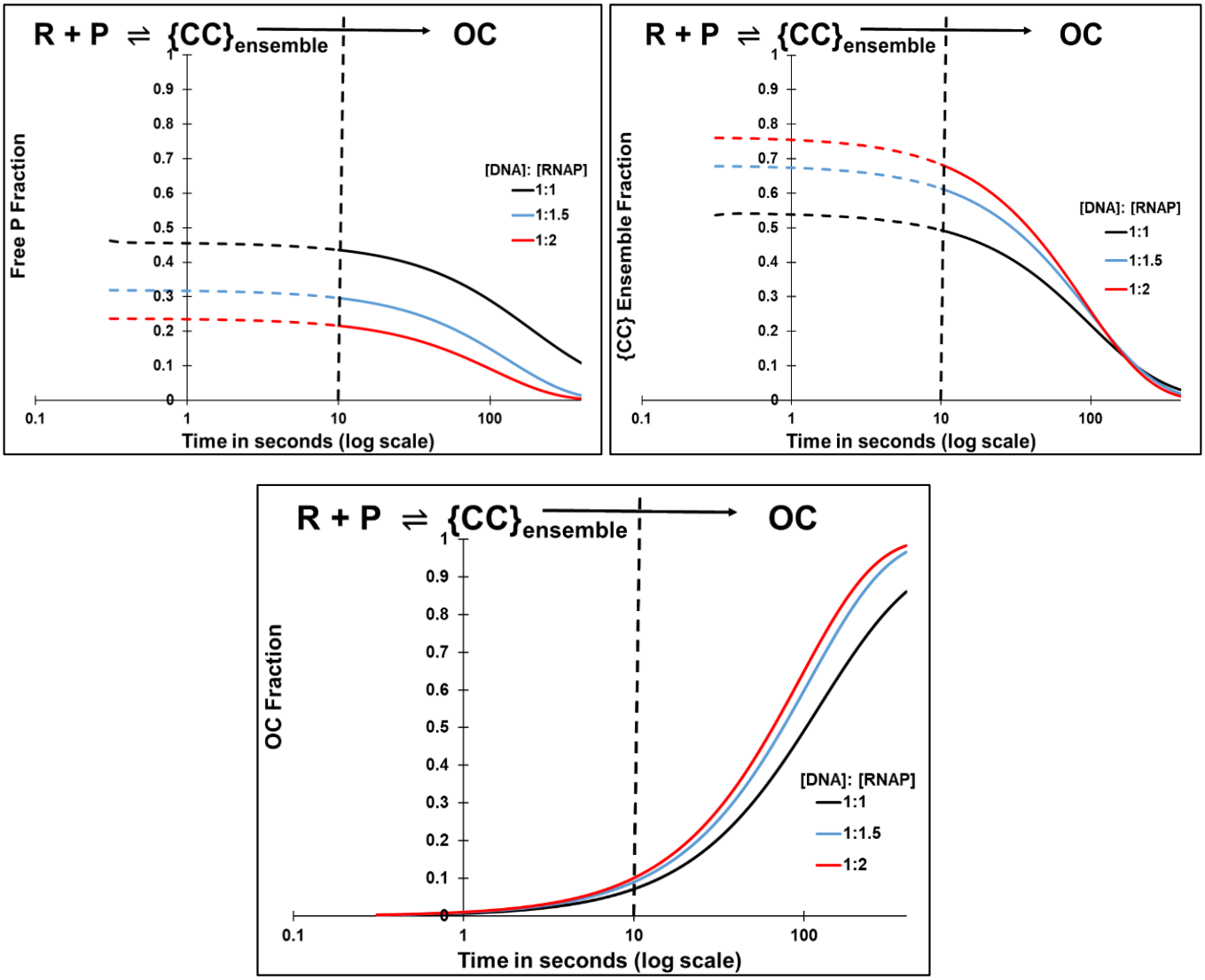
Table 6. The FRET efficiency values are converted to the distance using equation S21 and Forster radius values previously estimated for the same system at 19^0C, R_0 = 56 ± 1 A). The distances are schematically presented for each intermediates in Figure 8.

$$R = R_0 \left(\frac{1}{E} - 1\right)^{1/6}$$ -Equation S21

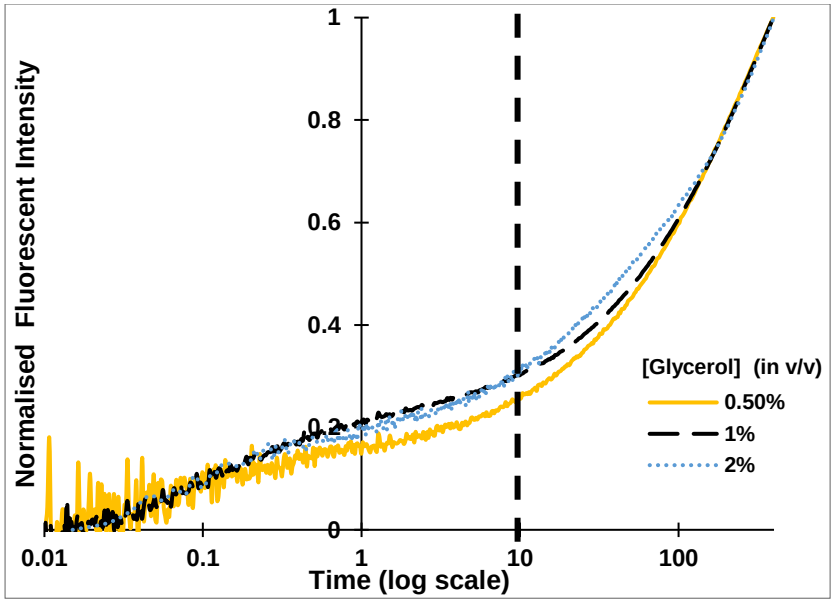
Figures:



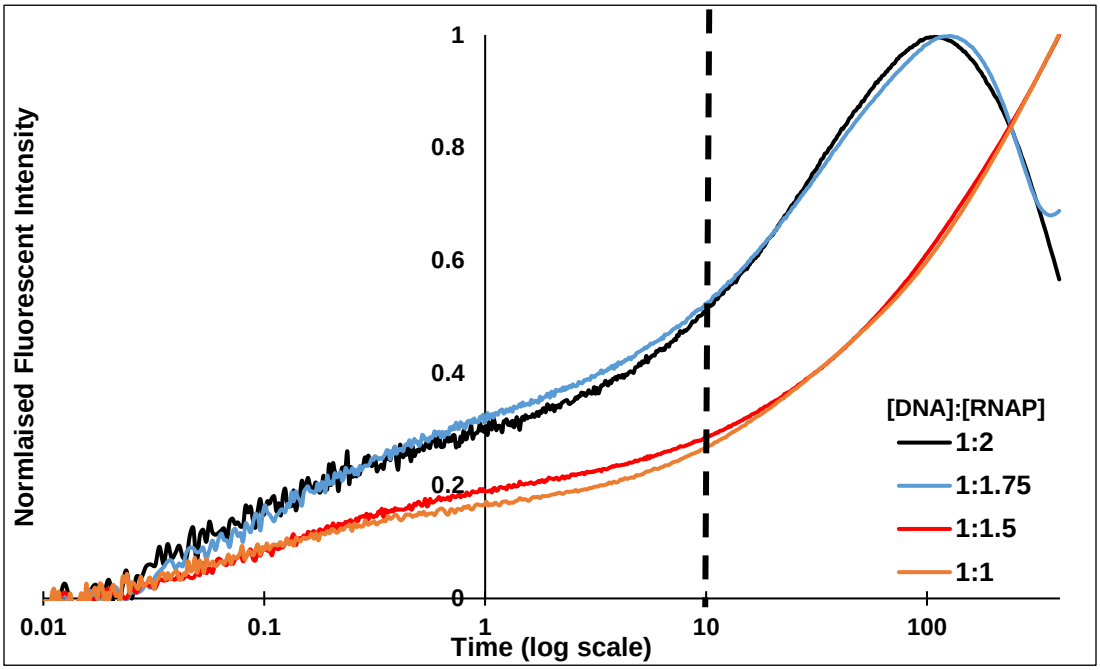
Supplemental Figure 1: Kinetics of OC formation monitored by MnO_4^- reactivity in log scale. Fast permanganate kinetics of OC formation (0.15s) plotted in log time scale for the data shown in Figure 2. Panel A and B corresponds to the thymine reactivity in non-template (nt) and template (t) strand respectively with the same data labels as described in Figure 2 captions.



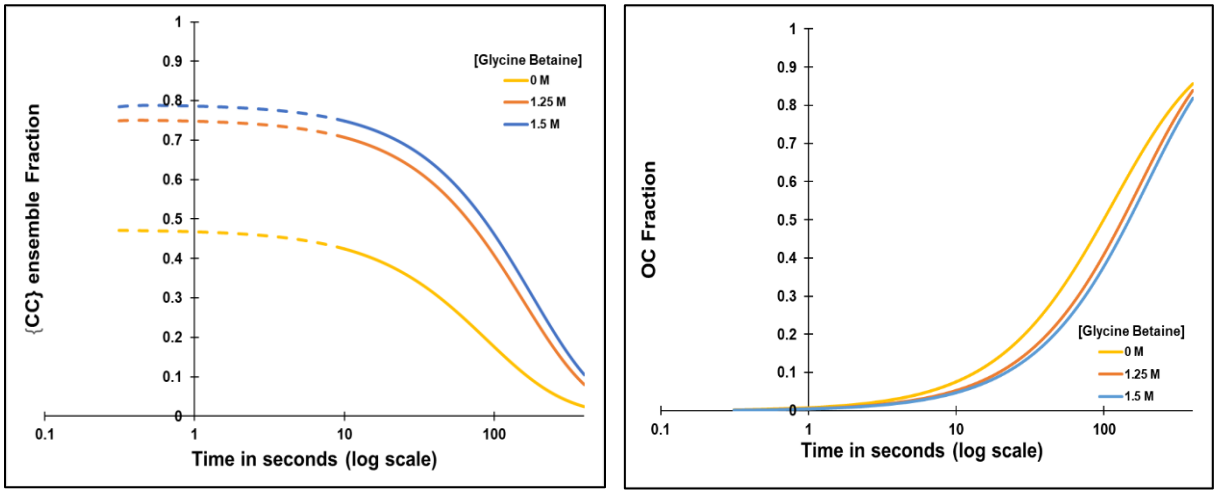
Supplemental Figure 2: Simulated kinetics of OC formation at different molar ratio of RNAP: DNA mixing. Predicted time course of fractional populations of free promoter DNA, closed complexes ensemble ({CC} and I_1) and OC formed as a function of time (log scale) covering the range from 0.01s to 400 s. The dashed vertical line at 10 s marks the onset of OC formation from the equilibrium mixture of {CC} and free promoter DNA (see text). Black line corresponds to 50 nM DNA mixed with 50nM RNAP, blue line is for 50nM DNA mixing with 75nM RNAP and red corresponds to 50nM DNA mixed with 100nM RNAP for all the three figures.



Supplemental Figure 3: OC formation FRET acceptor (Cy5) signal in buffer with glycerol concentration range (0.5-2% v/v). Control FRET detected OC formation kinetic experiments were performed to determine whether there is any effect of glycerol concentration, which was increased with incoming RNAP concentration in experiments in Figure 4A. Results in supplemental Figure 2 are FRET kinetics of mixing Cy3+14 Cy5-100 λP_R DNA with RNAP (1:1 molar ratio, 50 nM final concentration) at different glycerol concentrations in fluorescence buffer; 0.5% - yellow line, 1% - black line and 2%- blue line. These results verify that the FRET signal are same within the uncertainty at these glycerol concentrations. The dashed vertical line at 10 s marks the onset of OC formation from the equilibrium mixture of {CC} and free promoter DNA (see text).

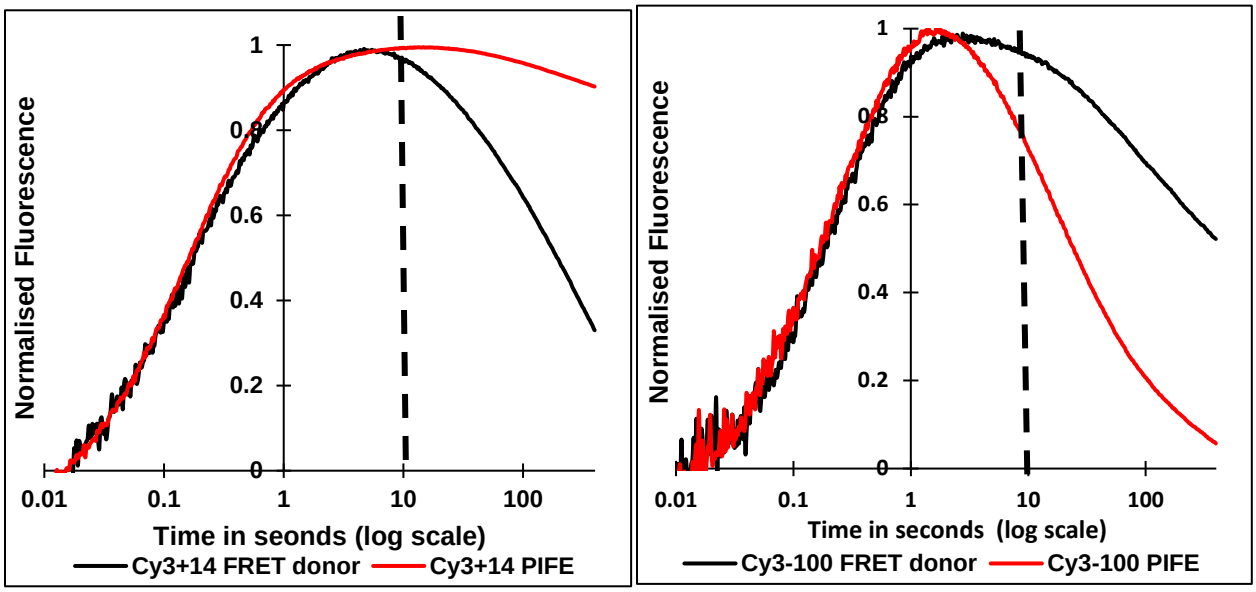


Supplemental Figure 4: Uncorrected OC formation Cy5-100 FRET signal obtained by RNAP:DNA mixing at >1:1 stoichiometric ratio. 50 nM Cy3+14 Cy5-100 λP_R DNA mixed with RNAP, 1:1 (orange line), 1:1.5 (red line), 1;1.75 (blue line) and 1:2 (black line) molar ratio.

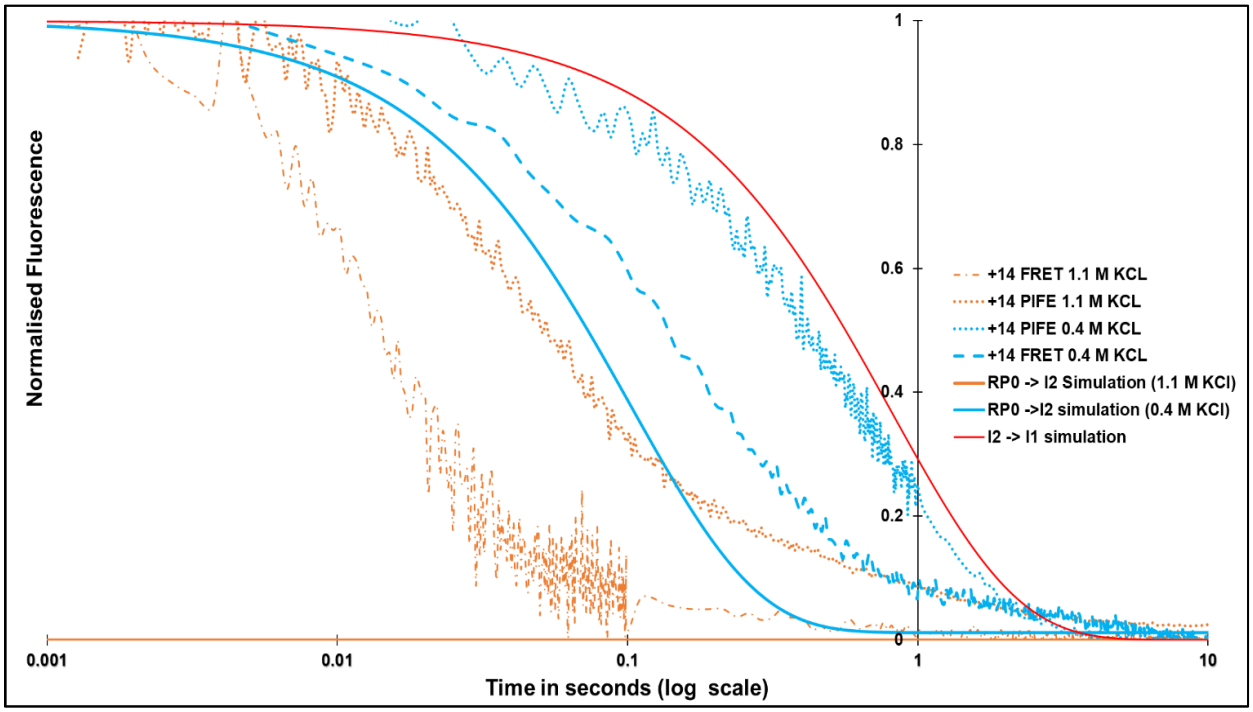


Supplemental Figure 5: Simulated OC formation kinetics in 1.25 M and 1.5 M glycine betaine. Time course of fractional populations of closed complexes ensemble ({CC} and I_1) and OC formed as a function of time (log scale) covering the range from 0.01s to 400 s at different glycine betaine concentration, 0 M corresponds to yellow line, 1.25 M orange and 1.5 M blue colored line. The increase in concentration of CC fraction at 1 sec from 0 to 1.5 M GB as shown in panel A left is similar to the fluorescence signal increase observed in figure 4B.

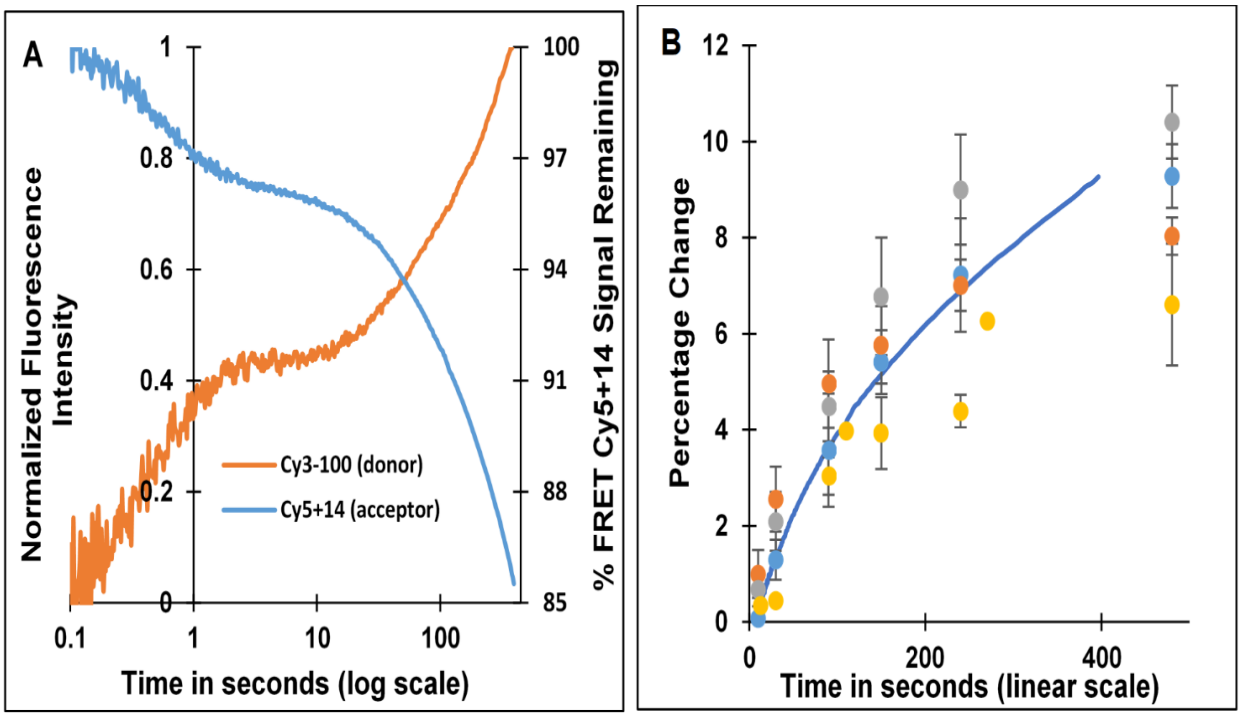
Supplemental Figure 6: Cy3+14 monitored kinetics of mixing E.coli RNAP and λP_R DNA (each at final concentration 50nM) in OC Formation.



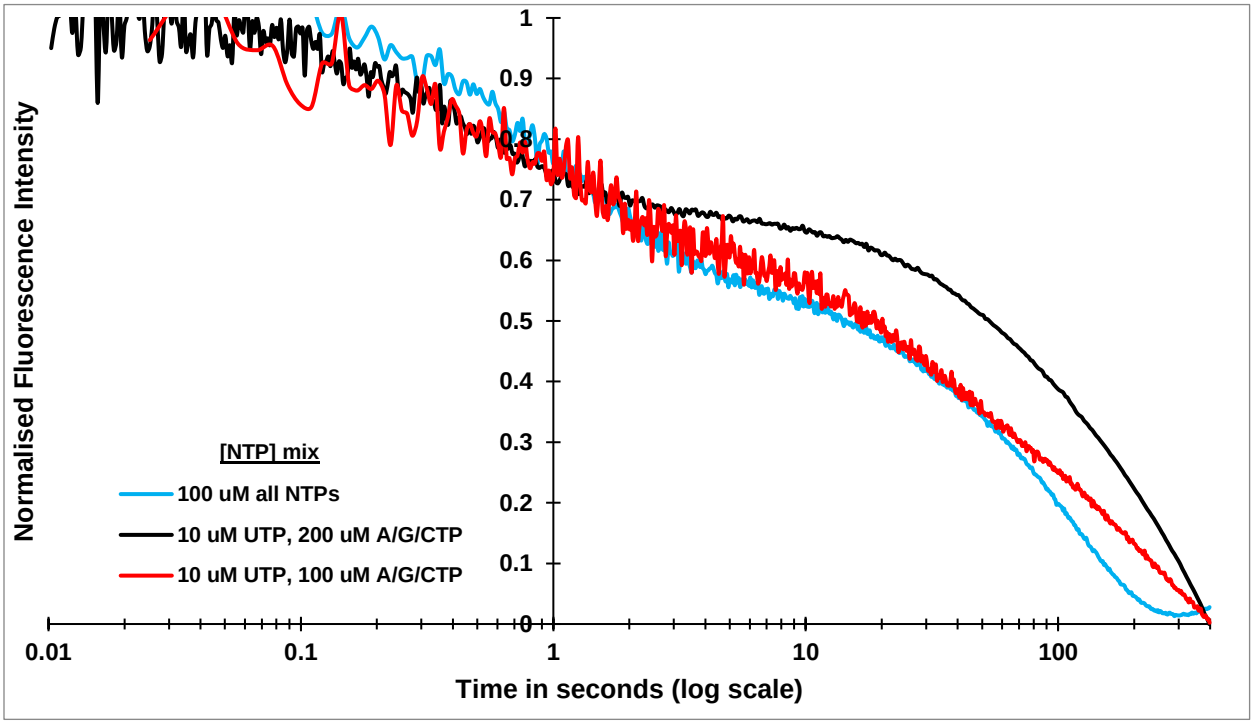
Black curve is Cy3 FRET donor signal from double dye DNA fragment-RNAP OC formation in comparison to the red curve of single dye Cy3 PIFE signal at +14 **(panel A, left)** and -100 position **(panel B, right)**. The vertical dashed line at 10 seconds corresponds to the onset of OC formation from the CC ensemble and free promoter DNA. (Figure 3).



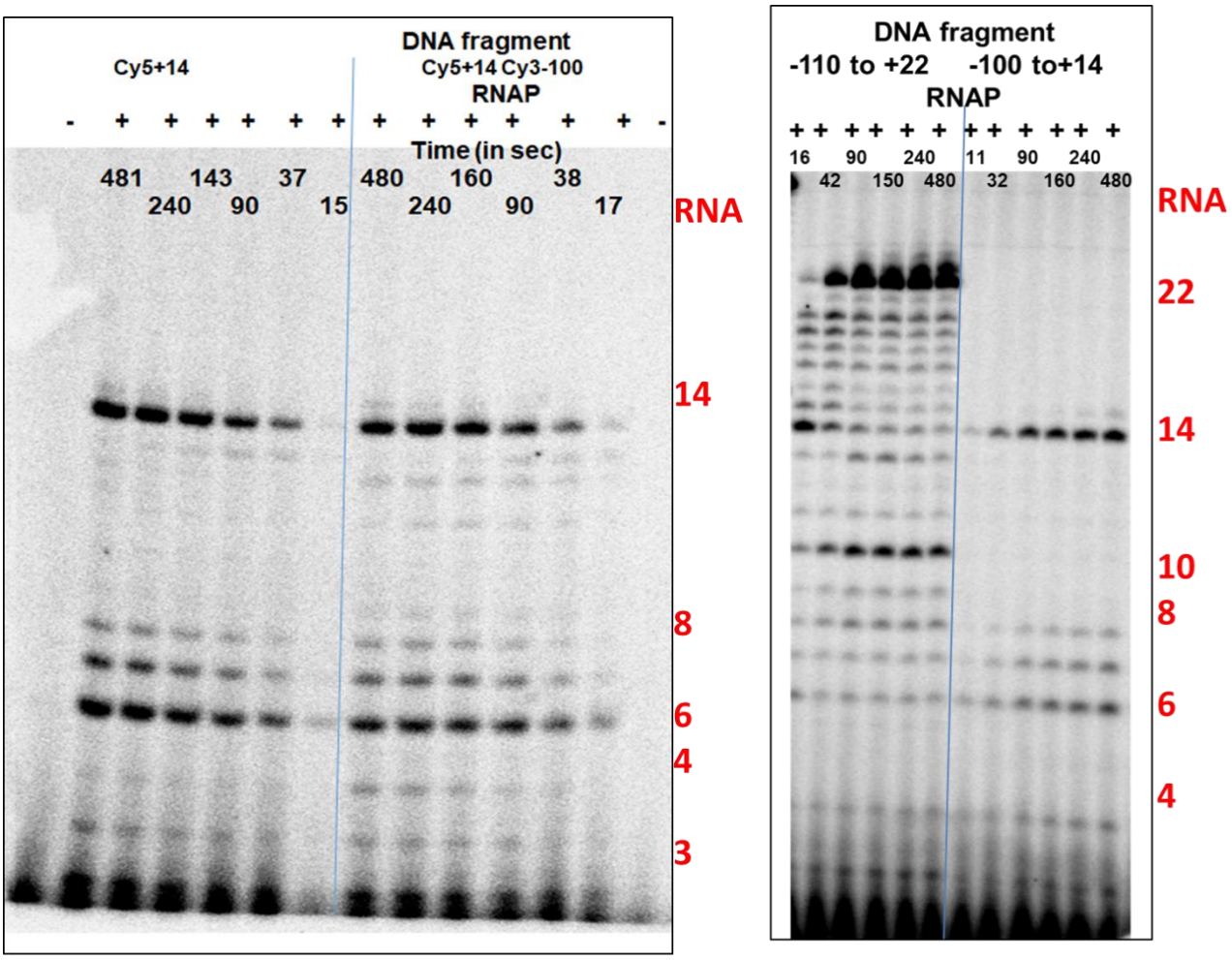
Supplemental Figure 7 Salt-upshift Unwrapping (FRET) and OC Dissociation (PIFE) Kinetics. Representative time courses (log scale) of reductions in Cy3(+14) PIFE (solid line) and Cy5(+14) FRET (dashed line) obtained after mixing OC with KCl at final concentrations of 50nM OC and of 1.1 M KCl (orange curve) and 0.4 M KCl (blue curve), respectively. Curves shown in the figure are the normalised average of a series of shots from one experiment, that is representative of various replicates done on different dates at same RNAP and promoter solutions (see Methods and Supplemental Table 3). The emission curves plotted alongside the two are the predicted, time-dependent reductions in OC: in blue, the amount of original OC (RP_o) resulting from $RP_o \rightarrow I_2$ conversion with rate constant k_{-3} at 1.1 M and 0.4 M KCl in orange and bue dotted line, respectively. The amount of intermediate OC (I_2) resulting from the DNA closing step ($I_2 \rightarrow$ free promoter DNA) with rate constant k_{-2} is plotted in red.



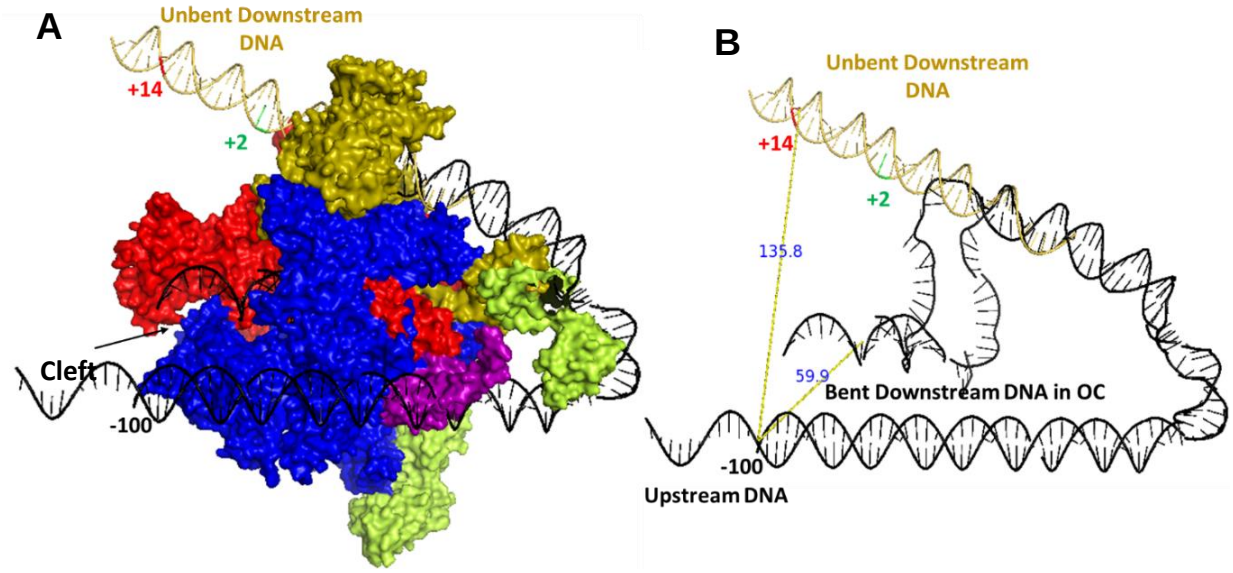
Supplemental Figure 8 FRET-detected unwrapping of λP_R Promoter DNA from RNAP in Transcription Initiation. A) Representative time course (log scale) of Cy5 FRET emission after rapid addition of NTPs (100 μM final concentrations) to pre-formed OC (1:1 [RNAP]:[DNA], 50 nM final OC concentration after mixing with NTPs) and excitation of Cy3 dye at 515nm. FRET curves are normalized by the intensity at 400 s: the un-normalized reduction in Cy5 acceptor signal is approximately 15 ± 2% (average of 5 replicates) occurring in two time intervals as shown. **B)** Blue curve is the percent change in FRET acceptor signal (Cy5+14) over time (linear scale) in the slower time phase (5s – 400 s). Circles represent the percentage of OC completing synthesis of a full length RNA transcripts as a function of time in a single- round transcription assay from double dye: Cy5(+14) Cy3(-100) (orange), single dye Cy5(+14) (blue), unlabelled DNA fragment of similar length (-100 to +14) (grey) and (-110 to +22) (yellow) OC (Supplemental figure 5).



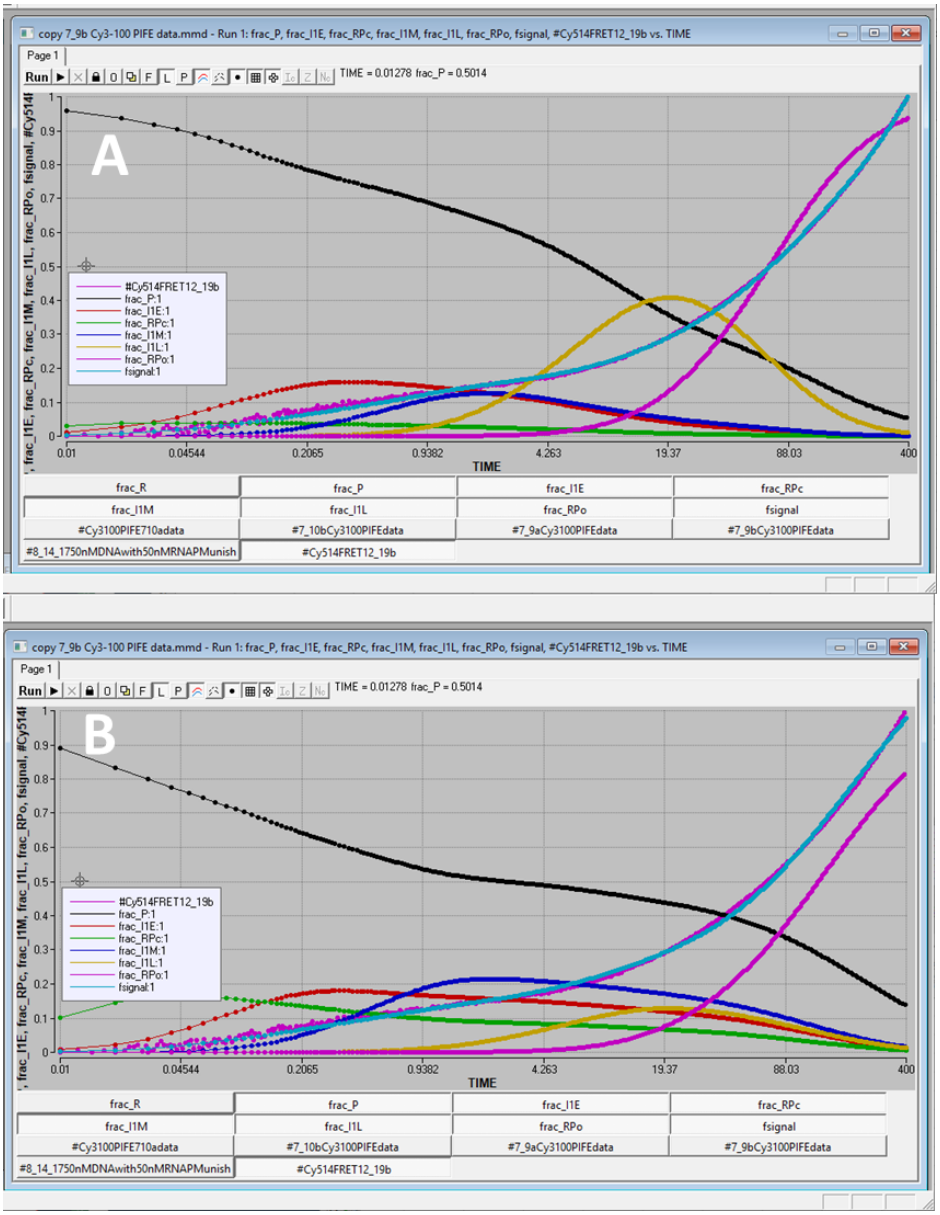
Supplemental Figure 9: Cy5-100 FRET monitored kinetics of initiation from E.coli RNAP-λP$_R$ DNA OC at various NTP concentrations. Blue curve is Cy5-100 FRET acceptor signal when Cy3+14 Cy5-100 λP$_R$ DNA-RNAP OC is mixed with 10 uM UTP and 200 uM rest of the NTPs (A/G/CTP). Red curve is Cy5-100 FRET acceptor signal change when Cy3+14 Cy5-100 λP$_R$ DNA-RNAP OC is mixed with 10 uM UTP and 100 uM rest of the NTPs (A/G/CTP). Black curve is the Cy5+14 FRET acceptor change on mixing OC with 100 uM NTP mix as shown in Supplemental Figure 8A.



Supplemental Figure 10 Single round transcription kinetics. Time course of synthesis of RNA transcripts formed from single round transcription quenched at different time points. Left panel: Template is dye-labelled DNA fragment (Cy5+14 Cy3-100; Cy3+14) and in Right Panel: Unlabeled DNA fragment with the same size duplex (-100 to +14) and (-110 to+22).



Supplemental Figure 11 Structural Model of Unbent downstream DNA into the cleft of RNAP but with bent/wrapped upstream promoter DNA. Panel A left: Structural model of E.coli RNAP-DNA OC with bent and wrapped upstream and bent downstream DNA into the cleft from Sreenivasan et al, 16 along with the unbent downstream DNA. DNA in golden color is the extended unbent downstream DNA projecting straight from -11 upstream position relative to the transcription start site. Color code of subunits of RNAP: beta' (blue), beta (red), alpha (green), sigma70 (wheat) and omega (magenta). DNA in OC (black) and extended unbent downstream DNA (golden color) with +2 (red) and +14 (green) labeled position. **Figure 6 panel B right** contains both downstream bent and unbent DNA without the subunits of RNAP protein with same color scheme shown in left panel A. The distance between upstream -100 and downstream +14 position is decreased by ~ 2 fold on bending the downstream DNA into the cleft. The distance between the unbent downstream +14 DNA and upstream -100 is 2.4 fold times the Forster distance between the FRET observed between these 2 upstream (-100) and downstream positions (+14) for OC at 19^{O}C obtained previously.



Supplemental Figure 12. Comparison of concentration distribution from two set of rates for fit of an individual experiment of acceptor data from Table S7. Panel A: faster rates, not RE, Panel B: slower rates and RE. Plots show individual acceptor experimental data, fit with rates and amplitudes from Table S7 and concentration distribution of all species.

References:

[1] Ruff, E. F., Record, M. T., Jr. , and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation, *Biomolecules 5*, 1035-1062.

[2] Estrem, S. T., Gaal, T., Ross, W., and Gourse, R. L. (1998) Identification of an UP element consensus sequence for bacterial promoters, *Proc Natl Acad Sci U S A 95*, 9761-9766.

[3] Hofer, B., Muller, D., and Koster, H. (1985) The pathway of E. coli RNA polymerase-promoter complex formation as visualized by footprinting, *Nucleic Acids Res 13*, 5995-6013.

[4] RT, K. (1987) The 0 "C Closed Complexes between Escherichia coli RNA Polymerase and Two Promoters, T7-A3 and ZacUV5, *Journal of Biological Chemistry 262*, 13654-13661.

[5] Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., and Record, M. T. J. (2002) Kinetic studies and structural models of the association of *E. coli* sigma 70 RNA polymerase with the LPR promoter: large scale conformational changes in forming the kinetically significant intermediates, *Journal of Molecular Biology 319*, 649-671.

[6] Saecker, R. M., Record, M. T. J., and deHaseth, P. L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase-promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis, *Journal of Molecular Biology 412*.

[7] Zuo, Y., and Steitz, T. A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Molecular Cell 58*, 534-540.

[8] Dienemann, C., Schwalb, B., Schilbach, S., and Cramer, P. (2019) Promoter Distortion and Opening in the RNA Polymerase II Cleft, *Mol Cell 73*, 97-106 e104.

[9] Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr., and Saecker, R. M. (2007) Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 104*, 7833-7838.

[10] Davis, C. A., Capp, M. W., Record, M. T., Jr., and Saecker, R. M. (2005) The effects of upstream DNA on open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 102*, 285-290.

[11] Ross, W., and Gourse, R. L. (2005) Sequence-independent upstream DNA-aCTD interactions strongly stimulate *Escherichia coli* RNA polymerase-*lac*UV5 promoter association, *Proceedings of the National Academy of Sciences 102*, 291-296.

[12] Sreenivasan, R., Heitkamp, S., Chhabra, M., Saecker, R. M., Lingeman, E., Poulos, M. A., McCaslin, D., Capp, M. W., Artsimovitch, I., and Record, M. T. J. (2016) Fluorescence resonance energy transfer characterization of DNA wrapping in closed and open *Escherichia coli* RNA polyerase-lP_R promoter complexes, *Biochemistry 55*, 2174-2186.

[13] Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., Saecker, R. M., and Record, M. T. J. (1998) DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA, *Journal of Molecular Biology 283*, 741-756.

[14] Rogozina, A., Zaychikov, E., Buckle, M., Heumann, H., and Sclavi, B. (2009) DNA melting by RNA polymerase at the T7A1 promoter precedes the rate-limiting step at 37 degrees C and results in the accumulation of an off-pathway intermediate, *Nucleic Acids Res 37*, 5390-5404.

[15] Sclavi, B., Zaychikov, E., Rogozina, A., Walther, F., Buckle, M., and Heumann, H. (2005) Real-time characterization of intermediates in the pathway to open complex formation by *Escherichia coli* RNA polymerase at the T7A1 promoter, *Proceedings of the National Academy of Sciences 102*, 4706-4711.

[16] Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002) Structure of the bacterial RNA polymerase promoter specificity s subunit, *Molecular Cell 9*, 527-539.

[17] Hwang, H., and Myong, S. (2014) Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions, *Chem Soc Rev 43*, 1221-1229.

[18] Heyduk, E., and Heyduk, T. (2018) DNA template sequence control of bacterial RNA polymerase escape from the promoter, *Nucleic Acids Res 46*, 4469-4486.

[19] Feklistov, A., Bae, B., Hauver, J., Lass-Napiorkowska, A., Kalesse, M., Glaus, F., Altmann, K. H., Heyduk, T., Landick, R., and Darst, S. A. (2017) RNA polymerase motions during promoter melting, *Science 356*, 863-866.

[20] Ko, J., and Heyduk, T. (2014) Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse, *Biochemistry Journal 463*, 135-144.

[21] Rammohan, J., Ruiz Manzano, A., Garner, A. L., Prusa, J., Stallings, C. L., and Galburt, E. A. (2016) Cooperative stabilization of Mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD, *Nucleic Acids Res 44*, 7304-7313.

[22] Hubin, E. A., Fay, A., Xu, C., Bean, J. M., Saecker, R. M., Glickman, M. S., Darst, S. A., and Campbell, E. A. (2017) Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA, *Elife 6*.

[23] Gries, T. J., Kontur, W. S., Capp, M. W., Saecker, R. M., and Record, M. T. J. (2010) One-step DNA melting in the RNA polyerase cleft opens the initiation bubble to form an unstable open complex, *Proceedings of the National Academy of Sciences 107*, 10418-10423.

[24] Henderson, K. L., Felth, L. C., Molzahn, C. M., Shkel, I., Wang, S., Chhabra, M., Ruff, E. F., Bieter, L., Kraft, J. E., and Record Jr., M. T. (2017) Mechanism of transcription initiation and promoter escape by *E. coli* RNA polymerase, *Proceedings of the National Academy of Sciences 114*, E3032-E3040.

[25] Gribskov, M., and Burgess, R. R. (1983) Overexpression and purification of the sigma subunit of Escherichia coli RNA polymerase, *Gene 26*, 109-118.

[26] Roe, J.-H., Burgess, R. R., and Record, M. T. J. (1984) Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λPR promoter, *Journal of Molecular Biology 176*, 495-521.

[27] Drennan, A., Kraemer, M., Capp, M. W., Gries, T., Ruff, E. F., Sheppard, C., Wigneshweraraj, S., Artsimovitch, I., and Record, M. T., Jr. (2012) Key roles of the downstream mobile jaw of *Escherichia coli* RNA polymerase in transcription initiation, *Biochemistry 51*, 9447-9459.

[28] Tsodikov, O. V., Craig, M. L., Saecker, R. M., and Record, M. T., Jr. (1998) Quantitative analysis of multiple-hit footprinting studies to characterize DNA conformational changes in protein-DNA complexes: application to DNA opening by Esigma70 RNA polymerase, *J Mol Biol 283*, 757-769.

[29] Roe, J.-H., Burgess, R. R., and Record Jr., M. T. (1985) Temperature dependence of the rate constants of the *Escherichia coli* RNA polymerase-Lambda PR promoter interaction. Assignment of the kinetic steps corresponding to protein conformational change and DNA opening, *Journal of Molecular Biology 184*, 441-453.

[30] Schoen, I., Krammer, H., and Braun, D. (2009) Hybridization kinetics is different inside cells, *Proc Natl Acad Sci U S A 106*, 21649-21654.

[31] Sanborn, M. E., Connolly, B. K., Gurunathan, K., and Levitus, M. (2007) Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA, *J Phys Chem B 111*, 11064-11074.

[32] Kuznetsov, S. V., Sugimura, S., Vivas, P., Crothers, D. M., and Ansari, A. (2006) Direct observation of DNA bending/unbending kinetics in complex with DNA-bending protein IHF, *Proc Natl Acad Sci U S A 103*, 18515-18520.

[33] Chen, Y. W., Wang, H., Hupert, M., Witek, M., Dharmasiri, U., Pingle, M. R., Barany, F., and Soper, S. A. (2012) Modular microfluidic system fabricated in thermoplastics for the strain-specific detection of bacterial pathogens, *Lab Chip 12*, 3348-3355.

[34] McClure, W. R. (1985) Mechanism and control of transcription initiation in prokaryotes, *Annu Rev Biochem 54*, 171-204.

[35] Tsodikov, O. V., and Record, M. T. (1999) General Method of Analysis of Kinetic Equations for Multistep Reversible Mechanisms in the Single-Exponential Regime: Application to Kinetics of Open Complex Formation between Eσ70 RNA Polymerase and λPR Promoter DNA, *Biophysical Journal 76*, 1320-1329.

[36] Kontur, W. S., Saecker, R. M., Davis, C. A., Capp, M. W., and Record, M. T., Jr. (2006) Solute probes of conformational changes in open complex (RPo) formation by Escherichia coli RNA polymerase at the lambdaPR promoter: evidence for unmasking of the active site in the isomerization step and for large-scale coupled folding in the subsequent conversion to RPo, *Biochemistry 45*, 2161-2177.

[37] Cowing, D. W., Mecsas, J., Record, M. T., and Gross, C. A. (1989) Intermediates in the Formation of the Open Complex by Rna-Polymerase Holoenzyme Containing the Sigma Factor Sigma-32 at the Groe Promoter, *Journal of Molecular Biology 210*, 521-530.

[38] Rao, L., Ross, W., Appleman, J. A., Gaal, T., Leirmo, S., Schlax, P. J., Record, M. T., Jr., and Gourse, R. L. (1994) Factor independent activation of rrnB P1. An "extended" promoter with an upstream element that dramatically increases promoter strength, *J Mol Biol 235*, 1421-1435.

[39] 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 Res 46*, 7284-7295.

[40] Rivetti, C., Guthold, M., and Bustamante, C. (1999) Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex, *EMBO 18*, 4464-4475.

[41] Doniselli, N., Rodriguez-Aliaga, P., Amidani, D., Bardales, J. A., Bustamante, C., Guerra, D. G., and Rivetti, C. (2015) New insights into the regulatory mechanisms of ppGpp and DksA on Escherichia coli RNA polymerase-promoter complex, *Nucleic Acids Res 43*, 5249-5262.

Chapter 4: Fluorescence Kinetic-Mechanistic Studies of Inhibition of OC Formation by the Drug Lipiarmycin

PREFACE

This Chapter is in preparation for publication with myself as the first author. My contribution to this research included: data collection, figure preparation and designing experiments. Malavika Abhineshababu Sridevi helped in performing replicates and collecting more data at other LpM concentration. Rezwana Karim and Michael Capp helped in purifying *E.coli* RNAP core and sigma-70 factor protein used in the experiments.

Introduction

The kinetics and thermodynamics of open complex formation by *E. coli* RNA polymerase (RNAP) and promoter DNA are regulated by promoter sequence, transcription factors (activators, repressors), drugs and other ligands, and solution conditions. Effects of promoter sequence as well as factors and ligands appear to be primarily on the reversible steps of forming the initial closed complex (RP_c) and/or converting it stepwise to the most advanced closed complex (I_{1L}) that undergoes the central mechanistic step of DNA opening, and not on the DNA opening step itself. As discussed previously, this situation is analogous to that seen in enzyme catalysis, where most regulation by inhibitors or activators occurs in the reversible steps of substrate binding and conformational changes that precede the catalytic step, and not in the catalytic step. In Chapter 3, we obtained kinetic evidence for three advanced closed complexes (designated I_{1E}, I_{1M} and I_{1L}) that follow the initial closed complex (RP_c) in the mechanism of open complex formation by RNAP at the λP_R promoter, and characterized these intermediates by FRET and PIFE. Effects of drugs, other ligands, promoter sequence and transcription factors on the rate of open complex formation, quantified by these methods, can now be interpreted in terms of this mechanism.

Closed complex intermediates RP_c, I_{1E}, I_{1M} and I_{1L} form sequentially in a series of rapidly-reversible steps occurring in the time range from <10 ms to ~30 seconds after mixing, prior to and during the irreversible conversion of the most advanced closed complex (I_{1L}) to the stable open complex RP_O (Mechanism 1).

$$R + P \underset{k_{-1}}{\overset{k_1}{\leftrightarrows}} RP_c \underset{k_{-2}}{\overset{k_2}{\leftrightarrows}} I_{1E} \underset{k_{-3}}{\overset{k_3}{\leftrightarrows}} I_{1M} \underset{k_{-4}}{\overset{k_4}{\leftrightarrows}} I_{1L} \overset{k_5}{\rightarrow} RP_o \qquad \text{Mechanism I}$$

From the kinetic-mechanistic analysis and the intrinsic fluorescence characteristics of the intermediates determined in Ch. 3, we proposed that the RNAP clamp opens in the conversion of

I_{1E} to I_{1M} in order to admit the downstream duplex in the conversion of I_{1M} to I_{1L}, but otherwise is closed to duplex DNA.

Different transcription activators and inhibitors modulate the transcription initiation process by targeting one or more of these rapidly reversible steps in the ensemble of closed complex intermediates. As an example, antibiotics like Lipiarmycin (LpM, also called as Fidaxomicin, Clostomicin B1, lipiarmicin, lipiarmycin A3, OPT 80, PAR 01, PAR 101,and tiacumicin B) and Myxopyronin (Myx) are are proposed to act on CC or OC intermediates that precede the stable open complex [1]. Lpm and Myx both inhibit the formation of open complexes by binding to the switch region at the base of the RNAP clamp. This switch region is part of the hinges that allow opening and closing of the clamps[2]. The switch region, highly conserved in both Gram-positive and Gram-negative bacteria, is also the binding site for corallopyronin and ripostatin [2], two other broad-spectrum antibacterial transcriptional inhibitors. It is important to understand the mechanism of action of these antibiotics and discover which closed complex intermediates are the primary drug targets.

LpM is a FDA approved drug used for the treatment of *C. difficile*-associated diarrhea in adults. It is a macrocyclic, narrow-spectrum antibiotic (Figure 1) that exhibits bactericidal activity [3]. Recently, cryo-EM structures were reported for LpM-*M.tuberculosis* RNAP holoeznyme complex [4] and a quaternary complex of *M. tuberculosis* RNAP holoenzyme, LpM, transcription factor RbpA and a synthetic promoter-like fragment with consensus -35, spacer, and -10 elements and a ss region downstream of -10 till -5 position relative to the transcription start site [5]. These indicate that LpM locks RNAP in an open clamp state, preventing clamp closing which in turn would prevent opening of downstream DNA in the active site cleft [5 4]. Mechanism 1 of this chapter identifies the closed-promoter intermediate I_{1M} as an open-clamp state of RNAP, and hence it is a likely target of LpM.

Footprinting studies on ternary RNAP-DNA-LpM closed complexes indicate that RNAP remains bound at the promoter but the downstream duplex including the -10 region is more accessible to nuclease attack than for RNAP-promoter closed complexes in the absence of LpM [6]. Consistent with this, RNAP-LpM closed complexes are unable to bend the downstream duplex into the active site cleft, and are inefficient in flipping the -10 element nucleotide base of non-template promoter DNA into the hydrophobic pocket of σ region 2 [7]. LpM was proposed to act on an I_1 intermediate for λP_R and on a RP_C or RP_I intermediate for lacUV5 promoter DNA [1], but much remained to be learned about which closed complex species is targeted by LpM.

The second order rate constant for binding LpM to RNAP to form a LpM-RNAP complex that inactivates RNAP was found to be approximately 2×10^4 $M^{-1}s^{-1}$ [8] which is only about 10^{-4} that expected from a diffusion-collision analysis. This second order rate constant was determined using an abortive initation assay for the fraction of RNAP not bound to LpM [8] . By contrast, the rate constant for binding promoter DNA to RNAP to form the initial closed complex (RP_C) intermediate is about 3×10^8 M^{-1} s^{-1} [9]. Binding of LpM to RNAP allows CC formation but inhibits OC formation as determined previously by permanganate footprinting assay for different promoters [6]. Which CC are formed in the presence of Lpm and which if any are inhibited is not known.

In the previous chapter, we developed real-time FRET and PIFE kinetics assays to probe the kinetics and mechanism of the conformational changes taking place in the upstream and downstream promoter region of DNA as an ensemble of closed-promoter intermediates forms from the initial closed-promoter complex (RP_c). In the first of these (I_{1E}) the upstream DNA is bent and wrapped on RNAP and the downstream duplex is bent to make contact with the top of the beta lobe. Large-scale movement of the beta side opens the clamp to form I_{1M}, in which the downstream duplex remains on the beta lobe. The downstream duplex is then bent deep into

the clamp to form the late closed-promoter intermediate I_{1L}. Clamp-closing and DNA opening then occur, converting I_{1L} to the initial open complex (I_2).

Here these FRET and PIFE assays are employed to determine the kinetics and mechanism of action of the antibiotic Lpm. Three types of order-of-addition experiments are performed. In one, Lpm at various concentrations (2.5-500 μM) is pre-incubated with RNAP to attain binding equilibrium (half saturation concentration at about 70 μM), and then rapidly mixed with λP_R promoter DNA. In another assay, RNAP is mixed with a solution of λP_R promoter DNA containing various concentrations of LpM (70 μM to 1.7 mM). Because the second order rate constant for binding LpM to RNAP is 4 orders of magnitude less than that for binding promoter DNA to RNAP to form RP_c, initial promoter binding occurs more rapidly in this experiment than LpM binding for the concentrations investigated. In a third type of kinetics assay, double mixing experiments were employed. In some cases, RNAP was first mixed with λP_R promoter DNA, and then 71 μM Lpm was added at various times. Alternatively RNAP was first mixed with 71 μM LpM for various times and λP_R promoter DNA was then added. These experiments provide support for the proposal that LpM binds to the open clamp state of RNAP and of closed-promoter complexes, targeting the closed complex intermediate I_{1M} to form the off-pathway adduct Lpm-I_{1M}. Because all members of the CC ensemble are in rapid equilibrium with one another, formation of Lpm-I_{1M} reduces the population of $I_{1L}$and therefore reduces the rate of open complex formation.

Materials and Methods

Buffers:

Storage Buffer (SB) for RNA polymerase holoenzyme and Fluorescence buffer (FB) for FRET and PIFE kinetics experiments were prepared as described previously in Chapter 2 and 3. Lipiarmycin (LpM) was purchased under the product name Fidaxomicin from ApexBio Technology LLC in 10 mM concentration present in 1 ml pure DMSO solvent.

Preparation of E. coli RNA Polymerase (RNAP) Holoenzyme and Fluorescent- and ^{32}P-labeled λP_R Promoter DNA:

RNAP core enzyme was overexpressed and purified as described previously, using *E. coli* BL21(λDE3) transformed with pVS10 or pIA900 [10]. σ^{70} was overexpressed and purified using *E. coli* M5219 transformed with plasmid pMRG8, as described previously [11] . Purified RNAP core enzyme and σ^{70} were incubated at a 1:2 molar ratio for 1 hour at 37°C in SB to reconstitute RNAP holoenzyme, and stored at -20°C until use. Different preparation of RNAP ranged from 50-90% activity in open complex formation as determined by filter binding assay [12].

Single-dye-labeled λP_R fragments [Cy3 (-100) or Cy5 (-100), Cy3 (+14) or Cy5 (+14)] for PIFE experiments and two-dye-labeled λP_R promoter DNA fragments [Cy3 (+14) Cy5 (-100) and Cy3 (-100) Cy5 (+14)] for FRET experiments were prepared by PCR as described previously in Chapter 2 and 3 using dye-labeled primers [13] .

Kinetics of Open Complex Formation from Free RNAP and Promoter DNA in the presence and absence of LpM by Stopped-Flow Fluorescence (FRET, PIFE):

Fluorescence kinetic experiments were performed at 19°C in a Kintek SF-300X stopped flow fluorimeter (Kintek Corporation, PA) equipped with a 150 Watt Hg-Xe lamp (Hamamatsu, Japan) by rapid-mixing of equal volumes (20 µL) of dye-labeled promoter DNA and RNAP stock solutions as described previously in Chapter 2.

Stocks of DNA, RNAP and LpM were prepared at twice the desired final concentration in FB to account for dilution upon mixing in the stopped flow. Experiments were performed at a 1:1 mole ratio of active RNAP to promoter DNA, with 50 nM final concentrations of each in the reaction cell. RNAP-LpM complex was formed by pre-incubating RNAP and LpM mixture in FB at 37^0C for at least 10 minutes.

Control experiments (5-6 shots) were performed first by mixing DNA solution in FB with RNAP in the absence of LpM to get fluorescence signal as previously obtained in Chapter 3. The DNA solution was then mixed with RNAP pre-incubated with different LpM concentration (2.5 μM – 1.7 mM) in FB. 6-8 shots were collected for each drug concentration. At 2.5 μM LpM concentration the amount of DMSO solvent in the mixture was 35.25 mM in the observation cell, and in 1.7 mM LpM it was 2.12 M. Control experiments with varying amount of DMSO (0.04 to 2.12 M final concentration) in FB of DNA and RNAP solution mixing to form OC was performed to test for pissible effects of DMSO in fluorescence FRET and PIFE kinetics as shown in Supplemental Figure 2.

The dye-labeled λP_R promoter constructs were excited in the observation cell at wavelengths of 515nm (Cy3) for single dye Cy3 PIFE experiments. In FRET experiments, Cy3 dye (FRET donor) was excited at 515 nm wavelength and fluorescence emission of Cy3 and Cy5 (FRET acceptor) were monitored with different PMT as a function of time. Fluorescence intensity was monitored from 10ms to 400s with 600 data points uniformly distributed on a log time scale.

Two types of double mixing experiment were performed. In one, RNAP was rapidly mixed with 71 μM LpM, sufficient to give about 90% binding of Lpm at equilibrium, held for either 0.03, 0.3, 3, 30, or 300 s, and then mixed with DNA. In the other type of double mixing experiment, RNAP was rapidly mixed with DNA, held for either 0.03, 0.3, 3, 30, or 300 s, and then rapidly mixed with 71 μ LpM. All three SF syringes were used in these experiments, and each syringe was filled with 3X the final concentration desired in the observation cell to account for the dilution upon mixing the solutions. The first mixing was in 1:1 volume ratio and the second mixing was in 2:1 volume ratio. Fluorescence intensity was collected in logarithmic time scale from 0.01 to 400 seconds with 600 data-points upon the beginning of the second mixing. Control experiments were performed first without LpM by mixing DNA/RNAP in FB. 3-4 shots were collected in each type of double-mixing experiments performed in the instrument.

When LpM was combined with DNA and mixed with RNAP, initial 5-6 shots were collected on mixing DNA solution in one syringe with the RNAP solution in FB present in the other syringe of the instrument to get control signal for 0 uM LpM condition. Then, RNAP solution is mixed with DNA mix containing different concentration of LpM in FB. 7-8 shots were collected for different concentration of LpM.

Control experiment of mixing double/single dye labeled DNA solution in one syringe with LpM in the other syringe of instrument was also performed and no change in fluorescence was seen within the time of observation (0.01 to 400 seconds) indicating that LpM doesn't interact with DNA alone to give change in fluorescence.

Data Analysis

Data collected from each shot was normalized to the maximum fluorescence signal change observed in the absence of LpM (0 uM LpM control experiment) and then averaged to reduce noise at earlier time points for further analysis, as described earlier in Chapter 3.

Fluorescence kinetics signal collected in the drug absence (0 uM LpM condition) in double mixing experiments was shifted by 0.01-0.1 fluorescence intensity units to match with the fluorescence signal observed in the presence of 71 uM LpM at earlier time points (10-20 msec). 4-5 individual shots of data collected from each type of double-mixing experiments were then averaged to reduce noise at earlier time points.

Results

Reductions in Both Early and Late Phase FRET Kinetics when Pre-equilibrated *E.coli* RNAP-Lpm Solution is Mixed with Cy3Cy5-labeled λP_R Promoter DNA

In Chapter 3 fluorescence FRET kinetics of OC formation with *E.coli* RNAP using Cy5 and Cy3 at downstream +14 and upstream -100 positions of λP_R promoter DNA fragment were

determined from 0.01 to 400 seconds. We established that the FRET signal observed in first kinetic phase (0.01 to 10 seconds) belongs to an ensemble of closed complex intermediates and the second phase (10 - 400 seconds) to OC formation. Here, on observing the FRET kinetics signal, both the amplitude of the plateau in the first kinetic phase and the rate of the second phase are greatly reduced in increasing LpM concentration (Figure 2). The FRET acceptor signal in the second kinetic phase at +14 and -100 position is reduced to about 50% of the signal observed in the absence of the drug at ~5-10 μM LpM concentration (Figure 2A and B). The amplitude of the plateau observed in the first kinetic phase is reduced to less than 10% of –LpM values at LpM concentration above 70 μM. When the FRET acceptor Cy5 probe is at -100 upstream position the signal in the is reduced more at same LpM concentration than when it is present at +14 downstream position, at 5 uM LpM Cy5+14 signal reduces by only ~20% whereas Cy5-100 is reduced by 50% (Figure 2A and B). Corresponding effects in the FRET donor Cy3 signal is shown in Figure S1. FRET donor Cy3 probe when present at downstream +14 position (Figure S1A) shows more larger reduction in the signal than at upstream -100 position in increasing LpM concentration (Figure S1B). At higher LpM concentration, Cy3+14 signal is reduced by 90% (Figure S1A) as compared to Cy3-100 signal which is reduced by only 50% at the final time-point of 400 seconds. FRET acceptor Cy5 signal being under the direct influence of Cy3 FRET donor, shows similar larger reduction in the signal at upstream -100 than at downstream +14 position in increasing LpM concentration (Figure 1).

Very Different Reductions in FRET Kinetics When Lpm-DNA Solutions are Mixed with RNAP

When *E.coli* RNAP is mixed with Cy3-Cy5 double dye labeled λP_R promoter DNA fragment-LpM mixture different FRET kinetics is observed from previous experiment in increasing LpM concentration (Figure 2 and Figure 3). Cy5-100 and Cy5+14 FRET acceptor shows consistent increase in signal by 10-20% in the first kinetic phase from 0.01 to 1 second as 0 μM LpM condition

and deviates from it in the second kinetic phase, after 1 to 400 seconds in increasing LpM concentration (Figure 3A and 3B).

Cy5-100 FRET acceptor signal in the second kinetic phase plateaus to ~20% in the presence of 250 μM and 300 μM LpM (Figure 3A). At 400 μM and 750 μM LpM concentration, Cy5 -100 FRET acceptor signal linearly decreases from 20% after 1 second to -20% till 400 seconds from the signal observed in the initial time-point (0.01 seconds) (Figure 3A). Corresponding FRET donor Cy3 probe at +14 downstream position increases up to 60-70% from the signal observed in 0 μM LpM condition from 0.01 to 5 seconds in the first kinetic phase in the presence of increasing LpM concentration (Figure S4A). In the second kinetic phase (>5 seconds to 400 seconds), Cy3+14 FRET donor signal decreases by 20% at 300 μM and 400 μM LpM and ~60% at 750 uM LpM concentration (Figure S4A).

Cy5+14 FRET acceptor in the second kinetic phase (>1 to 400 seconds) increases from 20% to 40% of the signal observed the absence of the drug in the presence of 71 μM LpM concentration (Figure 3B). At 150 μM LpM, Cy5+14 FRET acceptor signal increase by only 10% in the second kinetic phase. In the presence of 350 μM LpM, Cy5+14 FRET acceptor signal decreases by 10% from 1 to 10 seconds and then linearly increase by 30% after 10 till 400 seconds. At very high 1.7 mM LpM concentration, Cy5+14 FRET acceptor signal in decreases from 10% >1 second to -20% to 10 seconds and plateaus till 400 seconds (Figure 3B). Corresponding Cy3-100 FRET donor signal increases up to similar extent in the presence of 71 μM LpM as the signal observed in the absence of the drug (Figure S4B). In 150 uM LpM concentration, FRET donor -100 signal increases by 30% from 0.01 to 1 seconds and plateaus to that level till 400 seconds. At 350 μM LpM concentration, FRET donor signal increases by 50% from 0.01 to 1 second, then decreases by the same amount from 1 to 10 seconds followed by 30% increase and plateau from 10 to 400 seconds relative to 0 μM LpM condition (Figure S4B).

Reductions in +14 and -100 PIFE Kinetics when Pre-equilibrated *E.coli* RNAP-Lpm Solution is Mixed with Cy3-Cy5-labeled λP_R Promoter DNA

On mixing single Cy3 dye labeled λP_R promoter DNA fragment at either upstream -100 and downstream +14 position with pre-equilibrated mixture of E.coli RNAP and LpM, reduction in PIFE signal is observed (Figure 4).

Cy3+14 PIFE signal in the presence of increasing LpM concentration increases in two kinetic phases (Figure 4A). In the first kinetic phase from 0.01 to 1 second, the signal increases by 20% in 200 uM LpM to 40% in 0.5 μM LpM concentration. In the second kinetic phase after 1 second to 400 seconds the signal increases with a different slope till ~75% in 2.5 μM LpM and 40% in 200 μM LpM concentration relative to the signal observed in the absence of LpM (Figure 4A).

Cy3-100 PIFE signal in the presence of increasing LpM concentration also increase in two kinetic phases (Figure 4B). The signal increases with decreasing amplitude in the first kinetic phase from 0.01 to 10 seconds with increasing LpM concentration. At 15 μM LpM -100 PIFE signal increases by 75% in the first kinetic phase and by only 10% in the second kinetic phase from 10 to 400 seconds relative to 0 μM LpM condition (Figure 4B). In the presence of 45 μM and 75 μM LpM concentration, the signal increases by only ~50-52% from 0.01 to 10 seconds and up to 60% from 10 to 400 seconds with a different slope as compared to the signal observed in the absence of the drug. In 100 μM LpM concentration, -100 PIFE signal increases by 30% from 0.01 to 10 seconds and then up to 40% after 10 to 400 seconds (Figure 4B).

Very Different Reductions in -100 and +14 PIFE Kinetics When Lpm-DNA Solutions are Mixed with RNAP

When a mixture of single dye labeled λP_R promoter DNA fragment with Cy3 probe at either downstream +14 or upstream -100 position and increasing Lpm concentration was mixed with

E.coli RNAP different PIFE kinetics was observed from the previous order of Lpm-RNAP addition to DNA in Figure 4 (Figure 5).

In the presence of 750 μM Lpm, Cy3+14 PIFE signal increases with slower kinetics from 0.01 to ~30 seconds reaching to approximately to the same maximum value as observed in the absence of the drug and then decreases linearly by 25% after 30 seconds till 400 seconds (Figure 5A). At very high Lpm concentration, +14 PIFE signal increases by 65% in 1.1 mM and 75% in 1.4 mM Lpm from 0.01 to 30 seconds with slower kinetics and then decreases by 25% and 15% after 30 till 400 seconds respectively relative to 0 μM Lpm condition (Figure 5A).

Cy3-100 PIFE signal shows more drastic effects in the presence of increasing Lpm concentration as compared to Cy3+14 PIFE signal (Figure 5B and 5A). In 71 and 75 μM Lpm concentration, Cy3-100 PIFE signal increases by 40% from 0.01 to 1 seconds and plateaus to that level after 1 second till 400 seconds relative to 0 μM Lpm condition (Figure 5B). The signal increases from 0.01 to 1 seconds till 30% at 300 μM Lpm and 20% at 1.7 mM Lpm concentration.

+14 PIFE Kinetic Effects in Different Orders of Mixing of Lpm, λP_R Promoter DNA and RNAP

We performed 2 types of 3 syringe mixing experiments: first mixing λP_R promoter DNA with RNAP pre-incubating them for some time duration and then with Lpm (Figure 6). Second type involved mixing RNAP with Lpm first then with λP_R promoter DNA (Figure S5). λP_R promoter DNA was labeled with Cy3 probe at +14 downstream position.

In the first type of double mixing experiment, +14 PIFE signal is deviated more strongly from -Lpm when λP_R promoter DNA is pre-incubated with *E.coli* RNAP for 0.3 or more seconds before mixing with Lpm (Figure 6). I_{1M} intermediate is the most dominant closed complex intermediate formed at 0.3 seconds when λP_R promoter DNA interacts with *E.coli* RNAP to form OC based on

previous Chapter 3 mechanism. This suggests that Lpm acts on I1M closed complex intermediate formed having open clamps of RNAP.

In the second double mixing experiment, +14 PIFE signal is deviated from –Lpm condition when E.coli RNAP is pre-incubated with Lpm for 3 seconds or more (Figure S5). This confirms the slower kinetics of Lpm binding to RNAP than RNAP binding to DNA.

Discussion

Lpm Binds Only the Open-Clamp Conformations of RNAP (Free or as I_{1M} Closed-Promoter Intermediate) to Inhibit OC Formation

RNAP-Lipiarmycin complex inhibits heparin resistant OC formation for lacUV5, λP_R and galP2 promoter as detected by permanganate footprinting and electrophoretic mobility shift assays performed at 37^OC in 100 μM Lpm concentration [6]. *E.coli* σ^{70} RNAP holoenzyme *in-vitro* abortive transcription activity is abolished with inhibition concentrations IC_{50} of 53 μM for rrnAP3 and IC_{50} of 9-13 μM for lacUV5 promoter at 37^0C [5 6]. In the assay of mixing double dye Cy3-Cy5 labeled λP_R promoter DNA fragment with *E.coli* RNAP pre-incubated with increasing Lpm concentration 50% reduction in Cy5-100 and Cy5+14 FRET acceptor is seen at around 7 μM and 15 μM Lpm concentration (Figure 2).

In these experiments, pre-equilibrated mixture of 50 nM RNAP and RNAP bound with Lpm (RNAP-Lpm) fractions in higher μM Lpm concentration ranges were mixed with 50 nM fluorescently labeled DNA fragment, as shown in the schematics in Figure 2. Fractions of RNAP unbound with Lpm that have closed and partly closed clamps binds to DNA and later form OC. The other fractions of Lpm-RNAP with locked open clamps binds with DNA to form ternary closed promoter open clamp intermediate (Lpm-RNAP-DNA) complex to exhibit PIFE (Figure S1). However, this ternary Lpm-RNAP-DNA complex doesn't form a stable OC to show reduction in FRET signal in the first kinetic phase (Figure 2). Lpm-RNAP do form closed complex with lacUV5,

λP_R, gal P2 promoter DNA as shown previously by exonuclease footprinting, EMSA and permanganate footprinting assays performed at $37^{O}C$ in the presence of 100 uM Lpm concentration [6]. Cy3+14 and -100 PIFE signal is also not completely lost when RNAP-Lpm mixture is added to λP_R promoter DNA fragment in increasing Lpm concentration (Figure 4) suggesting that Lpm-RNAP complex do bind with DNA to form a ternary complex.

In Figure 1, If Lpm weren't able to affect the intermediate steps of DNA binding to the unbound fraction of RNAP to form OC, and there was always a constant population of Lpm-RNAP open clamp complex formed in the pre-equilibrated mixture of Lpm and RNAP. Then, one should have got a constant reduction in the FRET acceptor signal on mixing pre-equilibrated mixture of RNAP and Lpm with fluorescently labeled λP_R promoter DNA. However, this wasn't the case observed in Figure 1, the FRET acceptor signal proportionately reduced with increasing Lpm concentration.

In order to understand the effects of Lpm in the closed complex intermediates formed in the mechanism of OC formation we changed the order of Lpm addition to DNA and RNAP as shown in Figure 3. In this type of mixing all of the RNAP fractions binds with DNA initially with faster kinetics to form closed complex intermediates in the first kinetic phase from 0.01 to 1 seconds. The signal begins to deviate from 0 μM Lpm at around 1 seconds in the presence of increasing Lpm concentration indicating that Lpm is acting on one of the RNAP-closed promoter complex intermediate formed around this time. Based on the mechanism proposed in Chapter 3, I_{1M} with the open clamps seems to be the most likely intermediate to be targeted by Lpm. 3 syringe mixing experiment of pre-incubating DNA with RNAP first then with Lpm also suggested that Lpm binds to I_{1M} intermediate to form I_{1M}-LpM (Figure 6). The other type of 3 syringe mixing experiment validates the slower kinetics of Lpm binding to RNAP than DNA with RNAP

Reduced PIFE and FRET Signals of I_{1M}-Lpm complex Relative to I_{1M} Intermediate

LpM doesn't displace DNA from OC or act on it as seen in Figure S3. However, it does act on I_{1M} intermediate and show drastic effects in the upstream PIFE and FRET signal of promoter DNA than the downstream promoter region as seen in Figure 3, 5 and S4 in I_{1M}-Lpm complex formation. PIFE and FRET donor effects happen earlier in the upstream (~1 second) than the downstream region of promoter DNA (~30 seconds) (Figure 5 and S4) in I_{1M}-Lpm formation. This suggests that the upstream region of promoter DNA is affected first and destabilized more strongly followed by displacement of the downstream region from the active site cleft on I_{1M}-Lpm formation.

Exonuclease chemical footprint protection of the upstream promoter region is unaffected in the presence/absence of 100 μM Lpm [6]. The reduction in PIFE -100 signal observed could be due to the weakening of the contacts of the upstream promoter region with RNAP when Lpm acts on I_{1M} to form I_{1M}-LpM and its dissociation in higher Lpm concentration.

As the upstream promoter region is affected more strongly and first than the downstream region in I_{1M}-LpM complex, I_{1M}-LpM complex resembles an intermediate that can be formed earlier in between RP_C and I_{1E} in the kinetic scheme of mechanism of OC formation with bent but unwrapped upstream DNA and unbent downstream DNA (Chapter 3). The FRET acceptor signal in higher Lpm concentration (0.75 - 1.7 mM) in forming I_{1M}-LpM decreases to 20% of the initial signal observed at 0.01 seconds time-point (Figure 3). Previously, in Chapter 3 FRET acceptor signal begin to increase from 0.01 seconds only after I_{1E} formation. On getting a signal lesser than 0.01 second time-point suggests that I_{1M}-LpM have FRET characteristics of an intermediate that can be formed earlier than I_{1E} but after RP_C formation to exhibit PIFE in higher Lpm concentration.

It is highly plausible that Lpm-RNAP complex on binding with promoter DNA can form the earlier intermediate same as I_{1M}-LpM (Figure 7).

Formation of Off-Pathway Lpm-I_{1M} Complex Reduces Population Fraction of Subsequent I_{1L} and I_2 intermediates, thereby Inhibiting OC Formation

The isomerization rate constant of the DNA opening step (I_{1L} to I_2 transition) in the OC formation mechanism is the product of the population fraction of I_{1L} in the equilibrated $\{CC\}_{ensemble}$ and the intrinsic DNA opening rate k_{open} (Chapter 3). LpM drug inhibitor regulates the OC formation mechanism by reducing the population fraction of I_{1L} by acting on the I_{1M} intermediate formed at a step preceding I_{1L} intermediate formation. LpM on binding to I_{1M} intermediate locks the RNAP enzyme in an open clamp and forms an off-pathway trapped I_{1M}-LpM complex intermediate (Figure 7). The trapped I_{1M}-LpM complex intermediate having FRET and PIFE characteristics of an early intermediate may also form when RNAP-Lpm binds with DNA. On forming the off-pathway intermediate population fraction of subsequent intermediate formation is reduced thereby OC formation is inhibited.

Conclusion

Kinetic–Mechanistic studies of mixing fluorescently labeled DNA fragment with RNAP and increasing Lpm concentration in different orders aided in understanding the drug's mode of action in the mechanism of OC formation. It acts on the rapidly equilibrated early steps of closed complex intermediates formed preceding the rate limiting catalytic step of DNA opening. It, initially binds to a fraction of RNAP having open clamps to form Lpm-RNAP complex. The other Lpm unbound RNAP fractions interact with DNA to form an ensemble of closed complex intermediates in which one of the intermediate I_{1M} with open clamps becomes target for the drug again. I_{1M} on binding with Lpm forms an intermediate I_{1M}-Lpm, reducing the population fraction of subsequent closed complex intermediates I_{1L} thereby inhibiting OC formation. In low drug concentrations due to faster kinetics of DNA interacting with RNAP than the drug binding to RNAP enzyme, the fractions that

are able to advance from I_{1M} and are rescued from the drug's action irreversibly form stable OC with the closed clamps inert to its effects.

Figures

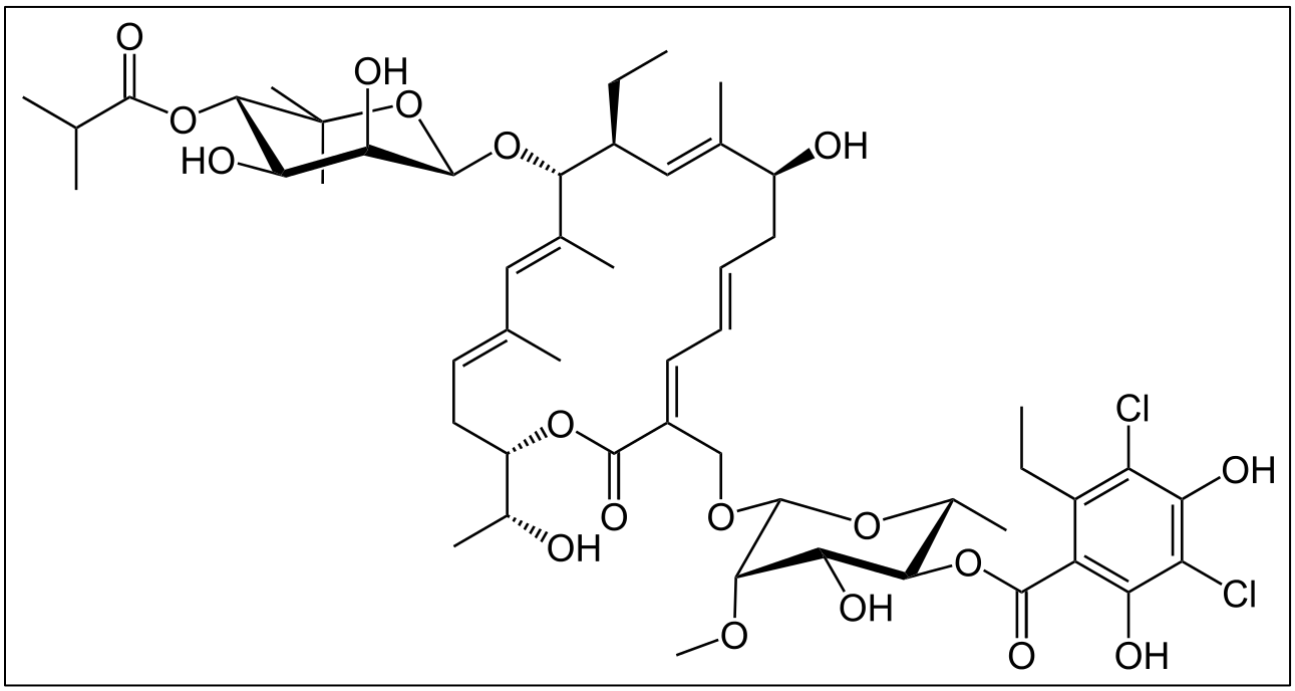


Figure 1 Chemical Structure of Lipiramycin (Lpm)

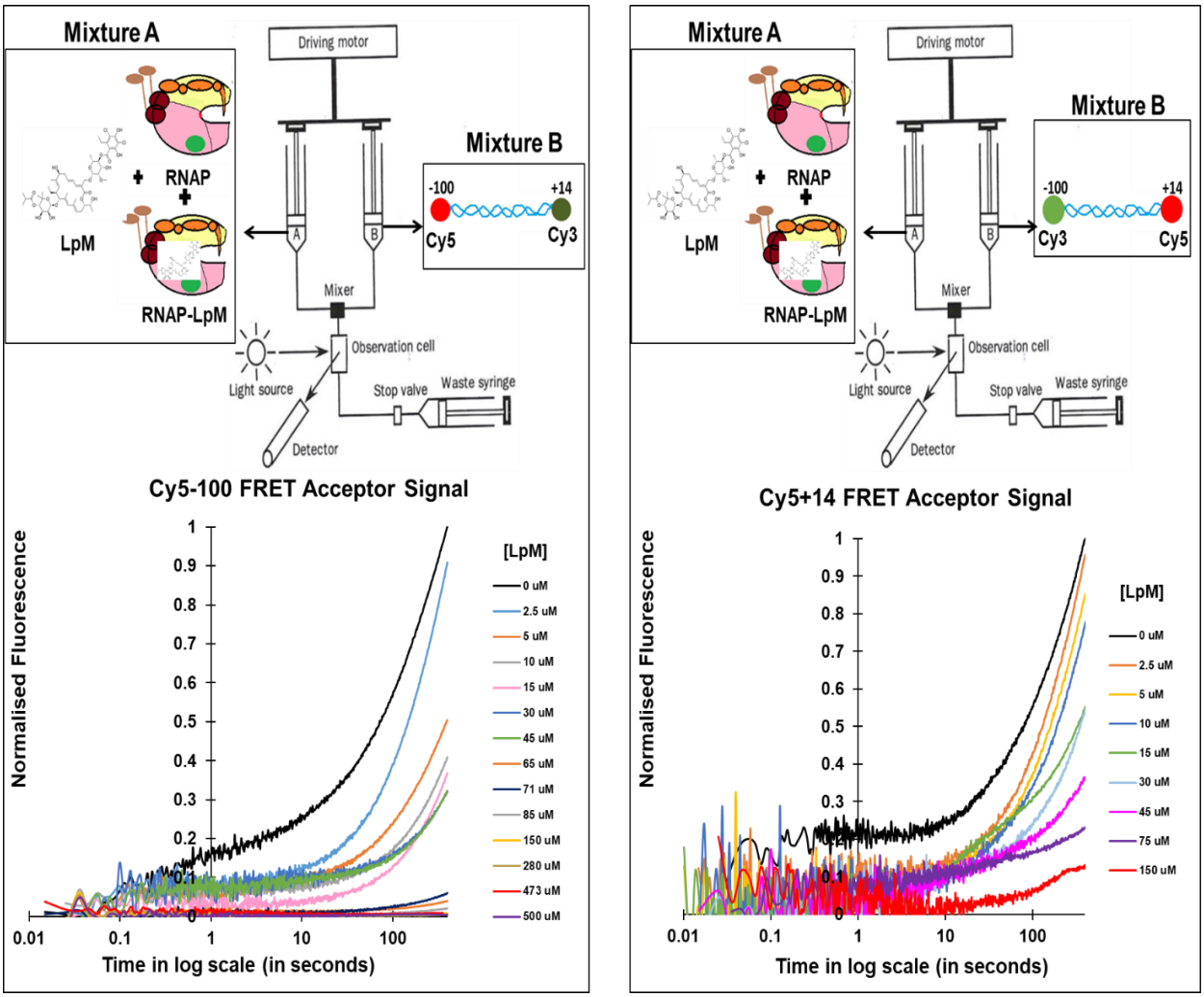


Figure 2. Time Courses of Normalized Cy5 Acceptor FRET after Mixing Cy3 Cy5-labeled λP_R promoter DNA with RNAP- Lipiarmycin (LpM) Mixture at Different LpM Concentrations at 19^{0}C. RNAP pre-equilibrated (37^{0}C, t $\geq$ 10 min) with different concentration of lipiarmycin (Lpm; Mixture A in the schematics above) is rapidly mixed with λP_R promoter DNA (Mixture B in the schematics above). The promoter DNA is labeled with the Cy3-Cy5 FRET pair in the orientations Cy5-100 Cy3+14 (left panel) and Cy5+14 Cy3-100 (right panel). Final concentrations of RNAP and promoter DNA are 50 nM; final Lpm concentrations are indicated on the figures. Change in Cy5 FRET acceptor signal upon excitation of Cy3 are measured on a logarithmic time scale from 0.01 to 400 seconds. Each trace is the average of 5-6 individually-normalized shots collected sequentially at one Lpm concentration. End points (400 s) of +Lpm time courses are scaled using

–Lpm control experiments (black trace; see Methods). Corresponding FRET donor signal is shown in Figure S1.

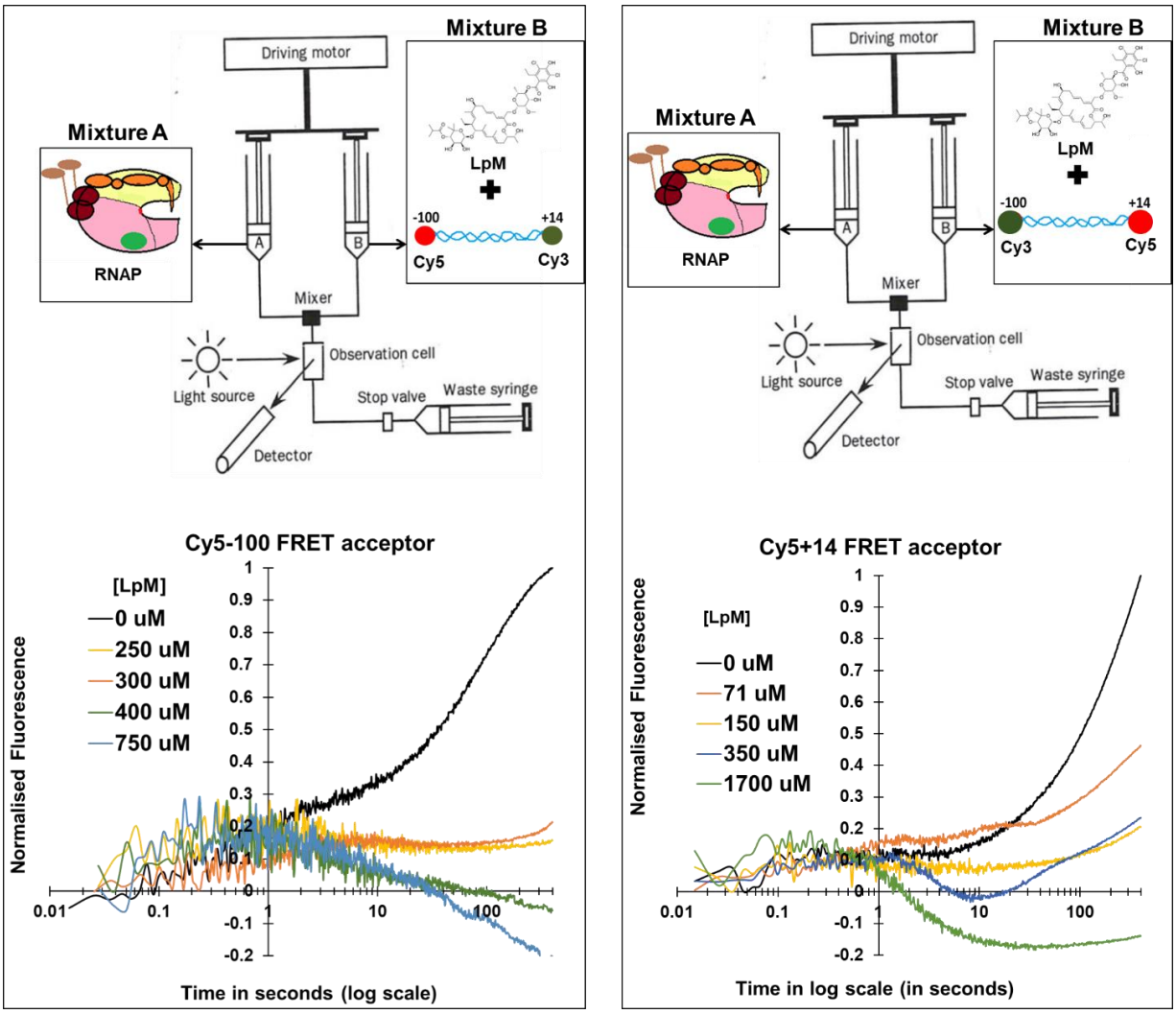


Figure 3 Time Courses of Normalized Cy5 Acceptor FRET after Mixing RNAP with Cy3Cy5-labeled λP_R promoter DNA – Lipiarmycin (LpM) Mixture at Different LpM Concentrations at 19^{0}C. RNAP (Mixture A in the schematics above) is rapidly mixed with λP_R promoter DNA containing various Lpm concentrations (Mixture B in the schematics above). The promoter DNA is labeled with the Cy3-Cy5 FRET pair in the orientations Cy5-100 Cy3+14 (panel A, left) and Cy5+14 Cy3-100 (panel B, right). Final concentrations of RNAP and promoter DNA are 50 nM; final Lpm concentrations are indicated on the figures. Changes in Cy5 FRET acceptor signal upon excitation of Cy3 are measured on a logarithmic time scale from 0.01 to 400 seconds. Each trace is the average of 5-6 individually-normalized shots collected sequentially at one Lpm

concentration. End points (400 s) of +Lpm time courses are scaled using –Lpm control experiments (black trace; see Methods). Corresponding FRET donor signal is shown in Figure S3.

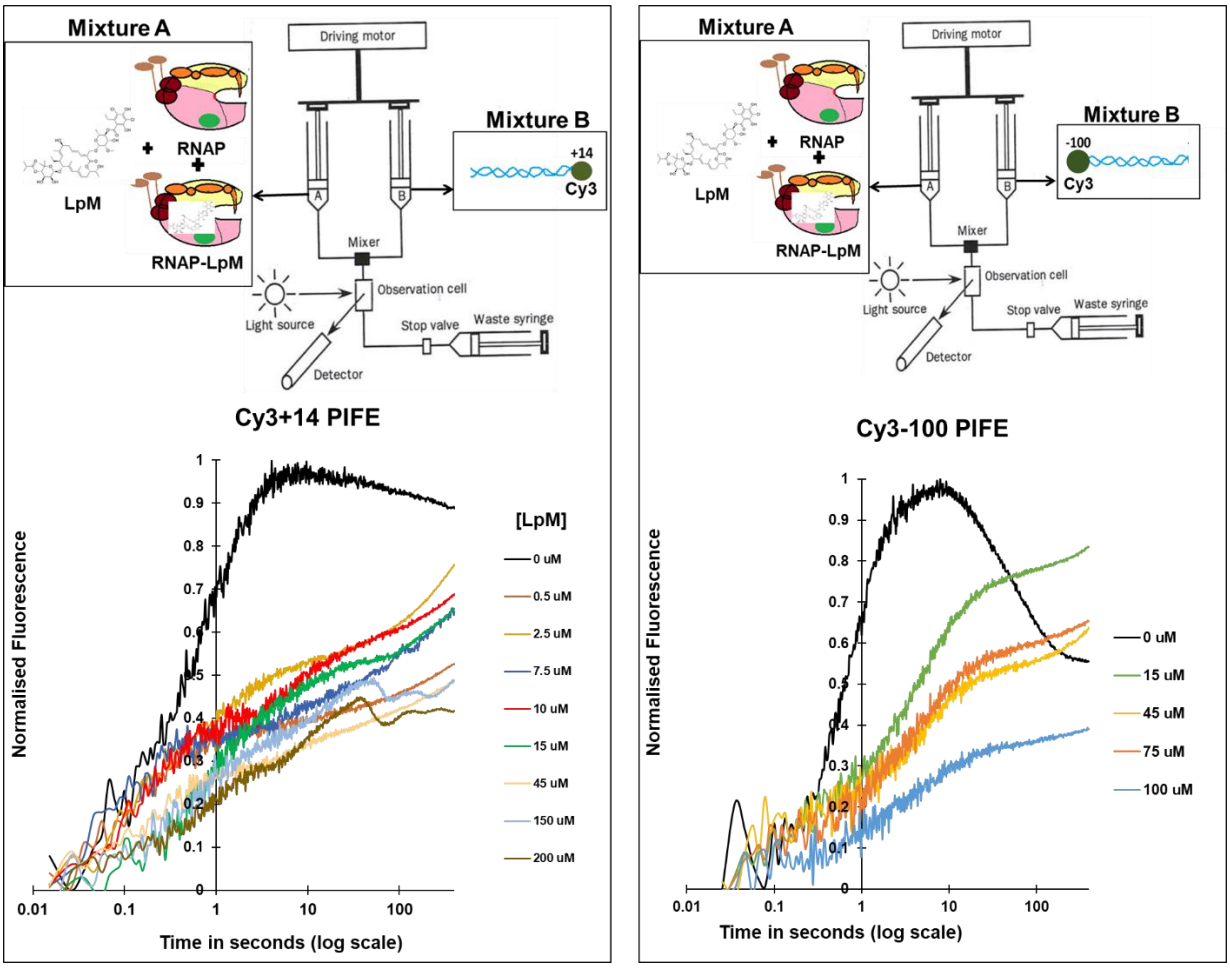


Figure 4. Time Courses of Normalized PIFE fluorescence after Mixing Cy3-labeled λP_R promoter DNA with RNAP- Lipiarmycin (LpM) Mixture at Different LpM Concentrations at 19^0C. RNAP pre-equilibrated (37^0C, t $\geq$ 10 min) with different concentration of lipiarmycin (Lpm; Mixture A in the schematics above) is rapidly mixed with λP_R promoter DNA (Mixture B in the schematics above). The promoter DNA is labeled with the Cy3 at +14 (left panel) and -100 (right panel). Final concentrations of RNAP and promoter DNA are 50 nM; final Lpm concentrations are indicated on the figures. Change in Cy3 PIFE are measured on a logarithmic time scale from 0.01 to 400 seconds. Each trace is the average of 5-6 individually-normalized shots collected sequentially at one Lpm concentration. End points (400 s) of +Lpm time courses are scaled using –Lpm control experiments (black trace; see Methods).

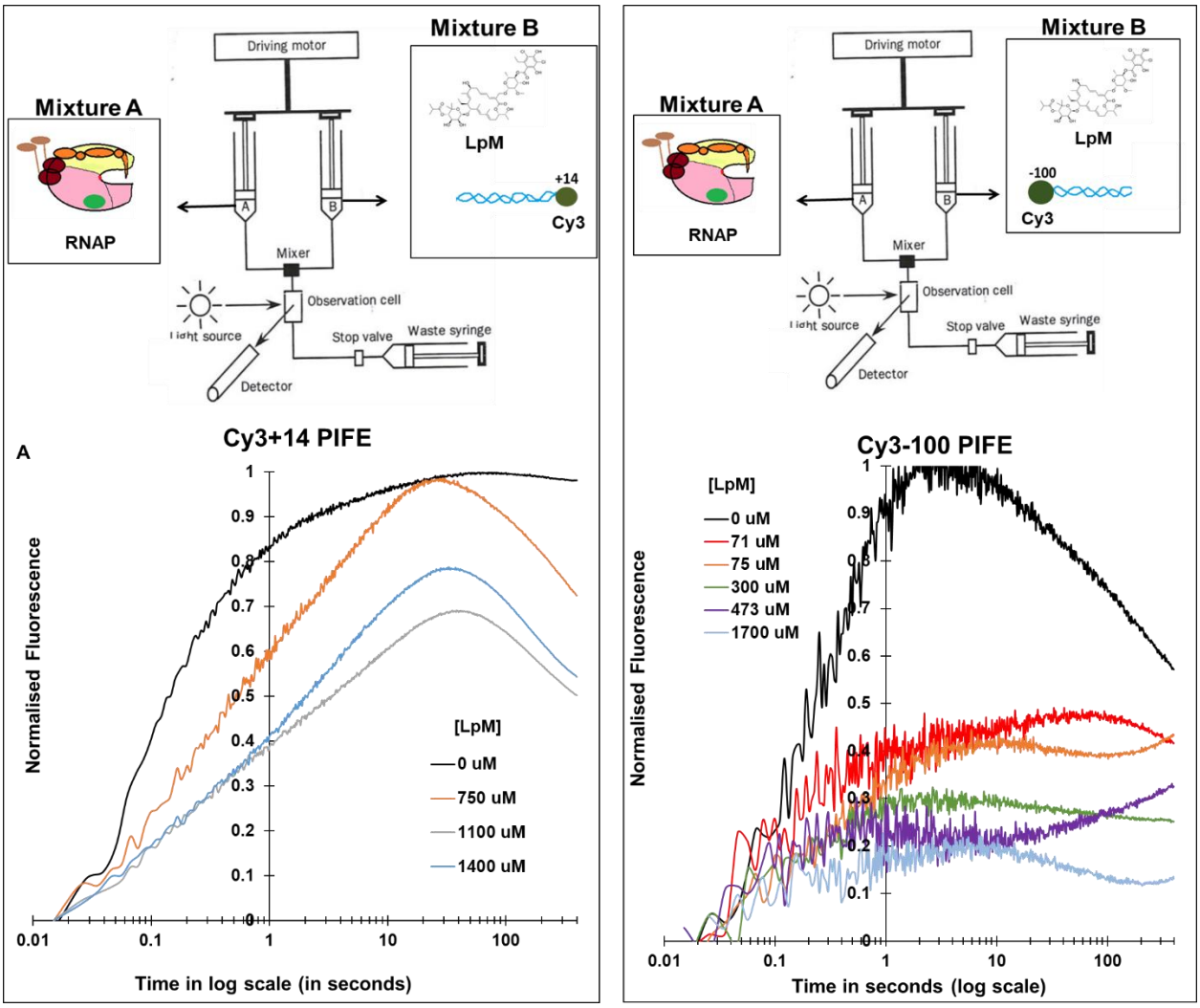


Figure 5 Time Courses of Normalized PIFE Fluorescence after Mixing RNAP with Cy3-labeled λP_R promoter DNA with RNAP- Lipiarmycin (LpM) Mixture at Different LpM Concentrations at 19^{0}C. RNAP (Mixture A in the schematics above) is rapidly mixed with λP_R promoter DNA containing various Lpm concentrations (Mixture B in the schematics above). The promoter DNA is labeled with the Cy3 at +14 (left panel) and -100 (right panel). Final concentrations of RNAP and promoter DNA are 50 nM; final Lpm concentrations are indicated on the figures. Changes in Cy3 PIFE are measured on a logarithmic time scale from 0.01 to 400 seconds. Each trace is the average of 5-6 individually-normalized shots collected sequentially at

one Lpm concentration. End points (400 s) of +Lpm time courses are scaled using –Lpm control experiments (black trace; see Methods).

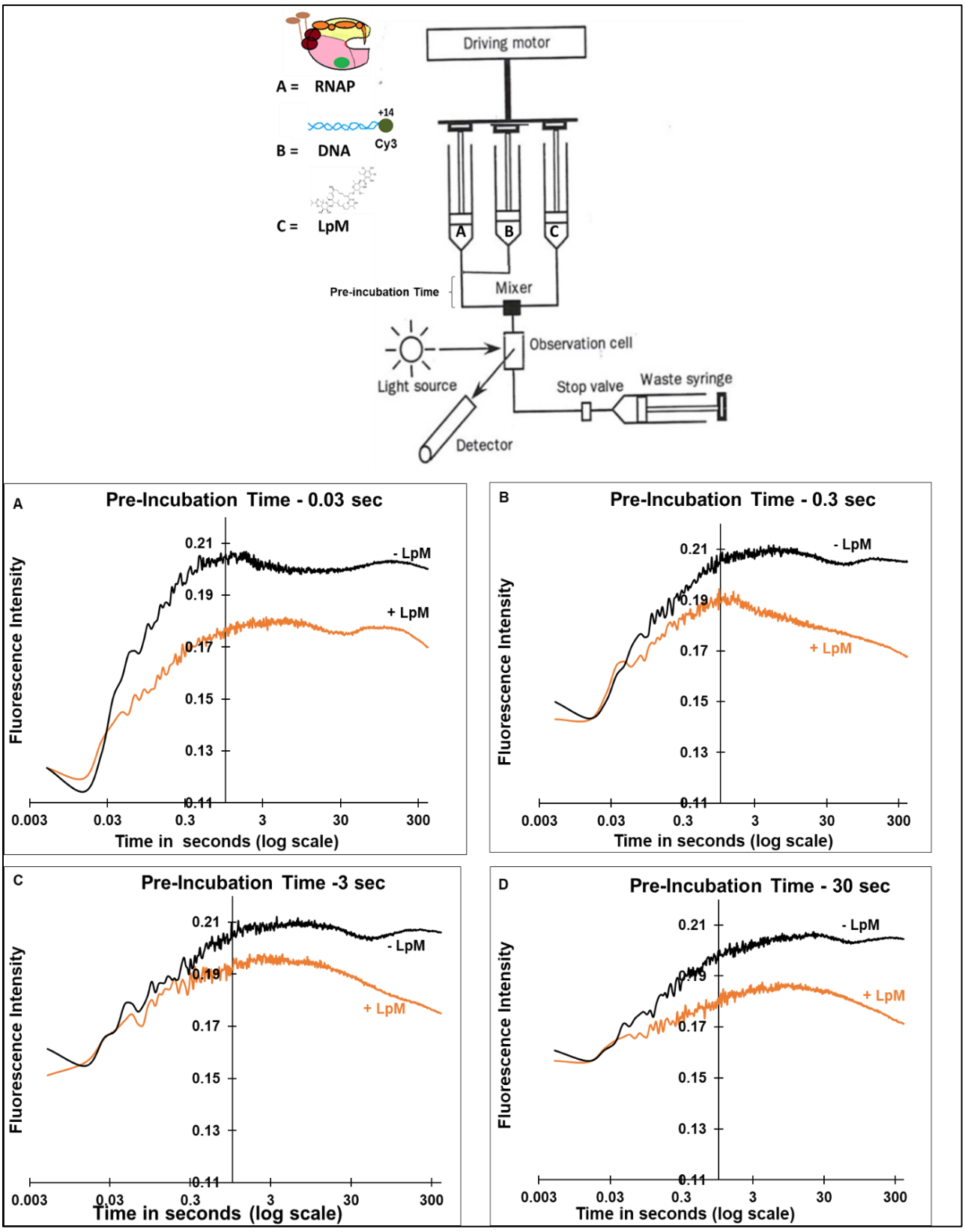

A = RNAP
B = DNA
+14
Cy3
C = LpM
Driving motor
A
B
C
Pre-incubation Time
Mixer
Observation cell
Light source
Stop valve
Waste syringe
Detector
A
Pre-Incubation Time - 0.03 sec
Fluorescence Intensity
0.21
0.19
0.17
0.15
0.13
0.11
- LpM
+ LpM
0.003
0.03
0.3
3
30
300
Time in seconds (log scale)
B
Pre-Incubation Time - 0.3 sec
Fluorescence Intensity
- LpM
+ LpM
Time in seconds (log scale)
C
Pre-Incubation Time -3 sec
Fluorescence Intensity
- LpM
+ LpM
Time in seconds (log scale)
D
Pre-Incubation Time - 30 sec
Fluorescence Intensity
- LpM
+ LpM
Time in seconds (log scale)

Figure 6 Three-Syringe Mixing of λP_R Promoter with RNAP and Subsequently with LpM, detected by Cy3 PIFE Fluorescence. RNAP (syringe A) is first mixed with Cy3+14 –labeled λP_R promoter DNA (syringe B) and incubated for 0.03s (panel A), 0.3 s (panel B), 3 s (panel C) or 30 s (panel D). After this time interval, the DNA-RNAP solution is first mixed with Lpm (final concentrations of 50 nM RNAP and DNA, 71 μM Lpm) and the Cy3+14 PIFE fluorescence kinetics signal is measured on a logarithmic time scale from 0.01 to 400 s.

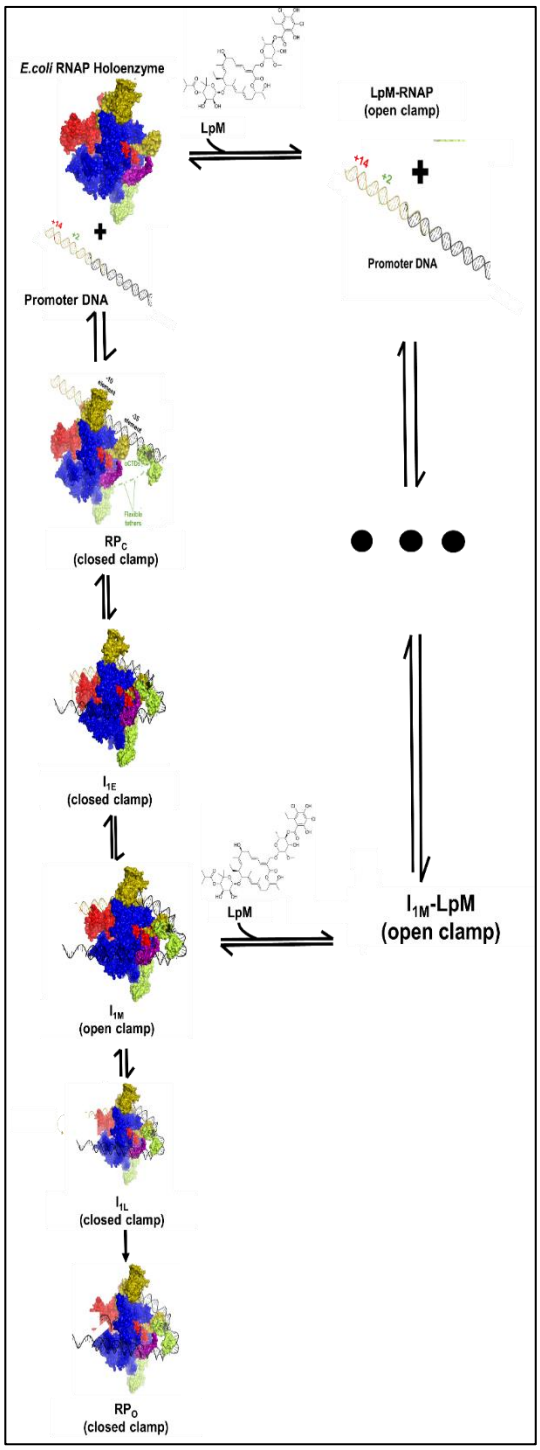
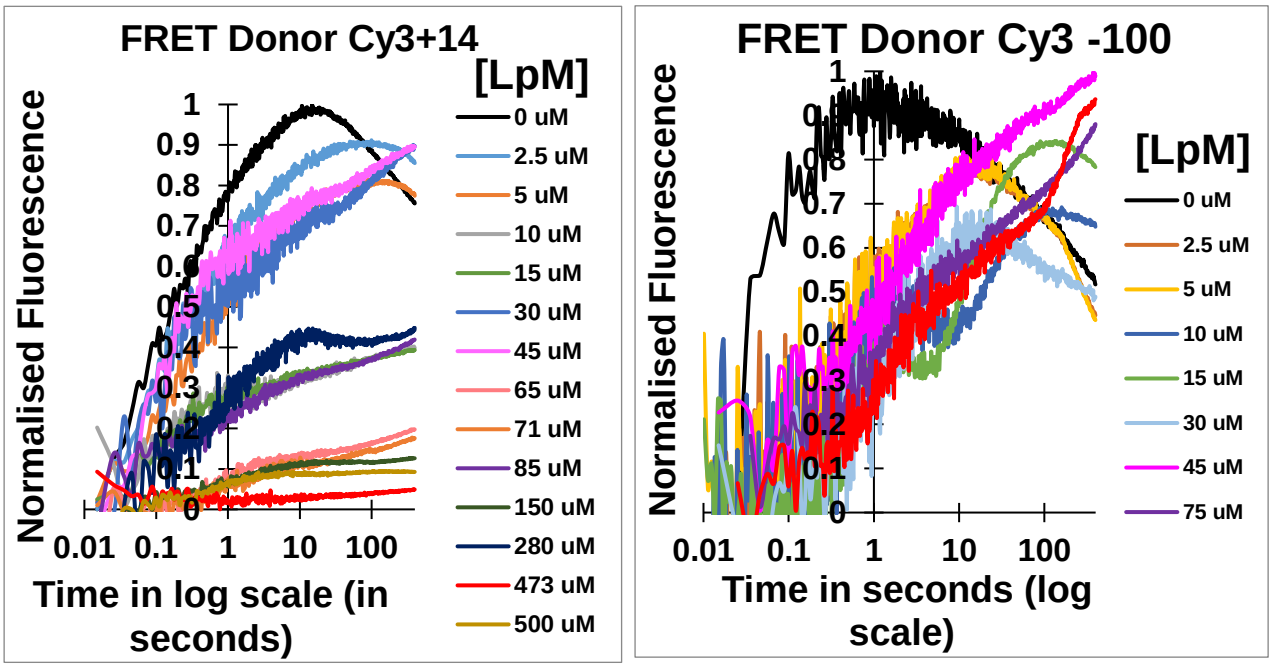
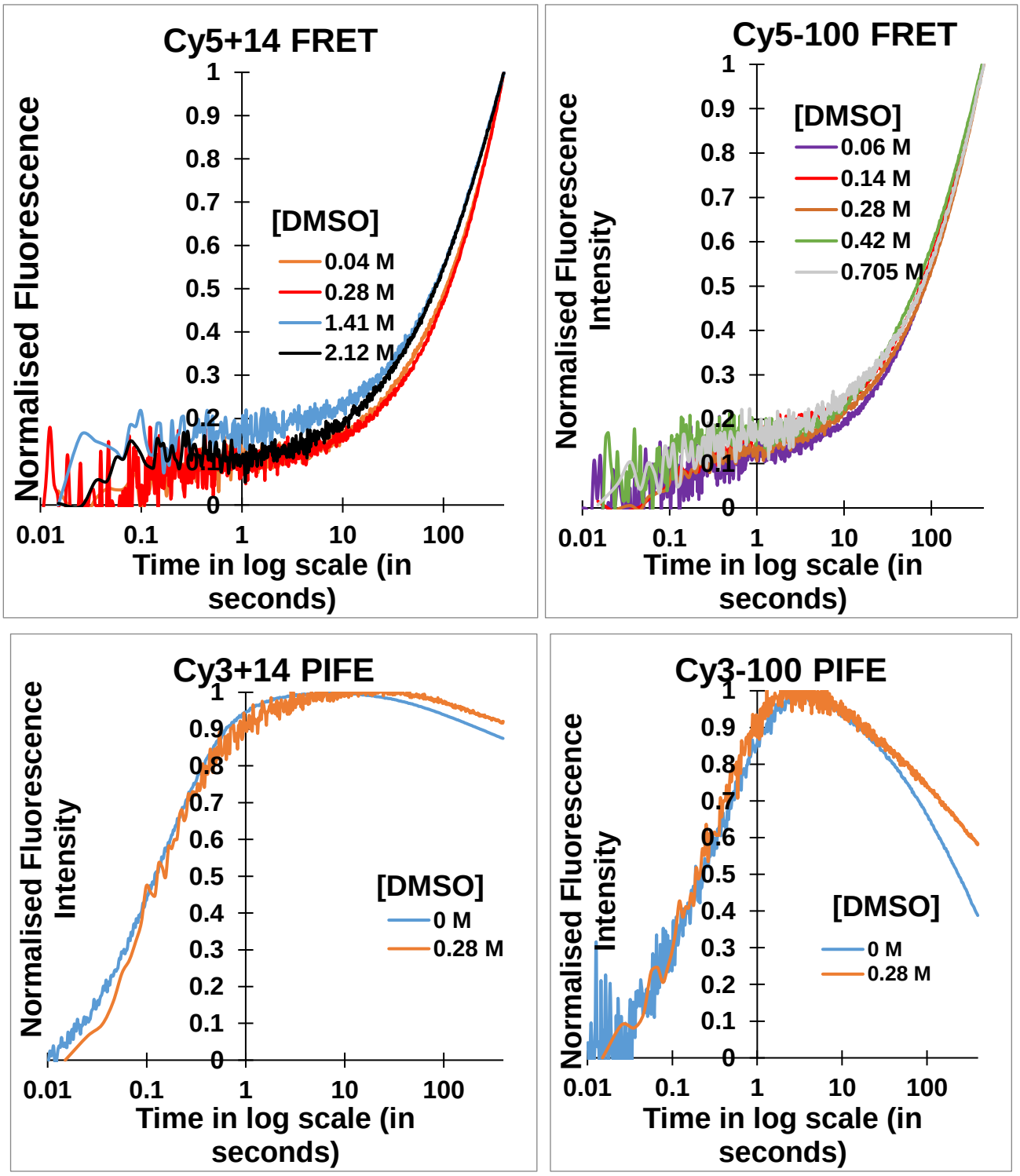


Figure 7 Schematic representation of LpM mode of action in transcription initiation mechanism. Pre-incubation of Lpm with RNAP (top line) converts a fraction of RNAP to a Lpm-RNAP complex in which the clamp is open. When Lpm and promoter DNA are added together to RNAP, closed promoter complexes form first, before Lpm binding, and Lpm appears to bind primarily to the open-clamp, closed-promoter intermediate I_{1M}, forming LpM-I_{1M}. This same Lpm-I_{1M} having FRET and PIFE characteristics of an early closed-promoter intermediate, may form directly by binding of promoter DNA to the open-clamp RNAP-Lpm complex. Formation of an off-pathway I_{1M}-Lpm reduces the population fraction of the subsequent I_{1L} intermediate in the CC ensemble and thereby inhibits OC formation.

Supplemental Figures

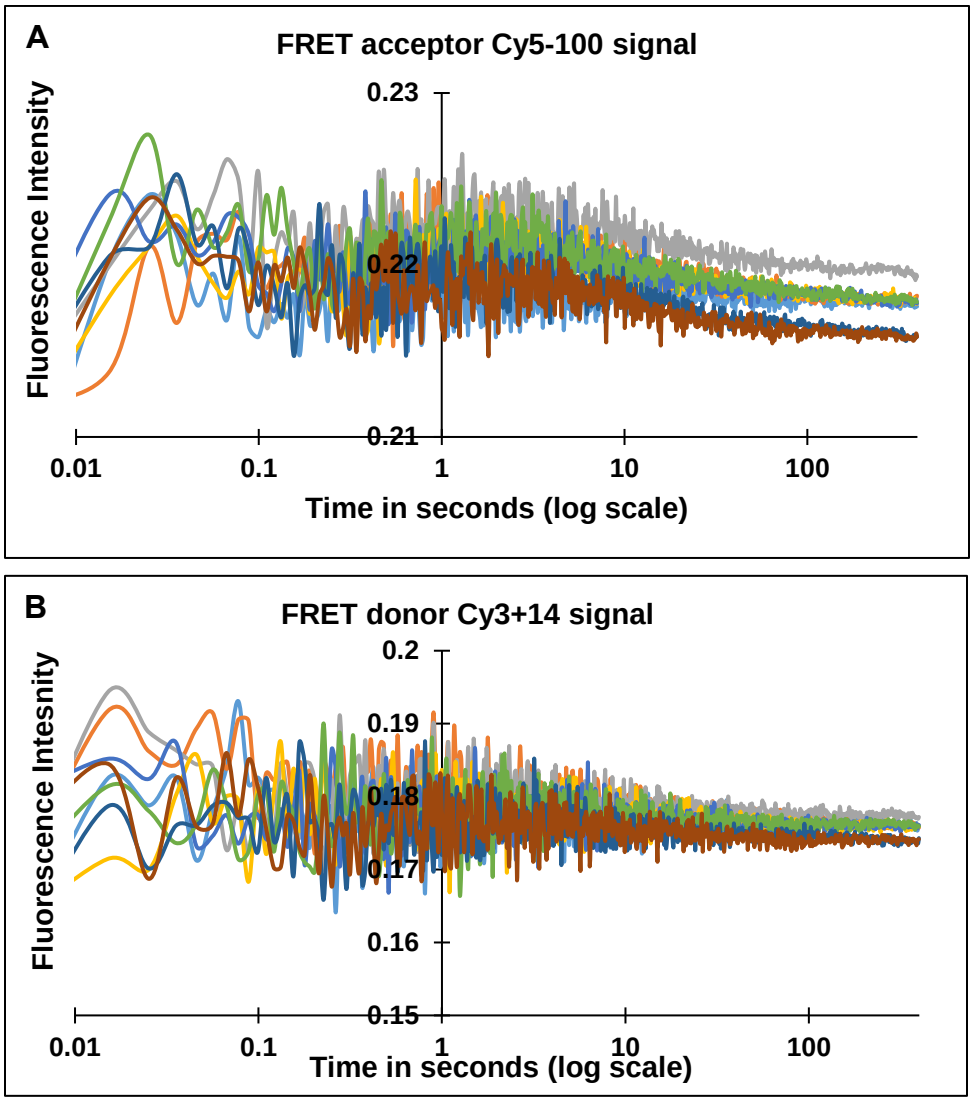


Supplemental Figure 1 Change in Cy3 FRET Donor Fluorescence after Mixing Cy3Cy5-labeled λP_R Promoter DNA with RNAP - Lipiarmycin (LpM) Mixture at Different LpM Concentrations at 19 ºC. RNAP pre-equilibrated (37 ºC, t ≥10 min) with different concentrations of lipiarmycin (Lpm) is rapidly mixed with λP_R promoter DNA. The promoter DNA is labeled with the Cy3-Cy5 FRET pair in the orientations Cy5-100 Cy3+14 (left panel) and Cy5+14 Cy3-100 (right panel) Final concentrations of RNAP and promoter DNA are 50 nM; final Lpm concentrations are indicated on the figures. Changes in Cy3 FRET donor signal upon excitation are measured on a logarithmic time scale from 0.01 to 400 seconds. Each trace is the average of 5-6 individually-normalized shots collected sequentially at one Lpm concentration. End points (400 s) of +Lpm time courses are scaled using - Lpm control experiments (black trace; see Methods). Corresponding FRET acceptor Cy5 signal is shown in Figure 1.

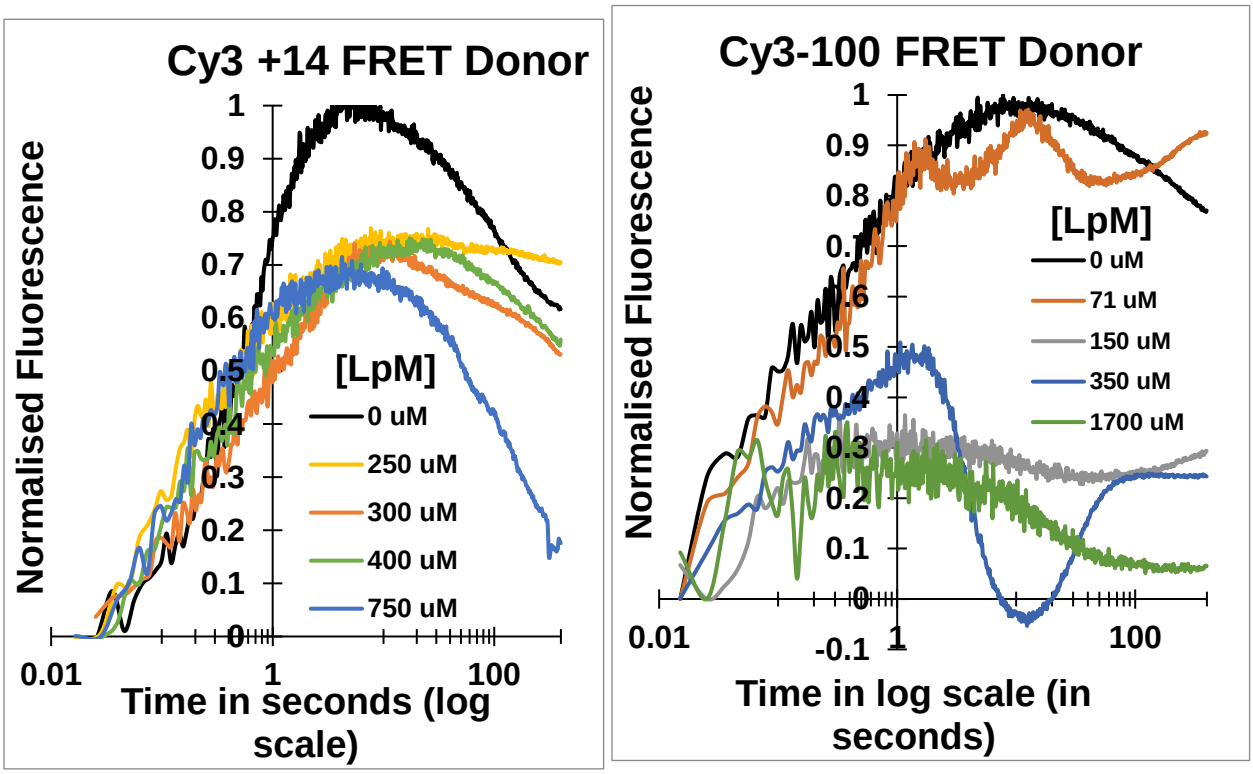


Supplemental Figure 2 Change in Cy5 FRET acceptor and Cy3 PIFE on mixing Cy3 and/or Cy5 labeled DNA fragment with RNAP in FB with different DMSO concentrations. λP_R promoter DNA fragment labeled with Cy3-Cy5 dye, Cy5+14 Cy3-100 (Panel A), Cy3+14 Cy5-100

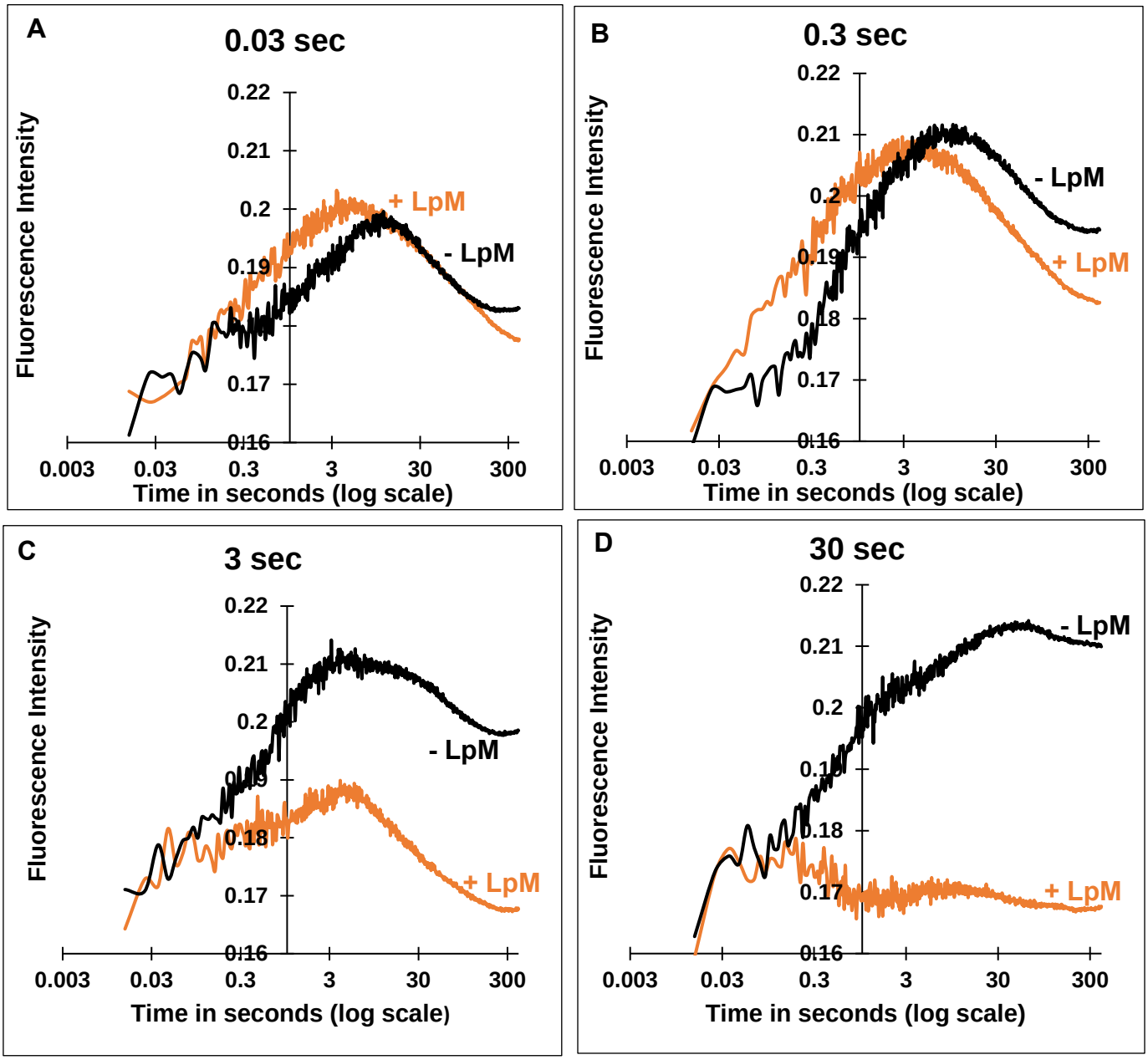
(Panel B) or single dye Cy3 labeled Cy3+14 (Panel C) and Cy3-100 (Panel D) were mixed with RNAP in buffer solution containing different DMSO concentrations (0 – 2 M). Changes in FRET acceptor Cy5 +14 in Panel A, Cy5-100 in Panel B Cy3 +14 PIFE in Panel C and Cy3-100 PIFE in Panel D signal were monitored from 0.01 s to 400 s. Each trace is a normalized average of 5-6 individual shots collected at an individual DMSO concentration.

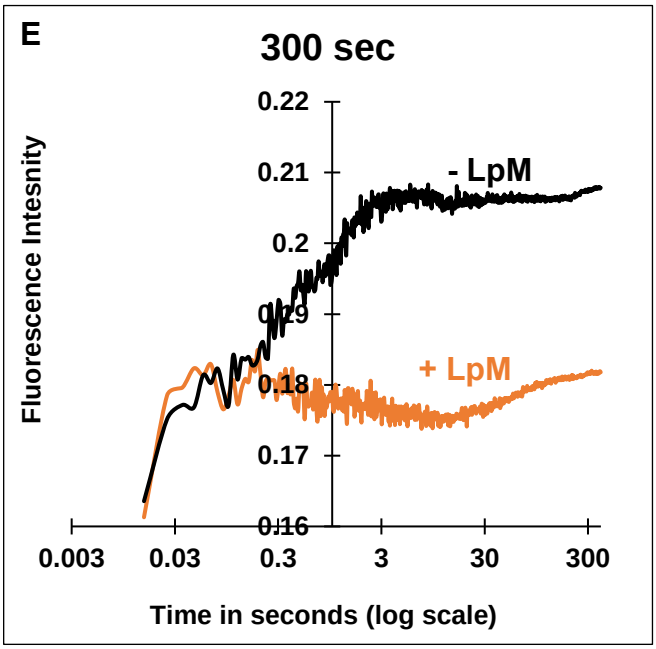


Supplemental Figure 3 Lack of Effect of Lpm on a Preformed DNA-RNAP OC. Representative time course (log scale) of FRET acceptor (Cy5-100) (panel A, top) and FRET donor (Cy3+14) (panel B, bottom) emission intensity traces after mixing Cy3+14-Cy5-100 labeled λP_R DNA-RNAP OC (50 nM final) with 15 uM LpM exciting FRET donor (Cy3) at 515nm.



Supplemental Figure 4. Time Courses of Normalized Cy5 Acceptor FRET after Mixing RNAP with Cy3 Cy5-labeled λP_R Promoter DNA - Lipiarmycin (LpM) Mixture at Different LpM Concentrations at 19 ºC. RNAP is rapidly mixed with λP_R promoter DNA containing various Lpm concentrations. The promoter DNA is labeled with the Cy3-Cy5 FRET pair in the orientations Cy5-100 Cy3+14 (left panel) and Cy5+14 Cy3-100 (right panel) Final concentrations of RNAP and promoter DNA are 50 nM; final Lpm concentrations are indicated on the figures. Changes in Cy5 FRET acceptor signal upon excitation of Cy 3 are measured on a logarithmic time scale from 0.01 to 400 seconds. Each trace is the average of 5-6 individually-normalized shots collected sequentially at one Lpm concentration. End points (400 s) of +Lpm time courses are scaled using - Lpm control experiments (black trace; see Methods). Corresponding FRET acceptor signal is shown in Figure 3.


A
0.03 sec
Fluorescence Intensity
0.22
0.21
0.2
0.19
0.17
0.16
+ LpM
- LpM
0.003
0.03
0.3
3
30
300
Time in seconds (log scale)
B
0.3 sec
Fluorescence Intensity
0.22
0.21
0.2
0.19
0.18
0.17
0.16
- LpM
+ LpM
0.003
0.03
0.3
3
30
300
Time in seconds (log scale)
C
3 sec
Fluorescence Intensity
0.22
0.21
0.2
0.19
0.18
0.17
0.16
- LpM
+ LpM
0.003
0.03
0.3
3
30
300
Time in seconds (log scale)
D
30 sec
Fluorescence Intensity
0.22
0.21
0.2
0.19
0.18
0.17
0.16
- LpM
+ LpM
0.003
0.03
0.3
3
30
300
Time in seconds (log scale)



Supplemental Figure 5 Double Mixing PIFE Fluorescence Kinetics Experiment of mixing first RNAP with/without 71 uM LpM then with fluorescently labeled DNA Cy3+14 fragment at different time-points. 50 nM RNAP is first mixed with 15 uM Lpm and incubated for different time points: 0.03 **(Supplemental Figure 5 panel A, top left)**, 0.3 **(Supplemental Figure 5 panel B, top right)**, 3 **(Supplemental Figure 5 panel C, mid left)**, 30 seconds **(Supplemental Figure 5 panel D, mid right)** and 300 seconds **(Supplemental Figure 5 panel E, bottom left).** RNAP-Lpm complex formed at the different time-points are then mixed with 50 nM Cy3+14 fluorescently labeled DNA fragment and PIFE fluorescence kinetics signal is observed from 0.01 to 400 seconds in 600 data points. The drift observed at later time-points +/- Lpm is due to the mixing effects.

References:

[1] Brodolin, K. (2011) Antibiotics trapping transcription initiation intermediates: To melt or to bend, what's first?, *Transcription 2*, 60-65.

[2] Srivastava, A., Talaue, M., Liu, S., Degen, D., Ebright, R. Y., Sineva, E., Chakraborty, A., Druzhinin, S. Y., Chatterjee, S., Mukhopadhyay, J., Ebright, Y. W., Zozula, A., Shen, J., Sengupta, S., Niedfeldt, R. R., Xin, C., Kaneko, T., Irschik, H., Jansen, R., Donadio, S., Connell, N., and Ebright, R. H. (2011) New target for inhibition of bacterial RNA polymerase: 'switch region', *Curr Opin Microbiol 14*, 532-543.

[3] Venugopal, A. A., and Johnson, S. (2012) Fidaxomicin: a novel macrocyclic antibiotic approved for treatment of Clostridium difficile infection, *Clin Infect Dis 54*, 568-574.

[4] Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y. W., Ebright, R. Y., Sineva, E., Gigliotti, M., Srivastava, A., Mandal, S., Jiang, Y., Liu, Y., Yin, R., Zhang, Z., Eng, E. T., Thomas, D., Donadio, S., Zhang, H., Zhang, C., Kapanidis, A. N., and Ebright, R. H. (2018) Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3), *Mol Cell 70*, 60-71 e15.

[5] Boyaci, H., Chen, J., Lilic, M., Palka, M., Mooney, R. A., Landick, R., Darst, S. A., and Campbell, E. A. (2018) Fidaxomicin jams Mycobacterium tuberculosis RNA polymerase motions needed for initiation via RbpA contacts, *Elife 7*.

[6] Tupin, A., Gualtieri, M., Leonetti, J. P., and Brodolin, K. (2010) The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site, *EMBO J 29*, 2527-2537.

[7] Andrey Feklistov, B. B., Jesse Hauver, Agnieszka Lass-Napiorkowska, Markus Kalesse, Florian Glaus, Karl-Heinz Altmann, Tomasz Heyduk, Robert Landick, and Seth A. Darst. (2017) RNA polymerase motions during promoter melting., *Science 356*, 4.

[8] Alexander, A. L. S. a. H. B. (1979) Initiation of transcription in vitro inhibited by lipiarmycin., *127*, 55-72.

[9] Ruff, E. F., Record, M. T., Jr. , and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation, *Biomolecules 5*, 1035-1062.

[10] Svetlov, V., and Artsimovitch, I. (2015) Purification of bacterial RNA polymerase: tools and protocols, *Methods in Molecular Biology 1276*, 13-29.

[11] Gribskov, M., and Burgess, R. R. (1983) Overexpression and purification of the sigma subunit of *Escherichia coli* RNA polymerase, *Gene 26*, 109-118.

[12] Roe, J.-H., Burgess, R. R., and Record, M. T. J. (1984) Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λPR promoter, *Journal of Molecular Biology 176*, 495-521.

[13] Sreenivasan, R., Heitkamp, S., Chhabra, M., Saecker, R. M., Lingeman, E., Poulos, M. A., McCaslin, D., Capp, M. W., Artsimovitch, I., and Record, M. T. J. (2016) Fluorescence resonance energy transfer characterization of DNA wrapping in closed and open *Escherichia coli* RNA polyerase-λP_R promoter complexes, *Biochemistry 55*, 2174-2186.

APPENDICES

Appendix I

I. Structural Modeling of Upstream DNA Unwrapping in OC to EC Transition Provides Precedence for Upstream Wrapping

Single molecule imaging by Wang et al of RNAP-promoter complexes detected conformational changes when the binary OC initiates transcription and forms a paused elongation complex (EC) [1]. The distance between two fluorescent dyes (one at -70 position of DNA and the other at either the ω CTD or β NTD) was measured as the position of the pause was increased from +20 to +60. Distances between the DNA-bound dye and either RNAP-bound dye increased linearly with slopes of ~1.7 Å/base step. Linear extrapolation to zero pause distance yields predictions for the differences in distances between -70 on DNA and the two RNAP positions between a hypothetical EC paused at the transcription start site and the binary OC of 38 +/- 2 Å between -70 and the N-terminus β subunit and 61 +/- 3 Å between -70 and the C-terminus ω subunit, interpreted by the authors using an unbent and linear upstream DNA model [1]. However, recent studies with TEC complexes of eukaryotic Pol I and Pol II revealed large static bends in the upstream DNA where it emanates from RNAP [2, 3], and cryo-EM structures of RNA Pol-III transcription initiation complex show that upstream DNA fully encircles the transcription factor TFIIIB to promote DNA opening [4]. Additionally, upstream DNA bending and wrapping of promoter DNA was observed in eukaryotic RNA Pol II Hop transcription initiation machinery [5-7].

Various bent-DNA EC models for *E. coli* RNAP were generated to explain the slopes observed by Wang et al. The model shown in Figure 1 of a partially wrapped DNA EC model, with a ~90˚ bend of the upstream DNA toward RNAP, yields predicted slopes of

1.5 Å/bp for the extensions of the distances from N-terminus β subunit and C-terminus ω subunit positions to -70 on the upstream DNA (Table 1A). These agree reasonably well with the observed values of 1.7 Å/bp, but intercept values for EC – OC differences calculated with this bent EC model and the linear upstream DNA model are inconsistent with experimental intercepts. However, use of a recently published OC model (Chapter 2)[8] with bent-wrapped upstream DNA provides very good agreement with the experimental intercepts (Table 1B, Figure 1).

II. Structural Modelling of Wrapped Upstream DNA in Transcription Activated Complex Implicates the Role of αCTDs to Facilitate Wrapping

A cryo-EM structure of *E.coli* RNAP holoenzyme transcription activated complex (TAC) interacting with class-I CAP dimer was recently obtained by Steitz et al'17 [9]. The upstream promoter DNA could be resolved until the -74 position relative to the transcription start site, showing that this upstream region bends and wraps around class-I dimer CAP protein by interacting with one of the αCTDs and σ region 4 of RNAP. Figure 2 shows a model utilizing this structure with extended upstream DNA to show upstream DNA to be interacting with the αCTDs with far upstream DNA predicted to be encircling the protein and contacting the αNTD region. We hypothesize that the flexible tethers of the αCTDs interact with extended upstream DNA and are contracted in TAC to draw the DNA closer to αNTD region to facilitate wrapping. Ongoing experiments are attempting to capture the extent of upstream wrapping in RNAP-DNA open and closed complexes through negative stain electron microscopy (EM) and single molecule fluorescence.

Methods

Modeling Trajectories of Upstream DNA in Complexes with E. coli RNAP

Open Complexes: Structural models of upstream DNA in open complexes are based on PDB 4YLN[10], used previously to create a structural model of an OC with upstream DNA bent and wrapped on RNAP with inputs from FRET and HO footprinting data [14, 11] . An OC model with unbent, unwrapped upstream DNA extending 100 bp from the catalytic site was built using a 62 bp linear DNA segment [8 11] to extend the upstream DNA in PDB 4YLN where it exits from RNAP (38 bp upstream of the catalytic side).

Elongation Complexes: To model the upstream DNA in an elongation complex (EC), previous EC model [12] and coordinates of transcribing elongating complex were taken from PDB 3LU0 and [13] as the starting point respectively. The EC model did not contain the αCTD and had upstream DNA up to 16 bp from the catalytic site interacting with the core RNAP assembly. To model the trajectory of the far upstream DNA in an EC inferred in Discussion section by subtracting the distance from -70 upstream position to two sites on RNAP i.e. (ω C terminal and β N terminal) from OC model with both wrapped and unwrapped upstream DNA, a bend in the range of 30 to 45 degrees is introduced at the upstream DNA where it exits from RNAP (25 to 35 bp upstream of the catalytic site). Upstream DNA exiting from core RNAP assembly were extended to 35 bp upstream of the catalytic site by taking structural coordinates of the bent DNA from PDB structure of αCTDs interacting with CAP and DNA (PDB structure- 1LB2) [14]. αCTDs are assumed to be interacting with the core RNAP assembly to bind and induce a bend of 30 to 45 degrees at the upstream DNA exiting from RNAP. We could not model αCTDs to be interacting with core RNAP assembly in EC model without σ^{70} (σ region 4), as was used previously in OC and CC model construction [8 11] .

DNA upstream of the bent region is extended to an additional 65 bp using linear DNA from OC model [8 11]. An EC model with linear (unbent) upstream DNA was built using a linear DNA segment of 65 bp taken from OC model [8 11] to extend the upstream DNA in TEC model and PDB 3LU0 [13] starting where it exits from RNAP, 16 bp upstream of the base in the catalytic site.

Modeling Trajectory of Far-Upstream DNA in TAC: Structural coordinates of ternary class-I transcription activated complex containing synthetic DNA scaffold (-74 to +18) having class-I dependent promoter bound to E.coli σ^{70} RNAP holoenzyme with a single αCTD (-47 to +10), and a CAP dimer (-50 to -72) were obtained from two transcription activation complexes (PDB structures 6B6H [9] and 3IYD [15]). The trajectory of far-upstream DNA (-75 to -89 region) beyond the bound CAP dimer interacting region was extended to be in contact with the second distal αCTD domain. The structural coordinates of the DNA from -75 to -89 region are obtained from PDB structure 1LB2 [14]. All structural models were made and manipulated using PyMOL Molecular Graphics System, version 1.7.0.3 (Schrödinger LLC, New York, NY).

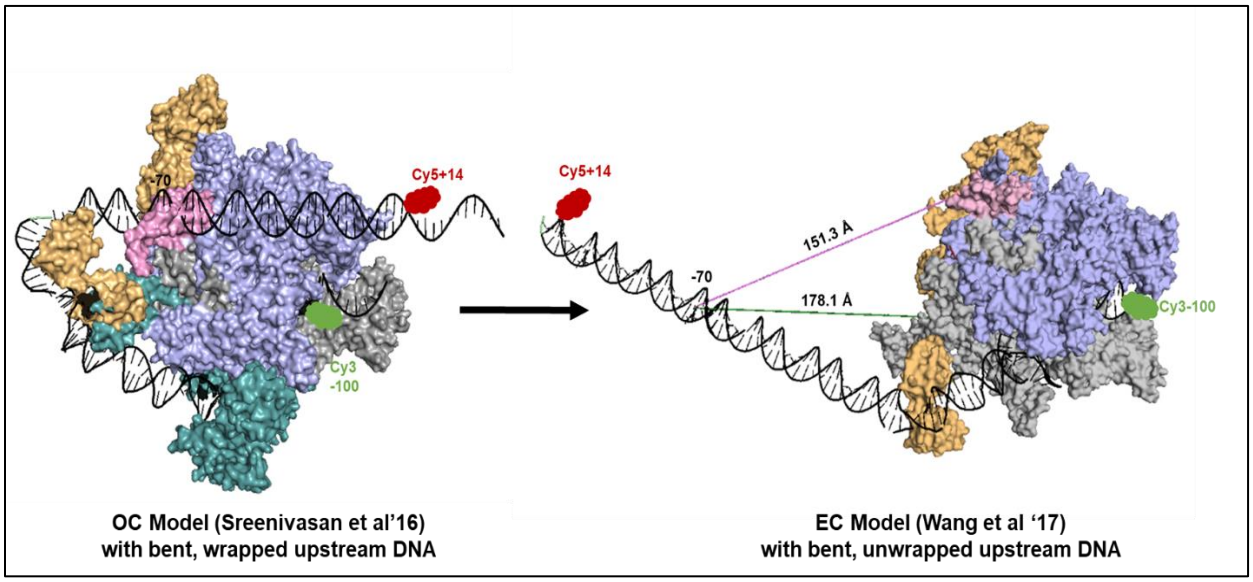


Figure 1 Structural Modelling of Upstream and Downstream DNA in the OC and the Elongation Complex (EC). The subunits of RNAP are β (grey), β' (blue), α (orange), σ^{70} (OC only; green), and ω (pink); promoter DNA is black. Cyanine dyes represented as green (Cy3) and red (Cy5) **Left:** OC model (Chapter 1) with bent and wrapped upstream DNA and bent downstream DNA. **Right**: Proposed model of EC consistent with differences between EC and OC in distance between -70 upstream position of promoter DNA and positions on ω C terminus and β N terminus of RNAP determined by Wang et al ('17) (yellow line, blue values).

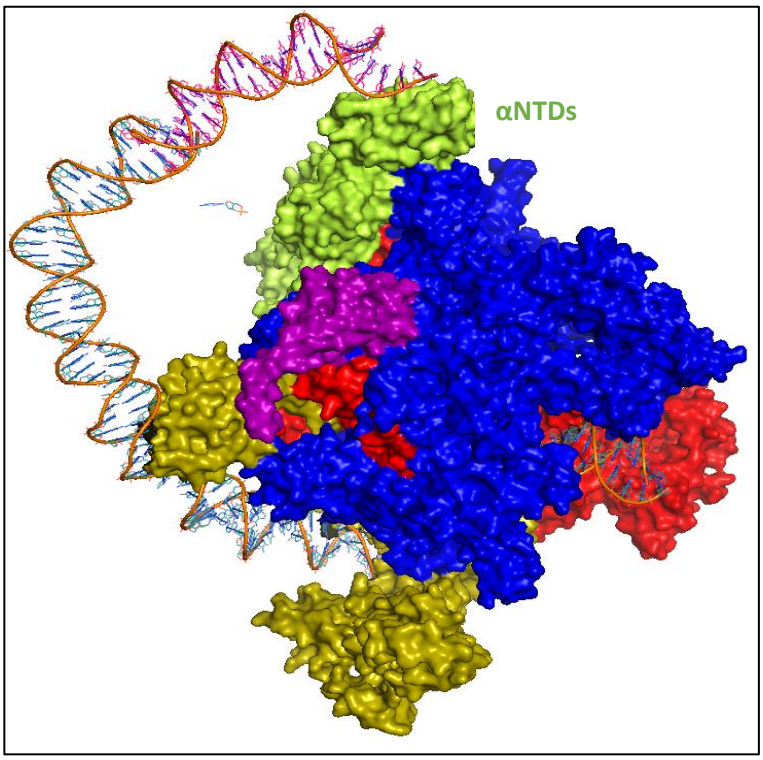


Figure 2: Structural modeling of upstream DNA in TAC Upstream DNA from TAC (Steitz et al'16) extended in the previous OC structural model to be encircled around the protein and touching αNTDs of the protein. Color code: the subunits of RNAP are β (red), β' (blue), αNTDs (green), σ^{70} (olive-brown), and ω (purple) and promoter DNA –Blue/pink orange strands.

Table 1:

A) Change in Distance from RNAP to -70 DNA in Hypothetical Conversion of OC to EC without Translocation (in Å)								
	Distance from -70 upstream position to points on RNAP in				Change in distance in OC->EC measured from DNA probe at -70 (relative to active site) to			
	OC		EC		Uncorrected		Corrected	
	ω C	β N	ω C	β N	ω C	β N	ω C	β N
	terminal							
Observed[a]	---	---	---	---	---	---	61 ± 3	38 ± 2
Predicted for bent, wrapped DNA in OC[b] to bent, unwrapped DNA in EC	29.8	102.4	146.5	173.6	117	71.2	59	36
Predicted: linear upstream DNA in OC and EC[a]	120.3	206.9	245.8	273.1	125.5	66.2	63	33.1
Predicted: bent, wrapped upstream DNA in OC[b] to linear upstream DNA in EC[a]	29.8	102.4	245.8	273.1	216	170.7	108	85
Predicted for linear upstream DNA in OC[a] to bent, unwrapped DNA in EC	120.3	206.9	146.5	173.6	26.2	33.3	13.1	17
a-Wang et al'17. b-Sreenivasan et al'16								

B) Change in Distance from RNAP to a Fixed Upstream Position of Promoter DNA in a Translocation Step of Elongation				
	Uncorrected		Corrected	
	For Probe on RNAP at			
	ω C term.	β N term.	ω C term.	β N term.
Observed[a]	---	---	1.7±0.3Å/bp	1.7±0.3 Å/bp
Predicted: Linear upstream DNA in EC	3.4 Å/bp	3.3 Å/bp	1.7 Å/bp	1.7 Å/bp
Predicted: Bent, unwrapped upstream DNA in EC	3 Å/bp	3 Å/bp	1.5 Å/bp	1.5 Å/bp
a-Wang et al'17 ***b-Calculated for DNA probe positions from -90 to -130 relative to active site***				

References:

[1] 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*, 1839-1852 e1821.

[2] Sadian, Y., Tafur, L., Kosinski, J., Jakobi, A. J., Wetzel, R., Buczak, K., Hagen, W. J., Beck, M., Sachse, C., and Muller, C. W. (2017) Structural insights into transcription initiation by yeast RNA polymerase I, *EMBO J 36*, 2698-2709.

[3] Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C., and Cramer, P. (2017) Structural Basis of RNA Polymerase I Transcription Initiation, *Cell 169*, 120-131 e122.

[4] Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., and Vannini, A. (2018) Structural basis of RNA polymerase III transcription initiation, *Nature 553*, 301-306.

[5] Forget, D., Robert, F., Grondin, G., Burton, Z. F., Greenblatt, J., and Coulombe, B. (1997) RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box, *Proc Natl Acad Sci U S A 94*, 7150-7155.

[6] Robert, F., Douziech, M., Forget, D., Egly, J. M., Greenblatt, J., Burton, Z. F., and Coulombe, B. (1998) Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF, *Mol Cell 2*, 341-351.

[7] Kim, T. K., Lagrange, T., Wang, Y. H., Griffith, J. D., Reinberg, D., and Ebright, R. H. (1997) Trajectory of DNA in the RNA polymerase II transcription preinitiation complex, *Proc Natl Acad Sci U S A 94*, 12268-12273.

[8] Sreenivasan, R., Heitkamp, S., Chhabra, M., Saecker, R. M., Lingeman, E., Poulos, M. A., McCaslin, D., Capp, M. W., Artsimovitch, I., and Record, M. T. J. (2016) Fluorescence resonance energy transfer characterization of DNA wrapping in closed and open *Escherichia coli* RNA polyerase-λP_R promoter complexes, *Biochemistry 55*, 2174-2186.

[9] Liu, B., Hong, C., Huang, R. K., Yu, Z., and Steitz, T. A. (2017) Structural basis of bacterial transcription activation, *Science 358*, 947-951.

[10] Zuo, Y., and Steitz, T. A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Molecular Cell 58*, 534-540.

[11] Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr., and Saecker, R. M. (2007) Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 104*, 7833-7838.

[12] Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000) A structural model of transcription elongation, *Science 289*, 619-625.

[13] Opalka, N., Brown, J., Lane, W. J., Twist, K. A., Landick, R., Asturias, F. J., and Darst, S. A. (2010) Complete structural model of Escherichia coli RNA polymerase from a hybrid approach, *PLoS Biol 8*.

[14] Benoff, B., Yang, H., Lawson, C. L., Parkinson, G., Liu, J., Blatter, E., Ebright, Y. W., Berman, H. M., and Ebright, R. H. (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex, *Science 297*, 1562-1566.

[15] Hudson, B. P., Quispe, J., Lara-Gonzalez, S., Kim, Y., Berman, H. M., Arnold, E., Ebright, R. H., and Lawson, C. L. (2009) Three-dimensional EM structure of an intact activator-dependent transcription initiation complex, *Proc Natl Acad Sci U S A 106*, 19830-19835.

Appendix II

Interpreting OC Dissociation Fluorescence +14 PIFE and FRET Salt Upshift Kinetics as a 2 Step Irreversible Mechanism and Determining the Forward and Backward Salt/Urea Rates Dependence for RP_O to I_2 formation.

In Chapter 3, we saw salt-dependent 2 phase decrease in Cy3+14 PIFE and Cy5+14 FRET acceptor signal till 10 seconds when pre-formed single Cy3 dye labeled and double Cy3-Cy5 labeled λP_R promoter DNA fragment-*E.coli* RNAP open complex (OC) is dissociated by mixing with high KCl salt concentration (0.4 M and 1.1 M). Analytically, the signal fits to double exponential equation 1.

$$Y = Y_0 + Pe^{-k_a t} + Qe^{-k_b t} \qquad \text{-Equation 1}$$

OC dissociation kinetics assay performed previously by filter binding kinetics in such high concentration ranges of KCl and Urea follow a 2 step irreversible mechanism. First irreversible step involves the salt dependent conversion of stable open complex (RP_O) to an unstable open complex intermediate I_2 followed by salt independent DNA closing step (I_2 to I_1) before the complete dissociation of promoter DNA from the RNAP enzyme (Mechanism 1) (Kontur et al, 05).

$$RP_O \xrightarrow{k_{-3}} I_2 \xrightarrow{k_{-2}} I_1 \qquad \text{-Mechanism I}$$

(Complete Dissociation of Promoter DNA from RNAP)

On solving the ordinary linear differential equations of the Mechanism I, one gets the following concentration kinetics dependence of RP_O (Equation 2), I_2 (Equation 3) and I_1 (Equation 4) (Book ref).

$$[RP_o]_t = [RP_o]_0\, e^{-k_{-3}t} \qquad \text{-Equation 2}$$

$$[I_2]_t = \frac{k_{-3}}{k_{-2}-k_{-3}}\left\{e^{-k_{-3}t} - e^{-k_{-2}t}\right\}[RP_o]_0 \qquad \text{-Equation 3}$$

$$[I_1]_t = \left(1 + \frac{k_{-3}e^{-k_{-2}t} - k_{-2}e^{-k_{-3}t}}{k_{-2}-k_{-3}}\right)[RP_o]_0 \qquad \text{-Equation 4}$$

$[RP_o]_0$ is the initial concentration of the stable open complex at the initial time-point t = 0.

The total fluorescence change (F) at time t observed in high salt/urea upshift OC dissociation assay following the Mechanism 1 can be expressed as the sum of products of fluorescence amplitudes of each species {(Ai), i = RP_O, I_2 and I_1} and the concentration of the species at time t {$[i]_t$, i = RP_O, I_2 and I_1} (Equation 5). (Similar to Chapter 3 analysis)

$$F = A_{RP_o}[RP_o]_t + A_{I_2}[I_2]_t + A_{I_1}[I_1]_t \qquad \text{-Equation 5}$$

On expressing equation 5 like a two exponential equation 1 using information from equations 2-4, one could correlate the kinetic parameters of equation 1 (Y_O,P,Q,k_a and k_b) to the fluorescence amplitudes and the concentration dependence of the reactant species following a 2 step irreversible mechanism as shown in Equation 6-12.

$$k_a = k_{-3} \text{ and } k_b = k_{-2} \qquad \text{-Equation 6}$$

$$Y_0 = A_{I_1} \qquad \text{–Equation 7}$$

$$P = \left[A_{RP_o} + \left(\frac{A_{I_2} * k_{-2}}{k_{-3} - k_{-2}}\right) - \left(\frac{A_{I_3} * k_{-3}}{k_{-3} - k_{-2}}\right)\right] \qquad \text{–Equation 8}$$

$$Q = \left[\left(\frac{A_{I_3} * k_{-3}}{k_{-3} - k_{-2}}\right) - \left(\frac{A_{I_2} * k_{-2}}{k_{-3} - k_{-2}}\right)\right] \qquad \text{-Equation 9}$$

On re-arranging the terms to correlate the fluorescence amplitude contribution of the species to the kinetic parameters of a 2 exponential equation one gets the following equations 10-12.

$$A_{RP_o} = \frac{P + Y_O + Q}{[RP_o]_o} \qquad \text{-Equation 10}$$

$$A_{I_2} = \frac{[Y_O + \frac{Q*(k_a - k_b)}{k_a}]}{[RP_o]_o} \qquad \text{-Equation 11}$$

$$A_{I_1} = \frac{Y_o}{[RP_o]_o} \qquad \text{-Equation 12}$$

To test the mathematical analysis described above, I fit the Cy3+14 PIFE salt upshift data obtained at 1.1 M KCl to equation 1 to get kinetic parameters (Y_O, P, Q, k_a and k_b), used Equations 10 to 12 to get the amplitude values, A_{RP_o}, A_{I_2} and A_{I_1} and then got a fluorescence simulated kinetics curve F using Equation 5 in Berkeley Madonna software. The fluorescence simulated kinetics curve F overlays beautifully on the experimental data (Figure 1A) validating the analysis and indicating that a two-step irreversible mechanism is being followed in the salt upshift OC dissociation. Fitting the experimental data performed at other salt concentrations 0.8 M, 0.7 M, 0.6 M and 0.5 M KCl to the calculated Fluorescence simulated curve F from above analysis also shows good agreement (Figure 1B, C and D) proving that the sum of two exponential equations can be modelled as a two-step irreversible mechanism.

Since the DNA closing step from $I_2 \xrightarrow{k_{-2}} I_1$ in Mechanism 1 is considered to be salt independent and $RP_O \xrightarrow{k_{-3}} I_2$ as salt dependent (Kontur et al'05), I fit the Cy3+14 PIFE and Cy5+14 FRET experimental data to a 2 exponential Equation 1 by fixing the second kinetic rate k_b or $k_{-2} = 1.23\ s^{-1}$ and got the other kinetic parameters shown in Figure 2A. The value of k_b or $k_{-2} = 1.23\ s^{-1}$ at 19^0C (assay temperature) is determined from the interpolation of the previous estimates of k_{-2} obtained at 37^0C and 10^0C by filter binding kinetics assay using Arrhenius equation. On plotting a log-log plot of log (salt

concentration) in X-axis and log of kinetic rate k_b in the Y-axis I get positive linear salt dependence value, $\mathrm{SK}_{k_{a,KCl}} = \frac{d[\ln(k_a)]}{\mathrm{d}[\ln(\mathrm{KCl})]}$ = 1.07 ± 0.16 for Cy3+14 PIFE and non-linear salt dependence for Cy5+14 FRET (Figure 2B and 2C). The FRET acceptor Cy5 signal at +14 downstream position is under the influence of the FRET Cy3 donor present at -100 upstream position. Cy5+14 FRET giving more non-linear salt dependence than Cy3+14 PIFE indicates the disruption of more electrostatic interactions taking place between the upstream promoter DNA and RNAP due to unwrapping of the upstream promoter region around the enzyme in the OC dissociation.

At 0.4 M or lower KCl concentration RP_O reversibly dissociates to form I_2 before the DNA closing step following Mechanism 2 (Kontur et al'05). The first kinetic rate (k_a) obtained from fitting the experimental data to an analytical sum of 2 exponents (Equation 1) can be expressed as a decay to equilibrium kinetic rate approximated as a sum of forward (k_3) and backward rate constants (k_{-3}) on following the Mechanism 2 . (Equation 12-13) (Similar to analysis shown in Chapter 3 for forward direction).

$$\mathrm{RP_O} \overset{\mathrm{K_3}}{\leftrightarrow} \mathrm{I_2} \overset{\mathrm{k_2}}{\rightarrow} \mathrm{I_1}$$ -Mechanism 2

(Complete Dissociation of Promoter DNA from RNAP)

$$\mathrm{K_3} = \frac{\mathrm{k_3}}{\mathrm{k_{-3}}}$$ -Equation 12

$$\mathrm{k_{decay\ to\ equilibrium}} \sim \mathrm{k_3} + \mathrm{k_{-3}} = \mathrm{k_{rate\ from\ sum\ of\ 2\ exponents}}$$ -Equation 13

The equilibrium constant values $\mathrm{K_3}$ and the ratio of $\frac{\mathrm{k_{-2}}}{\mathrm{k_{-3}}}$ have been previously estimated by filter binding salt upshift kinetics assay at 37^0C and 10^0C (Got the values in personal

communication with Dr. Wayne Kontur) (Kontur et al'05). Since we are in the high salt concentration regime 0.4 M to 1.1 M KCl, I assumed k_2 = 1.23 s^{-1} to be constant at different salt concentration and obtained interpolated values of k_3 , k_{-3} using Arrhenius equation at 19^{0}C (assay temperature) and then the equilibrium constant K_3 as shown in Table 1. The first analytic kinetic rate obtained in Figure 2 (k_a) is interpreted as a decay to equilibrium (Equation 13) and using the interpolated equilibrium constant K_3 (Equation 12) I estimated the forward (k_3) and backward (k_{-3}) kinetic rate of reversible conversion of RP_O to I_3 in these salt concentrations using +14 PIFE and FRET experimental data as shown in Table 2A and Table 2B. On plotting the log-log plot of ln([KCl]) in the X-axis and ln(k_3) and ln(k_{-3}) in the Y axis I get linear opposite salt dependence values of $SK^{PIFE}{}_{k_{3,KCl}} = \frac{d[\ln(k_3)]}{d[\ln(KCl)]} = -11.16 \pm 0.32$ and $SK^{PIFE}{}_{k_{-3,KCl}} = \frac{d[\ln(k_{-3})]}{d[\ln(KCl)]} = 1.2 \pm 0.2$ for +14 PIFE data (Figure 3A and B). In case of +14 FRET I get non-linear salt dependence for k_{-3} but linear salt dependence for k_{-3} $SK^{FRET}{}_{k_{3,KCl}} = \frac{d[\ln(k_3)]}{d[\ln(KCl)]} =$ -10.1 ± 0.66 (Figure 3C and D).

The opposite trends observed in forward (positive) and backward (negative) kinetic rates for the step of RP_O to I_3 formation on increasing salt concentration (Figure 3) implies the conversion of reversible to irreversible nature for this step with the increasing salt concentration. The absolute salt dependent values for the forward kinetic rates in both FRET and PIFE ($SK^{PIFE}{}_{k_{3,KCl}}$ and $SK^{FRET}{}_{k_{3,KCl}}$) is 10-11 fold more than the backward kinetic rate (Figure 3). This suggests the stabilization of the stronger electrostatic interactions taking place between the downstream promoter region and RNAP enzyme by the downstream mobile elements (DMEs) in I_2 to RP_O transition. On performing Cy3+14 PIFE

OC dissociation mediated by 0.8 M KCl upshift experiment with a variant RNAP lacking one of the DMEs β' jaw region- ΔJAW I get 2-3 fold reduction in the first kinetic rate as compared to WT RNAP as shown in Figure 4. These results verify the hypothesis that the nature of interactions taking place between the DMEs of the enzyme and the downstream promoter DNA in I_3 conversion to RP_O in forward direction is majorly electrostatic in nature.

The backward kinetic rate from Cy5+14 FRET of RP_O to I_3 conversion shows stronger non-linear positive salt dependence as compared to Cy3+14 PIFE ($SK^{FRET}{}_{k_{3,KCl}}$) (Figure 3A and 3C). Since, the FRET acceptor signal is under the direct influence of the FRET donor present in the upstream -100 position of promoter DNA. The non-linear salt dependence observed in FRET acceptor but not in PIFE indicates the disruption of more electrostatic interactions taking place between the RNAP protein and the upstream promoter DNA due to its unwrapping from the enzyme.

Previously, urea was also used a co-solute to initiate dissociation of stable open complex (RP_O) into I_3 intermediate and get information for the backward and forward kinetic rates salt dependence (Kontur et al'15). I performed Cy3+14 PIFE OC dissociation urea upshift experiment in 3 M and 4 M urea as described in chapter 3 and got 2 phase reduction signal from 0.01 to 10 seconds fitting well to double exponential equation 1 (Figure 5A-B). On plotting urea dependence of forward and backward kinetic rate of RP_O to I_3 kinetic step in this urea concentration regime, I get $SK^{PIFE}{}_{k_{-3,Urea}} = \frac{d[\ln(k_{-3})]}{d[Urea]} = 1.6 \pm 0.01$ and $SK^{PIFE}{}_{k_{3,Urea}} = \frac{d[\ln(k_3)]}{d[Urea]} = -1.8 \pm 0.1$ at 19^0C (Figure 6A-B) doing the similar analysis as described above. The values are very comparable to the previously obtained urea

dependent slopes of filter binding OC dissociation kinetics assay using urea upshift at 37^0C and 10^0C (Figure 6C), validating our analysis. Forward and backward kinetic rate of RP_O to I_2 conversion show opposite linear dependence with similar absolute values in increasing urea concentration to dissociate the OC. Similar to KCl salt dependence, the kinetic step of RP_O to I_2 conversion change from reversible to irreversible nature on increasing urea concentration.

De-convoluting the salt and urea dependence on the forward and backward kinetic rates of RP_O to I_3 conversion helped us to understand the nature of interactions taking place in the upstream and downstream promoter region of DNA with the enzyme in this kinetic step on doing side by side comparison of the +14 PIFE and FRET salt dependence log-log plots. The analysis also helped us to understand the change in reversible to irreversible nature of the kinetic step ($RP_O \rightarrow I_3$) on increasing salt/urea concentration in the OC dissociation upshift assay.

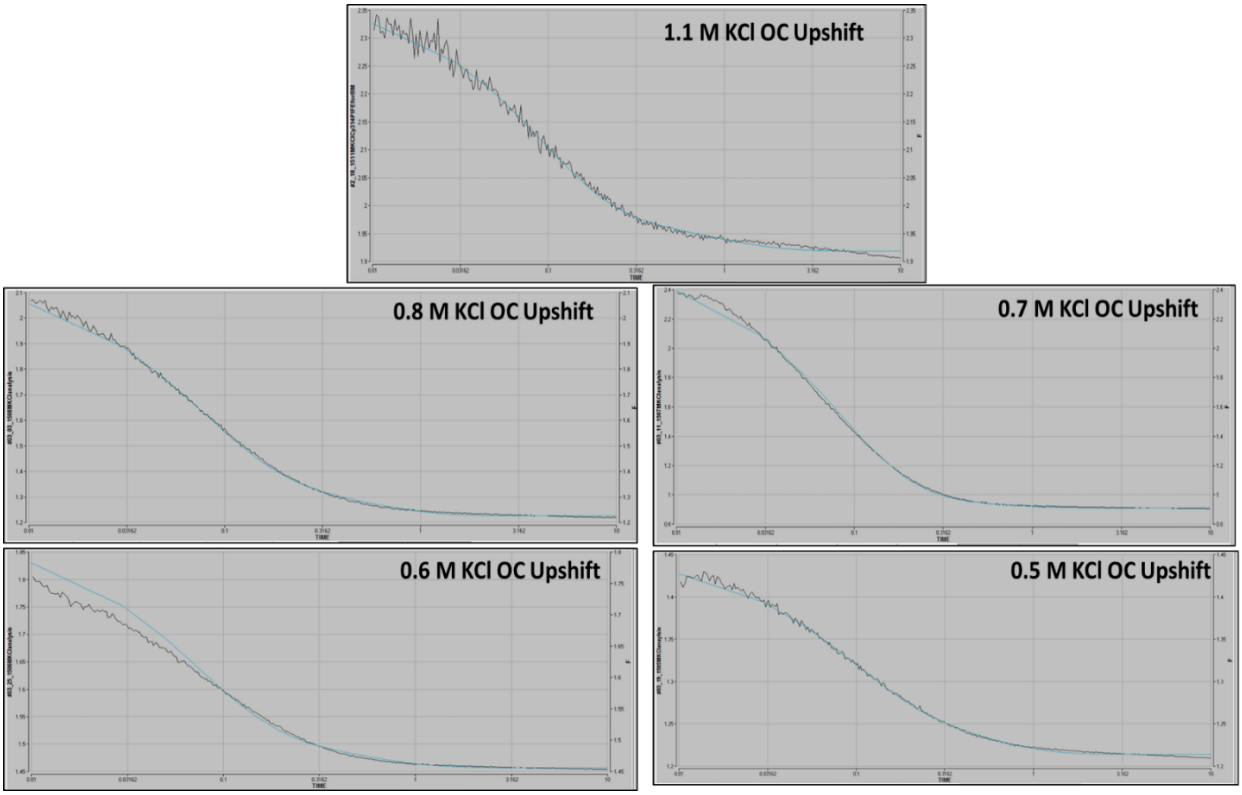
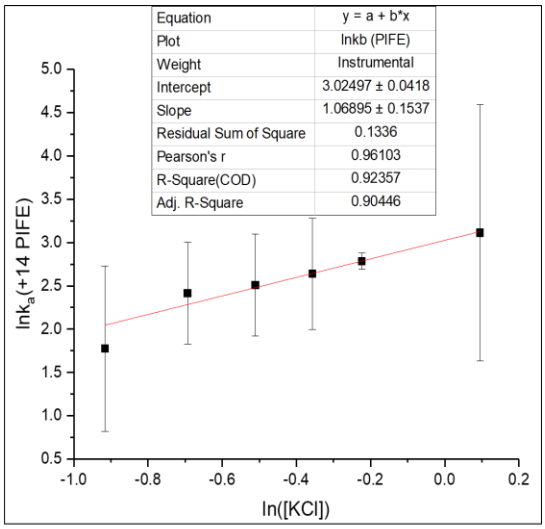


Figure 1 Overlay of Simulated Fluorescence Kinetics Curve to the Experimental Reduction Observed in Cy3+14 Salt Upshift OC Dissociation PIFE signal. Blue colored line is the calculated simulated fluorescence kinetics curve using the analysis described in the text overlaid onto the experimental Cy3+14 PIFE fluorescence reduction observed by dissociating OC with 1.1 M (panel A), 0.8 M (panel B), 0.7 M (panel C), 0.6 M (panel D)and 0.5 M KCl (panel E) concentrations. X-axis is the time in log scale from 0.01 to 10 seconds, left Y axis is the experimental fluorescence value at time t and right Y axis is the predicted simulated fluorescence value at time t.

$Y = Y_0 + Pe^{-k_a t} + Qe^{-k_b t}$					
Cy3+14 PIFE					
[KCl] (in M)	**P**	**Q**	**k_a (in s^{-1})**	**k_b (in s^{-1})**	**R-square**
1.1	0.9 ± 0.2	0.12 ± 0.03	22.5 ± 4.4	1.23	0.96 ± 0.05
0.8	1.07 ± 0.03	0.08 ± 0.004	16.2 ± 1.1		0.99 ± 0.0005
0.7	1.15 ± 0.06	0.05 ± 0.011	14 ± 1.9		0.99 ± 0.001
0.6	0.98 ± 0.05	0.12 ± 0.01	12.3 ± 1.8		0.99 ± 0.002
0.5	1.02 ± 0.06	0.12 ± 0.03	11.2 ± 1.8		0.98 ± 0.003
0.4	0.35 ± 0.1	0.69 ± 0.2	5.9 ± 2.6		0.99 ± 0.001
Cy5+14 FRET					
[KCl] (in M)	**P**	**Q**	**k_a (in s^{-1})**	**k_b (in s^{-1})**	**R-square**
1.1	0.97 ± 0.08	0.08 ± 0.02	52 ± 20	1.23	0.95 ± 0.03
0.8	0.95 ± 0.05	0.15 ± 0.03	8.45 ± 1.7		0.98 ± 0.003
0.6	0.32 ± 0.2	0.7 ± 0.2	5.8 ± 2.8		0.97 ± 0.005
0.4	0.7 ± 0.02	0.31 ± 0.01	4.5 ± 1.5		0.97 ± 0.02



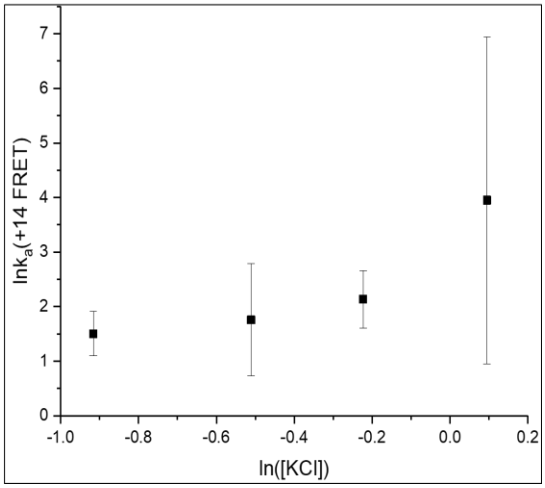
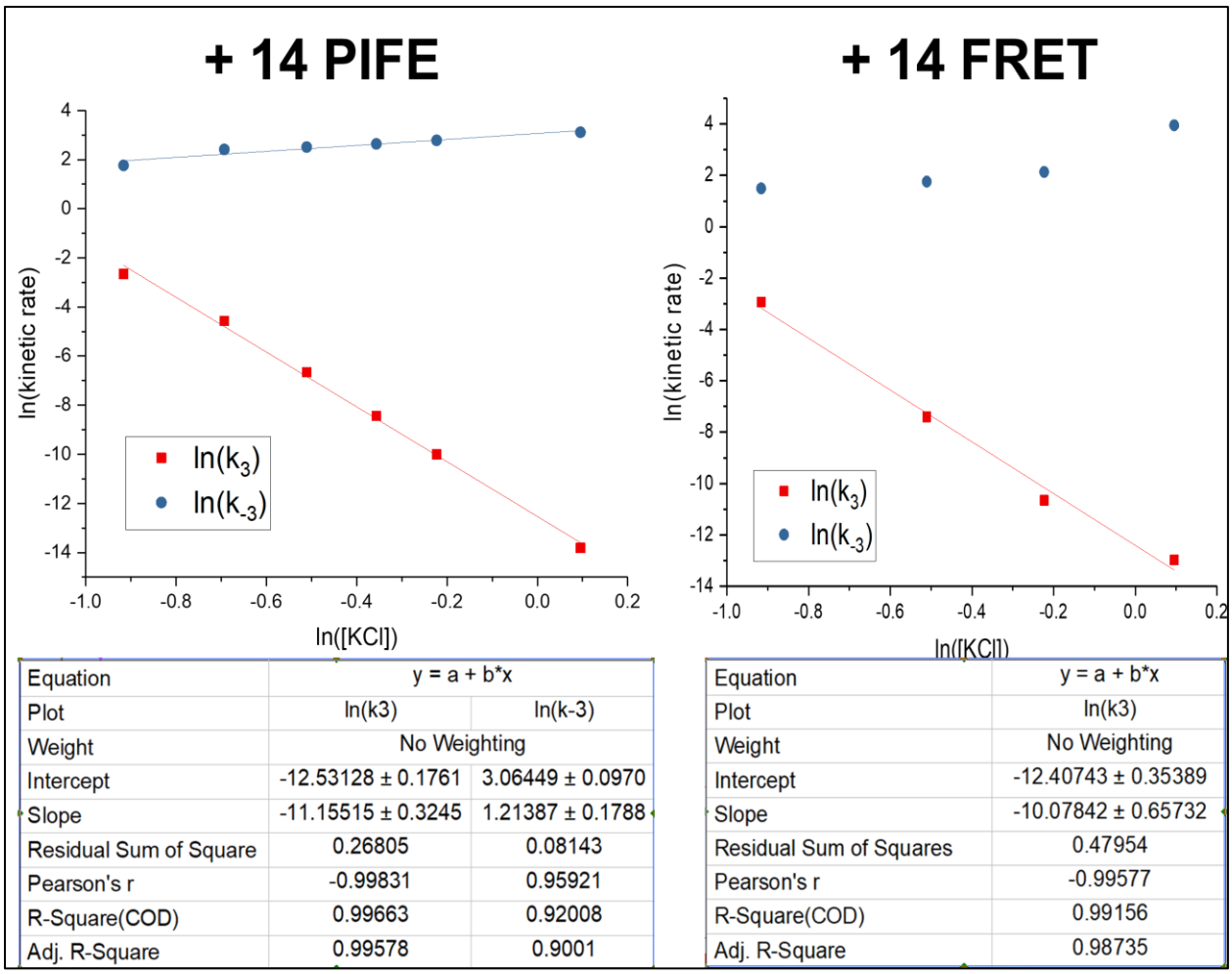


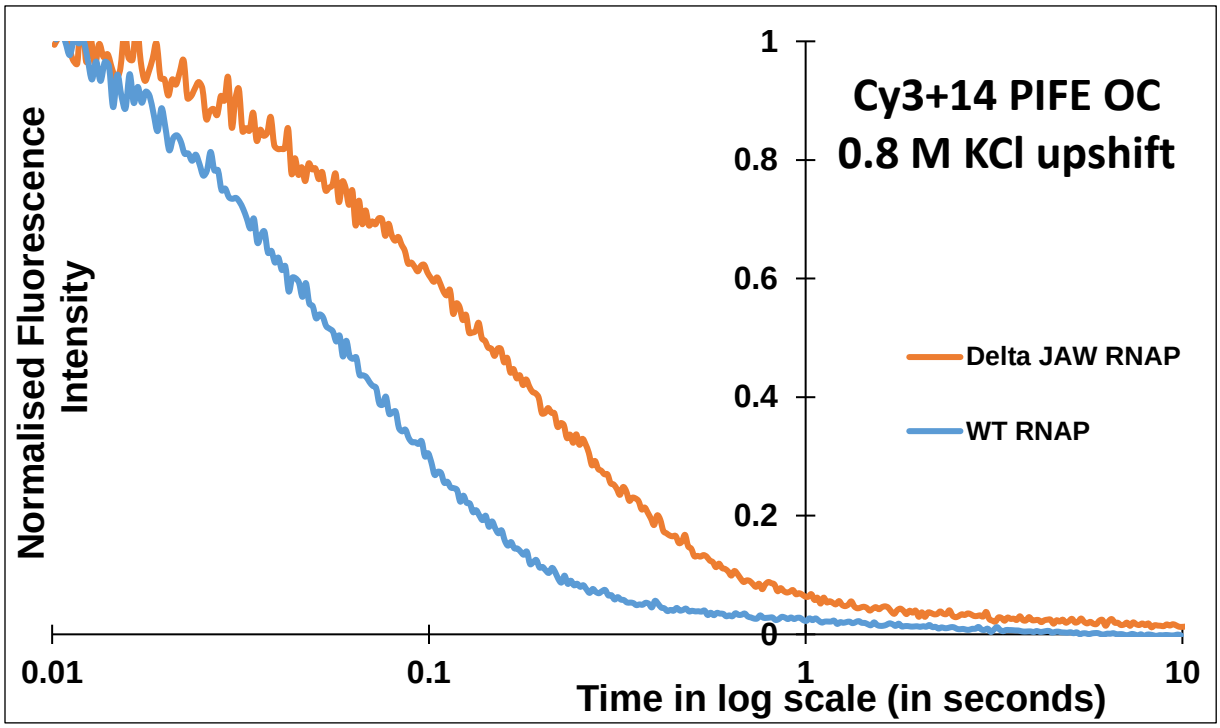
Figure 2 Fitting Cy3+14 PIFE and Cy5+14 FRET salt upshift experimental data to analytical sum of two exponents. Panel A Table gives the kinetic parameters obtained by fitting the experimental data to a sum of two exponents. **Panel B top right** is a log-log plot with ln[KCl] in the X-axis and ln(k1) Cy3+14 PIFE values from Panel A Table in the Y axis. The values fit to a linear trend with a slope of 0.92 ± 0.1 (R-square- 0.94) (in the inset). **Panel C bottom right** is a log-log plot with ln[KCl] in the X-axis and ln(k1) Cy5+14 FRET values in the Y axis showing non-linear salt dependence.



Equation	y = a + b*x	
Plot	ln(k3)	ln(k-3)
Weight	No Weighting	
Intercept	-12.53128 ± 0.1761	3.06449 ± 0.0970
Slope	-11.15515 ± 0.3245	1.21387 ± 0.1788
Residual Sum of Square	0.26805	0.08143
Pearson's r	-0.99831	0.95921
R-Square(COD)	0.99663	0.92008
Adj. R-Square	0.99578	0.9001

Equation	y = a + b*x
Plot	ln(k3)
Weight	No Weighting
Intercept	-12.40743 ± 0.35389
Slope	-10.07842 ± 0.65732
Residual Sum of Squares	0.47954
Pearson's r	-0.99577
R-Square(COD)	0.99156
Adj. R-Square	0.98735

Figure 3 Log-Log plots to show KCl salt dependence on the forward and backward kinetic rates of RP_O to I_3 formation upon OC dissociation using Fluorescence FRET (Cy5+14) and PIFE (Cy3+14) kinetics. Panel A left and Panel B right is a log-log plot with ln[KCl] in the X-axis, ln(k_3) (red box) and ln(k_{-3}) (blue circle) in the Y-axis obtained from +14 PIFE and +14 FRET data respectively (Table 2A). Data points showing linear trend are fit to a linear equation to get the slope and intercept parameters given in the table below the figure using Origin software.

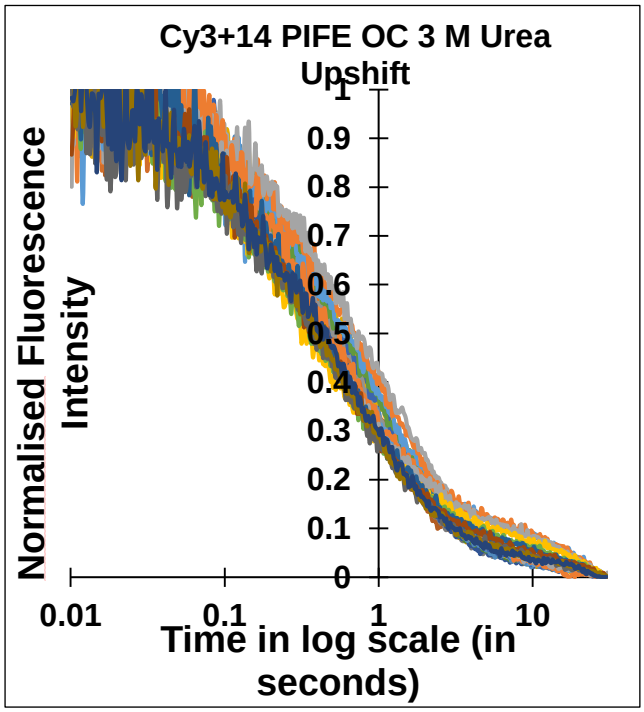


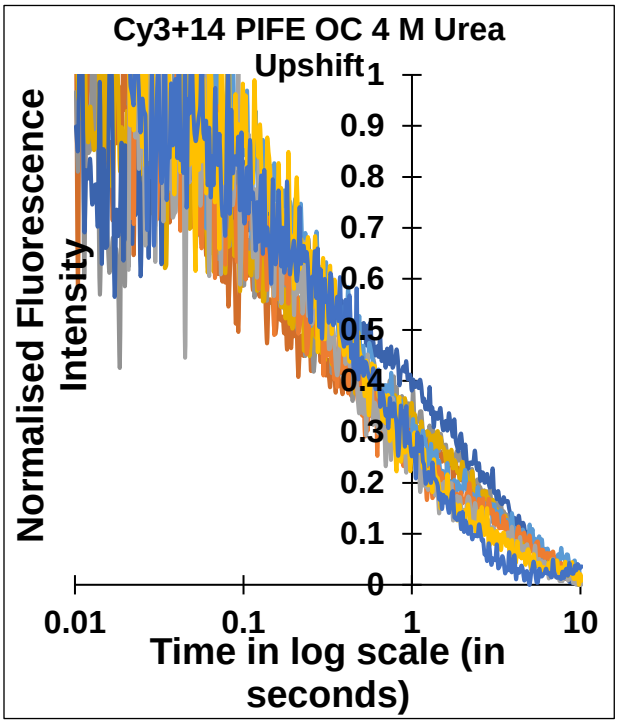
$$Y = Y_0 + Pe^{-k_1 t} + Qe^{-k_2 t}$$

Cy3+14 PIFE (0.8 M KCl upshift)

RNAP	P	Q	k_1 (in s^{-1})	k_2 (in s^{-1})	Number of shots Averaged	Number of Independent Determinations
WT RNAP	1.1 ± 0.16	0.08 ± 0.13	16 ± 2	1.23	20	4
ΔJAW	0.8 ± 0.02	0.2 ± 0.02	7.3 ± 0.3		14	2

Figure 4 Cy3+14 PIFE 0.8 M KCl upshift OC dissociation assay with WT and RNAP variant ΔJAW. Panel A top is the Cy3+14 PIFE signal observed in log time scale from 0.01 to 10 seconds when pre-formed open complex (50 nM Cy3+14 λP_R promoter DNA-50 nM RNAP variant (ΔJAW) (orange colored line) and WT RNAP (blue colored line)) is dissociated on mixing with 0.8 M KCl. **Table below** panel A figure gives the kinetic parameters obtained from fitting 20 shots and 14 shots of an experiment to a sum of 2 exponents fixing the second kinetic rate to be 1.23 s^{-1}. 4 and 2 independent replicates have been obtained for WT RNAP and ΔJAW RNAP OC dissociation in the presence of 0.8 M KCl giving similar kinetic trend and values.



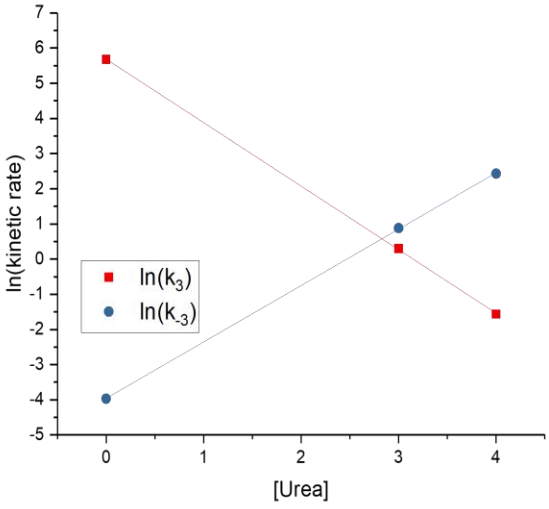


$$Y = Y_0 + Pe^{-k_1 t} + Qe^{-k_2 t}$$

Cy3+14 PIFE

[Urea] (in M)	P	Q	k_1 (in s^{-1})	k_2 (in s^{-1})	Number of shots Averaged	Number of Independent Determinations
3	0.64 ± 0.12	0.3 ± 0.13	1.65 ± 0.5	1.23	25	3
4	0.36 ± 0.03	0.6 ± 0.03	11.6 ± 1.7		8	2

Figure 5 Cy3+14 PIFE Urea upshift OC dissociation assay. Panel A and B top left and right is the Cy3+14 PIFE signal observed from each shot in log time scale from 0.01 to 10 seconds when pre-formed open complex (50 nM Cy3+14 λP_R promoter DNA-50 nM RNAP variant is dissociated on mixing with 3 and 4 M Urea respectively. **Table below** panel A and B figure gives the kinetic parameters obtained from fitting 25 and 8 shots and of the experiment performed to a sum of two exponents fixing the second kinetic rate to be 1.23 s^{-1}. 3 and 2 independent replicates have been obtained for RNAP OC dissociation in the presence of 3 and 4M Urea respectively giving similar kinetic trends and values.



Equation	y = a + b*x	
Plot	ln(k3)	ln(k-3)
Weight	No Weighting	
Intercept	5.68713 ± 0.039	-3.95873 ± 0.03
Slope	-1.80649 ± 0.01	1.60382 ± 0.013
Residual Sum of Squ	0.00158	0.00164
Pearson's r	-0.99997	0.99996
R-Square(COD)	0.99994	0.99993
Adj. R-Square	0.99989	0.99985

(in M^{-1})	37°C [a]	10°C [a]	19°C (+14 PIFE)
$dlnk_{-3}/d[Urea]$	1.1 ± 0.2		1.6 ± 0.01
$dlnk_3/d[Urea]$	-2.2 ± 0.2	-2.5 ± 0.2	-1.8 ± 0.1
$dlnK_3/d[Urea]$	-3.3 ± 0.2	-3.5 ± 0.1	-3.4 ± 0.1

[a]Previous Filter Binding Kinetics assay (Kontur et al,05)

Figure 6 Log-Log plots to show Urea co-solute dependence on the forward and backward kinetic rates of RP_O to I_3 formation upon OC dissociation using Cy3+14 PIFE kinetics. Panel A top left is a semi-log plot with [Urea] in the X-axis, ln(k_3) and ln(k_{-3}) in the Y-axis obtained from +14 PIFE data (Figure 5). **Table below** panel A gives the slopes and intercepts on fitting the data to a linear equation. **Table on the right side** compares the urea dependence slopes obtained from the panel A semi-log plot to the previously determined values for the OC dissociation urea upshift experiment done by filter binding kinetics assay at 37°C and 10°C.

Table 1

Calculated from Wayne's Data and Interpolated to 19^0C.			
[KCl] (in M)	**k_3 (in s^{-1})**	**k_{-3} (in s^{-1})**	**K_3**
1.1	**4.66×10^{-4}**	**10413.9**	**4.47×10^{-8}**
0.8	**2.83×10^{-3}**	**1012.6**	**2.8×10^{-6}**
0.7	**6.08×10^{-3}**	**392.0**	**1.55×10^{-5}**
0.6	**1.41×10^{-2}**	**134.6**	**1.05×10^{-4}**
0.5	**3.66×10^{-2}**	**39.7**	**9.23×10^{-4}**
0.4	**1.15×10^{-1}**	**9.6**	**1.20×10^{-2}**

Table 2

+14 PIFE						
[KCl] (in M)	$K_3 = k_3/k_3$ (Table 1)	$k_{DecaytoEquilibrium}$ (k_3+k_{-3}) (Figure 1,PIFE)	k_3 (in s^{-1})	k_{-3} (in s^{-1})	ln(k_3) (PIFE)	ln(k_{-3}) (PIFE)
1.1	4.47×10^{-8}	22.5 ± 4.4	1.01×10^{-6}	22.5	-13.7	3.11
0.8	2.8×10^{-6}	16.2 ± 1.1	4.54×10^{-5}	16.2	-10	2.78
0.7	1.55×10^{-5}	14 ± 1.9	2.17×10^{-4}	14	-8.43	2.64
0.6	1.05×10^{-4}	12.3 ± 1.8	1.29×10^{-3}	12.3	-6.65	2.51
0.5	9.23×10^{-4}	11.2 ± 1.8	1.03×10^{-2}	11.2	-4.57	2.41
0.4	1.20×10^{-2}	5.9 ± 2.6	0.7×10^{-1}	5.83	-2.66	1.76
+14 FRET						
[KCl] (in M)	$K_3 = k_3/k_{-3}$ (Table 1)	$k_{DecaytoEquilibrium}$ (k_3+k_{-3}) (Figure 1,FRET)	k_3 (in s^{-1})	k_{-3} (in s^{-1})	ln(k_3) (FRET)	ln(k_{-3}) (FRET)
1.1	4.47×10^{-8}	52 ± 20	2.32×10^{-6}	52	-13	3.95
0.8	2.8×10^{-6}	8.45 ± 1.7	2.37×10^{-5}	8.45	-10.6	2.13
0.6	1.05×10^{-4}	5.8 ± 2.8	6.1×10^{-4}	5.8	-7.4	1.75
0.4	1.20×10^{-2}	4.5 ± 1.5	5.34×10^{-2}	4.45	-2.93	1.5

Appendix III

I. A Preliminary Attempt to Understand +14 PIFE OC Formation Kinetics of Upstream Truncated DNA.

In Chapter 3, open complex (OC) formation +14 PIFE kinetics signal begin to increase from 0.01 seconds and peak to its maximum value >1 seconds, obtained by mixing 50 nM E.coli RNAP with 50 nM single dye fluorescently labelled λP_R promoter DNA fragment in the stopped-flow fluorimeter. The DNA fragment is of 114 bp spanning from -100 upstream to +14 downstream positions with respect to the transcription start site. I created DNA fragments containing λP_R promoter sequence from -60 and -47 upstream position till downstream +14 attached with Cy3 dye at this position using the primers whose sequences are shown in Table 1 and Table 2 following the PCR protocol described in Chapter 2 and Chapter 3. The DNA fragments are similar to Cy3+14 upstream truncated at -47 (UT-47) and -60 position (UT-60) with respect to the transcription start site from FL fragment. On mixing Cy3+14 labelled UT-47 and UT-60 λP_R promoter DNA fragment with *E.coli* RNAP in the stopped flow fluorimeter to form OC the PIFE signal plateau shows reduction in 32% for UT-60 and 50% for UT-47 as compared to FL DNA fragment (Figure 1). These effects are similar to the decrease in the fraction of OC observed previously by filter binding kinetics assay for λP_R and lacUV5 promoter DNA upstream truncated at -63/-65 and -47/-45 position with respect to the transcription start site [1 2 3].

We propose that Cy3+14 PIFE OC formation signal observed from 0.01 to 400 seconds for UT-60 and UT-47 is majorly due to the formation of heparin sensitive closed complex ensemble of intermediates. It is due to the slower isomerization kinetic rates of transition from ensemble to OC, estimated previously by filter binding kinetics assay [1 2]. In order to test the hypothesis we performed double mixing-3 syringe mixing

experiment in the stopped flow fluorimeter (Figure 2). In the first syringe fluorescently labelled promoter DNA fragment (FL, UT-60, UT-47) is first mixed in 1:1 volume with E.coli RNAP and incubated for different time-points (8, 80 and 320 seconds). After incubation for certain amount of time, E.coli RNAP-promoter DNA mixture is then challenged with heparin mixture (50 ug/ml final concentration) by mixing in 2:1 volume ratio and the fluorescence signal is begin to be observed in logarithmic time scale from 0.01 to 400 seconds. The final concentration in the observation cell is 50 nM Cy3+14 λP_R promoter DNA, 50 nM E.coli RNAP and 50 ug/ml heparin mixture.

When *E.coli* RNAP and λP_R promoter DNA of varying upstream length is incubated for 8 seconds to form closed complex intermediate before mixed with heparin to form OC, FL DNA fragment shows an increase in signal by 0.04 units from 0.01 to 10 seconds followed by a slight decrease of 0.01 units after 10 seconds to 40 seconds (Figure 2A). The increase in signal after heparin challenge indicates increase in formation of heparin resistant open complexes. However, in case of UT-60, Cy3+14 PIFE signal shows an increase with earlier kinetics from 0.01 to 1 seconds by 0.03 units, plateau till 10 seconds followed by decrease to the original initial signal after 10 seconds till 40 seconds (Figure 2A). In case of UT-47, Cy3+14 PIFE signal increase by only 0.01 units from 0.01 to 1 seconds followed by a decrease till 400 seconds. At 80 and 320 seconds, Cy3+14 PIFE signal in FL and UT-60 increase by 0.02 units till 10 seconds followed by a decrease by similar amount (0.03 units) till 400 seconds (Figure 2B and Figure 2C). In case of UT-47, 80 seconds pre-incubation of DNA and RNAP shows 0.02 units increase till 10 seconds followed by a slight decrease beginning from 100 seconds indicating formation of late heparin resistant OC (Figure 2B). However, when Cy3+14 UT-47 λP_R promoter DNA fragment is pre-incubated with E.coli RNAP for 320 seconds then mixed with heparin PIFE+14 signal increases by

only 0.01 units from 0.01 seconds to 1 seconds and begins to decrease after that. Much more determinations and different types of experiments are needed to interpret these results. However, overall UT-47 DNA fragment shows no development of +14 PIFE signal when formed complex with E.coli RNAP till 8 seconds (Figure 2A) and 320 seconds (Figure 2B) and later challenged with heparin indicating the signal observed in Cy3+14 PIFE for UT-47 in Figure 1 is due to the formation of closed complex ensemble of intermediates.

II. Displacement of Downstream DNA from the Active site cleft of RNAP by Lipiarmycin drug in I_{1M}-LpM formation is Independent of the contacts of σ^{70} NCR with the Spacer Region of Promoter DNA.

Recent Cryo-EM structure of *E.coli* σ^{70} RNAP–DNA open complex showed a novel interaction between R157 reside of non-conserved region (NCR) domain of σ^{70} and the phosphate backbone of -16 and -17 base position in the spacer region of promoter DNA [4]. RNAP-DNA OC formation using σ^{70} R157A and R157E mutated variants through biochemical fluorescence and electrophoretic mobility shift assays showed R157 to not affect the early but later steps of OC formation [4]. Our Cy3+14 PIFE downstream promoter region interaction assay and it's displacement from the active site using Lipiarmycin drug as described in Chapter 3 and Chapter 4 also show similar changes for these variants as compared to WT RNAP for both full length and upstream truncated promoter DNA fragment at -47 position (Figure 3). These results indicate that the interaction between R157 and phosphate backbone in the spacer region does not take place in I_{1M}-LpMformation but later in the mechanism of OC formation.

III. Transcription Profile of RNAP-DNA Open Complexes is more different when varying the downstream than the upstream lengths of promoter DNA.

I performed single round transcription kinetics assay as described in Chapter 3 using 2 types of DNA fragments varying in downstream and upstream lengths. The sequences of the primers used to prepare these fragments are given in Table 1 and Table 2. The fragments were prepared by PCR protocol as described in Chapter 3. In one set of experiments, I had DNA fragment spanning from -110, -80, -47 upstream position until +22 downstream position with respect to the transcription start site. I pre-formed the open complex by incubating the DNA fragments with *E.coli* RNAP at 37^{O}C for an hour. I added radiolabelled UTPs with 20 μM UTP and other 100 μM ATP and CTP with heparin mix (50 μg/ml) to pre-formed OC to perform transcription in single round for different time-points (10, 30, 90, 150, 240 and 480 seconds) and then terminated the reaction by adding urea buffer, as described in Chapter 3. On running a transcription gel I get abortive RNA transcripts from 3 to 9 mer, build-up of 10 mer at the point of escape and longer RNA transcripts of length 11 to 22 mer for all three type of DNA fragments -110 to +22, -80 to +22 and -47 to +22 (Figure 4A).

In another set of experiment, I had DNA fragment varying from -110, -80 and -47 upstream position to the +36 downstream position with pauses introduced at +16, +21,+22, +23 sites by having base ‘G’ at this position which is not provided in the NTP reaction mix. On running the transcription gel as shown in Figure 4B I get less abortive RNA transcripts compared to Figure 4A when the DNA is truncated at downstream position +22, slight build up at 10 mer but more at the first 2 pause sites i.e +16 and +22 position (Figure 4B). Qualitatively, truncating the upstream promoter region at -47 position show late kinetics to form full length RNA transcript as compared to the OC formed by full length promoter DNA fragment. The amount of RNAs formed qualitatively also seem less for OC formed with upstream truncated (UT-47) promoter DNA fragment than the full length DNA fragment (Figure 4A and 4B). However,

stronger differences are observed in the transcription profile on doing side-by-side comparison of Figure 4A and Figure 4B which have varying downstream promoter length to transcribe but similar upstream promoter region. Truncating the upstream promoter length didn't affect the stability of the OC, determined by previous OC filter binding dissociation kinetics assay [1]. Thus we don't see much stronger effects in the transcription profile of open complexes formed with varying upstream promoter DNA region compared to varying the downstream region of promoter DNA whose interactions are stabilized with RNAP in the later stages of OC formation [5].

To test the hypothesis that downstream truncation matters most than upstream truncation in producing RNA transcripts, I performed similar single round transcription kinetics experiment as described above with pre-formed OC comprising of λP_R promoter DNA fragment spanning from upstream -100 to downstream +14 position with/without the dye at different temperature 19^0C, 25^0C and 37^0C (Figure 5A and Figure 5B). I get 6-8 mer RNA transcripts and then 14 mer as the full length for this DNA fragment at any temperature. On performing transcription experiment with the promoter DNA fragment spanning from -100/-110 upstream till +22/+36 downstream position I get 6 mer, 7 mer , 10 mer, and other RNA transcripts < 10 mer till FL (Figure 5B). Truncating the downstream promoter region at +14 position completely changed the transcription profile of short abortives and long RNA transcripts formed with the same ITR region.

Table 1 DNA sequences:

Promoter	Upstream primer	Downstream primer	Duplex (bp) and position relative to start site (+1)[a]	Experiment used in
λP_R (-110 to +22)	Up1	Down1	132 (-110 to +22)	Single round transcription kinetics
λP_R (-80 to +22)	Up2		102 (-80 to +22)	
λP_R (-47 to +22)	Up3		69 (-47 to +22)	
λP_R (-110 to +36)	Up1	Down2	146 (-110 to +36)	
λP_R (-80 to +36)	Up2		116(-80 to +36)	
λP_R (-47 to +36)	Up3		86(-47 to +36)	
Cy3(+14) FL λP_R [b]	Up	Cy3+14 (Down)	114 (-100 to +14)	Fluorescent kinetics
Cy3(+14) UT-60 λP_R [b]	Up4		142 (-128 to +14)	
Cy3(+14) UT-47 λP_R [b]	Up3		164 (-100 to +64)	
[a] Dye-labeled DNA have 8 base (downstream) ssDNA termini				
[b] Sreenivasan et al'16				

Table 2 Primer Sequences:

	Primer	Sequence
1	Up	5’-GTACGAATTCGATATCCAGCTATGACCATGATTACGCCAAGC-3’
2	Up1	5'-CAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGT-3'
3	Up2	5’-TGCCTGCAGGTTTAAACAGTCGATAAATATCTAACACCGT-3’
4	Up3	5’ – ACACCGTGCGTGTTGACTATTTTACCTCTGGCG -3’
5	Up4	5’-CGATAAATATCTAACACCGTGCGTGTTGAC – 3
6	Down1	5’-CCATACAACCTCCTTACTACATGCAACCATTATCACCGCC-3’
7	Down2	5’-CGC GAT ACA TTA TTG GAT AGA ACC TCC TTA CTA CAT GCA AC-3’
8	Cy3 (+14)	5’-CCATACAA/iCy3/CCTCCTTACTACATGCAACCATTATCACCGCC-3’

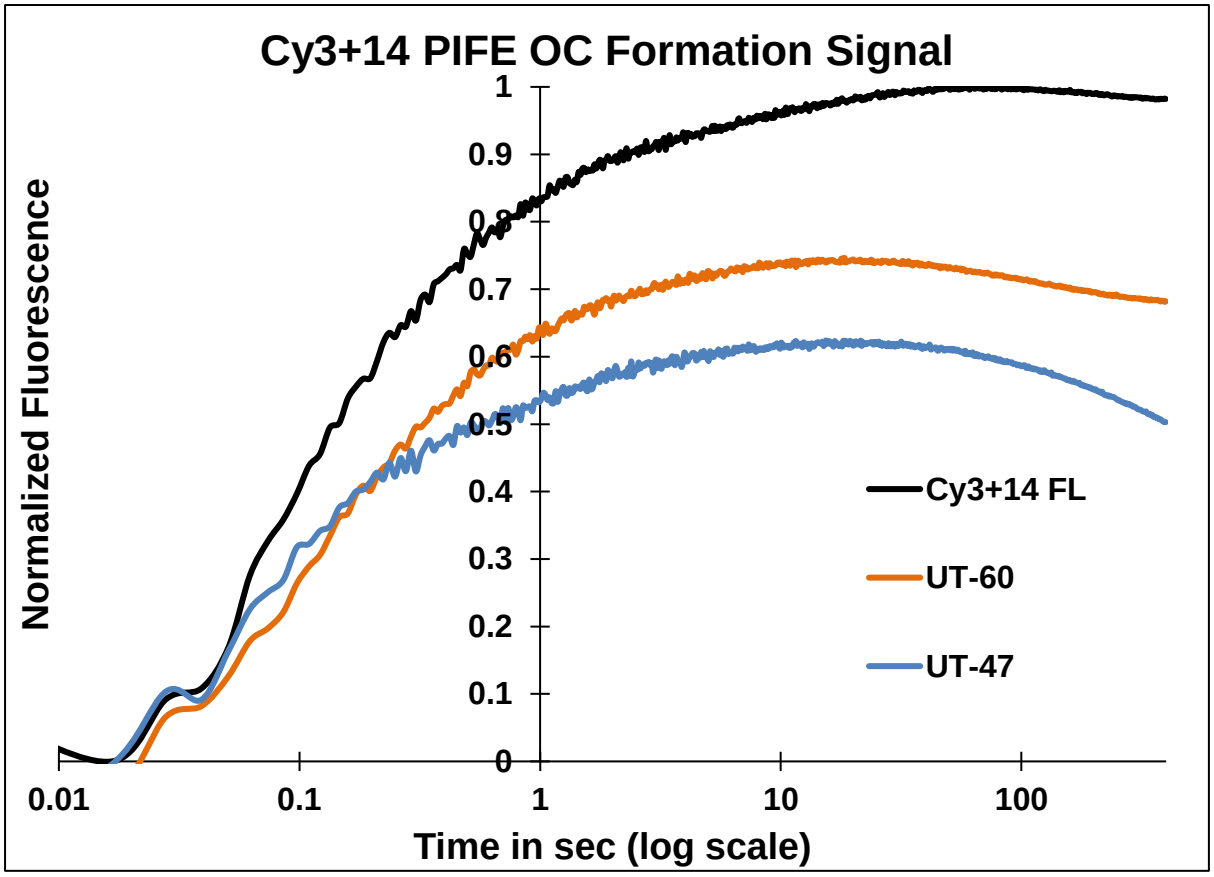


Figure 1: Cy3+14 PIFE open complex formation kinetics signal when mixed with upstream truncated DNA fragments. 50 nM *E.coli* RNAP is mixed with 50 nM λP_R promoter DNA fragment labeled with Cy3 dye at +14 downstream position and truncated at upstream -60 (blue curve) and -47 position (orange curve) with respect to the transcription start site to form OC. Change in Cy3+14 PIFE signal is monitored in logarithmic scale from 0.01 to 400 seconds upon mixing in the stopped-flow fluorimeter. Black curve is the +14 PIFE signal of a FL DNA fragment for comparison.

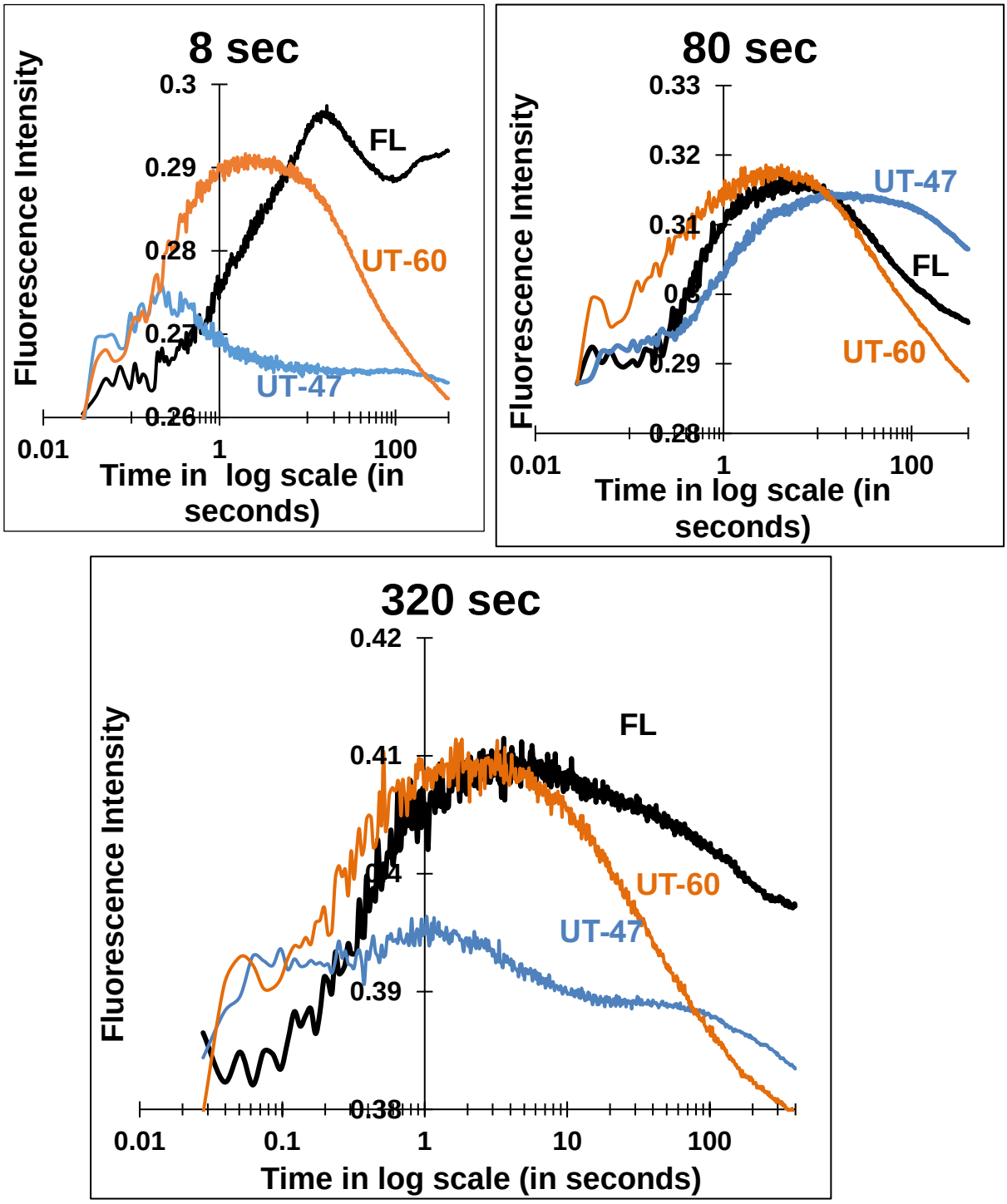


Figure 2: Cy3+14 PIFE open complex formation kinetics signal from double mixing/3 syringe-mixing-experiment when 50 nM λP_R promoter DNA with varying upstream length is mixed with 50 nM E.coli RNAP and then with 50 µg/ml Heparin mixture. 50 nM *E.coli* RNAP is mixed with 50 nM λP_R promoter DNA fragment labeled with Cy3 dye at +14 downstream position and truncated at upstream

-60 (blue curve) and -47 position (orange curve) with respect to the transcription start site to form OC and incubated for8, 80 and 320 seconds and then mixed with 50 μg/ml heparin. Change in Cy3+14 PIFE signal is monitored in logarithmic time scale from 0.01 to 400 seconds upon the second mixing in the stopped-flow fluorimeter. Black curve is the +14 PIFE signal of a FL DNA fragment for comparison.

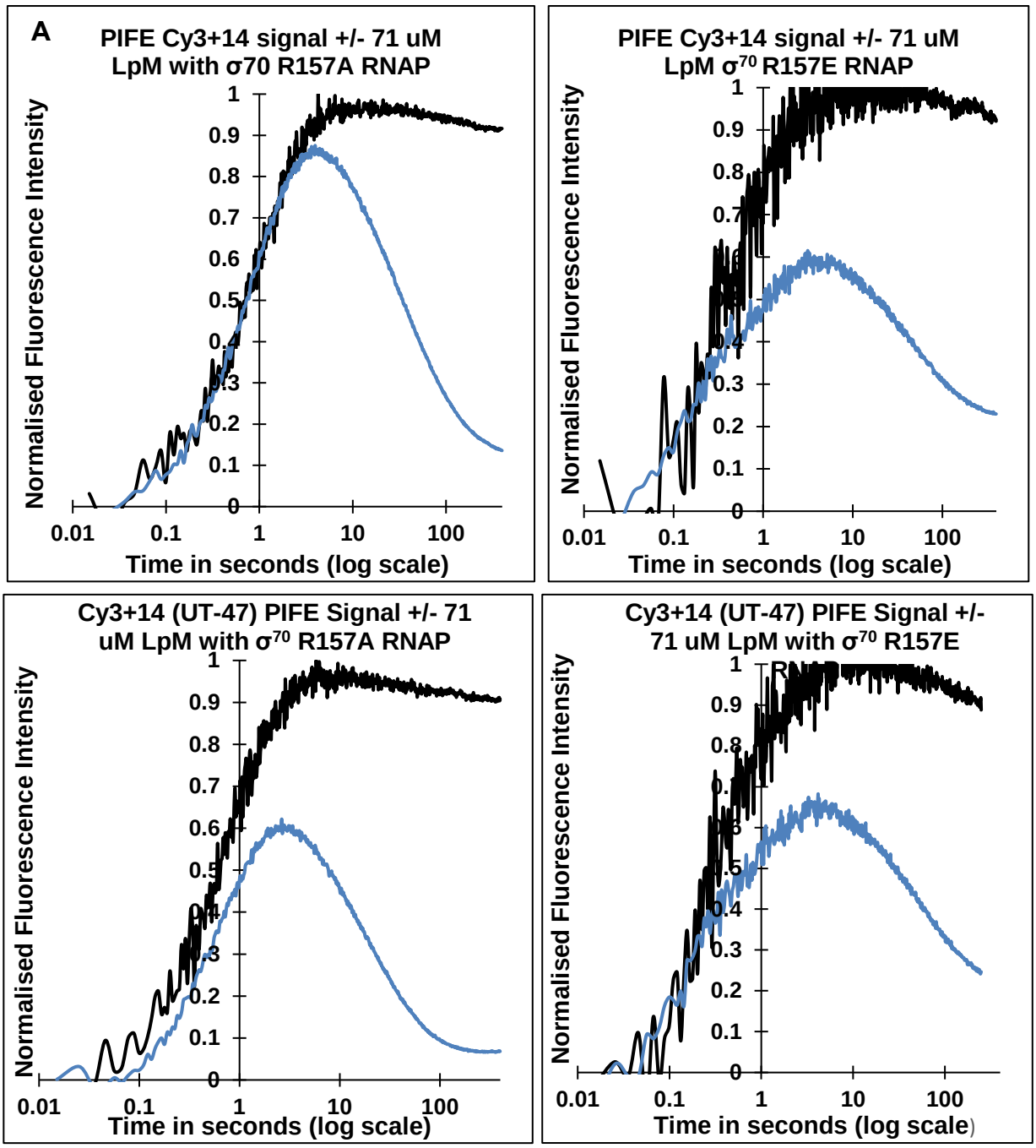


Figure 3 Downstream Cy3+14 OC Formation PIFE signal is unaffected when σ^{70} variant E.coli RNAP (R157A and R157E) mixed with Cy3+14 fluorescently labelled DNA fragment Full length and truncated at upstream -47 position in 71 uM LpM. Cy3+14 PIFE signal obtained on mixing 50 nM fluorescently labeled λP_R DNA fragment full length **(top Panel A and B)** or truncated at -47 upstream position **(bottom Panel C and D)** in 71 uM Lpm in one syringe with E.coli RNAP σ^{70} variant holoenzyme, R157A **(left Panel A and C)** and R157E **(right Panel B and D) ,** in the

other syringe of stopped flow kintek instrument at 19^0C. The fluorescence data is collected in logarithmic scale within 400 seconds in 600 points. Black curve is the signal when fluorescently labelled DNA is mixed with RNAP in the absence of LpM.

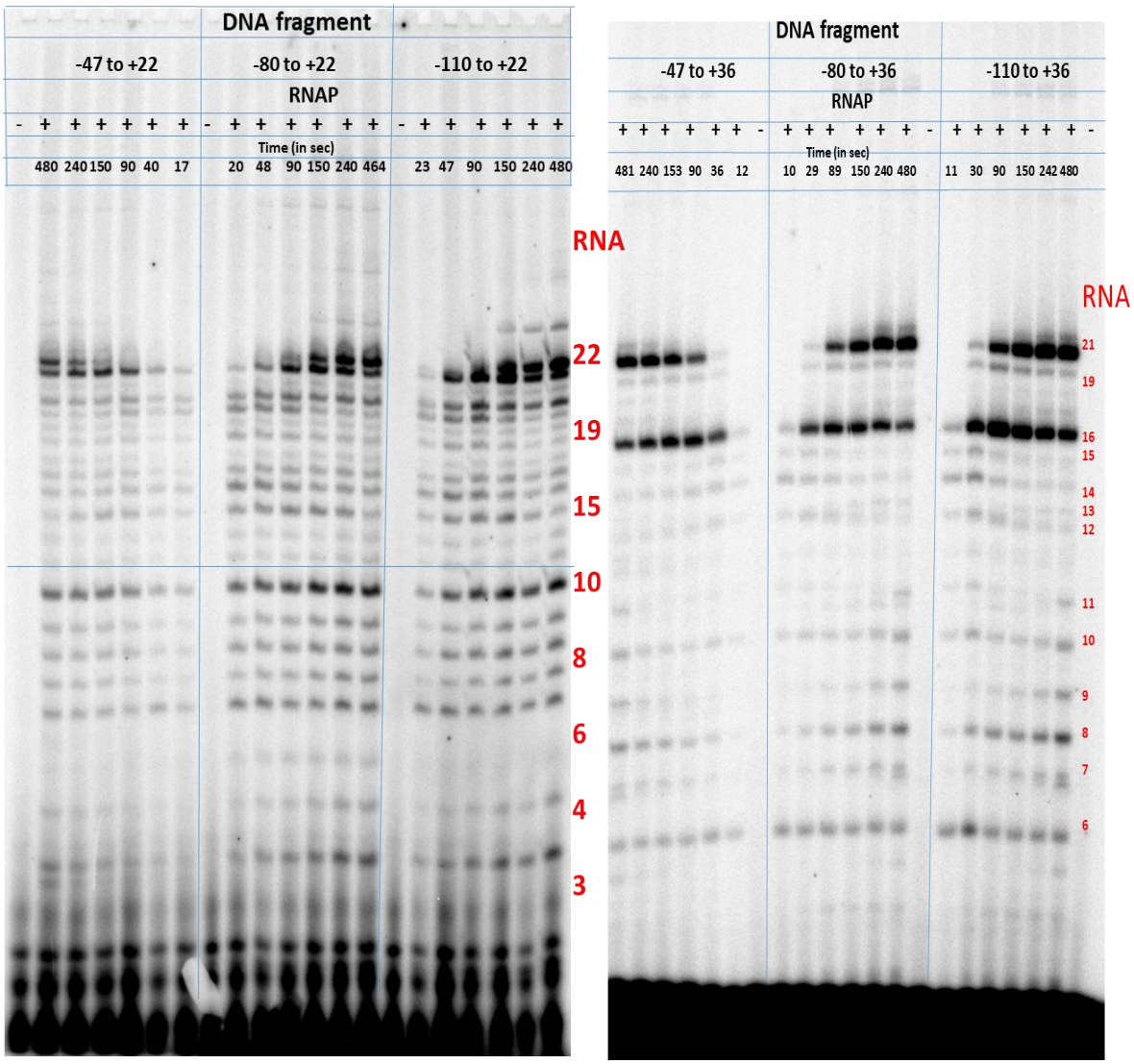


Figure 4 Single Round Transcription Kinetics Assay from Open complex with varying Upstream and Downstream Promoter DNA lengths. Transcription gel showing the different amount of RNA transcripts formed in single round at different time-points. In Panel A left, the promoter DNA spans from -110, -80 and -47 upstream position to +22 downstream position. In Panel B right, the downstream region of promoter DNA is till +36 with pause sites introduced at +16, +21 position with upstream region till -47, -80 and -110 position. In both the gels negative lane for each is a control without RNAP. The time at which the transcription reaction is terminated is mentioned in the top row.

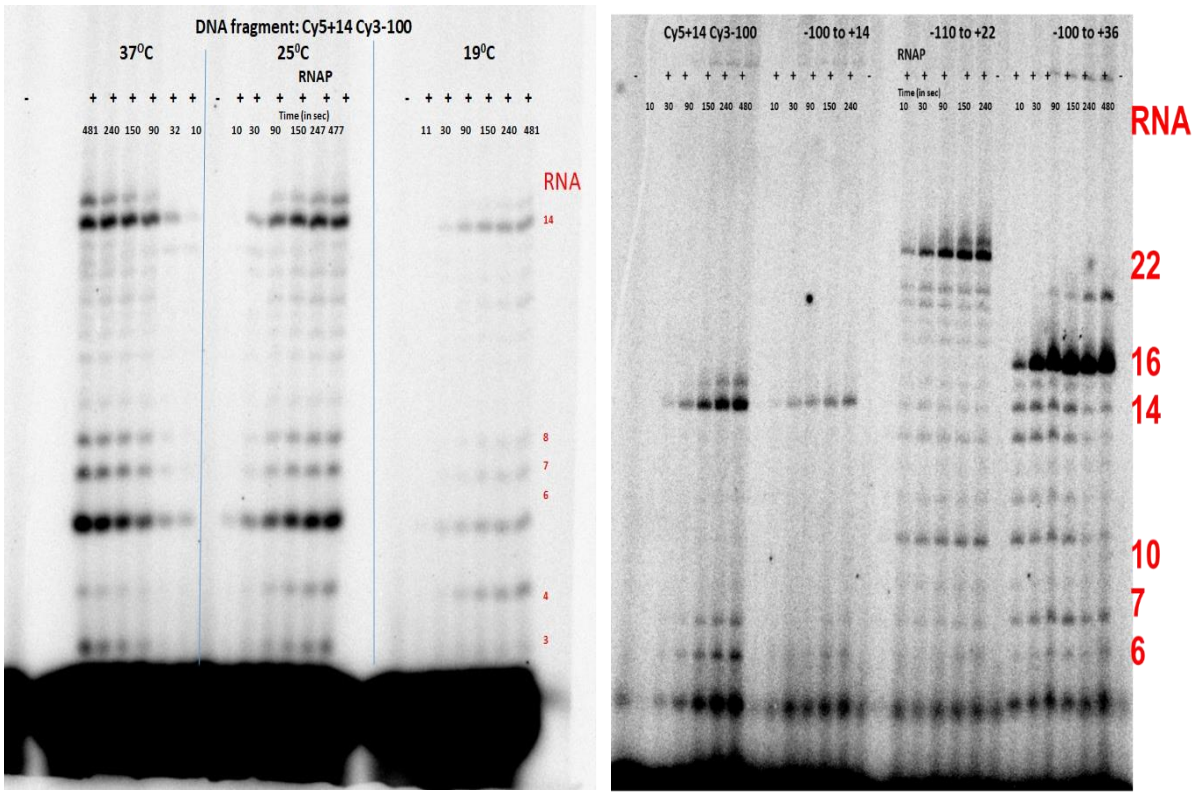


Figure 5 Single Round Transcription Kinetics Assay from Open complex with varying Downstream Promoter DNA lengths. Transcription gel showing the different amount of RNA transcripts formed in single round at different time-points. In Panel A left, λP_R promoter DNA fragment spans from -100 upstream to +14 downstream position and transcription assay is performed at 3 different tmepratures, 19, 25 and 37^0C. In Panel B right, the downstream region of promoter DNA is till +14, +22 and +36 with pause sites introduced at +16, +21 position with upstream region till -110 position. In both the gels negative lane for each is a control without RNAP. The time at which the transcription reaction is terminated is mentioned in the top row.

References:

[1] Davis, C. A., Capp, M. W., Record, M. T., Jr., and Saecker, R. M. (2005) The effects of upstream DNA on open complex formation by *Escherichia coli* RNA polymerase, *Proceedings of the National Academy of Sciences 102*, 285-290.

[2] Ross, W., and Gourse, R. L. (2005) Sequence-independent upstream DNA-aCTD interactions strongly stimulate *Escherichia coli* RNA polymerase-*lac*UV5 promoter association, *Proceedings of the National Academy of Sciences 102*, 291-296.

[3] Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., and Record, M. T. J. (2002) Kinetic studies and structural models of the association of *E. coli* sigma 70 RNA polymerase with the LPR promoter: large scale conformational changes in forming the kinetically significant intermediates, *Journal of Molecular Biology 319*, 649-671.

[4] 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 sigma(70) RNA polymerase and promoter DNA complex revealed a role of sigma non-conserved region during the open complex formation, *J Biol Chem 293*, 7367-7375.

[5] Ruff, E. F., Record, M. T., Jr. , and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation, *Biomolecules 5*, 1035-1062.