Fluorescence-Detected Conformational Changes in Duplex DNA in Open Complex Formation by *E. coli* RNA Polymerase: Upstream Wrapping and Downstream Bending Precede Clamp Opening and Insertion of the Downstream Duplex

*Raashi Sreenivasan[1], *Irina A. Shkel[2,3], *Munish Chhabra[1], Amanda Drennan[2], Sara Heitkamp[2], Hao-Che Wang[1], Malavika A. Sridevi[2], Dylan Plaskon[2], Christina McNerney[2], Katelyn Callies[2], Clare K. Cimperman[2], M. Thomas Record, Jr.[1,2,3]

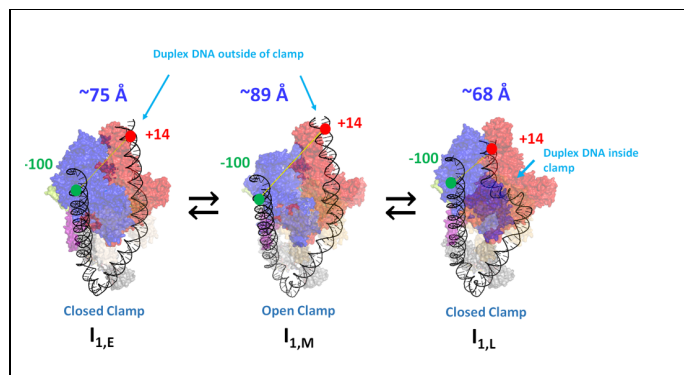Biophysics[1] Program and Departments of Biochemistry[2] and Chemistry[3]

University of Wisconsin – Madison

Madison, WI 53706

* Co-first authors

Abstract

FRET (fluorescence energy transfer) between far-upstream (-100) and downstream (+14) cyanine dyes showed extensive bending/wrapping of λP_R promoter DNA on *E. coli* RNA polymerase (RNAP) in closed and open complexes (CC, OC).

Here we determine the kinetics and mechanism of DNA bending/wrapping by FRET and of formation of RNAP contacts with -100 and +14 DNA by single-dye fluorescence enhancements (PIFE). FRET/PIFE kinetics exhibit two phases: rapidly-reversible steps forming a CC ensemble ({CC}c of four intermediates (initial (RP$_C$), early (I$_{1E}$), mid- (I$_{1M}$), late (I$_{1L}$)), followed by conversion of {CC} to OC via I$_{1L}$. FRET and PIFE are first observed for I$_{1E}$, not RP$_c$. FRET/PIFE together reveal large-scale bending/wrapping of upstream and downstream DNA as RP$_C$ advances to I$_{1E}$, reducing -100/+14 distance to ~75Å and making RNAP-DNA contacts at -100 and +14. We propose that far-upstream DNA wraps on the upper β'-clamp while downstream DNA contacts the top of the β-pincer in I$_{1E}$. Converting I$_{1E}$ to I$_{1M}$ (~1s time-scale) *reduces* FRET efficiency with little change in -100/+14PIFE, interpreted as clamp-opening that moves far-upstream DNA (on β') away from downstream DNA (on β) to increase the -100/+14 distance by ~14Å. FRET increases greatly in converting I$_{1M}$ to I$_{1L}$, indicating bending of downstream duplex DNA into the clamp and clamp-closing to reduce the -100/+14 distance by ~21Å. In the subsequent rate-determining DNA-opening step, in which the clamp may also open, I$_{1L}$ converts to the initial unstable OC (I$_2$). Implications for facilitation of CC-to-OC isomerization by upstream DNA and upstream-binding, DNA-bending transcription activators are discussed.

2

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.[1, 2] Hence the kinetics and equilibria of the steps of OC formation are highly regulated by promoter sequence, transcription factors, ligands, and solution variables. Monitored by abortive initiation and filter binding assays that detect only long-lived OC, the kinetics of OC formation usually are well-described by the two-step minimal Mechanism 1.[1-7]

$$R + P \underset{K_{CC}}{\rightleftarrows} \{CC\} \xrightarrow{k_{isom}} OC \qquad \qquad \text{Mechanism 1}$$

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.[4] Evidence for three likely members of $\{CC\}$ is provided by the different downstream boundaries of chemical and/or enzymatic CC footprints obtained at different promoters. These CC include the initial specific complex (RP_C; downstream footprint boundary at approximately -5)[8-14] and two more-advanced CC (designated I_{1E} and I_{1L} for "early" and "late") with downstream footprint boundaries at approximately +2/+7[15] and at +20.[8, 9, 11, 14] (All positions are numbered relative to the +1 start site.) The isomerization rate constant k_{isom} is the composite rate constant for conversion of $\{CC\}$ to OC via I_{1L}, the most advanced member of $\{CC\}$, and the subsequent DNA opening step.[4]

This interpretation of the parameters of Mechanism 1 differs subtly but significantly from that used originally. Most previous discussions of effects of changes in promoter sequence or length, transcription factors and solution variables on K_{CC} and k_{isom} have not considered changes in the composition of the equilibrated $\{CC\}$ ensemble with these variables. In general, any shift in the proportions of the different members of the equilibrated $\{CC\}$ ensemble will change both K_{CC} and k_{isom} (see ref. 4 and SI Eqs. S19-22). The large effects of upstream truncation of promoter DNA on k_{isom}[16, 17] are well-explained in this way (see ref. 4 and Discussion).

The initial closed complex RP_C is stabilized by some combination of specific interactions of the flexibly-tethered αCTDs with the promoter UP element (a 20 bp region containing A and T tracts immediately upstream of the -35 region)[18] and of regions 2 and 4 of the σ^{70} subunit with the promoter -10 and -35 elements.[19] After RP_C formation, the downstream duplex must bend by at least 90° into the open

clamp in order for 13-14 bp, including the -10 and discriminator regions and the transcription start site, to be opened (melted) by RNAP using binding free energy.[3-5, 20] Bacteriophage T7 RNAP also bends the downstream duplex in the process of DNA opening[21-24].

CryoEM studies indicate that ATP-independent promoter opening by eukaryotic RNA polymerases involves 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° by binding to the TATA element of transcription factors TBP (TATA-binding protein) together with TFIIA, TFIIB, and TAFs (TBP-associated factors).[25-30] After Pol II PIC assembly, cryoEM reveals three closed complexes, designated CC1, CC2, and CC^{dist}, which appear to be on-pathway intermediates in OC formation[31]. In CC1, the downstream promoter is positioned above the closed clamp of RNAP. In CC2 the clamp domain is open and in CC^{dist} the downstream duplex is bent into the open clamp. In the conversion of CC^{dist} to the final OC, the clamp closes and the DNA is opened[31].
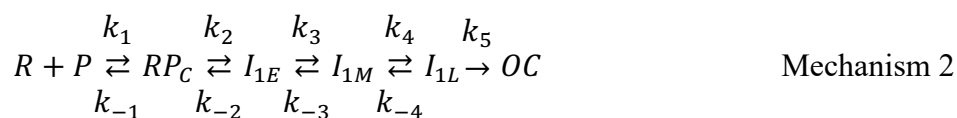
For *E. coli* RNAP, the isomerization of {CC} to OC is greatly facilitated by the presence of far-upstream DNA. For full-length (FL) λP$_R$ promoter DNA, real-time HO footprinting of {CC} revealed interactions with RNAP up to at least position -82.[15] Upstream truncation of λP$_R$ and lacUV5 promoters at positions between -63 (UT-63) and -42 (UT-42) reduces the isomerization rate constant k$_{isom}$ (Mechanism 1) for conversion of {CC} to OC by one to two orders of magnitude.[16, 17] For the UT-47 λP$_R$ truncation variant, DNase footprinting of {CC} early in the time course of OC formation revealed downstream boundaries of partial protection of the template and nontemplate strands of the downstream duplex at +2 and +7, compared to +20 for {CC} at FL λP$_R$.[15] 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 bending and wrapping of FL λP$_R$ upstream DNA in an earlier closed complex (I$_{1E}$) facilitates bending of the downstream duplex into the clamp/cleft in the most advanced member of the {CC} ensemble (I$_{1L}$), in which the duplex is poised to be opened by RNAP.[4, 15] The absence of upstream wrapping in UT-47 was proposed to greatly disfavor conversion of I$_{1E}$ to I$_{1L}$, thereby reducing k$_{isom}$.[4] Tests of these proposals are provided by the FRET and PIFE kinetics studies reported here.

Equilibrium FRET studies on low temperature (2 °C; closed) and higher temperature (19 °C; open) complexes revealed that far-upstream DNA in both complexes is highly bent and wrapped on RNAP.[32]

DNA backbone footprinting of RNAP- λP_R complexes revealed far-upstream protection to approximately -82 ({CC})[15] and -65 (OC)[15, 33], interpreted as upstream bending and wrapping, and AFM compaction measurements indicated extensive upstream wrapping in OC.[34] Upstream modulators of initiation that bend and in some cases wrap DNA like CAP[35], IHF[36-38] and HU [39-41] may therefore affect k_{isom} by altering the distribution of wrapped and unwrapped intermediates in the {CC} ensemble and/or the trajectory of bending and wrapping of upstream DNA on RNAP in these intermediates.

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 Cy3 and Cy5 fluorescence enhancements (PIFE)[42-44] induced by contacts between RNAP and these dyes at the -100 and +14 positions of promoter DNA. The PIFE kinetics complement the FRET kinetics by reporting on the development of these RNAP-promoter contacts in {CC} intermediates of OC formation. PIFE effects using Cy3 at positions near the TSS were previously used to study aspects of the kinetics of OC formation, initiation, and escape.[45-47] 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.[48-50]

Here we use FRET and PIFE in kinetic assays of OC formation to investigate large DNA conformational changes and changes in interactions of upstream and downstream 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.

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

Illustrations of these intermediates, deduced from the research reported here, are provided in the graphical abstract and in Discussion.

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

5

as {CC} advances. 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. The placement of the downstream duplex and the status of the RNAP clamp in the different CC intermediates deduced from the FRET and PIFE studies reported here relate well to cryoEM observations and proposals for pol II[31] and E. coli RNAP intermediates[46, 51]. 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 compared 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 increases in the isomerization rate constant k_{isom} when upstream DNA or upstream-binding transcription factors are present, as well as from changes in promoter sequence, result from changes in the {CC} population distribution to favor the advanced closed complex I_{1L}, in which the start site region of duplex promoter DNA is poised for DNA-opening in the RNAP clamp/cleft, and not primarily from increases in the rate constant of the subsequent DNA-opening step.[4, 52, 53]

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 °C), 0.1 M NaCl, 0.1 mM DTT, and 0.1 mM Na$_2$EDTA. 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$_2$EDTA, 0.05 mg/ml BSA, 0-2% glycerol and 0.02% Tween. Permanganate footprinting buffer (PFB) is 40 mM Tris (pH 8 at 19 °C), 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$_2$EDTA), 0.05% xylene cyanol (w/v) and 0.05% bromophenol blue (w/v).

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.[32, 54] σ^{70} was overexpressed and purified using *E. coli* M5219 transformed with plasmid pMRG8, as described previously[32]. 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. Other experiments were performed with preparations of WT RNAP holoenzyme.[53, 55] Holoenzyme activities, determined from promoter binding assays at high promoter concentration,[6] were in the range 50-90%. No significant differences between different RNAP preparations were observed in experiments reported here.

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.[32] ^{32}P-labeled λP_R promoter DNA fragments for MnO_4^- footprinting experiments were prepared by PCR as described previously.[52, 55] Sequences and lengths of the different primers and DNA constructs are described in the supporting information (Tables S1-2).

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 NaMnO4 for 50 ms, previously determined to be a concentration and time adequate for accurate detection of the rate of OC formation[52]. Samples were expelled, quenched with 500 µl ethanol, and immediately precipitated with ethanol and washed. Modified fragments were cleaved by 1 M piperidine at 90°C. Reactions were evaporated and resuspended in ULB and resolved on 8% acrylamide sequencing gels.

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-detected 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". Previous equilibrium FRET results on OC[32] were obtained at 19°C because the fluorescence of cyanine dyes decreases strongly with increasing temperature[32, 56]. This temperature is used for kinetics experiments as well because 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 OC is sufficiently slow to separate this kinetic phase from the earlier CC phase, and the final OC is sufficiently stable that its formation is irreversible.[5, 57]

Stocks of dye-labeled promoter DNA and RNAP were prepared at twice the desired final concentration in FB. Most experiments were performed at either 1:1 or 0.5:1 mole ratio of RNAP to promoter DNA (50 nM final DNA concentration), in order to avoid a competitive binding mode observed previously on this promoter DNA fragment in RNAP excess[32]. 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.

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 as a function of time. Fluorescence emission signals were monitored using a 565-625 nm band pass filter for Cy3 and a 660 nm long-pass filter for Cy5 (Omega Optical, Brattleboro, VT).   A monochromator excitation slit width of 1.56 mm or 3.14 mm was used. Fluorescence intensity was monitored from 10ms to 400s with 600 data points that typically were 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. FRET data collected from each shot were normalized as described in

supplemental and then averaged to reduce noise at earlier time points for further analysis. A total of 15 FRET experiments (each ~10 shots) were performed with Cy5 acceptor at -100, and another 12 experiments were performed with Cy5 at +14. In PIFE experiments, individual shots were also normalized and averaged. For comparative analysis of PIFE effects at -100 and +14, fluorescence increases relative to the initial (10 ms) signal were determined for each shot and averaged. A total of 34 PIFE experiments were performed: 15 monitoring +14 PIFE (10 with Cy3+14, 5 with Cy5+14) and 19 monitoring -100 PIFE (11 with Cy3-100, 8 with Cy5-100). Different experiments used independent dilutions (and in some cases independent preparations) of RNAP and DNA solutions. Analyses of these data in terms of a sum of exponentials and in terms of the proposed five-step mechanism are described in Supplemental Methods. Curves shown in Figs. 3 and 4 in Results are averages of series of normalized shots obtained in single experiments, selected as high S/N examples but otherwise representative of the full sets of independent experiments.

Salt-Upshift Dissociation Kinetic Assays Monitored By FRET and +14 PIFE

Dye-labeled OC, prepared as described above, were rapidly mixed with KCl (50nM final OC concentration; 0.4 M final KCl concentration) in FB in the stopped-flow fluorimeter. Fluorescence kinetic data collected in each shot were normalized, and results from series of 10 or more shots were averaged as described above. FRET and PIFE curves shown in Results are representative of 4 independent FRET experiments and 5 independent PIFE experiments.

Illustrating the Bending and Wrapping of Promoter DNA in Open Complex Formation

The PYMOL Molecular Graphics System, Version 2.3.2 (Schrödinger, LLC) was used to replace part or all of the DNA in the crystal structure of a transcription initiation complex (4YLN) by segments of DNA duplex (-100 to +14) to illustrate a plausible path for bending and wrapping of upstream DNA and bending of downstream DNA in the closed intermediates in OC formation as well as the final OC. These proposals for the approximate location of upstream and downstream duplex DNA in the bent-wrapped CC intermediates were developed to be semi-quantitatively consistent with the FRET distances and PIFE contacts determined in this research.

9

Results

Fast Permanganate Footprinting Kinetics of OC Formation

To test whether opening of the λP_R 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. Results are shown in Fig. 1. 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[13] 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.[52] Although MnO_4^- reaction kinetics are moderately slow, requiring a reaction time of 150 ms even at high
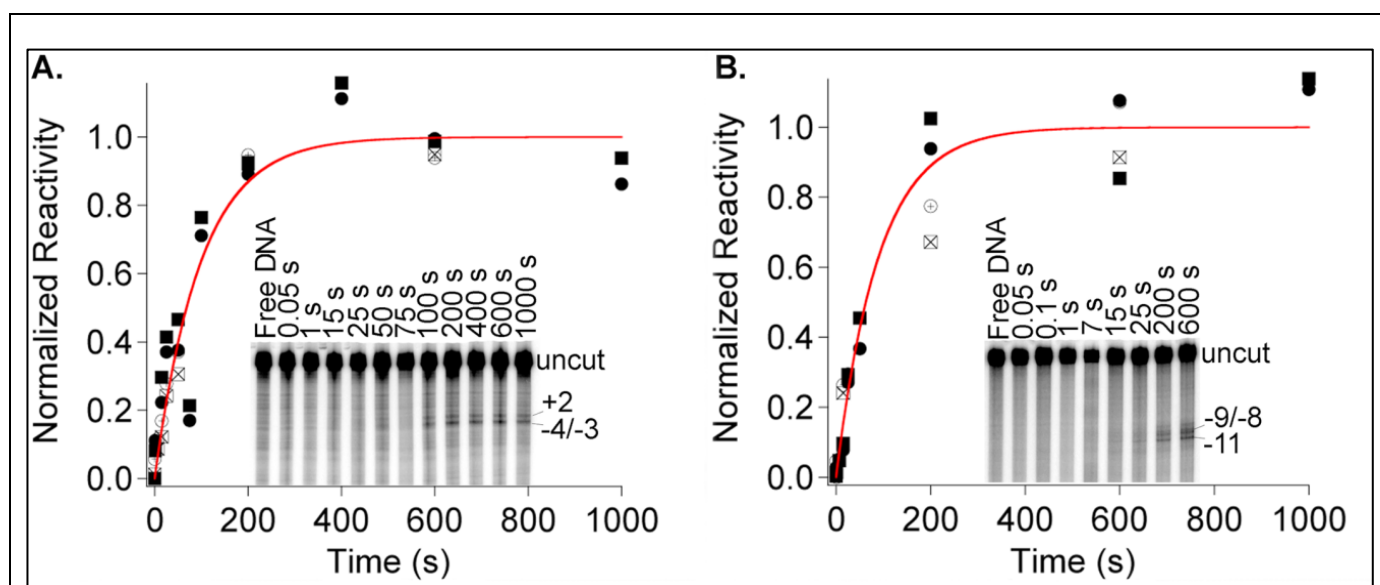

Figure 1. Kinetics of OC Formation Monitored by MnO_4^- Reactivity. Fast (0.15 s) MnO_4^- snapshots[52] monitor the time course of opening individual thymines in OC formation after mixing excess RNAP (55 nM) with λP_R promoter DNA (0.3 nM) at 19 °C. Representative gels are shown as insets. Kinetics of development of MnO_4^- reactivity are plotted for thymines on the nt strand (+2 ($\oplus$, ●) and -4/-3 ($\boxtimes$, ■) in panel **A** and for the t strand (-9/-8 ($\oplus$, ●) and -11 ($\boxtimes$, ■) in panel **B**. Rate constants k_{obs} for OC formation from these fits are the same within uncertainty (template strand $k_{obs} = 0.010 \pm 0.002$ s^{-1}; non-template strand $k_{obs} = 0.011 \pm 0.002$ 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.

MnO_4^- concentration, this is fast relative to the time scale of the CC-to-OC isomerization at 19°C ($1/k_{isom}$ ~ 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 nonpromoter) complexes that were present.

Gel lanes in the insets in Fig. 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, these MnO_4^- snapshots detect no reactive (open) thymines (Fig S1), indicating that only CC complexes are present. 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 for each strand (Figs. 1, S1) yield values of k_{obs}, the observed first order rate constant for the formation of MnO_4^--detected complexes at 55 nM excess RNAP, which are the same within the uncertainty (10-20%) for all thymines detected on both template and nontemplate strands.

Rate constants k_{obs} obtained from these MnO_4^- footprinting assays at 19 °C are compared with those obtained previously by filter binding[5] at 20 °C in a plot vs [RNAP] in Figure 2A, and were fit to the expected hyperbolic functional form:[1-7]

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

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} and k_{isom} obtained from this analysis (see Fig. 2 caption) agree within the uncertainty with those reported previously.[5]

Predicted Time Evolution of CC and OC Populations at RNAP and Promoter Concentrations of FRET and PIFE Experiments

The kinetic results (K_{CC}, k_{isom}) from the analysis in Fig. 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 very different conditions of the fluorescence experiments (1:1 or limiting (e.g. 0.5:1) RNAP; high (50 nM) DNA concentrations). The extent of conversion of P to {CC} in the rapidly

11

reversible first step of Mechanism 1 is predicted by K_{CC} and the kinetics of conversion of this rapidly-reversible mixture of CC and P to OC is predicted by k_{isom}. These predictions, given in Fig. 2B for 50 nM RNAP and promoter DNA on a logarithmic time scale, show that significant OC formation does not occur in the fluorescence kinetics experiments 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 formation of the ensemble of closed complex intermediates ({CC}) from free promoter DNA. Because the kinetics of OC formation in filter binding and MnO_4^- reactivity assays that detect only OC are single-exponential in excess RNAP, formation of the {CC} ensemble (Mechanism 1) and hence the individual steps of {CC} formation (Mechanism 2) are

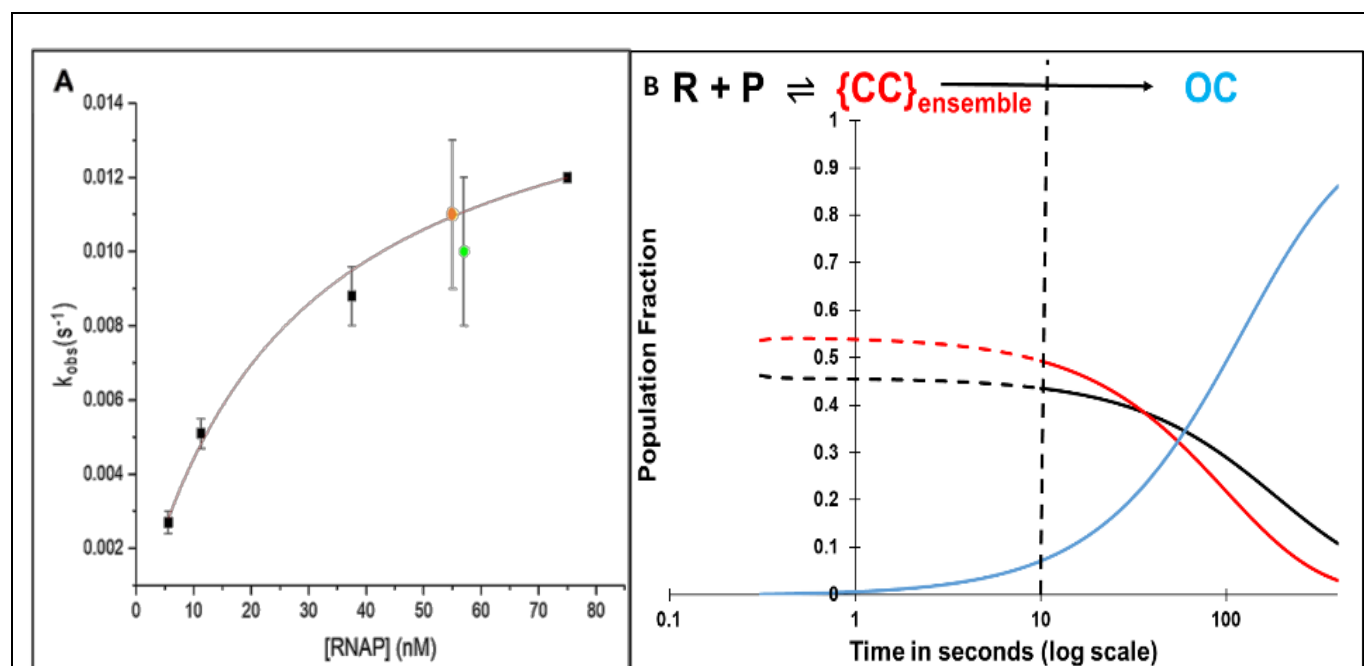**Figure 2. Panel A: Dependence of Rate Constant k_{obs} for λP_R OC Formation on RNAP Concentration in RNAP Excess.** First order rate constants k_{obs} (Eq. 1) for OC formation in excess RNAP are plotted as a function of RNAP concentration. ■ : k_{obs} values from filter binding assays at 20°C.[4] ●, ● : k_{obs} values from MnO_4^- footprinting of template and non-template strands, respectively at 19 °C (Fig 1). A fit of these data to hyperbolic Eq. 1 gives a composite {CC} binding constant $K_{\{CC\}} = (5 \pm 1) \times 10^7$ M^{-1} and an isomerization rate constant $k_{isom} = 0.014 \pm 0.003$ s^{-1} for conversion of {CC} to OC. **Panel B: Simulated Time Evolution of {CC} and OC Formation for FRET/PIFE Conditions (50 nM λP_R Promoter and RNAP).** Fractional populations of open complexes (OC) ——, closed complexes {CC} ——, and free promoter DNA ——, predicted from K_{CC} and k_{isom} as a function of time (log scale) assuming that the {CC} population equilibrates with free promoter DNA by 10 s (▬ ▬ ▬ and ▬ ▬ ▬) . The dashed vertical line at 10 s marks the onset of OC formation from the equilibrium mixture of {CC} and free promoter DNA determined from this simulation.

rapidly reversible[58]. Equilibrium between {CC} and P is therefore established in the first 10 s of the reaction, quantified by the equilibrium constant K_{CC}.

At 10 s, Fig. 2B predicts that approximately half of total promoter DNA is bound in closed complexes ({CC}) with RNAP while the other half is free (unbound) P. This {CC} ensemble (and free 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

FRET-Detected Kinetics of DNA Bending and Wrapping in Closed and Open Complex Formation
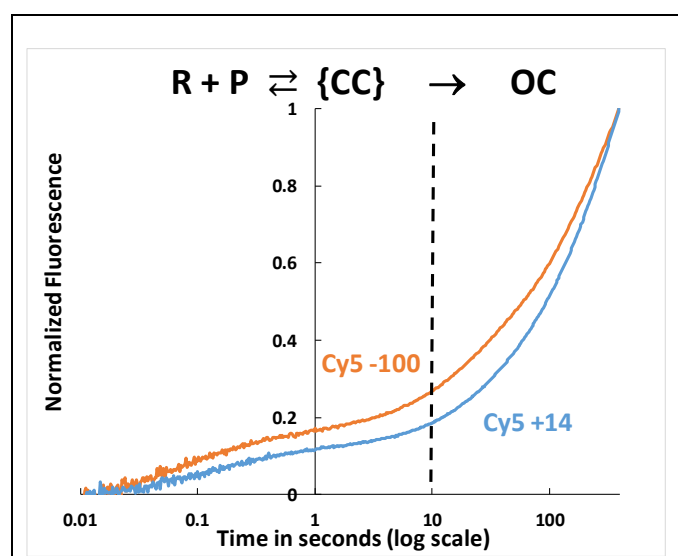
Previous FRET studies of equilibrium populations of closed (2°C) and open (19°C) complexes of RNAP and doubly-labeled (Cy3 and Cy5) λP_R promoter DNA revealed that upstream and downstream DNA regions are highly bent and upstream DNA is wrapped on RNAP in the low-temperature (2 °C) population of closed complexes and in the 19 °C open complex, 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 19° for Cy3(-100)Cy5(+14) and Cy3(+14)Cy5(-100). Representative FRET acceptor time-courses for both promoter fragments, plotted on a logarithmic time scale, are shown in Fig. 3.

Figure 3. FRET-detected Bending and Wrapping of Promoter DNA by RNAP as {CC} Ensemble Advances and Forms the Stable OC. Time course (log scale) of normalized FRET acceptor (Cy5) emission intensity after mixing Cy3-Cy5 dye-labelled λP_R DNA (50 nM final) with *E. coli* RNAP (50 nM final) at 19°C and exciting FRET donor (Cy3) at 515nm. Both dye orientations are shown: Cy5(+14) Cy3(-100) ——, Cy3(+14) Cy5(-100) ——. The vertical dashed line at 10 s corresponds to the onset of OC formation at these conditions.

The FRET acceptor kinetics span a time range of more than four orders of magnitude, from ~20 ms to at least 400 s. As predicted in Fig. 2B, two kinetic phases are observed. The faster phase, accounting for approximately 15% of the total FRET

increase, occurs in the first 10 s after mixing, and the slower second phase develops from 10 s to 400 s. By comparison with the simulations of Fig. 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 been converted to the {CC} ensemble, is only ~15% of the long-time (400 s) value, at which ~80% of promoter DNA is in an OC (Fig 2B). This indicates that the {CC} ensemble is on average less bent/wrapped, with a larger dye-dye distance, than the final OC. The two high S/N experiments plotted in Fig. 3 are otherwise representative of the full set of 27 FRET experiments. In the time range from 0.5 s to 200 s the normalized FRET signal for Cy5+14 is generally smaller than for Cy5-100.

RNAP-Induced Fluorescence Enhancements (PIFE) at Both -100 and +14 Exhibit Similar Rates to FRET Acceptor Increases but Decrease Late in {CC}. Phase and in OC Formation

Single-dye kinetics experiments (Fig. 4) show large increases in Cy3 and Cy5 fluorescence (PIFE effects) in the time range 10 ms to 1 s at both -100 and +14 positions of λP_R promoter DNA. -100 and +14 PIFE signals in Fig. 4 both develop on the same time scale as the FRET acceptor signal (t > 10 ms; Fig. 3). These +14 and -100 PIFE effects provide kinetic information regarding the involvement of downstream and upstream DNA in steps of OC formation. At each position, Fig. 4 shows that normalized PIFE effects of Cy3 and Cy5 are similar. PIFE effects for both dyes at both positions exhibit two kinetic phases, as in FRET assays (Fig. 3). In all four cases, PIFE effects are first detectable approximately 10-20 ms after mixing and increase to a maximum near the end of the {CC} phase. For +14 PIFE, a broad maximum in the time range 2 – 20 s is observed (Fig. 4A). For -100 PIFE the maximum is sharper and occurs earlier (1-3 s). For t > 3 s (-100 PIFE) and t > 20 s (+14 PIFE). PIFE signals decrease as the {CC} ensemble is completed and conversion to the stable OC begins, indicating that intrinsic -100 and +14 PIFE signal intensities of {CC} intermediates like I_{1E} and I_{1M} exceed those of the stable OC (see Discussion).

Position +14 is ~25 bp downstream of the footprint boundary for RP_c (typically -5) but within the +20 downstream boundary of the footprint observed for more advanced CC and OC.[3, 4] Position -100 is 20-40 bp upstream of footprint boundaries of all RNAP-λP_R promoter CC and OC investigated. Nevertheless, the PIFE results demonstrate that RNAP must contact[44] these dye positions on promoter DNA in the step(s) that advance the {CC} ensemble by bending and wrapping upstream and downstream promoter DNA on RNAP. The Discussion considers the implications of these contacts for the operation of

the RNAP-promoter biophysical machinery that uses binding free energy to bend the downstream duplex onto the top of the β pincer and then into the open clamp, before clamp-closing and opening of 13 base pairs including the transcription start site.
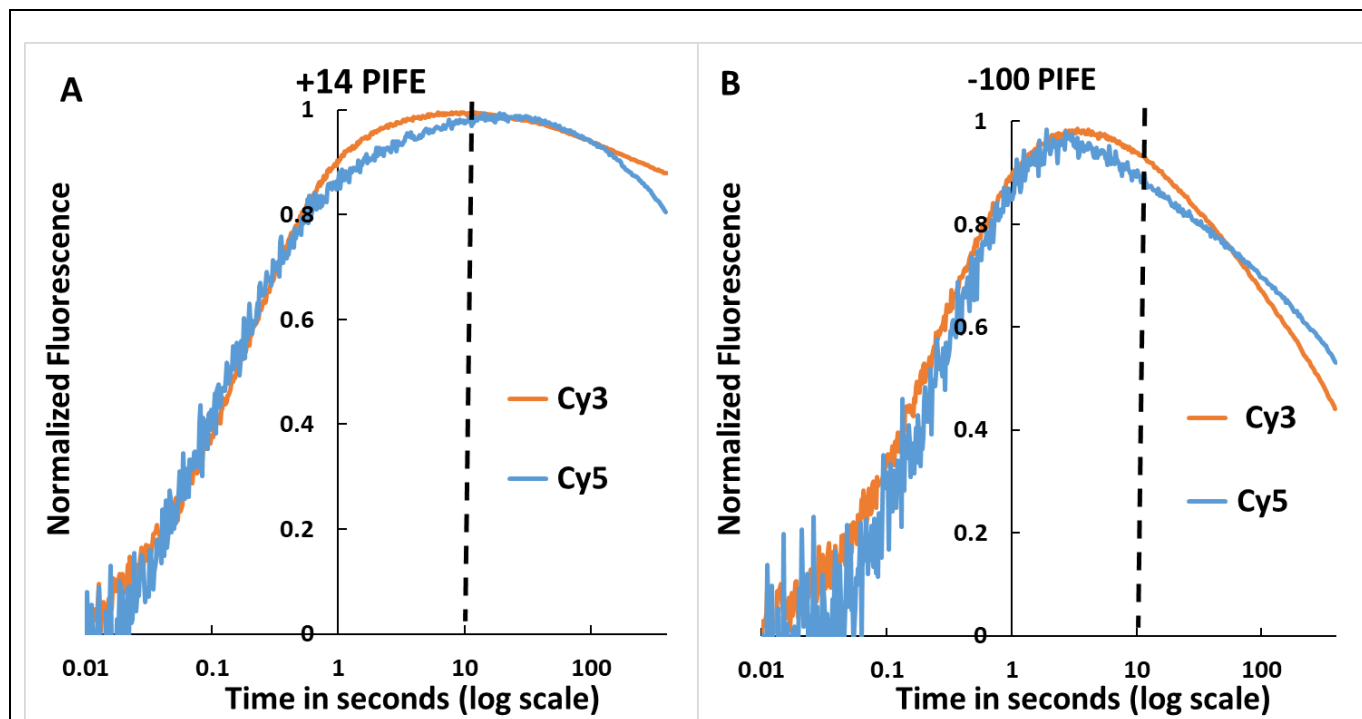

Figure 4. OC Formation Monitored by Single-Dye Cy3, Cy5 Fluorescence (PIFE) at Downstream (+14) and Far-Upstream (-100) Positions of λP$_R$ Promoter DNA. Representative time courses (log scale) of normalized single-dye fluorescence (PIFE) of Cy3- or Cy5-labeled λP$_R$ DNA after mixing with RNAP (final concentrations 50 nM RNAP and λP$_R$ DNA) at 19 °C. +14 PIFE (**Panel A)** and -100 PIFE (**Panel B**) from Cy3 —— and Cy5 ——. Cy3 was excited at 515 nm and Cy5 at 610 nm. The vertical dashed line at 10 s is the predicted onset of OC formation (see Fig. 2B).

Unwrapping in Dissociation of OC by a Salt-Upshift

Rapid salt upshifts were used to investigate unwrapping in OC dissociation. Mixing of a stable RNAP-λP$_R$ promoter OC with high KCl or urea rapidly destabilizes it, forming the unstable but [KCl]-insensitive open-promoter intermediate I$_2$, which undergoes DNA closing on a 1 s time scale[52, 53, 59, 60]. Upshifts of doubly- and singly-labeled OC to 0.4 M KCl at 19°C result in rapid reductions in Cy5+14 acceptor FRET and Cy3 +14 PIFE. In Fig. 5 the kinetics of these fluorescence changes are compared with one another and with simulations of the two-step irreversible conversion of the stable OC at 19 °C to closed promoter complexes via the unstable open-promoter intermediate I$_2$ and the DNA-closing step [52, 53, 59, 60]. Figure 5 shows that the FRET and PIFE fluorescence changes occur prior to the DNA closing

step of high-salt-induced dissociation. The FRET decrease occurs faster than the reduction in +14 PIFE. Both are slower than the conversion of RP_o to I_2 but faster than the conversion of I_2 to closed-promoter complexes. The finding that the FRET decrease occurs faster than the reduction in +14 PIFE indicates that KCl-driven unwrapping of upstream DNA occurs before the release of downstream contacts. This sequence of steps in response to a large KCl-upshift differs from the mechanism of OC formation at lower KCl, where upstream DNA wrapping and formation of +14 contacts occur together as the {CC} ensemble advances.
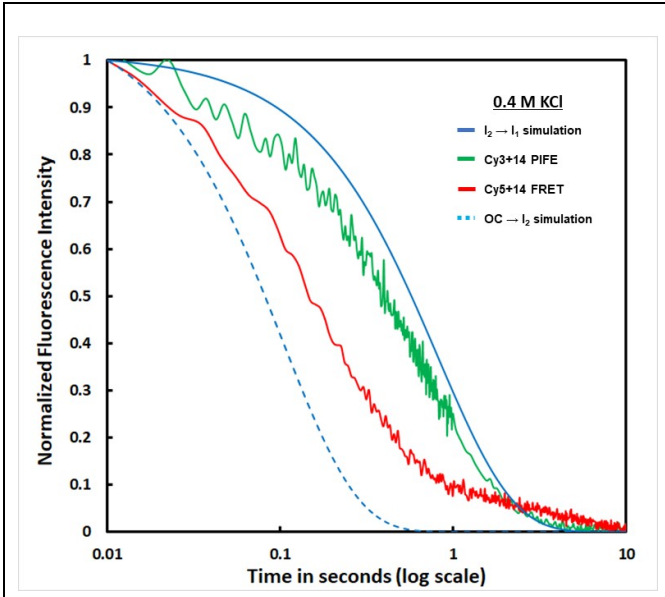

Figure 5. Unwrapping (detected by FRET) and Release of Downstream Contacts (detected by +14 PIFE) in OC Dissociation after an Upshift to 0.4 M KCl at 19 °C. Representative time courses (log scale) of reductions in Cy5+14 acceptor FRET —— and Cy3+14 PIFE —— after destabilizing the 19 °C OC with 0.4 M KCl. Blue curves are simulations based on 0.4 M KCl rate constants, interpolated to 19 °C from results at 10 °C and 37 °C.[59] These predict the time course of conversion of the initial OC to the unstable open intermediate I_2 (OC → I_2, rate constant k_{-3} = ~9.7 s^{-1}; – –) and subsequent DNA-closing, designated as I_2 → I_1 with rate constant k_{-2} = ~1.2 s^{-1}; ——.

Analysis and Discussion

A Five-Step Mechanism of OC Formation with a Key Intermediate CC (I_{1M}) Not Previously Observed by DNA Footprinting

OC formation, studied in RNAP excess by MnO_4^- footprinting and filter binding assays that only detect the stable OC (e.g. Fig. 1), occurs over two decades in time and exhibits single-exponential kinetics (Fig. S1). The RNAP concentration dependence of the kinetics is well described by two-step Mechanism 1 (parameters K_{CC}, k_{isom}). The single-exponential kinetics of OC formation in these experiments show that formation of the {CC} ensemble equilibrates rapidly (characterized by K_{CC}) on the time scale of the conversion of {CC} to OC (characterized by k_{isom}). Because of {CC} rapid equilibrium, no information is obtained in these kinetic studies about the number of CC intermediates or the kinetics of their interconversions.

On the other hand, four exponentials with rate constants $1/\tau_{i,obs}$ that span four decades (~ 15 s^{-1} to ~ 0.015 s^{-1}; Table S3) are needed to fit the FRET (Fig. 3) and PIFE (Fig. 4) kinetic data for OC formation,

as described in SI. This indicates the presence of significant populations of at least three intermediates in the {CC} ensemble with FRET and/or PIFE fluorescence signals which differ from free promoter DNA, from one another and from the final OC. Because the {CC} ensemble is in rapid equilibrium with reactants on the time scale of its isomerization to OC, each step of advancing this ensemble also rapidly equilibrates on the time scale of the next forward step (i.e. $k_{-1} > k_2$, $k_{-2} > k_3$, etc.). Consequently, the three rate constants $1/\tau_{1,obs}$, $1/\tau_{2,obs}$ and $1/\tau_{3,obs}$ in the four-exponential fit can be interpreted as decay-to equilibrium rate constants[61] for formation of three observable CC (designated I_{1E}, I_{1M}, I_{1L}; see SI Eqs. S3-S7). We deduce that rapid reversible formation of the initial specific CC (RP_C) from reactants is not directly detected by FRET, nor by -100 or +14 PIFE, because promoter DNA is not sufficiently bent in RP_C to give FRET and because contacts between RNAP and the promoter in RP_C do not extend to -100 or +14. The final (fourth) exponential is interpreted as irreversible conversion of the most advanced closed intermediate (I_{1L}) to the stable 19 °C OC in several steps, which are not separable because they follow the rate-determining DNA opening step ($I_{1L} \rightarrow I_2$, with rate constant k_5).

Given the need for four exponentials to fit these data, it is not surprising that a five-step mechanism (Mechanism 2) including RP_C, I_{1E}, I_{1M}, and I_{1L} is needed to obtain good fits to all the FRET acceptor and PIFE kinetic data using a common set of rate constants (see below). FRET and +14PIFE kinetic data can be adequately fit using a four step mechanism, but significantly different sets of rate constants are obtained for FRET and PIFE datasets. -100 PIFE kinetic data are not adequately fit by a four step mechanism.

Fits of the 5-step mechanism to the FRET and PIFE kinetic data of Figs. 3 and 4 are shown in Panels A-C of Fig. 6, using the best-fit amplitudes and rate constants for these individual data sets listed in Table S4. Clearly Mechanism 2 fits the experimental data with only minor deviations (Table S4). For comparison, Panels A-C of Fig. S2 show fits to the same FRET and PIFE time-courses using average FRET and PIFE signal amplitudes (Table S5) and average rate constants and obtained from fitting all 61 data sets to Mechanism 2. Use of these average amplitudes and rate constants also provides good fits with approximately two-fold larger deviations (Table S5).

The following sections discuss the average FRET acceptor and PIFE amplitudes of the intermediate CC and OC in 5-step Mechanism 2, their implications for the large-scale conformational changes in promoter DNA and RNAP in these steps, the average rate and equilibrium constants of

these steps, the simulated time-course of OC formation and the free energy vs progress diagram for this process.

FRET Evidence that Promoter DNA is Wrapped on RNAP in All Three I_1 Intermediates and that Far-Upstream (-100) and Downstream (+14) DNA are Farther Apart in I_{1M} than in I_{1E}, I_{1L}

Table 1 lists average Cy5 FRET acceptor signal amplitudes of {CC} intermediates for both Cy3(-100) Cy5(+14) and Cy3(+14)Cy5(-100), calculated from the average FRET acceptor fitting amplitudes in Table S5 and expressed relative to OC. Cy3 (FRET donor) fluorescence is dominated by PIFE and provides no FRET information. No Cy5+14 or Cy5-100 FRET is observed for the initial closed complex RP_C, showing that the distance between -100 and +14 exceeds the FRET detection limit (~100 Å). The three more advanced I_1 intermediates in the {CC} ensemble all exhibit significant FRET between -100 and +14, indicating that these positions, more than 300 Å apart in free promoter DNA, are less than ~100 Å apart in these intermediates. FRET amplitudes of the I_1
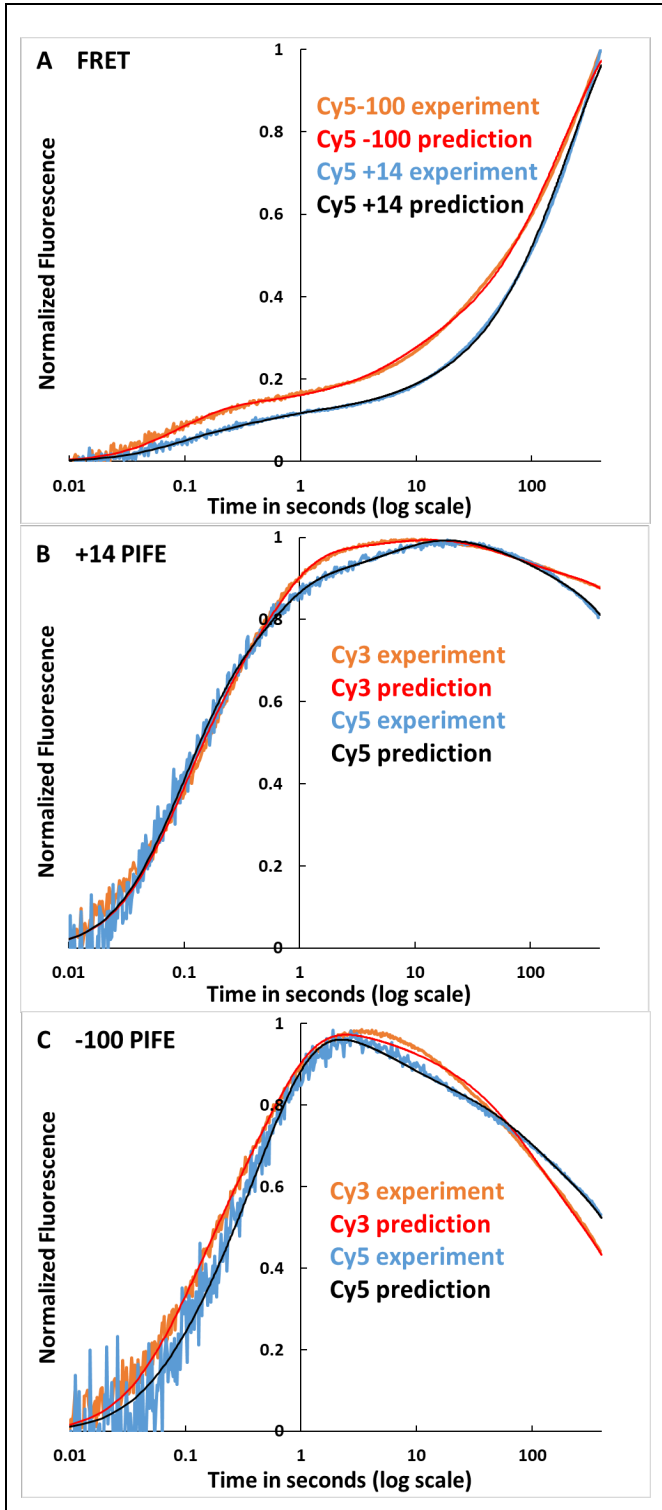
Figure 6. Comparison of Predicted and Observed FRET and PIFE Kinetics of OC Formation at λP_R Promoter. Panel A: FRET data (Fig. 3) for Cy5-100 —— and Cy5+14 ——. Panel B: +14 PIFE data (Fig 4A) for Cy3 —— and Cy5 ——. Panel C: -100 PIFE data (Fig. 4B) for Cy3 —— and Cy5 ——. Predictions use rate constants and amplitudes in Table S4, obtained from fitting these data sets to Eq. S7 for Mechanism 2. For comparison, fits in SI Fig. S2 use average rate constants (Table 3) and amplitudes (Table S5) for the entire data sets.

18

intermediates range from ~10-80% that of OC (Table 1). These observations are consistent with previous equilibrium FRET results which indicated that promoter DNA in the 2 °C {CC} ensemble is highly wrapped on RNAP, with a somewhat smaller FRET efficiency (i.e. larger -100/+14 distance) than the 19 °C OC[32].

Table 1. Relative FRET and PIFE Fluorescence Signals of RNAP-Promoter Complexes

	FRET Acceptor Amplitude (Between -100 and +14, Relative to OC)[a]		+14 PIFE Intensity (Relative to P)[b]		-100 PIFE Intensity (Relative to P)[b]	
	Cy5+14	Cy5-100	Cy3	Cy5	Cy3	Cy5
RP_c	0 ± 0.1	0 ± 0.1	1 ± 0.1	1 ± 0.1	1 ± 0.1	1 ± 0.1
I_{1E}	0.34 ± 0.13	0.56 ± 0.26	1.66 ± 0.32	1.15 ± 0.08	1.20 ± 0.11	1.06 ± 0.03
I_{1M}	0.13 ± 0.09	0.24 ± 0.13	1.53 ± 0.25	1.11 ± 0.06	1.19 ± 0.11	1.11 ± 0.04
I_{1L}	0.71 ± 0.24	0.79 ± 0.15	1.42 ± 0.16	1.12 ± 0.06	1.04 ± 0.04	1.02 ± 0.02
OC	1	1	1.27 ± 0.10	1.07 ± 0.03	1.04 ± 0.04	1.02 ± 0.02

[a] Average values of A_i/A_{OC} (Eq. S7) from fits of FRET kinetic data to Mechanism 2
[b] Average values of relative PIFE intensities ($\frac{a_i}{a_P}$), calculated from fits of PIFE kinetic data to Mechanism 2 using Eq. S17.

These average FRET acceptor signal amplitudes of {CC} intermediates are analyzed to obtain relative FRET amplitudes, FRET efficiencies and distances between dyes at -100 and +14 in Table 2, using the previously-determined OC FRET efficiency (0.32).[32] FRET efficiencies of I_{1E} , I_{1M} and I_{1L} are ~45%, ~18% and ~75% that of OC. Use of the previously-determined -100/+14 OC distance (63 Å)[32] as a reference yields a -100/+14 distance for I_{1E} of ~75 Å. The -100/+14 distance increases by ~14 Å in conversion of I_{1E} to I_{1M} before decreasing by ~21 Å in conversion of I_{1M} to I_{1L}. The -100/ +14 distance in I_{1L} (~68 Å) is ~5 Å greater than in the stable OC. Details of this analysis are provided in SI.

Although the uncertainties in these distances (Tables S5, 1, 2) are appreciable, comparable in some cases to the differences between the different I_1 intermediates, the trend is unambiguous. These uncertainties result from the imperfect reproducibility of the data for each position of the Cy5 acceptor and from the

marginally-significant difference in FRET signals for the two placements of the dye probes (Fig. 3). Separate analyses of Cy5-100 and Cy5+14 FRET acceptor data sets (Tables S4, S5) reveal the same progression of FRET fitting amplitudes for these intermediates, with I_{1M} exhibiting the smallest amplitude and therefore the greatest distance between -100 and +14 probe positions. The conversion of higher-FRET I_{1E} to low-FRET I_{1M} is the origin of the inflections in the Cy5+14 and Cy5-100 FRET acceptor time courses at about 1 s (Fig. 3).

Table 2. FRET-Determined Distances between -100 and +14 Probe Positions on λP_R Promoter DNA in CC Intermediates			
Species	**FRET Acceptor Relative Amplitude[a]**	**FRET Efficiency**	**Distance between -100 and +14**
I_{1E}	0.45 ± 0.27	0.14 ± 0.08	75 ± 9 Å
I_{1M}	0.18 ± 0.16	0.059 ± 0.052	89 ± 14 Å
I_{1L}	0.75 ± 0.30	0.24 ± 0.09	68 ± 6 Å
OC	1	0.32 [b]	63 Å [b]
[a] **Average of Cy5-100 and Cy5+14 results $\pm$ 1 SD (Table 1) expressed relative to OC.** [b] **From Ref. 32. FRET efficiencies and dye-dye distances for I_{1E}, I_{1M} and I_{1L} are calculated by comparison to this OC value.**			

These large changes in distance between -100 and +14 for the different intermediates, together with PIFE contact information at these positions, form the basis for a structural mechanism of OC formation, proposed below. They also provide a structural explanation of the very-large facilitation of isomerization in OC formation by upstream wrapping.

PIFE Indicates that RNAP-Promoter Contacts at +14 and Especially at -100, Absent in RP_C, Are Stronger in I_1 Intermediates than In OC

Table 1 also summarizes the RNAP-induced fluorescence enhancements (PIFE) determined for the different CC intermediates and OC, expressed relative to free promoter DNA (P). These are calculated as described in SI (Eqs. S8-17) from the average PIFE fitting amplitudes (Table S5), which span a ~10-fold range. Several general trends are clear from Table 1. Of most significance, PIFE intensities at -100

and +14 are larger for I_{1E} and I_{1M} intermediates than for OC for both Cy3 and Cy5, while RP_C exhibits no detectable PIFE intensity. The -100 PIFE intensity of I_{1L} is small, comparable to the -100 PIFE intensity of OC, while the +14 PIFE intensity of I_{1L} is larger, comparable to the -100 PIFE intensities of I_{1E} and I_{1M}. These rank orders of PIFE intensities indicate that relatively strong contacts[44] of RNAP with both far-upstream (-100) and downstream (+14) DNA are made in converting RP_C to I_{1E}, where -100/+14 FRET is also first observed. *These relatively strong contacts persist in I_{1M}, demonstrating that the decrease in FRET in converting I_{1E} to I_{1M} does not result from partial unwrapping of upstream or downstream DNA.*

Contacts of RNAP with far upstream (-100) DNA appear to be largely disrupted in conversion of I_{1M} to I_{1L} and OC, while contacts of RNAP with downstream (+14) DNA appear similar in all three I_1 intermediates and surprisingly are stronger in these intermediates than in the stable OC. In other trends, contacts of RNAP with +14 appear stronger than with -100, since the PIFE signal of each dye in each complex (I_1 intermediates, stable OC) is stronger at +14 than at -100. Also, the effect of RNAP-DNA interaction on dye fluorescence is dye-specific, with larger Cy3 PIFE intensities than Cy5 PIFE intensities at each position.

FRET/PIFE Insights into the Mechanism of Operation of the RNAP-Promoter Machine in OC Formation

a) Specific Contacts in RP_C Direct RNAP to Bend and Wrap Both Upstream and Downstream DNA in One Step to Form High-FRET/PIFE I_{1E}

The development of -100/+14 FRET and -100 and +14 PIFE in conversion of RP_C to I_{1E} demonstrates that both upstream and downstream promoter DNA are bent and wrapped on RNAP in this step, making contacts between RNAP and both -100 and +14 positions of promoter DNA and reducing the -100/+14 distance to ~75 Å (Table 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, resulting from or accompanied by contraction/folding of the flexible tethers linking α-CTD to α-NTD, as discussed below. Upstream bending allows far-upstream wrapping,

which, as proposed previously, allows the -100 position of the promoter to be contacted by a mobile downstream element (DME) or other region at the downstream end of the RNAP clamp[3, 15]. If only the upstream DNA but not the downstream DNA were bent and wrapped, the distance between -100 and +14 probe positions is estimated to be >130 Å, which is too large to exhibit FRET. Therefore bending and wrapping of upstream DNA must be accompanied by bending of the downstream duplex, presumably at the upstream end of the -10 region,[5, 19, 20, 51, 62] reducing the distance between +14 and -100 positions on
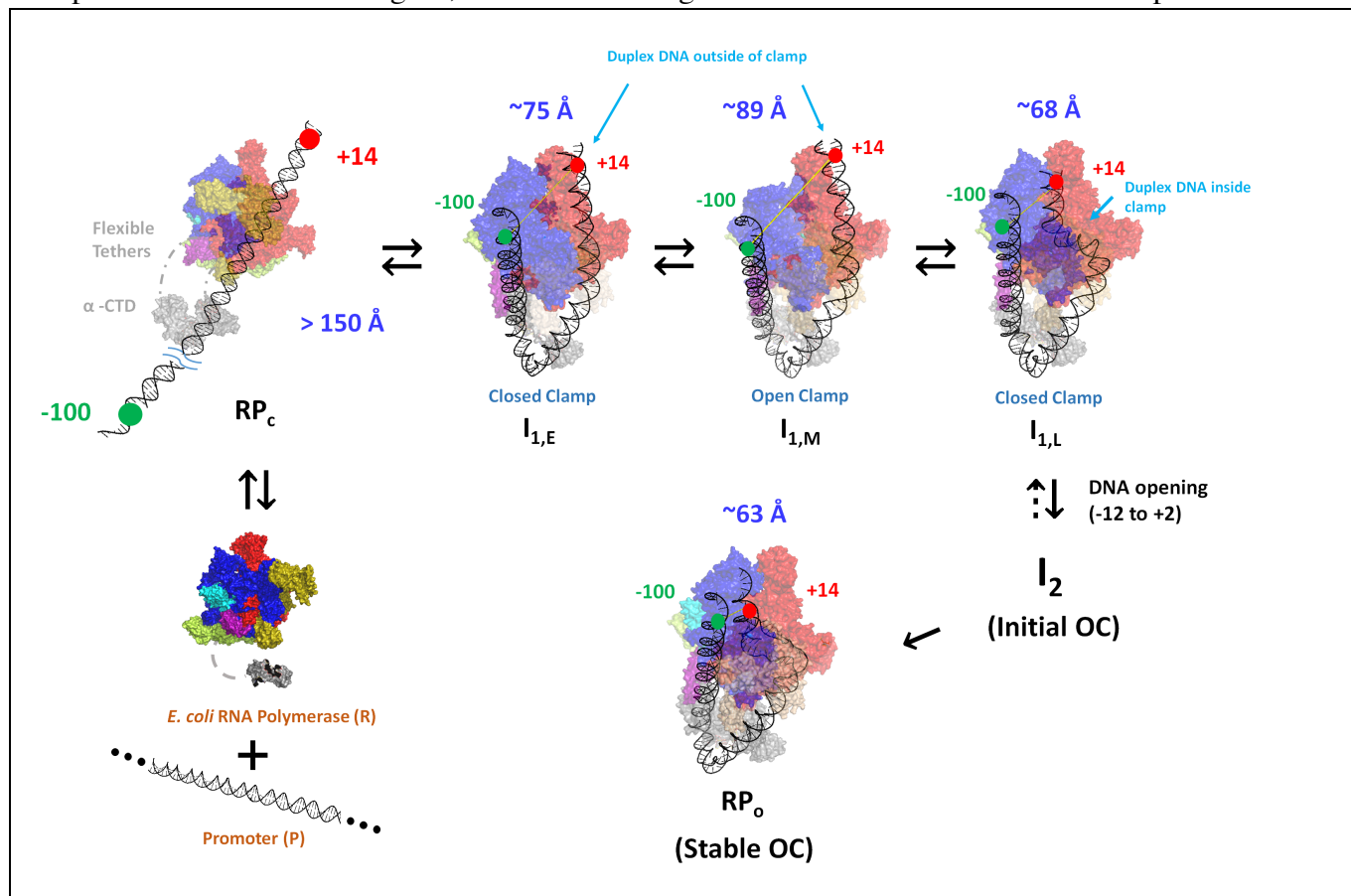

Figure 7. Illustrations of Bending and Wrapping of Promoter DNA in CC Intermediates and OC from FRET Distances and PIFE Contacts. Proposed mechanism of OC formation by RNAP determined by FRET and PIFE probes at -100 and +14 on promoter DNA. Unbound reactants (unbent promoter DNA, free RNAP) are at lower left. The absence of FRET and PIFE in the initial CC (RP_C; upper left) shows that promoter DNA is not yet bent and wrapped on RNAP. Promoter DNA in subsequent CC intermediates (I_{1E}, I_{1M}, I_{1L}) and in OC is highly bent and wrapped, making RNAP-DNA contacts at both -100 and +14. To explain the FRET distances and PIFE effects, we propose that concerted upstream bending/wrapping of upstream duplex DNA around the α subunits and onto the upper β' clamp and bending of downstream duplex DNA onto the top of the β clamp in I_{1E} (top left-center) trigger clamp opening to form I_{1M} (top right-center). Clamp-opening triggers descent of the downstream duplex into the clamp and clamp-closing to form I_{1L} (top right), which opens the initiation bubble to form the initial unstable OC (I_2). Stabilization of I_2 by binding of RNAP mobile elements to the downstream duplex yields the stable OC (bottom right).

promoter DNA to ~75 Å to give the FRET efficiency for I_{1E} calculated in Table 2. This downstream bending is not present in RP_C, because downstream footprints of RP_C complexes end at -5 while that of I_{1E} extends to at least +2/+7.[15] Hence the interactions of σ^{70} region 2 with the -10 region in RP_C are not sufficient by themselves to induce the large-scale bending of the downstream duplex observed in I_{1E}. If other interactions of the downstream duplex with RNAP are involved in the conversion of RP_C to I_{1E}, these might be with a RNAP element like the βSI1 sequence insertion or the NCD of σ^{70}.

A plausible illustration of I_{1E} is provided in Fig. 7. A -100/+14 dye-dye distance of ~75 Å is obtained by bending the upstream duplex around the α subunits, wrapping it on the outside of the β' clamp of RNAP, and bending the downstream duplex at the upstream end of the -10 region to bring +14 in contact with the top of the β pincer and sequence insertion βSI1. This model is similar in its placement of the downstream duplex to the CC1 pol II intermediate observed by cryoEM.[31] A recent cryoEM study of allosteric regulation of RNAP by TraR finds that this protein also interacts with the top of the β pincer and sequence insertion βSI1[63], which should affect formation of I_{1E} from RP_C and could also affect conversion of I_{1M} to I_{1L}.

b) RNAP Clamp-Opening Moves Wrapped Upstream DNA Away from Downstream DNA to Form Low-FRET, High-PIFE I_{1M}

The FRET efficiency of I_{1M} is less than that of I_{1E} (Table 2) while -100 and +14 PIFE signals of I_{1M} are similar to those of I_{1E} (Table 1). Analysis reveals that the conformational change that converts I_{1E} to I_{1M} increases the distance between the probes at -100 and +14 by ~14 Å (Table 2) without diminishing the PIFE contact of these positions with RNAP (Table 1). This increase in -100/+14 distance is readily explained if this step involves opening of the β' clamp. Movement of the β' clamp away from β in opening increases the distance between the tips of β and β' pincers by ~14 Å[64, 65]. The time scale of clamp opening in the absence of promoter DNA (1 s)[64, 65] is similar to that of the I_{1E} to I_{1M} conversion (Table 3). Fig. 7 therefore proposes that the far-upstream DNA is wrapped high on the outside of the upper portion of the β' clamp, and that in clamp opening the -100 DNA moves ~14 Å away from +14 DNA on the top of the β subunit. This movement of the clamp presumably distorts the -10 region of the promoter DNA greatly.

c) The Downstream Duplex Is Bent into the Open Clamp, which Closes to Form the High-FRET, Most Advanced CC, I_{1L}.   The FRET efficiency of I_{1L} is greater than that of I_{1E} and much greater than that of I_{1M} (Table 2), indicating that the upstream DNA has moved much closer to the downstream DNA in I_{1L}. We interpret this as a very large-scale set of conformation changes in which the downstream duplex is bent into the open clamp, followed by closing of the clamp (Fig 7). The -100/+14 distance is ~7 Å less than in I_{1E} and ~21 Å less than in I_{1M}. A simple interpretation of these changes in distance is that bending the downstream duplex into the clamp brings +14 DNA ~7 Å closer to -100 DNA, and that closing the clamp brings -100 DNA ~14 Å closer to +14 DNA. The PIFE intensity at -100 is smaller for I_{1L} than for I_{1M} or I_{1E} (Table 1) indicating less strong contacts with both these DNA positions in I_{1L} , while +14 PIFE intensity is similar for all three I1 intermediates.

d) Comparison with Other Mechanistic Proposals. A mechanism of open complex formation by E. coli σ^{70} RNAP was proposed based on fluorescence studies of interactions of designed short DNA oligomers with a -10 region or -35 and -10 regions (lacking upstream DNA), using RNAP drug complexes locked in closed and open clamp states. Probes of base flipping in the -10 element and of interactions of the +2 position in open complex formation were used in these studies[46]. Evidence was obtained for an early intermediate (called a recognition complex) in which closed-clamp RNAP is bound to somewhat bent, base-flipped but otherwise closed promoter DNA. This appears to correspond to the I_{1E} intermediate[4] of Fig. 7 (without the upstream DNA) in which the RNAP is closed and the downstream DNA is sufficiently bent to contact the top of the β pincer. In the next proposed intermediate, called RP_{I1}, the downstream duplex is more highly bent and the clamp of RNAP is open, similar to the I_{1M} intermediate characterized here. No evidence was obtained for the key late closed intermediate I_{1L}. The initiation bubble in the next proposed intermediate (designated RP_{I2}, based on a crystal structure of a fork-junction complex with open-clamp RNAP), appears similar to the initial, unstable OC previously identified by MnO_4^- footprinting and kinetic analysis and called I_2.[52, 59, 64] Another possible I_2 model is the open-clamp complex designated RP_{ip} observed by cryoEM with σ^{54} RNAP, stabilized by using a pre-melted heteroduplex DNA.[51] Since we find the clamp closes in converting I_{1M} to I_{1L}, the proposal that the clamp is open in I_2 indicates that the clamp opens when the initiation bubble opens in converting I_{1L} to I_2. If the strand backbones are held in the closed clamp in I_{1L}, as seems likely, clamp opening may be connected to bubble opening in $I_{1L} \rightarrow I_2$.

Closed-promoter intermediates formed by yeast pol II RNAP have been characterized by cryoEM. In its downstream interactions, the I_{1E} intermediate proposed here (Fig. 7) is similar to the yeast pol II CC1 intermediate, in which the downstream duplex is bent and located above the closed clamp. In its downstream interactions, open-clamp I_{1M} (Fig. 7) corresponds to the subsequent pol II CC2 complex. In the next step for both enzymes, the downstream duplex is bent into the open clamp, but the clamp is closed on the downstream duplex in E. coli σ^{70} RNAP intermediate I_{1L} (Fig. 7) but the clamp is open in yeast pol II intermediate CC^{dist}. Because the next step (DNA opening) for E. coli σ^{70} RNAP is rate-determining, no subsequent intermediate (like I_2) is observable, but only the final OC. Presumably a similar situation exists for yeast pol II.

Rate and Equilibrium Constants for Forming the Different Intermediates and their Use to Predict Time-Dependent Populations and the Free Energy vs. Progress Diagram

The time-evolution of populations of free promoter DNA (P), all four closed intermediates and OC for the reactant concentrations and conditions investigated here is shown in Fig. 8A. These predictions are based on the rate constants of Table 3, determined by fitting FRET and PIFE kinetic data to Mechanism 2.

a) Free Promoter DNA (P): Because the RP_C binding constant K_1 is modest ($\sim 7 \times 10^6\,M^{-1}$) and subsequent equilibrium constants for advancing RP_C to form the three closed I_1 intermediates are also modest (in the range $1 - 2$; Table 3), approximately half the population of promoter DNA remains unbound in the first 10 s of the reaction at the reactant concentrations simulated (50 nM). In these first 10 s, the {CC} forms as shown in Figs. 2B and 8A. After 10 s, formation of the very stable OC from the most advanced CC (I_{1L}) results in additional conversion of P to RP_C and I_1 species.

b) Initial Specific CC Intermediate (RP_C): Rate and equilibrium constants for RP_C formation (Table 3) are similar to published values for other promoters at similar conditions.[49, 50] The second order rate constant k_1 ($\sim 3.0 \times 10^8\,M^{-1}\,s^{-1}$) for RP_C formation is about 5% of the diffusion-collision limit, most simply interpreted as that only about 5% of collisions of R with P result in RP_C formation. About 10% of promoter DNA is predicted to form RP_C in the first 10-20 ms of the reaction at the concentrations simulated (50 nM R, P). No FRET or PIFE signal is observed in this time range (Figs. 3, 4), and indeed none is expected for RP_C because footprinting indicates that the downstream promoter DNA is not bent and that the PIFE probe positions (-100, +14) are not in contact with RNAP. [8-12] The lifetime ($1/k_{-1}$) of

RP_C is relatively short (< 25 ms). RP_C increases to a maximum (~15% of total promoter DNA) at ~40 ms and remains the primary CC up to ~100 ms.

    c) Wrapped CC Intermediate I_{1E}: The rate constant for advancement of RP_C to I_{1E} (~13 s^{-1}) is less than that for dissociation of RP_C (~ 44 s^{-1}), allowing RP_C to equilibrate with R and P on the time scale of its conversion to I_{1E}. I_{1E} is only marginally more stable than RP_C (K_2 = 1.8 (Table 3); ΔG^o_2 = - 0.3 kcal/mol), showing that favorable contacts of these wrapped regions of promoter DNA with RNAP are just sufficiently favorable to overcome the costs of DNA-bending. Starting at 30 ms, the population of I_{1E} (Fig. 8A) increases to a broad maximum (about 20% of total promoter DNA) at 300 ms and then decreases gradually at longer times as more advanced CC and OC form.

Table 3. Rate Constants and Equilibrium Constants for Formation of the Ensemble of Closed Complex Intermediates and DNA Opening at 19 ºC

	Kinetic Step	**Rate Constants (± Standard Deviation)**	**Equilibrium Constant[a]**
1	$R + P \rightleftharpoons RP_c$	k_1 = (3.0 ± 0.2) x 10^8 M^{-1}s^{-1}, k_{-1} = 44 ± 2 s^{-1}	K_1 = (6.8 ± 0.5) x 10^6 M^{-1}
2	$RP_c \rightleftharpoons I_{1,E}$	k_2 = 13 ± 1 s^{-1}, k_{-2} = 7.2 ± 0.4 s^{-1}	K_2 = 1.8 ± 0.2
3	$I_{1,E} \rightleftharpoons I_{1,M}$	k_3 = 2.5 ± 0.1 s^{-1}, k_{-3} = 1.7 ± 0.1 s^{-1}	K_3 = 1.5 ± 0.1
4	$I_{1,M} \rightleftharpoons I_{1,L}$	k_4 = 0.10 ± 0.01 s^{-1}, k_{-4} = 0.10 ± 0.01 s^{-1}	K_4 = 1.0 ± 0.1
5	$I_{1,L} \rightarrow OC$	k_5 = 0.040 ± 0.002 s^{-1} [b]	

[a] These four equilibrium constants yield a predicted $K_{\{CC\}}$ = (5.5 ± 0.8) x 10^7 M^{-1} from Eqs. S21-22, in agreement with the experimental $K_{\{CC\}}$ = (5 ± 1) x 10^7 M^{-1} (Fig. 2A).
[b] This k_5, multiplied by the fraction of the equilibrated {CC} that is I_{1L} calculated from K_1, K_2, K_3 and K_4 above (Eqs. S19-20), gives a predicted k_{isom} = 0.013 ± 0.001 s^{-1} the same within uncertainty as the experimental k_{isom} = 0.014 ± 0.003 s^{-1} (Fig. 2

d) Open-clamp Wrapped CC Intermediate I_{1M}: The rate constant for advancement of I_{1E} to I_{1M} (~ 2.5 s^{-1}) is less than that for reversal of I_{1E} to RP_C (~ 7 s^{-1}), allowing I_{1E} to equilibrate with RP_C on the time scale of its conversion to I_{1M}. I_{1M} is somewhat favored thermodynamically over I_{1E} ($K_3 = 1.5$ (Table 3) and $\Delta G^o_3 = - 0.2$ kcal/mol). The population of I_{1M} increases from ~100 ms to a broad maximum near 2 s (Fig. 8A) and then decreases gradually at longer times as more advanced CC and OC form. As in the previous step ($RP_C \rightarrow I_{1E}$), the cost of large conformational changes in conversion of I_{1E} to I_{1M} must be compensated by favorable interactions to result in a marginally favorable ΔG^o for this step.

e) Closed-clamp Late CC Intermediate I_{1L}: The rate constant for conversion of I_{1M} to I_{1L} (~0.1 s^{-1}) is much smaller than that for reversal of I_{1M} to I_{1E} (~2 s^{-1}), allowing equilibration of I_{1M} and I_{1L} on the time scale of conversion of I_{1M} to I_{1L}. The population of I_{1L} increases from ~1 s to a broad maximum (~15% of total promoter DNA) at ~25 s and decreases gradually at longer times as I_{1L} slowly converts to
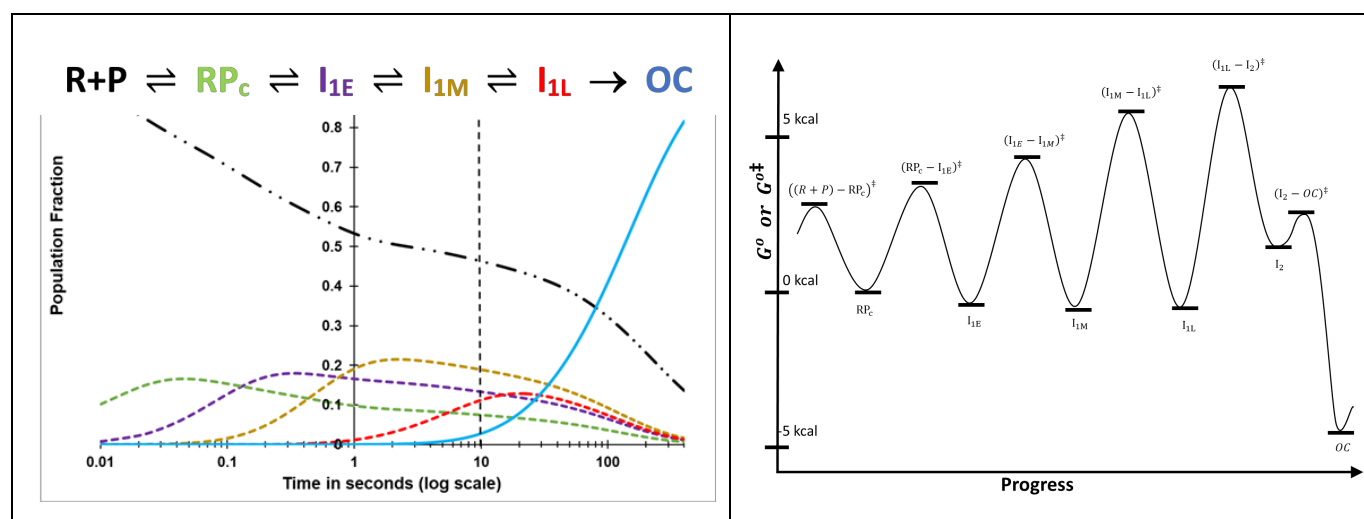

Figure 8. A) Time Evolution of Populations of Unbound Promoter DNA, Intermediates in the {CC} Ensemble and the Stable λP$_R$ Promoter OC. Simulations of population fractions of reactant, intermediates and product vs time (0.01 s to 400 s) for 5-step Mechanism 2 at 50 nM final concentrations of RNAP and promoter DNA, 19 ^{o}C, using rate constants (Table 3) from analysis of FRET and PIFE kinetic data. Free promoter DNA ▬ . . ▬ ; closed complex intermediates RP_C ▬ ▬ ▬ , $I_{1,E}$ ▬ ▬ ▬ , $I_{1,M}$▬ ▬ ▬ , $I_{1,L}$ ▬ ▬ ▬ ,OC ▬▬ . The dashed vertical line at 10 s marks the onset of OC formation from the equilibrium mixture of {CC} and free promoter DNA predicted from filter binding and MnO$_4^-$ kinetic data (Fig. 2B).   **B) Standard Free Energy (G^o) vs. Progress Diagram for OC Formation with FL λP$_R$ Promoter DNA.** G^o values for CC intermediates and the stable OC are obtained from the equilibrium constants of Table 3 ($\Delta G^o = $ -RT lnK_i) and are expressed relative to RP_C, which is arbitrary assigned G$^o = 0$ kcal. The G^o value for I_2 is obtained from $K_5 = k_5/k_{-5}$, where $k_{-5} = 1.2$ s-1 (Fig. 5). Activation free energies G^{o‡} relative to the same reference are calculated from the relationship $\Delta G^{o\ddagger} = $ - RTln(k/k$_{max}$) where k$_{max}$ is the (maximum) rate constant for the hypothetical situation $\Delta G^{o\ddagger} = 0$.[4] For purposes of illustration, we choose k$_{max} = 5$ x 10^3 s^{-1} for all steps.

OC in the DNA-opening step (Fig. 8A). The equilibrium constant $K_4 = 1.0$ (Table 3) for conversion of I_{1M} to I_{1L} so these two intermediates have the same standard free energies ($\Delta G_4^o = 0$).

f) Stable Wrapped OC: The FRET amplitude of the end-product 19 °C OC is somewhat larger than that of I_{1L}, indicating that opening the bubble to form the initial unstable OC (I_2) and subsequent formation of stabilizing downstream interactions in the conversion of I_2 to the stable 19 °C OC reduce the distance between +14 and -100. Intrinsic PIFE signals at -100 and +14 in the stable OC are similar to those of I_{1L} and smaller than those of the earlier I_1 intermediates (Table 1). Only overall changes in FRET and PIFE in converting I_{1L} to the stable OC are observed. While no information is obtained about the individual steps converting I_{1L} to I_2 and then to the stable OC, previous quantitative kinetic studies of the steps of OC dissociation induced by KCl-upshift (as in Fig. 5) can be used to obtain this information.

g) Free Energy vs. Progress Diagram Fig. 8B shows a standard free energy (G^o) vs. progress diagram for Mechanism 2, based on the rate and equilibrium constants of Table 3. RP_C is taken as the point of reference and G_{RPc}^o is set equal to zero. Activation free energies $G^{o\ddagger}$ relative to the same reference are calculated from the relationship $\Delta G^{o\ddagger} = - RT\ln(k/k_{max})$ where k_{max} is the (maximum) rate constant for the hypothetical situation $\Delta G^{o\ddagger} = 0$.[4] For purposes of illustration, we choose $k_{max} = 5 \times 10^3$ s^{-1} for all steps. The rapid-equilibrium condition for each step is indicated by the lower barrier for reversal of each intermediate, relative to going forward to the next intermediate. The highest $G^{o\ddagger}$ value is for the transition state for the rate-determining DNA opening step that converts I_{1L} to I_2. Barriers surrounding I_2 show the opposite pattern to those for previous intermediates, with a much lower barrier for forward conversion of I_2 to the stable OC than for reversal of I_2 to I_{1L}.

Fig. 8B reinforces the finding (Table 3) that the steps starting with RP_C and forming the more advanced I_1 intermediates are all only marginally favorable for this promoter and conditions, with equilibrium constants K_i between 1 and 2. Hence the net effect of bending and wrapping upstream and downstream promoter DNA to interact with the β' and β subunits, respectively, only marginally favorable, and all I_1 species are present at significant levels when the equilibrium distribution is established. This finding is consistent with HO footprinting results which showed that far-upstream contacts are substantially weaker than downstream contacts, and with the absence of a far-upstream DNase footprint, explicable because weak protein-DNA contacts can be displaced by binding of DNase.[15]

How Far-Upstream DNA Facilitates Conversion of {CC} to OC: Insights from the Five-Step Mechanism

a) Upstream-DNA Greatly Increases the Isomerization Rate Constant k_{isom} at the λP_R Promoter with only a Modest Effect on K_{CC}. The presence of upstream DNA (-40 to -100) in the FL λP_R promoter affects the kinetics of OC formation similarly to a strong, upstream-binding Type II transcription factor.[16, 17] The isomerization rate constant of the FL λP_R promoter under the conditions investigated here is ~50 times larger than that of a λP_R promoter truncated upstream at position -47 (UT-47) [17] Similarly, k_{isom}^{FL} for the lacUV5 promoter is 10- to 30-fold larger than k_{isom}^{UT} for UT-42, UT-45, and UT-63.[16] K_{CC} is either unaffected or increased by upstream truncation. [16, 17] These effects of the presence or absence of upstream DNA on k_{isom} of λP_R and lacUV5 promoters are as large as or larger than effects of promoter sequence changes or addition of upstream-binding factors.

From Eq. S19-20, upstream DNA could increase k_{isom} by increasing the fraction (f_{I1L}) of the {CC} ensemble that is I_{11} and/or by increasing the intrinsic DNA opening rate constant k_5. *Several lines of evidence indicate that profound differences in f_{I1L} are the origin of the much greater k_{isom}^{FL} as compared to k_{isom}^{UT} [4].* Real-time footprinting of {CC} ensembles during OC formation revealed that downstream contacts of RNAP with FL λP_R extend to +20 as compared to +2/+7 (partial protection) for λP_R UT-47[15] and to -5 for RP_C at other promoters. To interpret these results we proposed that the UT-47 {CC} ensemble is composed primarily of RP_C and the subsequent CC intermediate (I_{1E}), with relatively small concentrations of the more advanced CC (I_{1L} and probably also I_{1M}) that are a major part of the FL {CC} ensemble[4]. We deduce that formation of these advanced CC from I_{1E} is favored by the presence of upstream DNA for the FL promoter and greatly disfavored for the UT promoter variants. Insights from Mechanism 2 and Fig. 7 help clarify why wrapping of FL upstream DNA is needed to advance the {CC} ensemble beyond I_{1E} and form a significant population of I_{1L}, the only CC intermediate that can undergo the DNA opening step.

b) Predicted Large Differences in the Fraction of Advanced CC (f_{I1L}) between FL and UT-47 λP_R Promoter DNA. Fractional populations of the different CC as a function of time in OC formation with full-length (FL) λP_R promoter DNA are shown in Fig. 8A. After equilibration (at > 10 s), the {CC} ensemble formed from FL λP_R promoter DNA is relatively advanced (33% I_{1L}, 33% I_{1M}) with smaller percentages of RP_C (12%) and I_{1E} (22%). This population distribution is consistent with real-time footprinting of the {CC} ensemble, which shows strong protection downstream to +20, characteristic of

I_{1L} but not of I_{1E} or RP_C[15]. By contrast, for UT-47 λP_R , the analysis given in SI predicts that I_{1L} and perhaps also I_{1M} are greatly destabilized (Fig. 9). The predicted population distribution in the equilibrated {CC} ensemble for UT-47 is approximately 60% I_{1E} , 33% RP_C, 6% I_{1M} and only 0.6 % I_{1L}. This small population of I_{1L} in the equilibratedUT-47 {CC} ensemble is the likely origin of its small isomerization

rate constant.

c) Proposed Structural Basis of these Very Different {CC} Population Distributions. For FL λP_R the kinetic data and above analysis show that one or both steps of conversion of I_{1E} to I_{1L} is/are greatly favored by the interactions of bent-wrapped upstream DNA with RNAP, as previously proposed[4]. If conversion of I_{1E} to I_M is the step that is much more favorable for FL λP_R than for UT-47, then bending and wrapping of upstream DNA on RNAP favors clamp opening. This might involve compaction or folding of the α-CTD tethers to bend and wrap the upstream DNA in I_{1E}, causing the assembly of UP element DNA, α-CTD, and compacted tethers to interact with the body of RNAP. A structural analogy would be the lac repression complex, in which the flexible tethers connecting the DNA binding domains (DBD) to the core repressor fold into helices, interact with operator DNA, and cause the DNA-DBD assembly to interact with the repressor core[66-69].
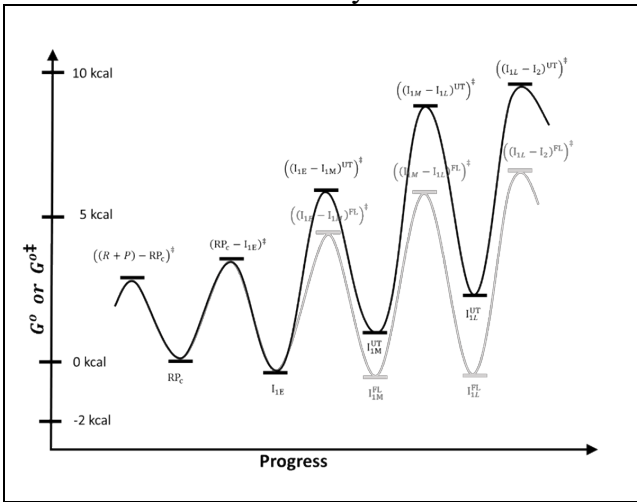

Figure 9. Proposed Standard Free Energies (Gº) vs. Progress Diagram for OC Formation with UT-47 λP_R Promoter DNA and Comparison with FL λP_R Promoter DNA. Standard free energies (Gº) for CC intermediates formed by UT-47 λP_R promoter DNA are obtained from equilibrium constants as described in SI and are expressed relative to RP_C, which is arbitrary assigned Gº = 0 kcal. Effects of truncation on the steps converting I_{1E} to I_{1L} are assumed to be entirely on the forward rate constant of these steps. Activation free energies $G^{o\ddagger}$ relative to the same RP_C reference are calculated from the relationship $\Delta G^{o\ddagger} = - RT\ln(k/k_{max})$ where k_{max}, the (maximum) rate constant for the hypothetical situation $\Delta G^{o\ddagger} = 0$, is arbitrarily assigned the value $k_{max} = 5 \times 10^3$ s^{-1} for all steps. Diagram from Fig. 8B for the full-length (FL) promoter is shown for comparison.

If conversion of I_{1M} to I_{1L} is the step that is much more favorable for FL λP_R than for UT-47, then upstream wrapping favors the entry and descent of the downstream duplex into the open clamp. A molecular explanation for this, proposed previously[3, 15], is that upstream wrapping could allow far-upstream DNA to contact downstream mobile elements (DME) of

RNAP and move them away from the downstream clamp/cleft. This proposal and the interaction of far upstream DNA with DME would explain the significant -100 PIFE signals of I_{1E} and I_{1M} (Table 1), and the reduction in -100 PIFE in I_{1L} and OC when the DME move to interact with the downstream duplex in the stable OC [4].

We think it likely that both these steps ($I_{1E} \rightarrow I_{1M}$ and $I_{1M} \rightarrow I_{1L}$) are affected by contacts of the upstream-wrapped DNA with the outside of the clamp and with the DME at the downstream end of the clamp (Fig. 7), and therefore that both conversions are facilitated by upstream wrapping. An example of this is shown in the comparison of free energy vs. progress diagrams for UT-47 and FL λP_R in Fig. 9 (details provided in SI). Fig. 9 predicts that the much smaller k_{isom} for UT-47 results from the relative instability of late CC intermediates $(G_{I1L}^{o,UT} > G_{I1M}^{o,UT} > G_{I1E}^{o,UT})$, in contrast to the situation for FL λP_R where late CC intermediates are more stable than earlier ones $(G_{I1L}^{o,FL} < G_{I1M}^{o,FL} < G_{I1E}^{o,FL})$.

Conclusions

The FRET and PIFE fluorescence kinetics studies reported here reveal for the first time when upstream and downstream promoter DNA are bent and wrapped on RNAP in the mechanism of OC formation, as well as why upstream DNA (-40 to -100) is necessary for efficient isomerization of the {CC} ensemble to OC. Information in the promoter sequence (presumably primarily the UP-element, -35 and -10 regions), read by interactions with the α-CTD and σ^{70} regions 4 and 2 in the initial CC (RP$_C$), directs concerted bending and wrapping of both upstream and downstream DNA in conversion of RP$_C$ to the more advanced CC, designated I_{1E}, in which the distance between -100 and +14 positions on promoter DNA is reduced from >300 A before binding of RNAP to ~75 A. Interactions between promoter DNA and RNAP in I_{1E}, which we propose are primarily between the upstream-wrapped DNA and the hinge of the clamp, result in opening of the clamp (on a ~1 s timescale for the conditions investigated) to form the intermediate I_{1M}. Clamp opening moves the far-upstream DNA, which we propose is wrapped high on the back of the β' clamp, ~14 A further away from the downstream duplex, which we propose is bound on the top of the β clamp in I_{1E} and I_{1M}. Subsequently, on a ~10 s timescale, the downstream duplex is bent into the clamp, which closes to reduce the distance between -100 and +14 positions to ~68 A in I_{1L}, the most advanced CC. Opening of the initiation bubble by strand binding free energy appears to occur in a single kinetic step at this promoter to form the initial unstable OC (I_2), which is rapidly stabilized by downstream interactions with DME. Each step of forming and advancing the {CC} ensemble rapidly

equilibrates on the time scale of the next step. The DNA opening step is rate-determining for OC formation because its forward rate constant ($k_5 = 0.04$ s^{-1}) is smaller than those of previous steps and because it is effectively irreversible since stabilization of I_2 by downstream interactions is more rapid than DNA-closing for the promoter and conditions investigated (Fig. 8B).

Accession Codes

P0A7Z4, subunit α; P0A8V2, subunit β; P0A8T7, subunit β'; P0A800, subunit ω; P00579 subunit σ.

Contents of SI:

The SI text describes the methods used in the following analyses. 1) the normalization of fluorescence kinetic data; 2) fitting methods, including the analysis of FRET and PIFE kinetics as a sum of four exponentials and fitting these data to five-step Mechanism 2; 3) consistency checks on the rate constants of Mechanism 2; 4) conversion of fluorescence fitting amplitudes to relative fluorescence intensities; 5) estimation of distance between upstream (-100) and downstream (+14) DNA in wrapped RNAP-promoter complexes from FRET amplitudes; 6) interpretation of rate and equilibrium constants of the classic two-step Mechanism 1 (k_{isom}, K_{CC}, k_a) using Mechanism 2; and 7) calculations for free energy vs. progress diagram and {CC} population distribution for formation at UT-47 λPR. SI figures include: 1) the MnO$_4^-$ footprinting data of Fig. 1 plotted on a logarithmic time scale; 2) comparison of the FRET and PIFE data of Figs. 3-4 with the prediction based on the average rate constants and signal amplitudes of Tables 1 and 3. SI Tables include: 1) promoter and 2) primer sequences used in PCR synthesis of the different dye-labeled promoters investigated here; 3) comparison of rate constants ($1/\tau_{i,obs}$) from the four-exponential fit with predictions from the rate constants of Mechanism 2; 4) individual rate constants and signal amplitudes obtained from analysis of all FRET and PIFE experiments; and 5) average FRET acceptor and PIFE fitting amplitudes of CC intermediates and the stable OC.

References

[1] McClure, W. R. (1985) Mechanism and control of transcription initiation in prokaryotes, *Annual Review of Biochemistry 54*, 171-204.

[2] M. T. Record, Jr., W. S. Reznikoff, M. L. Craig, K. L. Mcquade and P. J. Schlax. (1996) Escherichia coli RNA Polymerase (E sigma70), Promoters, and the Kinetics of the Steps of Transcription Initiation., *Second Edition of Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, 792-821.

[3] Saecker, R. M., Record, M. T. Jr., 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*, 754-771.

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

[5] Saecker, R. M., Tsodikov, O. V., Mcquade, K. L., Schlax Jr, P. E., Capp, M. W., and Record, M. T. Jr. (2002) Kinetic studies and structural models of the association of E. coli σ70 RNA polymerase with the λPR promoter: large scale conformational changes in forming the kinetically significant intermediates, *Journal of Molecular Biology 319*, 649-671.

[6] Roe, J.-H., Burgess, R. R., and Record, M. T. Jr. (1984) Kinetics and mechanism of the interaction of Escherichia coli RNA polymerase with the λPR promoter, *Journal of Molecular Biology 176*, 495-522.

[7] Buc, H., and McClure, W. R. (1985) Kinetics of open complex formation between Escherichia coli RNA polymerase and the lac UV5 promoter. Evidence for a sequential mechanism involving three steps, *Biochemistry 24*, 2712-2723.

[8] Cowing, D. W., Mecsas, J., Record, M. T. Jr., 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 l 210*, 521-530.

[9] Mecsas, J., Cowing, D. W., and Gross, C. A. (1991) Development of RNA polymerase-promoter contacts during open complex formation, *Journal of Molecular Biology 220*, 585-597.

[10] Kovacic, R. (1987) The 0 degree C closed complexes between Escherichia coli RNA polymerase and two promoters, T7-A3 and lacUV5, *Journal of Biological Chemistry 262*, 13654-13661.

[11] Schickor, P., Metzger, W., Werel, W., Lederer, H., and Heumann, H. (1990) Topography of intermediates in transcription initiation of E.coli, *EMBO Journal 9*, 2215-2220.

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

[13] 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 Research 37*, 5390-5404.

[14] 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.

[15] 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.

[16] Ross, W., and Gourse, R. L. (2005) Sequence-independent upstream DNA–αCTD interactions strongly stimulate Escherichia coli RNA polymerase-lacUV5 promoter association, *Proceedings of the National Academy of Sciences 102*, 291-296.

[17] 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.

[18] Estrem, S. T., Gaal, T., Ross, W., and Gourse, R. L. (1998) Identification of an UP element consensus sequence for bacterial promoters, *Proceedings of the National Academy of Sciences 95*, 9761-9766.

[19] Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex, *Science 296*, 1285-1290.

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

[21] Turingan, R. S., Liu, C. H., Hawkins, M. E., and Martin, C. T. (2007) Structural confirmation of a bent and open model for the initiation complex of T7 RNA polymerase, *Biochemistry 46*, 1714-1723.

[22] Ujvari, A., and Martin, C. T. (2000) Evidence for DNA bending at the T7 RNA polymerase promoter, *Journal of Molecular Biology 295*, 1173-1184.

[23] Tang, G. Q., and Patel, S. S. (2006) T7 RNA polymerase-induced bending of promoter DNA is coupled to DNA opening, *Biochemistry 45*, 4936-4946.

[24] Tang, G. Q., and Patel, S. S. (2006) Rapid binding of T7 RNA polymerase is followed by simultaneous bending and opening of the promoter DNA, *Biochemistry 45*, 4947-4956.

[25] Blair, R. H., Goodrich, J. A., and Kugel, J. F. (2012) Single molecule FRET shows uniformity in TBP-induced DNA bending and heterogeneity in bending kinetics, *Biochemistry 51*, 7444.

[26] Louder, R. K., He, Y., López-Blanco, J. R., Fang, J., Chacón, P., and Nogales, E. (2016) Structure of promoter-bound TFIID and insight into human PIC assembly, *Nature 531*, 604.

[27] Schilbach, S., Hantsche, M., Tegunov, D., Dienemann, C., Wigge, C., Urlaub, H., and Cramer, P. (2017) Structures of transcription pre-initiation complex with TFIIH and Mediator, *Nature 551*, 204-209.

[28] He, Y., Fang, J., Taatjes, D. J., and Nogales, E. (2013) Structural visualization of key steps in human transcription initiation, *Nature 495*, 481.

[29] He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016) Near-atomic resolution visualization of human transcription promoter opening, *Nature 533*, 359.

[30] Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016) Transcription initiation complex structures elucidate DNA opening, *Nature 533*, 353.

[31] Dienemann, C., Schwalb, B., Schilbach, S., and Cramer, P. (2019) Promoter distortion and opening in the RNA polymerase II cleft, *Molecular Cell 73*, 97 - 106.

[32] 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. Jr. (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.

[33] 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.

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

[35] Liu, B., Hong, C., Huang, R. K., Yu, Z., and Steitz, T. A. (2017) Structural basis of bacterial transcription activation, *Science 358*, 947-951.

[36] Rice, P. A., Yang, S.-W., Mizuuchi, K., and Nash, H. A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn, *Cell 87*, 1295-1306.

[37] Holbrook, J. A., Tsodikov, O. V., Saecker, R. M., and Record, M. T. Jr. (2001) Specific and non-specific interactions of integration host factor with DNA: Thermodynamic evidence for disruption of multiple IHF surface salt-bridges coupled to DNA binding, *Journal of Molecular Biology 310*, 379-401.

[38] VanderMeulen, K., 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, *Journal of Molecular Biology 377*, 9-27.

[39] Swinger, K. K., Lemberg, K. M., Zhang, Y., and Rice, P. A. (2003) Flexible DNA bending in HU-DNA cocrystal structures, *EMBO Journal 22*, 3749-3760.

[40] Koh, J., Saecker, R. M., and Record, M. T. Jr. (2008) DNA Binding Mode Transitions of Escherichia coli HU alpha beta: Evidence for Formation of a Bent DNA - Protein Complex on Intact, Linear Duplex DNA, *Journal of Molecular Biology 383*, 324-346.

[41] Koh, J., Shkel, I., Saecker, R. M., and Record, M. T. Jr. (2011) Nonspecific DNA Binding and Bending by HU alpha beta: Interfaces of the Three Binding Modes Characterized by Salt-Dependent Thermodynamics, *Journal of Molecular Biology 410*, 241-267.

[42] Hwang, H., and Myong, S. (2014) Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions, *Chemical Society Reviews 43*, 1221-1229.

[43] Lerner, E., Ploetz, E., Hohlbein, J., Cordes, T., and Weiss, S. (2016) A Quantitative Theoretical Framework For Protein-Induced Fluorescence Enhancement–Förster-Type Resonance Energy Transfer (PIFE-FRET), *Journal of Physical Chemistry B 120*, 6401-6410.

[44] 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 (PIFE), *Biophysical Journal 117,* 66-73.

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

[46] 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.

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

[48] Rammohan, J., Ruiz Manzano, A., Garner, A. L., Prusa, J., Stallings, C. L., and Galburt, E. A. (2016) Cooperative stabilization of Mycobacterium tuberculosis rrnA P3 promoter open complexes by RbpA and CarD, *Nucleic Acids Research 44*, 7304-7313.

[49] 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*, e22520.

[50] Jensen, D., Manzano, A. R., Rammohan, J., Stallings, C. L., and Galburt, E. A. (2019) CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the Mycobacterium tuberculosis RNA polymerase, *Nucleic Acids Research 47,* 6685-6698.

[51] Glyde, R., Ye, F., Jovanovic, M., Kotta-Loizou, I., Buck, M., and Zhang, X. (2018) Structures of bacterial RNA polymerase complexes reveal the mechanism of DNA loading and transcription initiation, *Molecular Cell 70*, 1111-1120.

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

[53] Kontur, W. S., Capp, M. W., Gries, T. J., Saecker, R. M., and Record, M. T. Jr. (2010) Probing DNA binding, DNA opening, and assembly of a downstream clamp/jaw in Escherichia coli RNA polymerase– λPR promoter complexes using salt and the physiological anion glutamate, *Biochemistry 49*, 4361-4373.

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

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

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

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

[58] Tsodikov, O. V., and Record, M. T. Jr. (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 E sigma(70) RNA polymerase and lambda P-R promoter DNA, *Biophysical Journal 76*, 1320-1329.

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

[60] Ruff, E. F., Drennan, A. C., Capp, M. W., Poulos, M. A., Artsimovitch, I., and Record, M. T. Jr. (2015) E. coli RNA Polymerase Determinants of Open Complex Lifetime and Structure, *Journal of Molecular Biology 427*, 2435-2450.

[61] Bernasconi, C. F. (1976) *Relaxation Kinetics*, Academic Press Inc, New York.

[62] 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.

[63] Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A. Y., Campbell, E. A., Gourse, R. L., Ross, W., and Darst, S. A. (2019) E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation and dynamics, *Elife 8*, e49375.

[64] Chakraborty, A., Wang, D., Ebright, Y. W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B. T., Knight, J., Weiss, S., and Ebright, R. H. (2012) Opening and closing of the bacterial RNA polymerase clamp, *Science 337*, 591-595.

[65] Duchi, D., Mazumder, A., Malinen, A. M., Ebright, R. H., and Kapanidis, A. N. (2018) The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp, *Nucleic Acids Research 46*, 7284-7295.

[66] Sengupta, R., Capp, M. W., Shkel, I. A., and Record, M. T. Jr. (2017) The mechanism and high-free-energy transition state of lac repressor-lac operator interaction, *Nucleic Acids Research 45*, 10.

[67] Bell, C. E., and Lewis, M. (2000) A closer view of the conformation of the Lac repressor bound to operator, *Nature Structural Biology. 7*, 209-214.

[68] Spronk, C., Folkers, G. E., Noordman, A., Wechselberger, R., Van Den Brink, N., Boelens, R., and Kaptein, R. (1999) Hinge-helix formation and DNA bending in various lac repressor-operator complexes, *EMBO Journal. 18*, 6472-6480.

[69] Kalodimos, C. G., Biris, N., Bonvin, A., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes, *Science 305*, 386-389.

For Table of Contents Only
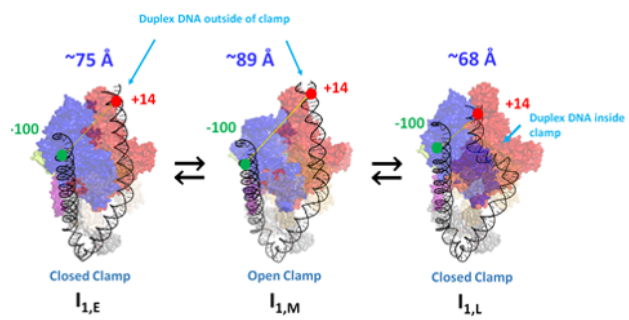
Supplemental Information

Fluorescence-Detected Conformational Changes in Duplex DNA in Open Complex Formation by *E. coli* RNA Polymerase: Upstream Wrapping and Downstream Bending Precede Clamp Opening and Insertion of the Downstream Duplex

*Raashi Sreenivasan[1], *Irina A. Shkel[2,3], *Munish Chhabra[1], Amanda Drennan[2], Sara Heitkamp[2], Hao-Che Wang[1], Malavika Abhineshababu Sridevi[2], Dylan Plaskon[2], Christina McNerney[2], Katelyn Callies[2], Clare K. Cimperman[2], M. Thomas Record, Jr.[1,2,3]

Biophysics[1] Program and Departments of Biochemistry[2] and Chemistry[3]

University of Wisconsin – Madison

Madison, WI 53706

* Co-first authors

Normalization of Fluorescence Kinetic Data

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

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

Visual inspection of the set of 10-15 normalized shots in each 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.[1] To avoid this, kinetics experiments were performed at concentration ratios of active RNAP to promoter DNA in the range 0.5:1 to 1:1.

Fitting Methods:

Analysis of FRET and PIFE Kinetics as Sums of Exponentials:

FRET and PIFE OC formation kinetic data (each an average of a series of ~10 normalized shots) collected over the time range 10 ms – 400 s were fit with Berkeley Madonna[2] to a sum of four exponentials and a linear correction for long-term drift[3] (Eq. S2):

$$F_{norm} = F_o + \sum_{i=1}^{4} A_i^\tau \left(1 - e^{-\frac{t}{\tau_{i,obs}}} \right) + m_{drift} t \qquad \text{Eq. S2}$$

where the $\tau_{i,obs}$ (i = 1 − 4) are time constants and A_i^τ are amplitudes. F_o is the extrapolated fluorescence at t = 0 (negative because of the normalization at 10 ms; Eq. S1) and m_{drift} is the slope of the drift correction, which generally was small and varied from experiment to

experiment. We found that three exponentials in Eq. S2 were insufficient while five were too many, supporting the use of Mechanism 2 with three detectable intermediates, differing in FRET and PIFE from the final open complex. Rate constants $1/\tau_{i,obs}$ were obtained from the fits to Eq S2. Average values of $1/\tau_{i,obs}$ calculated from all FRET and PIFE experiments reported here and listed in Table S3, span four orders of magnitude, in decade increments from ~ 20 s^{-1} to ~ 0.02 s^{-1}. No significant differences in $1/\tau_{i,obs}$ values were detected for the various types of FRET and PIFE experiments.

Because filter binding and MnO$_4^-$ kinetics of OC formation are single-exponential (see for example Figs. 1 and S1), the closed complex ensemble {CC} in Mechanism 1 must rapidly equilibrate with P and R on the time scale of isomerization of {CC} to OC, and therefore each step of forming and advancing the CC ensemble in Mechanism 2 must equilibrate rapidly on the time scale of the next step. For this reason values of $1/\tau_{i,obs}$ ($i \leq 3$) are interpretable as decay-to-equilibrium rate constants[4] for the steps in Mechanism 2 detected by FRET and PIFE that form the three I$_1$ intermediates.

$$1/\tau_{1,pred} = \frac{K_1 C_{eq}}{1 + K_1 C_{eq}} k_2 + k_{-2} \qquad \text{Eq. S3}$$

$$1/\tau_{2,pred} = \frac{K_1 C_{eq} K_2}{1 + K_1 C_{eq} + K_1 C_{eq} K_2} k_3 + k_{-3} \qquad \text{Eq. S4}$$

$$1/\tau_{3,pred} = \frac{K_1 C_{eq} K_2 K_3}{1 + K_1 C_{eq} + K_1 C_{eq} K_2 + K_1 C_{eq} K_2 K_3} k_4 + k_{-4} \qquad \text{Eq. S5}$$

where $C_{eq} = ([P]_{eq} + [R]_{eq})$.

The smallest rate constant $1/\tau_{4,obs}$ is interpreted as the rate constant for the subsequent, effectively-irreversible DNA opening step that converts the most advanced I$_1$ intermediate (I$_{1L}$) to I$_2$.

$$1/\tau_{4,pred} = \frac{K_1 C_{eq} K_2 K_3 K_4}{1 + K_1 C_{eq} + K_1 C_{eq} K_2 + K_1 C_{eq} K_2 K_3 + K_1 C_{eq} K_2 K_3 K_4} \, k_5 \qquad \text{Eq. S6}$$

In fitting to Mechanism 2, initial order-of-magnitude estimates of rate constants were obtained

from the corresponding $1/\tau_{i,obs}$ values (i ≤ 3) assuming $k_i \sim k_{-i} \sim 1/(2\tau_{i-1,obs})$ and $k_5 \sim 1/\tau_{4,obs}$.

Semi-quantitative comparisons of values of $1/\tau_{i,obs}$ obtained from fitting FRET and PIFE data to

Eq S2 with values obtained from equilibrium and rate constants from fits of these data to

Mechanism 2 are given in Table S3. In the time range from ~1 s to ~10 s which is of most

interest for these comparisons, $C_{eq} \approx [P]_{tot}$ based on the simulations in Figs. 2B and 7 which

show that [P] = [R] $\approx$ 0.5[P]$_{tot}$ in this time range.

Fitting to Mechanism 2:

The normalized fluorescence F_{norm} is interpreted as a sum of contributions from the I_1

intermediates and OC in Eq. S7.

$$F_{norm} = A_{I_{1E}} f_{I_{1E}} + A_{I_{1M}} f_{I_{1M}} + A_{I_{1L}} f_{I_{1L}} + A_{OC} f_{OC} + m_{drift} t \qquad \text{Eq.S7}$$

where $f_i = \frac{[i]}{[P]_{tot}}$ is the time-dependent population fraction of species i (i = I_{1E}, I_{1M}, I_{1L}, or OC) with

intrinsic signal amplitude A_i and m_{drift} is the slope of the linear drift correction. Justification for

this interpretation of F_{norm} is provided in the next section (Eqs. S8-14).

Values of F_{norm} vs. time from the high signal-to-noise (high S/N) Cy5 FRET acceptor

experiments in Fig. 3, which exhibit little if any drift when fit to Eq. S2, were fit interactively to

Mechanism 2 using Kintek Explorer[5,6] with the constraints described below to determine the

nine rate constants k_i and the FRET acceptor amplitudes A_i of the three I_1 intermediates and the

stable OC. Analyses in which a possible FRET contribution from RP$_C$ was included revealed

that A_{RP_C} is small and not significantly different from zero, consistent with it being the initial, not-

yet-wrapped closed complex.

As initial estimates of FRET amplitudes to use in fitting to Mechanism 2, we assumed a monotonic increase in FRET and arbitrarily set $A_{I_{1E}} = 0.3, A_{I_{1M}} = 0.4,\ A_{I_{1L}} = 0.8$ and $A_{OC} = 1.2$. The choice $A_{OC} = 1.2$ is based on the normalization F_{norm} = 1.0 (Eq. S1) at t = 400 s, where from the simulation in Fig 2B the population is approximately 80% OC. Initial estimates of rate constants for steps 2-5 of Mechanism 2, in which RP$_C$ is converted to OC via the series of I$_1$ intermediates, were obtained from the corresponding $1/\tau_{i,obs}$ values as described above. These initial estimates provide rapid equilibration of each step on the time scale of the next. Initial estimates of rate constants for the first step were k$_1$ ~6 x 10^8 M^{-1}s^{-1} and k$_{-1}$ ~60 s^{-1}, consistent with rapid equilibrium of this step on the time scale of the next, and consistent with an equilibrium constant K$_1$ = K$_{RPc}$ of ~10^7 M^{-1} as obtained from K$_{CC}$ (Fig. 2B) and initial estimates of the subsequent equilibrium constants (K$_i$ ≈1-2 for 2 ≤ i ≤ 4)) using Eq.S21-S22.

Reasonable fits to F_{norm} were obtained with fitted values of rate constants that differed from the initial estimates while maintaining a hierarchy so that each step equilibrated rapidly on the time scale of the next. Relative amplitudes of intermediates I$_{1E}$ and I$_{1M}$ inverted in these fits from the initial estimates so $A_{I_{1M}} < A_{I_{1E}} < A_{I_{1L}} \le A_{OC}$. However these initial fits failed to satisfy various constraints. In particular, the fraction of P near the end of the {CC} phase (t ~ 1 s) and/or the fraction of OC at late in OC formation (400 s) did not agree with the simulation in Fig 2B. Also, fitted rate constants for one or more intermediate steps differed significantly from expectations based on $1/\tau_{i,obs}$ values.

To refine the fit, pairs of rate constants k$_{-i}$, k$_{i+1}$ were linked by constraining each k$_{-i}$ to be at least two-to-three times greater than k$_{i+1}$ in order to have each step equilibrate on the time scale of the next. Then in interactive fitting k$_1$ and subsequent forward rate constants were varied to obtain a fraction of free promoter DNA (i.e. [P]/[P]$_{total}$) at 1 s of approximately 0.5 and a fraction of promoter DNA present as OC (i.e. [OC]/[P]$_{total}$) at 400 s of approximately 0.8, as expected from the simulation in Fig. 2B. Optimization of these fits yields the rate constants and

species amplitudes in Table S4, which provide high-quality fits to the data of Fig. 3 as shown in Fig. 6. In these optimized fits, the rank order of FRET amplitudes is $A_{I_{1M}} < A_{I_{1E}} < A_{I_{1L}} \leq A_{OC}$ for both Cy5-100 and Cy5+14 FRET experiments. These results differ from our initial assumption in fitting these data of a monotonic increase in FRET amplitude in the steps of OC formation.

To analyze the remaining FRET and PIFE data sets, many of which exhibited lower S/N ratios and/or larger drift than in the experiments analyzed above, rate constants were allowed to vary by up to ±10% from those obtained in the above analysis, and amplitudes of intermediates and OC were allowed to vary without any restriction in Berkeley Madonna[2] fitting. Global fitting of all FRET or PIFE kinetic data obtained for the same probe position(s) was not feasible because of the lower S/N ratio of many experiments, differences in the normalized signal at longer times (t> 50 s) in PIFE experiments, even after drift-correction, and because of differences in time points of data acquisition. High-quality fits over the entire time-range were obtained for more than 90% of the 61 data sets. Average amplitudes and uncertainties for each species detected are reported in Table S5. More than 75% of individual FRET acceptor experiments exhibit the same rank order of FRET amplitudes ($A_{I_{1M}} < A_{I_{1E}} < A_{I_{1L}} \leq A_{OC}$) and the others show only a single deviation from this order. Experimental FRET and PIFE kinetic curves from Figs. 3 and 4, corrected for linear long-time drift where necessary, are compared with those predicted from average rate constants (Table 3) and amplitudes (Table S5) in Fig S2.

Consistency Checks on the Rate Constants of the 5-Step Mechanism

Rate constants and equilibrium constants for the steps in Mechanism 2 (Table 3) pass several consistency tests.

a) Use of these rate and equilibrium constants in decay-to-equilibrium analyses of the reversible steps of Mechanism 2 (Eqs. S3-S7), predict the rate constants ($1/\tau_{i,obs}$ values) obtained from empirical 4-exponential fits to the FRET and PIFE kinetic data (Table S3).

b) These rate and equilibrium constants are quantitatively consistent with the filter binding and MnO_4^- results of Fig. 2. Eqs. S21-22 predict the composite {CC} binding constant K_{CC} (Fig. 2A) of Mechanism 1 from the equilibrium constants K_{RPc}, K_2, K_3 and K_4 of Mechanism 2 (in Table 3). Eqs. S19-20 predict the isomerization rate constant k_{isom} of Mechanism 1 from the DNA opening rate constant k_5 and equilibrium constants K_2, K_3 and K_4 of Mechanism 2 (in Table 3). Predicted and experimental values of K_{CC} and k_{isom} agree within the uncertainty, as shown in Table 3 footnotes.

c) Values of equilibrium constants for the steps converting RP_C to I_{1L} in the range 1 – 2 are consistent with real-time hydroxyl radical footprint of the equilibrated {CC} ensemble under similar conditions to those investigated here.[7] These exhibit a downstream boundary at +20, like that of the stable OC, indicating a dominant population of I_1 (i.e. post-RP_C) intermediates and hence values of $K_i \geq 1$ for at least some of these steps ($1 < i < 4$). Also, downstream protection for this {CC} is significantly weaker than for OC, and the far-upstream footprint is also of modest intensity, indicating that these contacts are not highly stable like those of the -10 and -35 elements and therefore that values of K_i for these steps are not much greater than 1. Rapid equilibrium also requires that values of K_i for these steps cannot be much greater than 1.

d) 5-step Mechanism 2 allows a quantitative interpretation of the large effect of upstream truncation (e.g. UT-47) and consequent loss of far-upstream wrapping on k_{isom} and the small effect of upstream truncation on K_{CC}[8], using Table 3 rate constants for steps that are proposed to be unaffected by truncation, as discussed in a subsequent section.

Conversion of Fluorescence Fitting Amplitudes to Relative Fluorescence Intensities

Here we obtain Eq. S7, used to fit normalized FRET acceptor and PIFE kinetic data, and interpret species amplitudes A_i obtained from these fits in terms of their fluorescence intensities a_i. The unnormalized fluorescence signal F, after correction for drift (if any) at long times, is the sum

$$F = F_{init} + f_P a_P + f_{RPc} a_{RPc} + f_{I_{1E}} a_{I_{1E}} + f_{I_{1M}} a_{I_{1M}} + f_{I_{1L}} a_{I_{1L}} + f_{OC} a_{OC} \qquad \text{Eq. S8}$$

where a_i is the fluorescence intensity of species i and $f_i = \frac{[i]}{[P]_{tot}}$ is its concentration fraction. In Eq. S8, i represents P, RP$_C$, I$_{1E}$, I$_{1M}$, I$_{1L}$, or OC and $\sum_i f_i = 1$.

For the present FRET and PIFE applications, $a_P = a_{RPc}$ and $a_{I_{1E}}, a_{I_{1M}}, a_{I_{1L}}$ and a_{OC} all are larger than a_P and a_{RPc}. We define intensity differences Δa_i between each species i and free promoter P:

$$\Delta a_i = a_i - a_P > 0, \ i = I_{1E}, I_{1M}, I_{1L}, OC \qquad \text{Eq. S9}$$

Because $f_P + f_{RPc} + f_{I_{1E}} + f_{I_{1M}} + f_{I_{1L}} + f_{OC}$ = 1, therefore

$$F = F_{init} + a_P + \Delta a_{I_{1E}} f_{I_{1E}} + \Delta a_{I_{1M}} f_{I_{1M}} + \Delta a_{I_{1L}} f_{I_{1L}} + \Delta a_{OC} f_{OC}$$

$$\cong a_p + \Delta a_{I_{1E}} f_{I_{1E}} + \Delta a_{I_{1M}} f_{I_{1M}} + \Delta a_{I_{1L}} f_{I_{1L}} + \Delta a_{OC} f_{OC} \qquad \text{Eq. S10}$$

because the background fluorescence F_{init} is small in comparison to the fluorescence a_P of P ($F_{init} \ll a_P$).

The minimum fluorescence F$_{min}$ (Eq. S1) is the fluorescence at t$_0$, the initial time (~10 ms) of data collection. Because no significant amount of any fluorescence-detected (i.e. I$_1$) intermediate is present at time t$_0$ (see Fig. 7), therefore $f_{I_{1E}}(t_0) \cong f_{I_{1M}}(t_0) \cong f_{I_{1L}}(t_0) \cong f_{OC}(t_0) \cong 0$ and from Eq. S10:

$$F_{min} \cong a_P \qquad \text{Eq. S11}$$

The maximum fluorescence F_{max} (Eq. S1) occurs at time t_{max}.

$$F_{max} \cong a_P + \Delta a_{I_{1E}} f_{I_{1E}}(t_{max}) + \Delta a_{I_{1M}} f_{I_{1M}}(t_{max}) + \Delta a_{I_{1L}} f_{I_{1L}}(t_{max}) + \Delta a_{OC} f_{OC}(t_{max})$$ Eq.S12

Therefore the normalized fluorescence F_{norm} is given by

$$F_{norm} = \frac{F - F_{min}}{F_{max} - F_{min}} = \frac{\Delta a_{I_{1E}} f_{I_{1E}} + \Delta a_{I_{1M}} f_{I_{1M}} + \Delta a_{I_{1L}} f_{I_{1L}} + \Delta a_{OC} f_{OC}}{\Delta a_{I_{1E}} f_{I_{1E}}(t_{max}) + \Delta a_{I_{1M}} f_{I_{1M}}(t_{max}) + \Delta a_{I_{1L}} f_{I_{1L}}(t_{max}) + \Delta a_{OC} f_{OC}(t_{max})}$$ Eq. S13

which is the same as Eq. S7 with

$$A_i = \frac{\Delta a_i}{\Delta a_{I_{1E}} f_{I_{1E}}(t_{max}) + \Delta a_{I_{1M}} f_{I_{1M}}(t_{max}) + \Delta a_{I_{1L}} f_{I_{1L}}(t_{max}) + \Delta a_{OC} f_{OC}(t_{max})}$$ Eq. S14

where i represents I_{1E} , I_{1M}, I_{1L} , or OC.

From Eqs. S11-12,

$$\frac{F_{max} - F_{min}}{F_{min}} = \frac{F_{max}}{F_{min}} - 1 \cong \frac{\Delta a_{I_{1E}} f_{I_{1E}}(t_{max}) + \Delta a_{I_{1M}} f_{I_{1M}}(t_{max}) + \Delta a_{I_{1L}} f_{I_{1L}}(t_{max}) + \Delta a_{OC} f_{OC}(t_{max})}{a_P}$$ Eq. S15

and from Eqs. S14-15

$$A_i = \frac{\Delta a_i}{a_P} \left(\frac{F_{max}}{F_{min}} - 1\right)^{-1}$$ Eq. S16

The average value designated $\langle \frac{F_{max}}{F_{min}} - 1 \rangle$ is obtained by first averaging factors $\left(\frac{F_{max}}{F_{min}} - 1\right)$ for the series of shots in each experiment and then averaging over all PIFE experiments of a given type. These $\langle \frac{F_{max}}{F_{min}} - 1 \rangle$ values and average fitting amplitudes $\langle A_i \rangle$ are reported in Table S5, together with uncertainties in these quantities. Average PIFE intensities $\langle \frac{a_i}{a_P} \rangle$ relative to promoter DNA are obtained from these quantities by Eqs. S17 (the average of Eq. S16) and reported in Table 1.

$$\langle \frac{a_i}{a_P} \rangle = 1 + \langle A_i \rangle \langle \frac{F_{max}}{F_{min}} - 1 \rangle$$ Eq. S17

Estimation of Distance between Upstream (-100) and Downstream (+14) DNA in Wrapped RNAP-promoter Complexes from FRET Amplitudes

Cy5+14 and Cy5-100 FRET fitting amplitudes (Table S5) were normalized by the OC fitting amplitude to obtain the values listed in Table 1. These are averaged and converted to FRET efficiency values of intermediates I_{1E}, I_{1M} and I_{1L} in Table 2, using the published OC FRET efficiency (0.32).[1] These FRET efficiencies are converted to distances using Eq. S18:

$$R = R_0 \left(\frac{1}{E} - 1\right)^{1/6} \hspace{4cm} \text{Eq. S18}$$

with the Forster radius previously estimated for this DNA at 19 °C (R_0 = 56 ± 1 A). Illustrations of intermediates with these FRET distances are shown in Fig. 8.

Interpretation of Rate and Equilibrium Constants of the Classic Two-Step Mechanism 1 (k_{isom}, K_{CC}, k_a) Using Mechanism 2

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 isomerization rate constant k_{isom} and on the composite equilibrium constant for formation of the {CC} ensemble, K_{CC}. Interpretation of these effects using relationships between these parameters of Mechanism 1 and the rate and equilibrium constants (Table 3) of 5-step Mechanism 2 provides new insight into how promoter sequence, upstream truncation and factors regulate the rate of OC formation and transcription initiation.

Because each step of {CC} ensemble formation equilibrates rapidly on the time scale of the next, k_{isom} of Mechanism 1 is related to k_5, the rate constant of the DNA-opening step in Mechanism 2, by

$$k_{isom} = f_{I1L} \, k_5 \hspace{4cm} \text{Eq. S19}$$

where f_{I1L} is the fraction of the equilibrated {CC} ensemble that is I_{1L}:

$$f_{I1L} = \left(\frac{[I_{1L}]}{[\{CC\}]}\right)_{eq} = \frac{K_{I1E} \, K_{I1M} \, K_{I1L}}{\left(1 + K_{I1E} \left(1 + K_{I1M} \left(1 + K_{I1L}\right)\right)\right)} \hspace{2cm} \text{Eq. S20}$$

where $\qquad [\{CC\}] = [RP_c] + [I_{1E}] + [I_{1M}] + [I_{1L}]$

The {CC} ensemble is equilibrated at t ≥ 10 s (the onset of OC formation) for the conditions investigated here (Fig 7).

For this rapid equilibrium situation, K_{CC} of Mechanism 1 is related to the equilibrium constant for forming the initial closed complex, K_{RPc}, by

$$K_{CC} = \left(\frac{[\{CC\}]}{[R][P]}\right)_{eq} = \frac{K_{RPc}}{f_{RPc}} \qquad \text{Eq. S21}$$

where f_{RPc} is the fraction of the equilibrated {CC} that is RP_c:

$$f_{RPc} = \left(\frac{[RP_C]}{[\{CC\}]}\right)_{eq} = \frac{1}{\left(1+K_{I1E}\left(1+K_{I1M}\left(1+K_{I1L}\right)\right)\right)} \qquad \text{Eq. S22}$$

Eqs. S19-22 show that $k_{isom} < k_5$ and $K_{CC} > K_{RPc}$. The overall second order rate constant for OC formation (k_a) is interpreted using these two mechanisms as

$$k_a = K_{CC}k_{isom} = K_{RPc}K_{I1E}\, K_{I1M}\, K_{I1L}\, k_5 \qquad \text{Eq. S23}$$

In the equilibrated {CC} ensemble, population fractions of I_{1E} and I_{1M} intermediates are

$$f_{I1LE} = \left(\frac{[I_{1E}]}{[\{CC\}]}\right)_{eq} = \left(\frac{[I_{1E}]}{[RP_C]}\right)_{eq} f_{RPc} = \frac{K_{I1E}}{\left(1+K_{I1E}\left(1+K_{I1M}\left(1+K_{I1L}\right)\right)\right)} \qquad \text{Eq. S24}$$

$$f_{I1LM} = \left(\frac{[I_{1,M}]}{[\{CC\}]}\right)_{eq} = \left(\frac{[I_{1M}]}{[I_{1E}]}\right)_{eq} f_{1E} = \frac{K_{I1M}\, K_{I1E}}{\left(1+K_{I1E}\left(1+K_{I1M}\left(1+K_{I1L}\right)\right)\right)} \qquad \text{Eq. S25}$$

Effects of promoter sequence, upstream DNA length and transcription factors on the parameters of Mechanism 1 (k_{isom}, K_{CC}, and k_a) can arise from effects of these variables on the distribution of complexes in {CC} (i.e. on f_{RPc} and f_{I1L}) as well as from effects of these variables on K_{RPc} and k_5. Based on the limited evidence available to date, and by analogy with regulation of enzyme catalysis, we proposed previously that the 100-fold or greater range of k_{isom} for different upstream promoter truncations, different promoter sequences, and/or addition of

transcription factors results primarily from an equally wide range in the fraction f_{I1L} of the equilibrated {CC} ensemble that is I_{1L}, the intermediate that is poised for DNA opening, and not primarily from changes in the DNA opening rate constant k_5 with these variables.[9] An example is provided in the Discussion. This proposal is supported by the finding that, for λP_R and T7A1 promoters and wild-type and variant RNAP, k_{-5}, the DNA closing rate, is insensitive to promoter identity and to salt or solute concentration, and primarily a function of temperature[10-12]

A fundamental analogy exists between Mechanism 2 for open complex formation in the RNAP clamp and mechanisms of enzyme action. Regulation of the velocity of product formation in noncooperative and cooperative enzyme-catalyzed reactions is often 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 convert it to the more advanced members of the CC ensemble (I_{1E}, I_{1M}, I_{1L}). These reversible steps are analogous to reversible substrate binding steps in enzyme kinetic mechanisms, and in-cleft DNA opening-closing is the analog of the catalytic step.

Calculations for Free Energy vs. Progress Diagram and {CC} Population Distribution for Formation at UT-47 λP_R.

Assuming the same DNA-opening rate constant (k_5) for UT-47 and FL λP_R at 19 °C, then from the experimental k_{isom}^{UT} it follows from Eq. S19 that $f_{I1L}^{UT} = 0.006$. Hence, less than 1% of the equilibrated UT-47 {CC} ensemble is I_{1L} as compared to 33% for FL, explaining the 60-fold difference in isomerization rate constants. The remaining 99% of the UT-47 {CC} ensemble is a mixture of RP_C, I_{1E}, and possibly I_{1M}. The large reduction in k_{isom} for UT-47 indicates a large reduction in K_{I1L} and/or K_{I1M}.

For the progress diagram in Fig. 9, we assume that $K_{I1E}^{UT} = K_{I1E}^{FL} = 1.8$ and that $K_{I1M}^{UT} = K_{I1L}^{UT} << 1$. In this case, we calculate $K_{I1M}^{UT} = K_{I1L}^{UT} = 0.098$ as compared to $K_{I1M}^{FL} = 1.5$ and $K_{I1L}^{FL} = 1.0$ for FL (Table 3). From these equilibrium constants, the fully-formed UT-47 {CC} ensemble is predicted to be 33% RP_C, 60% I_{1E}, 6% I1M, and only 0.6% I_{1L}, consistent with the {CC} ensemble footprinting result for UT-47. For comparison, the {CC} ensemble for FL λP_R (Fig. 8A) is 12% RP_C, 22% I_{1E}, 33% I_{1M} and 33% I_{1L}.

References:

1) 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.,Jr. (2016) Fluorescence resonance energy transfer characterization of DNA wrapping in closed and open *Escherichia coli* RNA polymerase-λP_R promoter complexes, *Biochemistry 55*, 2174-2186.

2) Macey, R., Oster, G., Zahnley, T., Berkeley Madonna User's Guide, University of California, Department of Molecular and Cellular Biology, Berkeley, CA 94720, http://www.berkeleymadonna.com

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

4) Bernasconi, C.F., Relaxation Kinetics, New York : Academic Press, 1976.

5) Johnson, K., Simpson, Z. and Blom, T., Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data. Analytical Biochemistry (2009) 387, 20-29.

6) Johnson, K., Simpson, Z. and Blom, T., FitSpace Explorer: An algorithm to evaluate multidimensional parameter space in fitting kinetic data. Analytical Biochemistry (2009) 387, 30-41.

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

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

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

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

13) Drennan, A., Kraemer, M., Capp, M., Gries, T., Ruff, E., 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.

SI Figures
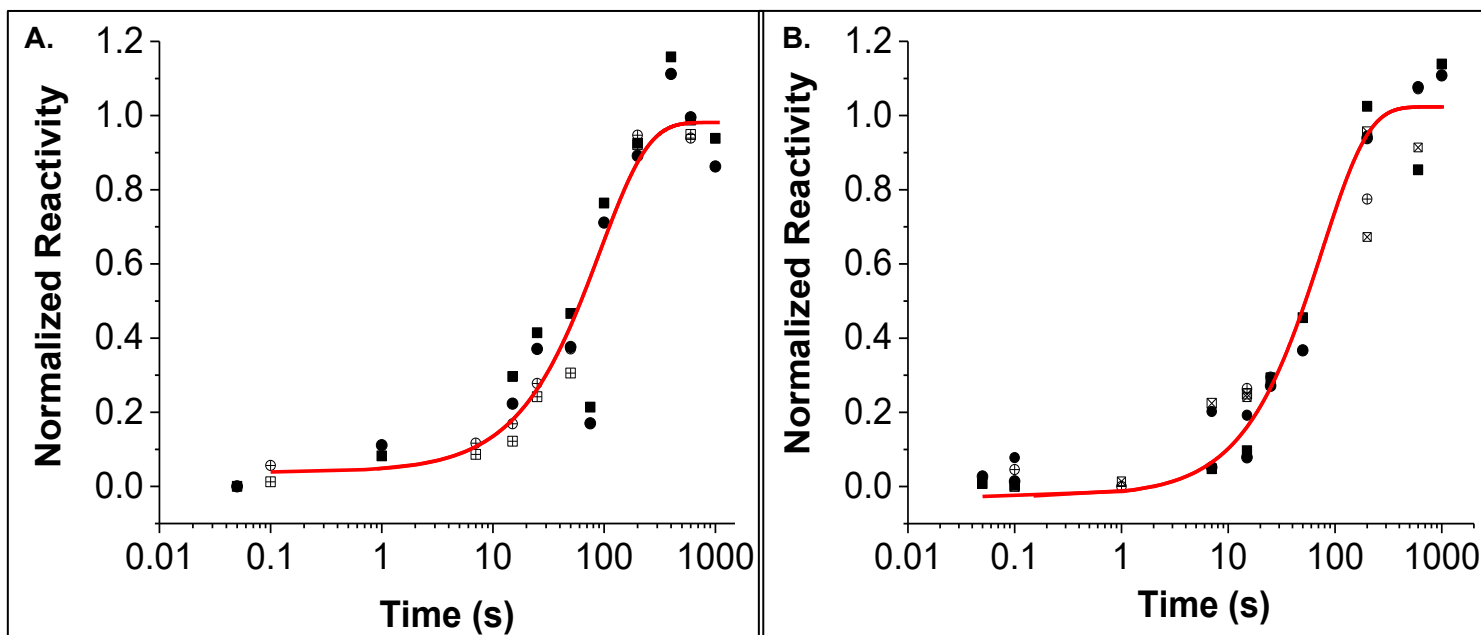
Figure S1: MnO$_4^-$-Detected Kinetics of OC Formation on a Log Time Scale. Results in Fig. 1 of the text for the kinetics of OC formation monitored by fast permanganate footprinting are replotted on a log time scale to ascertain if the entire time course is described by single-exponential kinetics. Symbols for different thymines and fitted curves are the same as in Fig. 1 and correspond to rate constants k_{obs} = 0.011 ± 0.002 s^{-1} (non-template strand, Panel A) and k_{obs} = 0.010 ± 0.002 s^{-1} (template strand, Panel B). These comparisons reveal that all reactive thymines on each strand exhibit very similar single-exponential kinetics of OC formation.

A FRET

Cy5-100 experiment
Cy5-100 prediction
Cy5 +14 experiment
Cy5 +14 prediction

Normalized Fluorescence

Time in seconds (log scale)

B +14 PIFE

Cy3 experiment
Cy3 prediction
Cy5 experiment
Cy5 prediction

Normalized Fluorescence

Time in seconds (log scale)

C -100 PIFE

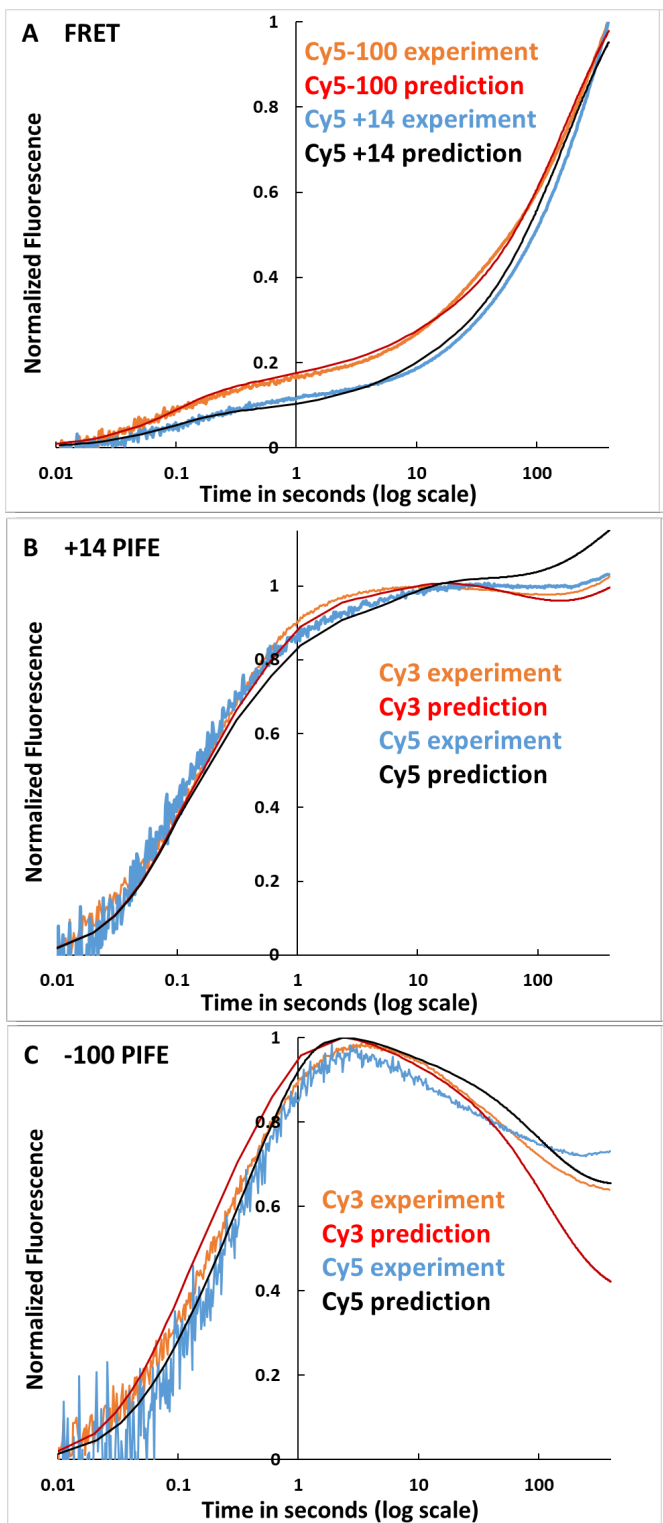Cy3 experiment
Cy3 prediction
Cy5 experiment
Cy5 prediction

Normalized Fluorescence

Time in seconds (log scale)

Figure S2: Comparison of Fitted (Predicted) and Observed FRET and PIFE Kinetics of OC Formation. Comparison of FRET and PIFE kinetic data (as in Figs. 3,4, but after a linear drift correction) with predictions using average rate constants (Table 3) and amplitudes (Table S5).

Panel A: **FRET** for Cy5 -100 —— ; Cy5 +14 ——.

Panel B: **PIFE** for Cy3 +14 —— ; Cy5 +14 ——.

Panel C: PIFE for Cy3 -100 —— ; Cy5 -100 ——.

Elimination of signal drift from the +14 PIFE data (Figs. 4, 6) reveals a small increase in +14 PIFE at t >50 s. A similar but smaller effect is observed for -100 PIFE. The late +14 PIFE increase results from the conversion of remaining free P and RP_C (no PIFE intensity) to OC (modest PIFE intensity), which more than counterbalances the effect of converting higher-PIFE I_1 intermediates to OC.

16

Table S1. Promoter DNA Constructs

Promoter	Template	Upstream primer	Downstream primer	Duplex (bp) and position relative to start site (+1)[a]	Experiment used in
λP$_R$ (-59 to +34)[b]	Unlabeled λP$_R$	Up	Down	192 (-128 to +64)	Permanganate footprinting kinetics
Cy3(+14) λP$_R$ [c]	Unlabeled λP$_R$	Up	Cy3(+14) Down	142 (-128 to +14)	Fluorescence kinetics
Cy5(+14) λP$_R$ [c]	Unlabeled λP$_R$	Up	Cy5(+14) Down	142 (-128 to +14)	
Cy3(-100) λP$_R$ [c]	Unlabeled λP$_R$	Cy3(-100) Up	Down	164 (-100 to +64)	
Cy5(-100) λP$_R$ [c]	Unlabeled λP$_R$	Cy5(-100) Up	Down	164 (-100 to +64)	
Cy3(+14)-Cy5 (-100) λP$_R$ [c]	Unlabeled λP$_R$	Cy5(-100) Up	Cy3(+14) Down	114 (-100 to +14)	
Cy5 (+14)- Cy3 (-100) λP$_R$ [c]	Unlabeled λP$_R$	Cy3(-100) Up	Cy5(+14) Down	114 (-100 to +14)	
[a] Dye-labeled DNAs have 12 base (upstream) and/or 8 base (downstream) ssDNA termini					
[b] SI Ref. 13 [c] SI Ref. 1					

Table S2. PCR Primer Sequences

Primer	Sequence
Up	5'-GTACGAATTCGATATCCAGCTATGACCATGATTACGCCAAGC-3'
Up1	5'-CAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGT-3'
Down	5'-CAGGACCCGGGAAGCTTTTAATTAACACTCTTATACATTATTCC-3'
Down1	5'-CCATACAACCTCCTTACTACATGCAACCATTATCACCGCC-3'
Down2	5'-CCTCCTTACTACATGCAACCATTATCACCGCC-3'
Cy3 (+14)	5'-CCATACAA/iCy3/CCTCCTTACTACATGCAACCATTATCACCGCC-3'
Cy5 (+14)	5'-CCATACAA/iCy5/CCTCCTTACTACATGCAACCATTATCACCGCC-3'
Cy3(-100)	5'-CAGCTATGACCA/iCy3/TGATTACGCCAAGC-3'
Cy5 (-100)	5'-CAGCTATGACCA/iCy5/TGATTACGCCAAGC-3'

Table S3. Comparison of Observed and Predicted $\tau_{i,obs}$ Values

	Observed $1/\tau_{i,obs}$ from FRET, PIFE Fits ($\pm$ 1 SD)	Predicted $1/\tau_{i,pred}$ from Rate Constants (Table 3)[a]
$1/\tau_{1,obs}$	16 $\pm$ 12 s^{-1}	10.5 $\pm$ 0.6 s^{-1}
$1/\tau_{2,obs}$	2.6 $\pm$ 1.5 s^{-1}	2.4 $\pm$ 0.1 s^{-1}
$1/\tau_{3,obs}$	0.21 $\pm$ 0.18 s^{-1}	0.12 $\pm$ 0.01 s^{-1}
$1/\tau_{4,obs}$	0.020 $\pm$ 0.019 s^{-1}	0.010 $\pm$ 0.002 s^{-1}

[a] **Predicted using Eqs. S3-S6 and the finding from the simulation in Fig. 7 that in the time range of most interest (1 s -10 s), C_{eq} = ([P]$_{eq}$ + [R]$_{eq}$) $\approx$ [P]$_{tot}$, the total promoter concentration.**

Table S4. Rate Constants, Amplitudes of FRET and PIFE Kinetics Experiments (Fig. 6)

	FRET				PIFE			
	Cy5 +14		Cy5 -100		Cy3 +14	Cy5 +14	Cy3 -100	Cy5 -100
	Initial[a]	Refined[b]	Initial[a]	Refined[b]	Refined[b]	Refined[b]	Refined[b]	Refined[b]
k_1 (M^{-1} s^{-1})	2.8×10^8	3.0×10^8	3.1×10^8	2.8×10^8	3.4×10^8	3.0×10^8	3.4×10^8	2.8×10^8
k_{-1} (s^{-1})	40	48	44	43	42	43	40	44
k_2 (s^{-1})	13	14	15	13	14	14	12	14
k_{-2} (s^{-1})	6.7	7.5	7.4	7.1	7.1	7.1	7.0	6.6
k_3 (s^{-1})	2.2	2.3	2.5	2.4	2.5	2.6	2.4	2.5
k_{-3} (s^{-1})	1.7	1.9	1.9	1.7	1.6	1.9	1.6	1.6
k_4 (s^{-1})	0.10	0.098	0.088	0.096	0.096	0.090	0.11	0.097
k_{-4} (s^{-1})	0.10	0.11	0.088	0.091	0.098	0.10	0.092	0.098
k_5 (s^{-1})	0.040	0.040	0.036	0.038	0.041	0.039	0.042	0.044
A_{I1E}	0.36	0.37	0.61	0.68	2.4	2.9	2.1	1.5
A_{I1M}	0.25	0.29	0.20	0.20	2.0	1.7	2.5	3.0
A_{I1L}	0.51	0.69	1.2	1.1	1.5	2.1	0.83	0.33
A_{OC}	1.2	1.3	1.2	1.2	1.1	1.2	0.67	0.79
RMS [c]	0.014	0.0081	0.01	0.009	0.014	0.023	0.019	0.054

[a] Initial Kintek fit with no drift correction

[b] Refined Berkley Madonna fit using Eq. S7 with a linear drift correction as in Eq. S2 and Kintek average rate constants $\pm$ 10%.

[c] RMS (root mean square) error calculated by Berkeley-Madonna for fits in Fig 6.

Table S5. Average FRET Acceptor and PIFE Fitting Amplitudes $\langle A_i \rangle$ of CC Intermediates and OC and Conversion Factors $\langle \frac{F_{max}}{F_{min}} - 1 \rangle$ [a]

	Average FRET Acceptor Fitting Amplitude $\langle A_i \rangle$		Average PIFE Fitting Amplitude $\langle A_i \rangle$			
	Cy5+14	Cy5-100	Cy3+14	Cy5+14	Cy3-100	Cy5-100
[b]RP$_C$	0	0	0	0	0	0
I$_{1E}$	0.40 ± 0.15	0.67 ± 0.3	2.5 ± 0.8	2.3 ± 0.8	2.4 ± 0.4	1.6 ± 0.5
I$_{1M}$	0.15 ± 0.1	0.28 ± 0.16	2.0 ± 0.6	1.7 ± 0.6	2.2 ± 0.5	2.7 ± 0.5
I$_{1L}$	0.83 ± 0.3	0.94 ± 0.2	1.6 ± 0.2	1.9 ± 0.2	0.5 ± 0.5	0.6 ± 0.7
OC	1.2 ± 0.08	1.2 ± 0.07	1.0 ± 0.1	1.1 ± 0.2	0.3 ± 0.3	0.6 ± 0.4
$\langle \frac{F_{max}}{F_{min}} - 1 \rangle$			0.27 ± 0.1	0.066 ± 0.028	0.085 ± 0.044	0.039 ± 0.014
RMS [c]	0.025	0.012	0.017	0.044	0.086	0.073

[a] Eq. S17

[b] The absence of significant FRET and PIFE in RP$_C$ was judged from comparisons of fits in which this term was included or excluded (see text).

[c] RMS (root mean square) error calculated by Berkeley-Madonna for fits in Fig S2.