

Transcription initiation at a consensus bacterial promoter proceeds via a ‘bind-unwind-load-and-lock’ mechanism

Abhishek Mazumder^{1*}, Richard H Ebright², Achillefs N Kapanidis^{1*}

¹Biological Physics Research Group, Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, United Kingdom; ²Waksman Institute and Department of Chemistry, Rutgers University, Piscataway, United States

Abstract Transcription initiation starts with unwinding of promoter DNA by RNA polymerase (RNAP) to form a catalytically competent RNAP-promoter complex (RPo). Despite extensive study, the mechanism of promoter unwinding has remained unclear, in part due to the transient nature of intermediates on path to RPo. Here, using single-molecule unwinding-induced fluorescence enhancement to monitor promoter unwinding, and single-molecule fluorescence resonance energy transfer to monitor RNAP clamp conformation, we analyse RPo formation at a consensus bacterial core promoter. We find that the RNAP clamp is closed during promoter binding, remains closed during promoter unwinding, and then closes further, locking the unwound DNA in the RNAP active-centre cleft. Our work defines a new, ‘bind-unwind-load-and-lock’, model for the series of conformational changes occurring during promoter unwinding at a consensus bacterial promoter and provides the tools needed to examine the process in other organisms and at other promoters.

***For correspondence:**

abhishek.mazumder@physics.ox.ac.uk (AM);

kapanidis@physics.ox.ac.uk (AK);
kapanidis@physics.ox.ac.uk (ANK)

Competing interest: The authors declare that no competing interests exist.

Funding: See page 15

Preprinted: 28 March 2021

Received: 06 May 2021

Accepted: 06 October 2021

Published: 11 October 2021

Reviewing Editor: Maria Spies, University of Iowa, United States

© Copyright Mazumder et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

Transcription initiation is the first and most highly regulated step in gene expression (*Mazumder and Kapanidis, 2019; Ruff et al., 2015*). During transcription initiation, RNA polymerase (RNAP), together with the transcription initiation factor σ , unwinds ~13 bp of promoter DNA to form a ‘transcription bubble’, and places the template-strand ssDNA of the unwound transcription bubble in contact with the RNAP active centre, yielding a catalytically competent RNAP-promoter transcription-initiation complex (RPo; *Mazumder and Kapanidis, 2019*). High-resolution structures of RPo define the contacts that RNAP and σ make with promoter DNA, as well as the conformation and interactions of unwound template-strand ssDNA engaging the RNAP active centre (*Bae et al., 2015; Zuo and Steitz, 2015; Zhang et al., 2012*). Structural and biochemical experiments suggest that transcription-bubble formation is initiated by unwinding the DNA base pair at the upstream end of the transcription bubble – breaking the base pair, and unstacking and flipping the non-template-strand base of the broken base pair, and inserting the unstacked and flipped non-template-strand base into a protein pocket of σ – followed by propagation of the unwinding in a downstream direction (*Mazumder and Kapanidis, 2019; Ruff et al., 2015; Zhang et al., 2012; Feklistov and Darst, 2011*). However, the mechanism by which DNA is unwound and loaded into the RNAP active-centre cleft has remained controversial (reviewed in *Mazumder and Kapanidis, 2019*). In crystal structures of RNAP σ holoenzyme, the RNAP active-centre cleft is too narrow to accommodate double-stranded DNA (dsDNA) (<20 Å), and σ obstructs access of dsDNA to the RNAP active-centre cleft (*Mekler et al., 2002; Murakami et al., 2002b; Murakami et al., 2002a; Vassylyev et al., 2002*). As a result, there is no unobstructed path by which dsDNA can access the RNAP active-centre cleft in RNAP holoenzyme (*Mekler et al., 2002; Murakami et al., 2002b*;

Murakami et al., 2002a; Vassilyev et al., 2002; Young et al., 2002; Murakami and Darst, 2003). Accordingly, it has remained unclear where, when, and how promoter DNA is unwound and loaded into the RNAP active-centre cleft.

To address these questions, two classes of models have been proposed. One class of models, termed 'load-unwind' models, propose that: (i) the RNAP active-centre cleft opens, through the swinging outward of one wall, termed the 'clamp', of the active-centre cleft, allowing loading of promoter DNA into the active-centre cleft as dsDNA; (ii) promoter DNA unwinds inside the active-centre cleft; and (iii) the active-centre cleft closes, through the swinging inward of the clamp, during or after DNA unwinding (**Mazumder and Kapanidis, 2019; Ruff et al., 2015; Mekler et al., 2002; Murakami et al., 2002b; Murakami et al., 2002a; Young et al., 2002; Darst et al., 2002; Chakraborty et al., 2012**). The other class of models, termed 'unwind-load' models, proposes that: (i) promoter DNA unwinds outside the RNAP active-centre cleft and (ii) unwound-promoter DNA loads into the active-centre cleft as single-stranded DNA (**Mazumder and Kapanidis, 2019; Vassilyev et al., 2002**). Some versions of the unwind-load model postulate that opening and closing of the RNAP clamp is required for DNA unwinding and DNA loading to occur (**Feklistov et al., 2017; Boyaci et al., 2019**). Other versions of the unwind-load model postulate that opening and closing motions of the RNAP clamp is not required for DNA unwinding and DNA loading (**Chen et al., 2020**).

Some previous results have been interpreted as supporting *load-unwind* models, including results from static DNA footprinting of trapped putative on-pathway intermediates in formation of RPo, suggesting the presence of dsDNA inside the active-centre cleft (**Spassky et al., 1985; Duval-Valentin and Ehrlich, 1986; Cowing et al., 1989; Schickor et al., 1990**), results from kinetic DNA footprinting suggesting the existence of on-pathway intermediates having dsDNA inside the active-centre cleft (**Davis et al., 2007**), fluorescence resonance energy transfer (FRET) results showing clamp opening and closing in RNAP and in trapped putative on-pathway intermediates in formation of RPo (**Chakraborty et al., 2012; Lin et al., 2018; Duchi et al., 2018; Mazumder et al., 2021**), functional correlations between inhibition of clamp opening and closing with inhibition of formation of RPo (**Chakraborty et al., 2012; Lin et al., 2018; Duchi et al., 2018**), and crystal and cryo-EM structures of trapped putative on-pathway intermediates in formation of RPo having dsDNA inside an open active-centre cleft (**Glyde et al., 2018**).

Other previous results have been interpreted as supporting *unwind-load* models, including time-resolved footprinting experiments suggesting that promoter unwinding occurs outside the active-centre cleft and precedes rate-limiting conformational changes in RNAP (**Rogozina et al., 2009**), and cryo-EM structures of trapped putative on-pathway intermediates in formation of RPo containing partly unwound DNA inside a closed active-centre cleft (**Chen et al., 2020**).

However, the previous results either have relied on analysis of artificially trapped complexes that have not been firmly established to correspond to *bona fide* on-pathway intermediates, or have relied on analysis of ensemble kinetics for which the identities and orders of appearance of intermediates have not been firmly established (**Chakraborty et al., 2012; Spassky et al., 1985; Duval-Valentin and Ehrlich, 1986; Cowing et al., 1989; Schickor et al., 1990; Davis et al., 2007; Lin et al., 2018; Duchi et al., 2018; Mazumder et al., 2021; Glyde et al., 2018; Rogozina et al., 2009**). Moreover, the previous results have been complicated by differences in the species source of the RNAP analysed, differences in the σ factors analysed, and differences in the promoter sequences analysed (**Chakraborty et al., 2012; Spassky et al., 1985; Duval-Valentin and Ehrlich, 1986; Cowing et al., 1989; Schickor et al., 1990; Davis et al., 2007; Lin et al., 2018; Duchi et al., 2018; Mazumder et al., 2021; Glyde et al., 2018; Rogozina et al., 2009**). Here, we report the use of single-molecule kinetic studies to define the pathway of DNA unwinding and DNA loading, without the assumptions that complicate analysis of artificially trapped complexes, and without the heterogeneity, population averaging, and time averaging that complicate analysis of ensemble kinetics. We used single-molecule promoter unwinding-induced fluorescence enhancement (smUIFE) to monitor DNA unwinding in solution in real time during formation of RPo, and we used single-molecule fluorescence resonance energy transfer (smFRET) to monitor opening and closing of the RNAP clamp in solution in real time during formation of RPo. In all experiments, we analysed *Escherichia coli* RNAP σ^{70} holoenzyme at a consensus bacterial core promoter, comprising a consensus -35 element, a consensus -10 element, and a consensus $-35/-10$ spacer.

Results

Single-molecule UIF

Previous work indicates that a promoter derivative having the fluorescent probe Cy3 site-specifically incorporated in the transcription-bubble region exhibits an ~2 -fold increase in fluorescence emission intensity upon promoter unwinding during RPo formation and exhibits an ~ 2-fold decrease in fluorescence emission intensity upon promoter rewinding during promoter escape (Feklistov *et al.*, 2017; Ko and Heyduk, 2014; Koh *et al.*, 2018). These changes in fluorescence emission intensity provide a powerful approach to monitor promoter unwinding and rewinding in solution during transcription initiation (Feklistov *et al.*, 2017; Ko and Heyduk, 2014; Koh *et al.*, 2018). Here, we have adapted this approach to enable detection of promoter unwinding in solution at the single-molecule level in real time, and we designate our adaptation of the approach 'smUIFE', to underscore the similarity to the established method of single-molecule protein-induced fluorescence enhancement (Hwang and Myong, 2014; Figure 1A).

First, we analysed a promoter DNA fragment having Cy3 incorporated at a site at the downstream edge of the transcription bubble (non-template-strand position +2) of a consensus σ^{70} -dependent bacterial promoter (lacCONS-[+ 2 Cy3]; Figure 1—figure supplement 1A). Upon adding the Cy3-containing promoter DNA fragment to *E. coli* RNAP σ^{70} holoenzyme immobilised on a coverslip mounted in a total-internal-reflection fluorescence (TIRF) microscope, we detected the appearance of fluorescence signal from single fluorescent species, indicating binding of single molecules of Cy3-containing promoter DNA fragment to surface-immobilised single molecules of RNAP holoenzyme (Figure 1A, bottom). Control experiments show that: (i) Cy3-containing promoter DNA fragments bound exclusively to immobilised RNAP molecules (Figure 1—figure supplement 2A), and the majority (~60%) of the resulting complexes were resistant to challenge with heparin (Figure 1—figure supplement 2B); (ii) binding of Cy3-containing non-promoter DNA fragments to immobilised RNAP holoenzyme molecules did not result in heparin-resistant complexes (Figure 1—figure supplement 2C); and (iii) binding of Cy3-containing promoter DNA fragments to immobilised RNAP core enzyme molecules did not result in heparin-resistant complexes (Figure 1—figure supplement 2D). These results confirm that sequence-specific, heparin-resistant, complexes were formed only between RNAP holoenzyme and Cy3-containing promoter DNA fragments.

We extracted intensity vs. time trajectories for the formation of RNAP-promoter complexes, and identified different classes of trajectories. A large class of trajectories (~45%) was characterised by the appearance of a fluorescence intensity of ~200 counts, followed by an increase in fluorescence intensity to ~450 counts, followed by either a decrease of the intensity to ~200 counts or a disappearance of the intensity (Figure 1B, top left; Figure 1—figure supplement 3, middle). Control experiments show that (i) binding of Cy3-containing non-promoter DNA fragments to immobilised RNAP holoenzyme did not result in ≥ 2 -fold fluorescence enhancement (Figure 1—figure supplement 3C); and (ii) binding of Cy3-containing promoter DNA fragments to immobilised RNAP core molecules did not result in ≥ 2 -fold fluorescence enhancement events for the overwhelming majority of the time trajectories (~96 %; Figure 1—figure supplement 3D). These results indicate that the detected ~2 -fold fluorescence enhancement events represent sequence-specific complexes between RNAP holoenzyme and promoter DNA fragments. Therefore, we assigned the states with no intensity as species that lack promoter DNA or that have promoter DNA with photobleached Cy3; the states with an intensity level of ~200 counts as RNAP-promoter complexes having dsDNA at the Cy3 incorporation site; and states with an intensity level ~450 counts as RNAP-promoter complexes having single-stranded DNA at the Cy3 incorporation site. This assignment yields a reaction sequence comprising: binding of double-stranded promoter DNA to RNAP (~200 counts; 'pre-unwinding state'), followed by promoter unwinding (~450 counts; 'unwound state'), followed by promoter rewinding (~200 counts) or probe photobleaching (~0 counts). Upon addition of an NTP subset enabling synthesis of transcripts up to 11 nt in length (ATP, UTP, and GTP), ~ 49 % of trajectories showed formation of a stable unwound-promoter state (with fluorescence intensity similar to that of the sub-population exhibiting promoter unwinding), as expected for the formation of initial transcribing complexes (RPitc ≤ 11 ; Figure 1—figure supplement 3B, top; see also Materials and methods). Upon subsequent addition of a separate NTP subset (GTP and CTP), enabling synthesis of transcripts up to 14 nt in length, ~ 27 % of the resulting complexes showed formation of a stable rewound-promoter state (Figure 1—figure supplement 3B, bottom), as expected for promoter escape and formation of a transcription elongation

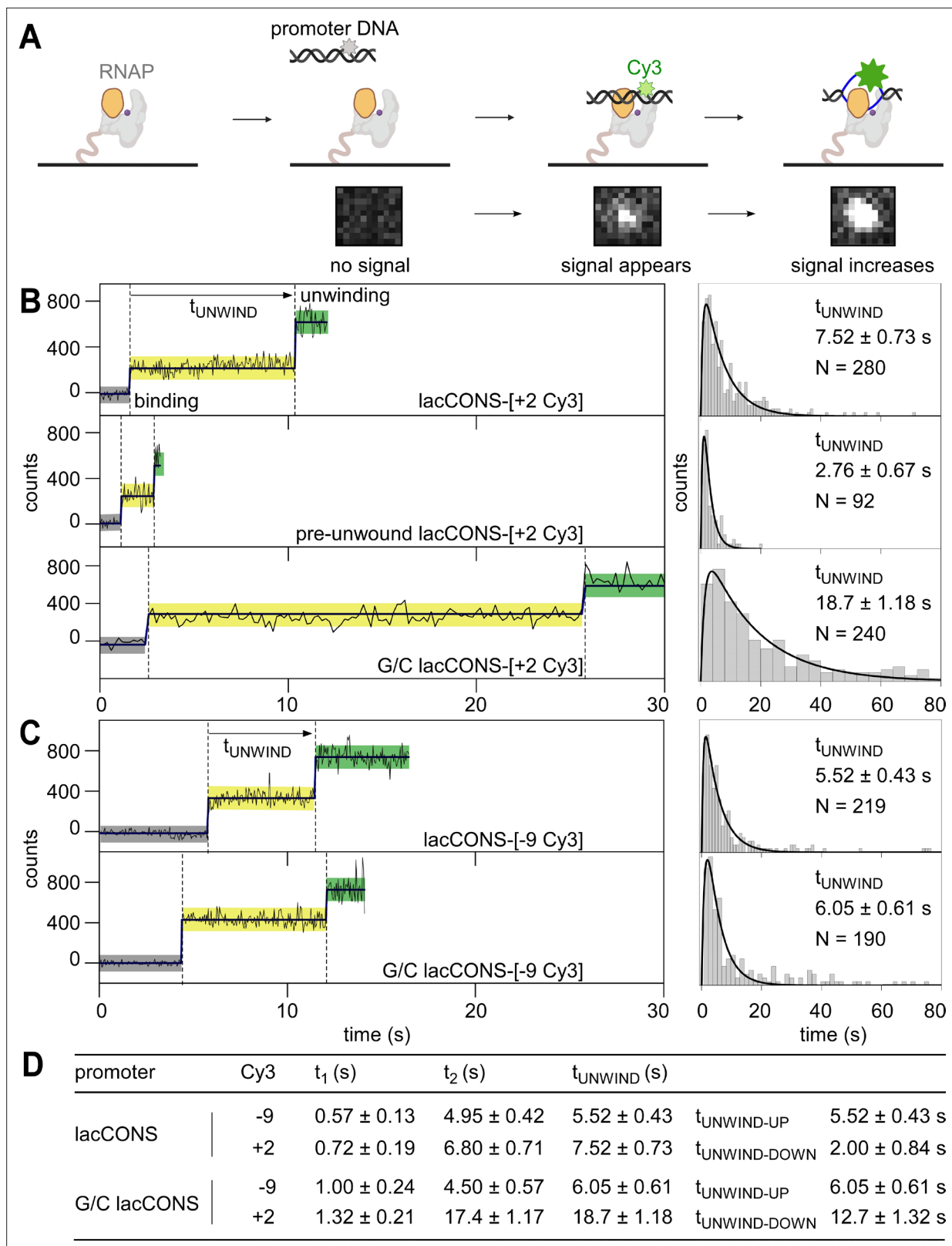


Figure 1. Single-molecule promoter unwinding-induced fluorescence enhancement (smUIFE): DNA unwinding in the upstream part of the transcription bubble precedes DNA unwinding in the downstream part of the transcription bubble. (A) (Top) Design of experiment monitoring promoter unwinding in real time. Grey, RNAP; orange, RNAP clamp; purple dot, RNAP active-centre; black, ds-DNA; blue, ss-DNA; light green, Cy3 on ds-DNA; dark green, Cy3 on ss-DNA. (Bottom) A cropped area ($0.94 \mu\text{m} \times 1.034 \mu\text{m}$) of the field of view, showing appearance and enhancement of fluorescence signal from

Figure 1 continued on next page

Figure 1 continued

binding of single Cy3-labelled promoter fragment to an immobilised RNAP molecule. **(B)** (Left) Time trajectories of intensity from Cy3 on downstream segment of promoter bubble. Black, raw intensity; dark blue, idealised intensity; hidden Markov model (HMM)-assigned states: no promoter (black bars), closed promoter (light yellow bars) and open promoter (green bars). Frame rates: 50 ms, top and middle; 200 ms, bottom. Laser powers: 0.60 mW, top and middle; 0.15 mW, bottom. (Right) Dwell-time histograms of promoter state before unwinding, t_{UNWIND} . **(C)** (Left) Time trajectories of intensity from Cy3 on upstream segment of promoter bubble. Colours as in B. Frame rates: 50 ms. Laser powers: 0.6 mW. (Right) Dwell-time histograms of promoter state before unwinding, t_{UNWIND} . **(D)** Table comparing unwinding times for different promoter constructs.

The online version of this article includes the following source data and figure supplement(s) for figure 1:

Source data 1. Data for single-molecule UIFE experiments in **Figure 1**.

Figure supplement 1. Sequence of consensus lac-promoter fragments used in the study.

Figure supplement 2. Characterisation of complexes between RNAP and Cy3-labelled DNA fragments.

Figure supplement 3. Representative intensity vs. time trajectories from.

Figure supplement 4. Fluorescence intensities from Cy3-labelled promoter fragments pre- and post-unwinding.

complex. The results show that RNAP-promoter complexes formed in these experiments are transcriptionally active and are competent for promoter escape (see also Materials and methods).

Focusing on the class of trajectories that showed promoter unwinding after binding (as opposed to other classes of trajectories that did not show promoter unwinding after binding; **Figure 1—figure supplement 3A**), we analysed trajectories further to define the kinetics of promoter unwinding (see Materials and methods). Using hidden Markov model (HMM; *van de Meent et al., 2014*), we extracted dwell times for the pre-unwinding states and plotted a dwell-time histogram (**Figure 1B, top right**). The dwell-time histogram showed a peaked distribution, indicating a non-Markovian process having more than one rate-limiting steps before the formation of the unwound state (*Floyd et al., 2010*). We fitted this histogram to a two-exponential function (see Materials and methods for a detailed explanation; *Floyd et al., 2010*), and estimated the total time spent in the pre-unwinding state to be ~ 7.5 s (**Figure 1B, top right; Figure 1D**).

To assess whether the assay reports accurately on the kinetics of promoter unwinding, we analysed altered promoter derivatives predicted to unwind *more quickly* or *more slowly* than lacCONS. To accelerate unwinding, we lowered the energy barrier for unwinding by introducing a non-complementary sequence at positions -10 to -4 relative to the transcription start site (pre-unwound lacCONS-[+ 2 Cy3]; **Figure 1—figure supplement 1**). Time trajectories for pre-unwound lacCONS-[+ 2 Cy3] were qualitatively similar to those for lacCONS-[+ 2 Cy3] in terms of the intensity increase (**Figure 1B, middle left, S4A**) and the shape of the dwell-time histogram (**Figure 1B, middle right**); however, the trajectories showed significantly shorter dwell times in the pre-unwinding state (~ 2.8 s vs. ~ 7.5 s; **Figure 1B, middle right vs. top right**). To decelerate the process, we raised the energy barrier for unwinding by introducing a G/C-rich sequence at positions -4 to $+1$ relative to the transcription start site (G/C lacCONS-[+ 2 Cy3]; **Figure 1—figure supplement 1**). Time trajectories for G/C lacCONS-[+ 2 Cy3] showed significantly longer dwell times in the pre-unwinding state (~ 18.8 s vs. 7.5 s; **Figure 1B, bottom right vs. top right; Figure 1D**). Taken together, these results show that dwell times in the pre-unwinding state depend on energy barriers for promoter unwinding, consistent with expectation that the dwell times report on the kinetics of promoter unwinding.

Next, to determine whether promoter unwinding occurs in one step, or in more than one step, we assessed whether unwinding of the upstream half of the transcription bubble (positions -11 to -5) coincides with, or does not coincide with, unwinding of the downstream half of the transcription bubble (positions -4 to $+2$). To probe unwinding of the upstream half of GC promoter bubble, we performed analogous smUIFE experiments using promoter derivatives having Cy3 incorporated in the upstream half of the transcription bubble, at template-strand position -9 (lacCONS-[-9 Cy3] and G/C lacCONS-[-9 Cy3]; and **Figure 1—figure supplement 1**). The resulting trajectories were similar in terms of intensity increases (**Figure 1—figure supplement 4B**) and dwell-time-distribution shapes to those obtained with promoter DNA fragments having Cy3 incorporated in the downstream edge of the transcription bubble, but the dwell times in the pre-unwinding state were significantly shorter: ~ 5.5 s vs. ~ 7.5 s for lacCONS-[-9 Cy3] vs. lacCONS-[+ 2 Cy3] (**Figure 1C, top right vs. Figure 1B, top right; Figure 1D**) and ~ 6.0 s vs. ~ 18.8 s for G/C lacCONS-[-9 Cy3] vs. G/C lacCONS-[+ 2 Cy3] (**Figure 1C, bottom right vs. Figure 1B, bottom right; 1D**). We conclude that, for the promoters

analysed, unwinding of the upstream half of the transcription bubble occurs faster than unwinding of the downstream half of the transcription bubble, and we conclude that we can estimate from our data both the reaction time required for unwinding of the upstream half of the transcription bubble, $t_{\text{UNWIND-UP}}$ (from the lifetime of the pre-unwinding state when Cy3 is incorporated at position -9; ~ 5.5 s for lacCONS), and the reaction time required for the subsequent unwinding of the downstream half of the transcription bubble, $t_{\text{UNWIND-DOWN}}$ (from the difference in lifetimes between the pre-unwinding state when Cy3 is incorporated at position +2 and the pre-unwinding state when Cy3 is incorporated at position -9; ~ 2.0 s for lacCONS; **Figure 1D**). Our results confirm previous results (**Boyaci et al., 2019; Chen et al., 2020; Rogozina et al., 2009; Suh et al., 1993; Helmann and deHaseth, 1999; Lim et al., 2001; Auner et al., 2003**), indicating that promoter unwinding proceeds in a step-wise fashion, in an upstream-downstream direction, and suggest that the step for upstream unwinding is slower than downstream unwinding for lacCONS (~ 5.5 s vs. ~ 2.0 s), whereas the step for downstream unwinding is slower than upstream unwinding for G/C lacCONS (~ 6.0 s vs. ~ 13.0 s).

smUIFE in the presence of an inhibitor that prevents RNAP clamp opening

To determine whether promoter unwinding is affected by preventing opening of the RNAP clamp, we repeated our experiments in presence of myxopyronin (Myx; **Chakraborty et al., 2012; Feklistov et al., 2017; Mukhopadhyay et al., 2008; Belogurov et al., 2009; Dulin et al., 2018**), an RNAP inhibitor that prevents RNAP clamp opening (**Chakraborty et al., 2012; Duchi et al., 2018; Mukhopadhyay et al., 2008**; state with an smFRET efficiency, $E^* \sim 0.36$ in **Figure 2—figure supplement 1D**)

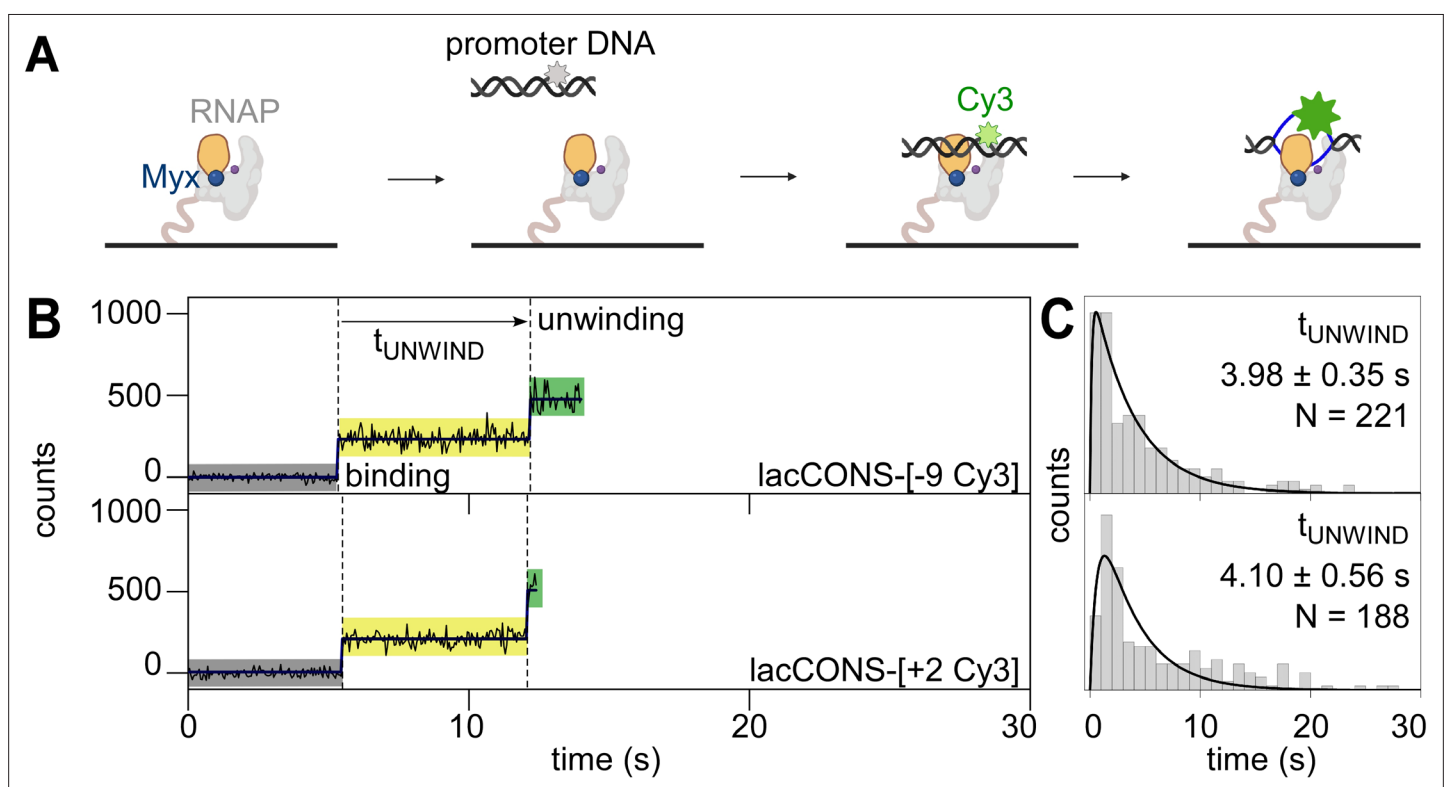


Figure 2. Single-molecule promoter unwinding-induced fluorescence enhancement (smUIFE) in the presence of an inhibitor that prevents RNAP clamp opening: preventing RNAP clamp opening does not prevent DNA unwinding. **(A)** Design of promoter unwinding experiment in presence of myxopyronin (Myx). Blue sphere, Myx; rest as in 1 A. **(B)** (Left) Time trajectory of intensity from Cy3 on upstream (top) and downstream (bottom) segment of promoter bubble. Colours as in 1B. Frame rates: 50 ms. Laser powers: 0.60 mW. (Right) Dwell-time histograms of promoter state before unwinding, t_{UNWIND} .

The online version of this article includes the following source data and figure supplement(s) for figure 2:

Source data 1. Data for single-molecule UIFE experiments in presence of Myxopyronin in **Figure 2**.

Figure supplement 1. Characterisation of complexes formed between RNAP and promoter fragments in presence of myxopyronin (Myx).

and that allows formation of heparin-sensitive RNAP-promoter complexes but prevents formation of heparin-resistant RNAP-promoter complexes (Mukhopadhyay et al., 2008; Figure 2—figure supplement 1A–B). We reasoned that, if clamp opening is obligatory for promoter unwinding, preventing clamp opening by addition of Myx should either prevent or delay downstream unwinding. We first performed smUIFE experiments in the presence of Myx using promoter DNA fragment lacCONS-[+2 Cy3], which has Cy3 incorporated at the downstream edge of the transcription bubble. The results showed a fluorescence intensity enhancement of ~2.4 -fold, consistent with unwinding of the downstream edge of the transcription bubble (Figure 2B, bottom; Figure 2—figure supplement 1C; see Feklistov et al., 2017) and showed a dwell time between initial binding and intensity enhancement of ~4.1 s (Figure 2C, bottom). We next performed smUIFE experiments in the presence of Myx using promoter DNA fragment lacCONS-[−9 Cy3], which has Cy3 incorporated in the upstream half of the transcription bubble. The results were essentially identical: a fluorescence intensity enhancement of ~2.4 -fold (Figure 2B, top; Figure 2—figure supplement 1C, top) and a dwell time between initial binding and intensity enhancement of ~4.0 s (Figure 2C). For both promoter derivatives, the observed fluorescence intensity enhancements were similar in the absence and presence of Myx (Figure 1—figure supplement 4 vs. Figure 2—figure supplement 1C), and the observed dwell times between initial binding and intensity enhancement times were shorter – not longer – in the presence of Myx than in the absence of Myx (~3.5 s shorter for transcription-bubble downstream edge and ~1.5 s shorter for transcription-bubble upstream half; Figure 2B–C vs. Figure 1B–C, top panels). We suggest that Myx prevents formation of heparin-resistant open complexes either by preventing loading of unwound DNA into active-centre cleft or by preventing subsequent locking of the RNAP clamp. We infer – contrary to the models in which clamp opening is obligatory for promoter unwinding – that preventing RNAP clamp opening does not prevent promoter unwinding, and increases, not decreases, the kinetics of promoter unwinding.

Single-molecule FRET

To assess *directly* whether RNAP clamp motions occur during RPo formation, we performed smFRET measurements using an RNAP derivative containing Cy3B, serving as a fluorescence donor, incorporated at the tip of the RNAP clamp and Alexa647, serving as a fluorescence acceptor, incorporated at the tip of the opposite wall of the RNAP active-centre cleft, monitoring smFRET efficiency (E^*) during RPo formation (Figure 3A). In previous smFRET studies using the same probes, we showed that the RNAP clamp interconverts between open ($E^* \sim 0.2$), partly-closed ($E^* \sim 0.3$), and closed ($E^* \sim 0.4$) conformations in solution (Chakraborty et al., 2012; Lin et al., 2018; Duchi et al., 2018; Mazumder et al., 2021). To monitor RNAP clamp motions during the formation of RPo, we immobilised promoter DNA molecules (biotin-lacCONS; Figure 1—figure supplement 1, Figure 3—figure supplement 1) on a coverslip mounted on a TIRF microscope, started recording, and added the doubly labelled RNAP (Figure 3A). Binding of doubly labelled RNAP molecules to DNA was monitored by detecting the simultaneous appearance of donor and acceptor fluorescence emission signals on the surface, and RNAP clamp motions in RNAP-promoter complexes at and after binding were monitored by quantifying E^* (Figure 3—figure supplement 2A). Control experiments show that: (i) doubly labelled RNAP molecules bound exclusively to immobilised promoter DNA molecules (Figure 3—figure supplement 1A); (ii) the majority (~87%) of the resulting complexes were resistant to challenge with heparin (Figure 3—figure supplement 1B); and (iii) doubly labelled RNAP molecules did not bind stably to immobilised non-promoter DNA molecules (Figure 3—figure supplement 2C). These results indicate that sequence-specific, heparin-resistant RNAP-promoter complexes were only formed between the doubly labelled RNAP holoenzyme molecules and promoter DNA fragments.

In order to determine the RNAP clamp conformations immediately upon initial binding of RNAP to promoter DNA, we plotted distributions of E^* values for the first five frames (0.5 s) after initial binding. The resulting distributions could be fitted to a Gaussian function with mean E^* of ~0.4, indicating that the initial binding of RNAP to promoter DNA involved RNAP with a closed clamp (Figure 3—figure supplement 2B). Next, we examined full E^* time trajectories, for up to ~60 s, following initial binding, seeking E^* changes potentially consistent with RNAP clamp opening. We detected no E^* changes – not even transient E^* changes, within the ~100 ms temporal resolution of the analysis – that potentially could be assigned as consistent with clamp opening (E^* of ~0.2; Figure 3B, area highlighted in cyan; Figure 3—figure supplement 2A).

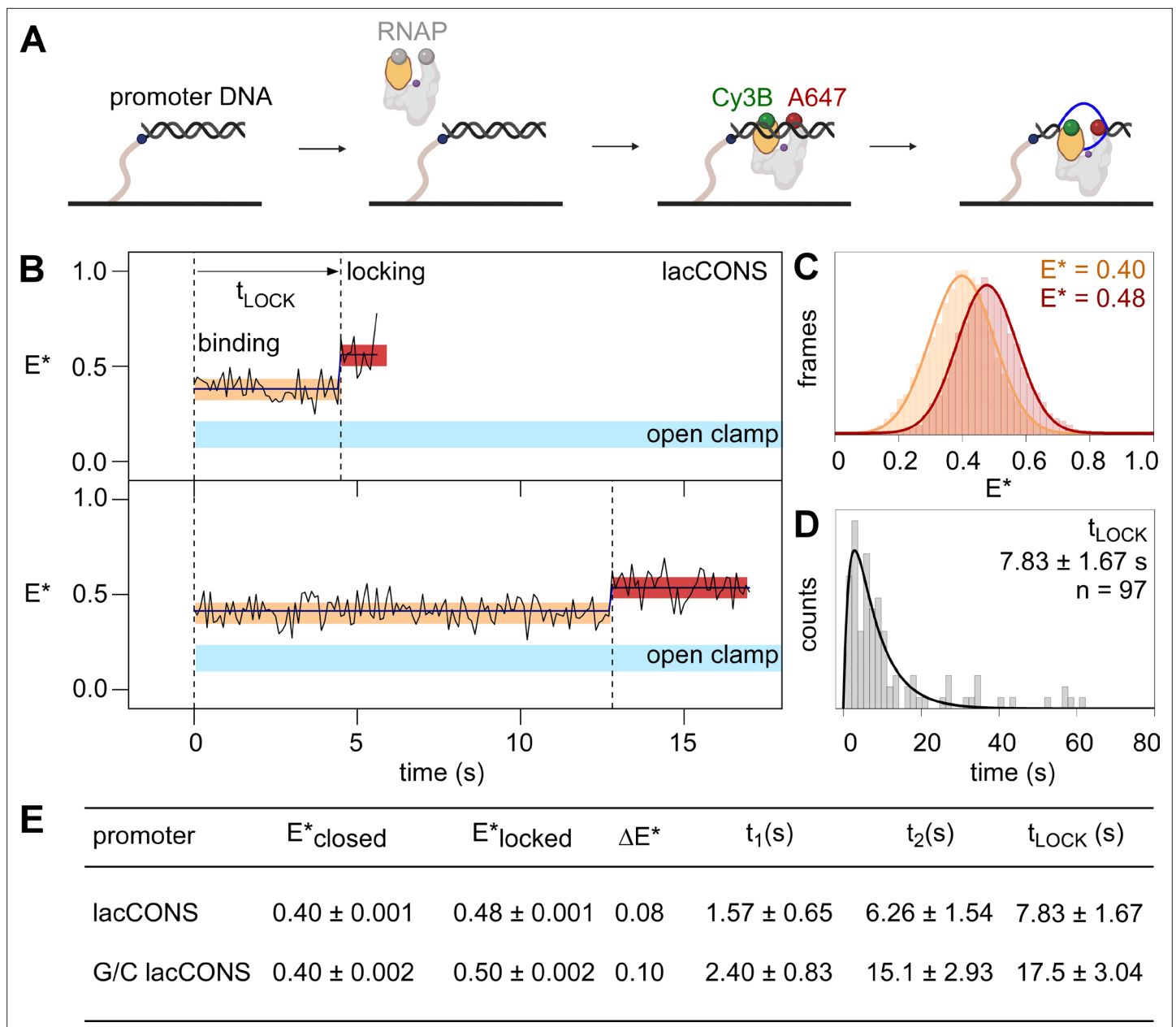


Figure 3. Single-molecule fluorescence resonance energy transfer (smFRET): DNA unwinding occurs without RNAP clamp opening and is followed by RNAP clamp locking. **(A)** Design of experiment monitoring clamp status in real time. Black, ds-DNA; orange, RNAP clamp; grey, rest of RNAP; purple dot, RNAP active-centre; blue, ss-DNA; green, Cy3B; and red, Alexa647. **(B)** Representative time trajectories of E^* for experiments with a lacCONS promoter fragment, showing hidden Markov model (HMM)-assigned closed-clamp state (orange), locked-clamp state (red), and interstate transition (dark-blue). Expected range of E^* values for an open-clamp state is highlighted in light blue. Frame rate: 100 ms. Laser powers: 200 μW in red and 500 μW in green. **(C)** HMM-assigned histograms and Gaussian fits of E^* for full-time trajectories from experiments with a lacCONS promoter fragment. **(D)** Dwell-time histograms of time before transition to the locked-clamp state, t_{LOCK} , for experiments with a lacCONS promoter fragment. **(E)** Table showing mean E^* ; difference in E^* (ΔE^*) between closed-clamp or locked-clamp states and time to transition to a locked-clamp state after initial binding for the lacCONS and lacCONS-GC promoter fragments.

The online version of this article includes the following source data and figure supplement(s) for figure 3:

Source data 1. Data for single-molecule FRET experiments in **Figure 3**.

Figure supplement 1. Characterisation of complexes formed between clamp-labelled RNAP and immobilised promoter fragments.

Figure supplement 2. Single-molecule fluorescence resonance energy transfer (smFRET): initial binding of RNAP to surface-immobilised promoter DNA fragments take place via a closed-clamp conformation.

Figure 3 continued on next page

Figure 3 continued

Figure supplement 3. Classification for different time trajectories of E^* of clamp-labelled RNAP molecules bound to immobilised lacCONS-promoter fragments in real time.

Figure supplement 4. Single-molecule fluorescence resonance energy transfer (smFRET) data showing binding of clamp-labelled RNAP to immobilised biotin-G/C lacCONS promoter fragments.

Next, we examined full E^* time trajectories, following initial binding, seeking any E^* changes potentially consistent with any RNAP clamp motions (see Materials and methods and **Figure 3—figure supplement 3** for full classification). Intriguingly, a large fraction of trajectories (~44%) started at $E^* \sim 0.40$, indicative of the previously defined closed-clamp state, then transitioned to $E^* \sim 0.48$, indicative of a new, more tightly closed, clamp state, and then returned to $E^* \sim 0.40$ or photobleached (**Figure 3B–C; Figure 3—figure supplement 3B, top**). We refer to the new, more tightly closed, clamp state with $E^* \sim 0.48$, as the ‘locked-clamp’ state. We further analysed E^* time trajectories to determine the time between initial binding of RNAP to promoter DNA and appearance of the locked-clamp state. The corresponding dwell-time histogram showed a peaked distribution and was fitted to a two-exponential function, yielding ~ 7.8 s, as the time between initial binding of RNAP to promoter DNA and appearance of the locked-clamp state (**Figure 3D**), a time that, within error, is identical to the ~ 7.5 s time between initial binding and unwinding of the downstream half of the transcription bubble.

Analogous experiments with a promoter derivative having G/C-rich sequence at positions -4 to $+1$ relative to the transcription start site, G/C lacCONS, yielded similar results: that is, initial binding by RNAP with a closed-clamp state ($E^* \sim 0.41$; **Figure 3—figure supplement 4A; B, left**); no trajectories showing transitions – not even transient transitions, within the ~ 400 ms temporal resolution of the analysis – to an open-clamp state during a period up to ~ 200 s after initial binding; a large fraction of trajectories (~48%), showing a transition to a locked-clamped state ($E^* \sim 0.50$, **Figure 3—figure supplement 4A; B, right**), and a time between initial binding and appearance of the locked-clamped state matching the time required for unwinding of the downstream half of the transcription bubble (**Figure 3E, Figure 3—figure supplement 4C**).

We conclude that the RNAP clamp is in a closed state upon promoter binding, that the RNAP clamp does not open – not even transiently, within the temporal resolution of our analysis – between promoter binding and promoter unwinding, and that the RNAP clamp closes further – ‘locks’ – after promoter unwinding.

Discussion

Taken together, our results indicate that RPo formation by *E. coli* RNAP σ^{70} holoenzyme at a consensus bacterial core promoter proceeds through a ‘bind-unwind-load-and-lock’ mechanism, in which the RNAP clamp is closed upon promoter binding, remains closed during unwinding of promoter DNA

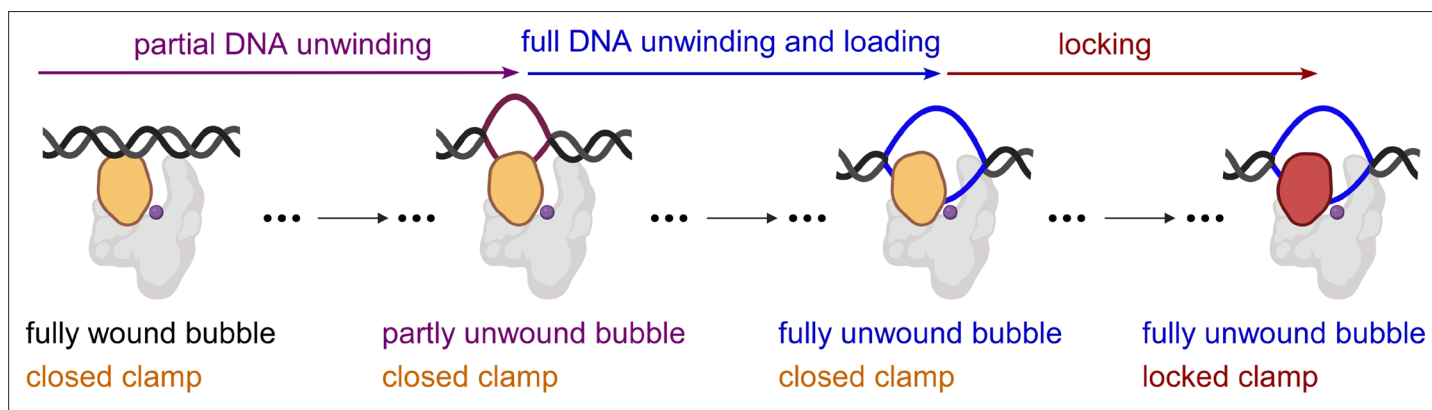


Figure 4. ‘Bind-unwind-load-and-lock’ mechanism for the formation of RPo. Orange, closed clamp; red, locked clamp; grey, rest of RNAP; purple dot, RNAP active-centre; black, ds-DNA; magenta, partly unwound bubble and blue, fully unwound bubble.

– which proceeds in an upstream-to-downstream direction – and then closes further, locking the unwound DNA in the RNAP active-centre cleft (**Figure 4**).

As described in the introduction, it has been the predominant view that structural considerations necessitate RNAP clamp opening for loading of dsDNA into the RNAP active-centre cleft but do not necessitate RNAP clamp opening for loading of ssDNA into the RNAP active-centre cleft (**Mazumder and Kapanidis, 2019; Ruff et al., 2015; Mekler et al., 2002; Murakami et al., 2002b; Murakami et al., 2002a; Young et al., 2002; Darst et al., 2002; Chakraborty et al., 2012**). Under that predominant view, our finding that the RNAP clamp remains closed throughout promoter unwinding at a consensus bacterial promoter implies that it is ssDNA, not dsDNA, that enters the RNAP active-centre-cleft, and thus that promoter unwinding occurs at least in part outside, not inside, the RNAP active-centre cleft (**Figure 4**).

In contrast to the predominant view, it recently has been speculated, based on molecular-dynamics simulations, that part of the RNAP cleft potentially could become accessible to both dsDNA and ssDNA without clamp opening, due to motions, putatively occurring on the nanosecond time scale, of the RNAP β lobe ('gate opening'; **Unarta et al., 2021**). Under this alternative view, our finding that the RNAP clamp remains closed throughout promoter unwinding at a consensus bacterial promoter potentially could be consistent with either ssDNA or dsDNA entering the RNAP active-centre cleft, and thus potentially could be consistent with promoter unwinding that occurs either outside or inside the RNAP active-centre cleft. Nevertheless, in view of the absence of experimental support for the alternative view, and in view of the mismatch between the nanosecond time scale of the limiting reaction in the alternative view and the millisecond-to-second time scales of promoter unwinding, we tend to disfavour this alternative view.

The consensus bacterial promoter employed in this study has been used in previous work to define transcription mechanisms, including RNAP clamp closure upon RPo formation (**Chakraborty et al., 2012**), DNA scrunching in transcription-start-site selection (**Winkelman et al., 2016**), DNA scrunching in initial transcription (**Kapanidis et al., 2006**), and RNAP pausing in initial transcription (**Duchi et al., 2016; Winkelman et al., 2020**), all of which subsequently were validated for other promoters. Nevertheless, we emphasise that it remains to be determined whether the bind-unwind-load-and-lock pathway defined here for the consensus bacterial promoter is also used by *E. coli* RNAP σ^{70} holoenzyme at other promoters and by other *E. coli* RNAP holoenzymes. Consistent with the possibility that a similar mechanism is used by *E. coli* RNAP σ^{70} holoenzyme at other promoters, a recent cryo-EM study of transcription initiation by *E. coli* RNAP σ^{70} holoenzyme at another promoter, *rpsT* P2, identified structural states having a partly unwound transcription bubble and a closed RNAP clamp (**Chen et al., 2020**). We note that the smUIFE and smFRET methods reported in this work could be applied to analyse transcription initiation by any *E. coli* RNAP holoenzyme at any promoter and potentially could be adapted to analyse transcription initiation by any RNAP – bacterial, archaeal, or eukaryotic – at any promoter. Extension of these assays using more complex labelling and imaging schemes should allow simultaneous observation of promoter unwinding and RNAP clamp conformational changes in the same molecule, enabling a more detailed view of the coupling of the molecular motions involved in promoter unwinding. Finally, development of massively parallel single-molecule assays interrogating large libraries of promoter sequences would enable a comprehensive analysis of general and specific features of this first step in transcription initiation for different promoter sequences.

Materials and methods

Preparation of reagents: RNAP, oligodeoxyribonucleotides, Myx RNAP derivatives

For experiments in **Figures 1 and 2**, hexahistidine-tagged *E. coli* RNAP holoenzyme was prepared using co-expression of genes encoding RNAP β' , β , α , ω , and σ^{70} subunits to afford an RNAP σ^{70} holoenzyme derivative as follows: single colonies of *E. coli* strain BL21(DE3) (Millipore) co-transformed with plasmid pV10 (**Belogurov et al., 2007**) and plasmid pRSFduet-sigma (**Hudson et al., 2009**) were used to inoculate 20 ml LB broth (**Sambrook and Russell, 2001**) containing 100 $\mu\text{g/ml}$ ampicillin and 50 $\mu\text{g/ml}$ kanamycin and cultures were incubated 16 hr at 37°C with shaking. Culture aliquots (2 \times 10 ml) were used to inoculate LB broth (2 \times 1 L) containing 100 $\mu\text{g/ml}$ ampicillin and 50 $\mu\text{g/ml}$ kanamycin; cultures were incubated at 37°C with shaking until $\text{OD}_{600} = 0.6$; IPTG was added to 1 mM; and

cultures were further incubated 3.5 hr at 37°C with shaking. Cells were harvested by centrifugation (4000 × *g*; 20 min at 4°C), re-suspended in 20 ml buffer A (10 mM Tris-HCl, pH 7.9, 200 mM NaCl, and 5% glycerol), and lysed using an EmulsiFlex-C5 cell disrupter (Avestin). The lysate was cleared by centrifugation (20,000 × *g*; 30 min at 4°C), precipitated with polyethyleneimine (Sigma-Aldrich) as in *Niu et al., 1996*, and precipitated with ammonium sulphate as in *Niu et al., 1996*. The precipitate was dissolved in 30 ml buffer A and loaded onto a 5 ml column of Ni-NTA-agarose (Qiagen) pre-equilibrated in buffer A; the column was washed with 50 ml buffer A containing 10 mM imidazole, and eluted with 25 ml buffer A containing 200 mM imidazole. The sample was further purified by anion-exchange chromatography on Mono Q 10/100 GL (GE Healthcare; 160 ml linear gradient of 300–500 mM NaCl in 10 mM Tris-HCl, pH 7.9, 0.1 mM EDTA, and 5% glycerol; flow rate = 2 ml/min). Fractions containing hexahistidine-tagged *E. coli* RNAP σ^{70} holoenzyme were pooled, concentrated to ~2 mg/ml using 30 kDa MWCO Amicon Ultra-15 centrifugal ultrafilters (EMD Millipore), and stored in aliquots at –80°C.

For experiments in **Figure 3** and **Figure 3—figure supplements 1–4**, fluorescently labelled, hexahistidine-tagged *E. coli* RNAP holoenzyme (hereafter ‘labelled-RNAP’) with Cy3B and Alexa647 at positions 284 on the β' subunit, and 106 on the β subunit, respectively, was prepared using in vivo reconstitution methods as described (*Lin et al., 2018*).

Nucleic acids

Oligodeoxyribonucleotides were purchased from IBA Lifesciences, dissolved in nuclease-free water (Ambion) to a final concentration of 100 μ M and stored at –20°C. Oligodeoxyribonucleotides used in **Figure 1**, **Figure 1—figure supplements 2–4**, and **Figure 2—figure supplement 1** were labelled with Cy3 *N*-hydroxysuccinimidyl ester (Fisher Scientific) as described (*Mukhopadhyay et al., 2003*). Oligodeoxyribonucleotides were annealed by mixing two complementary strands in a ratio of 1:1 in hybridisation buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA) and by heating for 5 min at 95°C, followed by cooling to 25°C in 2°C steps with 1 min per step using a thermal cycler (Applied Biosystems).

Myx was prepared as described (*Mukhopadhyay et al., 2008*).

Single-molecule fluorescence microscopy: smUIFE experiments

For experiments in **Figures 1 and 2**, **Figure 1—figure supplements 2–4**, and **Figure 2—figure supplement 1**, observation wells for real-time experiments were prepared as described (*Lin et al., 2018*). Briefly, a biotin-PEG-passivated glass surface was prepared, functionalised with Neutravidin (Sigma-Aldrich), and treated with biotinylated anti-hexahistidine monoclonal antibody (Penta-His Biotin Conjugate; Qiagen), yielding wells with (biotinylated anti-hexahistidine monoclonal antibody)-Neutravidin-biotin-PEG-functionalised glass floors. Hexahistidine-tagged RNAP σ^{70} holoenzyme molecules were immobilised in observation wells with (biotinylated anti-hexahistidine monoclonal antibody)-Neutravidin-biotin-PEG-functionalised glass floors, as follows: aliquots (30 μ l) of 0.1 nM hexahistidine-tagged RNAP σ^{70} holoenzyme in KG7 buffer (40 mM HEPES-NaOH, pH 7.0, 100 mM potassium glutamate, 10 mM MgCl₂, 1 mM dithiothreitol, 100 μ g/ml bovine serum albumin, and 5 % glycerol) were added to the observation chamber and incubated 2–4 min at 22 °C, solutions were removed, wells were washed with 2 × 30 μ l KG7, and 30 μ l KG7 imaging buffer (KG7 buffer containing 2 mM Trolox, 1 mg/ml glucose oxidase, 40 μ g/ml catalase, and 1.4% w/v D-glucose) at 22 °C was added.

For experiments in **Figure 1—figure supplement 2**, Cy3-labelled promoter (or non-promoter) fragments were manually added (final concentration of 2 nM) to the observation wells containing immobilised RNAP σ^{70} holoenzyme (or RNAP core enzyme) molecules and incubated for 5 min; wells were then washed with KG7 and movies were recorded. Next, observation wells containing RNAP-promoter complexes were supplemented with KG7 containing 250 μ g/ml heparin (Sigma-Aldrich) solution, incubated for 1 min; wells were then washed with KG7 and movies were recorded.

For experiments monitoring RNAP-promoter complex formation reactions in real time (**Figures 1 and 2**, **Figure 1—figure supplement 3**, and **Figure 2—figure supplement 1**), observation wells containing immobilised RNAP were supplemented with KG7 imaging buffer, recordings were started, and Cy3-labelled promoter (or non-promoter) fragments were manually added (using a pipette) to the observation wells during the recording so as to yield a final concentration of 2 nM of Cy3-labelled

DNA in the wells. For experiments in **Figure 2—figure supplement 1A–C**, same procedures were followed, except that KG7 imaging buffer was supplemented with 20 μM Myx.

For experiments in **Figure 1—figure supplement 3B** top panel, observation wells containing immobilised RNAP-promoter complexes (formed with LC-[+ 2-Cy3]; sequence in **Figure 1—figure supplement 1**) were supplemented with KG7 buffer containing an NTP subset (500 μM ApU and 100 μM each of ATP, GTP, and UTP) directing synthesis of RNA up to position +11 (RPitc ≤ 11), incubated for 5 min at 22 $^{\circ}\text{C}$, solutions were removed, wells were washed with $2 \times 30 \mu\text{l}$ KG7, and 30 μl KG7 imaging buffer at 22 $^{\circ}\text{C}$ was added. Movies of complexes corresponding to stalled RPitc ≤ 11 molecules were recorded. For experiments in **Figure 1—figure supplement 3B** bottom panel, observation wells containing immobilised stalled RPitc ≤ 11 molecules were supplemented with KG7 imaging buffer, the recording of a movie was started, and an NTP subset (100 μM each of GTP and CTP) directing extension of stalled complexes with 11-mer RNA to a 14-mer RNA was added in real time. Extension of RNA up to +11 results in stalled initial transcribing complexes (*Dulin et al., 2018*) and extension of RNA to +14 results in promoter escape and leads to formation of stalled elongation complexes for the lacCONS promoter used (Wang, Mazumder et al., in preparation).

Single-molecule fluorescence microscopy: smFRET experiments

For experiments in **Figure 3** and **Figure 3—figure supplements 1–4**, observation wells of biotin-PEG-passivated glass surface were prepared and functionalised with Neutravidin (Sigma-Aldrich) to yield Neutravidin-biotin-PEG-functionalised glass floors. Biotin-tagged promoter DNA fragments were then immobilised in these observation wells as follows: aliquots (30 μl) of 0.05 nM biotin-tagged promoter fragments in KG7 were added to the observation chamber and incubated 1 min at 22 $^{\circ}\text{C}$, solutions were removed, wells were washed with $2 \times 30 \mu\text{l}$ KG7, and 30 μl KG7 imaging buffer at 22 $^{\circ}\text{C}$ was added.

For experiments in **Figure 3—figure supplement 1**, clamp-labelled RNAP molecules were manually added (final concentration of 2 nM) to the observation wells containing immobilised promoter DNA fragments, incubated for 5 min, wells were washed with KG7, and movies were recorded. Next, observation wells containing RNAP-promoter complexes were supplemented with KG7 containing 250 $\mu\text{g/ml}$ heparin solution, incubated for 1 min, wells were washed with KG7, and movies were recorded.

For experiments monitoring RNAP-promoter complex formation reactions in real time (**Figure 3** and **Figure 3—figure supplements 2–4**), observation wells containing immobilised promoter (or non-promoter) fragments were supplemented with KG7 imaging buffer, recordings were started, and labelled-RNAP molecules (with Cy3B and Alexa647 at positions 284 on the β' subunit, and 106 on the β subunit of RNAP σ^{70} holoenzyme) were manually added (using a pipette) to the observation wells during the recording so as to yield a final concentration of 2 nM of labelled RNAP in the wells.

Single-molecule fluorescence microscopy: data collection

Single-molecule fluorescence experiments were performed using a custom-built objective-type TIRF microscope (*Holden et al., 2010*). Light from a green laser (532 nm; Samba; Cobolt) and a red laser (635 nm; CUBE 635-30E, Coherent) was combined using a dichroic mirror coupled into a fiber-optic cable focused onto the rear focal plane of a 100 \times oil-immersion objective (numerical aperture 1.4; Olympus) and was displaced off the optical axis, such that the incident angle at the oil-glass interface of a stage-mounted observation chamber exceeded the critical angle, thereby creating an exponentially decaying evanescent wave (*Axelrod et al., 1983*). Alternating-laser excitation (ALEX; *Kapanidis et al., 2006*; *Lee et al., 2005*) was implemented by directly modulating the green and red lasers using an acousto-optical modulator (1205 C, Isomet). Fluorescence emission was collected from the objective, was separated from excitation light using a dichroic mirror (545 nm/650 nm, Semrock) and emission filters (545 nm LP, Chroma; and 633/25 nm notch filter, Semrock), was focused on a slit to crop the image, and then was spectrally separated (using a dichroic mirror; 630 nm DLRP, Omega) into donor and emission channels focused side-by-side onto an electron-multiplying charge-coupled device camera (EMCCD; iXon 897; Andor Technology). A motorised x/y-scanning stage (MS-2000; ASI) was used to control the sample position relative to the objective.

For experiments in **Figures 1–2**, **Figure 1—figure supplements 2–4**, and **Figure 2—figure supplement 1**, frame rates were either 50 or 200 ms, and laser powers were either 0.60 or 0.15 mW

at 532 nm. For experiments in **Figure 3** and **Figure 3—figure supplements 2–3**, frames were either 100 or 200 ms long, and laser powers were either 200 μ W in red, 500 μ W in green, or 80 μ W in red, 200 μ W in green, respectively. For experiments in **Figure 3—figure supplement 4**, frame rates were 400 ms long, and laser powers were 50 μ W in red, 150 μ W in green. All data acquisition was carried out at 22 °C.

Single-molecule fluorescence microscopy: data analysis for smUIFE experiments

For experiments in **Figures 1–2**, **Figure 1—figure supplements 2–4**, and **Figure 2—figure supplement 1**, localisation of single fluorescence emitters was detected and background-corrected fluorescence intensity vs. time trajectories from each localisation were extracted and curated to exclude trajectories exhibiting very high fluorescence intensities ($I_{\text{Cy3}} > 750$ photon counts), trajectories exhibiting multiple binding events during the observation window and trajectories exhibiting photoblinking. Initial inspection of time trajectories for experiments with lacCONS-[+ 2 Cy3] revealed three main classes of molecules. Class-I molecules started with the appearance of signal having intensity of ~ 200 photon counts, which remained stable until its disappearance ($\sim 46\%$; **Figure 1—figure supplement 3**, top). Class-II molecules also started with appearance of a signal having intensity of ~ 200 photon counts, followed by intensity increase to ~ 450 photon counts after some time, followed by either signal disappearance or intensity decrease to ~ 200 photon counts ($\sim 45\%$; **Figure 1B**, **Figure 1—figure supplement 3**, middle). Class-III molecules started with appearance of a signal having intensity of ~ 450 photon counts, followed by intensity decrease to ~ 220 photon counts, and subsequent signal disappearance, or intensity increase to ~ 450 photon counts ($\sim 9\%$; **Figure 1—figure supplement 3**, bottom). Similar observations were made for all other Cy3-labelled promoters studied.

Based on intensity levels, we assigned states with no signal to absence of a promoter DNA or to a bleached probe; states with ~ 200 photon counts to a promoter DNA that is dsDNA in the Cy3 vicinity; and states with ~ 450 counts to a promoter that is unwound in the Cy3 vicinity. We thus assigned events in Class-I molecules to binding of a promoter DNA molecule to RNAP, followed by dissociation of complexes, indicating formation of non-specific complexes; these events are also consistent with bleaching occurring prior to any intensity increase. We assigned events in Class-II molecules to binding of a promoter DNA molecule to RNAP, followed by promoter unwinding, followed by bleaching or promoter rewinding. Finally, we assigned events in Class-III molecules to binding of an unwound-promoter DNA molecule to RNAP, followed by promoter rewinding, followed by bleaching or promoter unwinding, indicating formation of complexes where the initial unwinding event is missed. We further curated intensity time trajectories up to the point of initial ~ 2 fold intensity enhancement for Class-II molecules only, since they showed an unambiguous bubble-opening event after initial binding.

Next, photon counts (I_{Cy3}) for the set of curated intensity vs. time trajectories were divided by 1000 to obtain I_{Cy3}^* ; the corresponding I_{Cy3}^* vs. time trajectories were analysed to identify the different fluorescence intensity states using HMM as implemented in the MATLAB (MathWorks) software package ebFRET (*van de Meent et al., 2014*) using a three-state model (one of which corresponds to a negligible I_{Cy3}^* state). The intensity I_{Cy3}^* , for each of the other two HMM-derived states, was extracted from ebFRET and converted to I_{Cy3} values, which in turn were binned and plotted as I_{Cy3} count histograms (**Figure 1—figure supplement 4**, **Figure 2—figure supplement 1C**). These histograms were fitted using Gaussian distributions in Origin to define the mean fluorescence intensity that corresponds to each state (**Figure 1—figure supplement 4**, **Figure 2—figure supplement 1**).

Dwell times for each fluorescence intensity states were extracted from HMM fits to I_{Cy3}^* vs. time trajectories, were binned, and were plotted as distribution histograms in Origin. For experiments in **Figures 1–2**, dwell-time distributions corresponding to the pre-unwinding state resembled peaked distributions indicating presence of at least two sub-steps and were fit to a two-exponential function of the form $y = A*(e^{-x/t_1} - e^{-x/t_2})$, where y represents counts of dwells in the closed-bubble state following the initial binding event and before the signal increase event, x represents time, and t_1 and t_2 represent the lifetime of the individual sub-steps. From the fit, we estimate the lifetimes corresponding to the two sub-steps, as well as the total time spent in the first closed-bubble state (as the sum of time spent in the two sub-steps; **Figure 2B**, right panels).

We note that it is possible that the transition involves more than two steps, but our current analysis and available data sets cannot specify the number and duration of each sub-step; consequently, we interpret qualitatively the shape of the dwell-time histogram as an indication of a multi-step process, and focus on the total time spent before the intensity enhancement event, since that can be interpreted with certainty as the time taken before unwinding in the vicinity of Cy3.

Single-molecule fluorescence microscopy: data analysis for smFRET experiments

For experiments in **Figure 3** and **Figure 3—figure supplements 1–4**, localisations in donor-emission (green) and acceptor-emission (red) channels were detected using the peak-finding algorithm of the MATLAB (MathWorks) software package Twotone, as described (**Holden et al., 2010**). Peaks detected in both emission channels (i.e., peaks for molecules containing both donor and acceptor probes) were fitted with two-dimensional Gaussian functions to extract background-corrected intensity vs. time trajectories for donor-emission intensity upon donor excitation (I_{DD}), acceptor-emission intensity upon donor excitation (I_{DA}), and acceptor-emission intensity upon acceptor excitation (I_{AA}), as described (**Holden et al., 2010**). Intensity vs. time trajectories were curated to exclude trajectories exhibiting $I_{DD} < 100$ or > 1000 counts or $I_{AA} < 200$ or > 1000 counts, trajectories exhibiting multiple-step donor or acceptor photobleaching, trajectories exhibiting donor or acceptor photo-blinking and portions of trajectories following donor or acceptor photobleaching. Intensity vs. time trajectories were used to calculate trajectories of apparent donor-acceptor smFRET efficiency (E^*) as described (**Holden et al., 2010; Kapanidis et al., 2004; Lee et al., 2005**):

$$E^* = I_{DA} / (I_{DD} + I_{DA})$$

E^* time trajectories were analysed globally to identify E^* states by use of HMM as implemented in MATLAB (MathWorks) software package ebFRET (**van de Meent et al., 2014**).

HMM analysis of E^* time trajectories revealed four types of molecules. The most abundant were Class-I molecules (~45 % of all events) that started with $E^* \sim 0.40$, remained at $E^* \sim 0.40$, and then bleached (**Figure 3—figure supplement 3A**), and Class-II molecules (~44%) that started with $E^* \sim 0.40$, transitioned to $E^* \sim 0.48$, and then bleached or returned to $E^* \sim 0.40$ (**Figure 3—figure supplement 3B, top**). More rarely, we observed Class-III molecules (~9 % of all events) that started with $E^* \sim 0.48$, transitioned to $E^* \sim 0.40$, followed by bleaching or transition to $E^* \sim 0.48$; **Figure 3—figure supplement 3C**, and, very rarely, we observed Class-IV molecules (~2%) that started with $E^* \sim 0.40$, switched between $E^* \sim 0.40$ – 0.20 and bleached (**Figure 3—figure supplement 3D**).

We refer to the new $E^* \sim 0.48$ state observed here as the ‘locked-clamp’ conformation, with Class-II molecules providing information on the time taken to form the ‘locked-clamp’ conformation; in contrast, Class-I molecules represent RNAP molecules that either bleach before the transition to the locked-clamp state or bind non-specifically to DNA; and Class-III molecules represent molecules where the initial clamp locking is missed.

E^* time trajectories for Class-II molecules were fitted to a two-state HMM model, E^* -values from the fitted model were extracted, binned, and plotted using Origin (Origin Lab), and were fitted to Gaussian distributions using Origin (**Figure 3B**, right panels and **Figure 3—figure supplement 3B**, left panel; coloured curves). The resulting histograms provide population distributions of E^* states and, for each E^* state, define mean E^* (**Figure 3B**, right panels and **Figure 3—figure supplement 4B**, right panel; coloured bars and inset). Dwell-time distributions corresponding to time spent before the first transition from a closed-clamp state to a locked-clamp state were extracted from HMM fits to E^* vs. time trajectories, and were binned and plotted as distribution histograms in Origin (**Figure 3** and **Figure 3—figure supplement 4C**). Dwell-time histograms obtained in this manner exhibited the shape of a peaked distribution, indicating presence of at least two sub-steps, were fitted to a two-exponential function of the form $y = A*(e^{-x/t_1} - e^{-x/t_2})$, where y represents counts of dwells in the closed-clamp state following the initial binding event and before the clamp-locking event, x represents time, and t_1 and t_2 represent the lifetimes of the individual sub-steps. From the fit, we estimate the time spent in the two sub-steps, as well as the total time before transition to ‘locked-clamp’ conformation (t_{LOCK}) from the sum of times corresponding to the two sub-steps. Similar to the case for fluorescence enhancement experiments described previously, we note that there may be more than two steps which contribute significantly to these dwell times, and we could not infer an accurate model in terms of the number and duration of each sub-step involved from this dataset. Therefore, we avoid

assigning the two sub-steps to specific conformations or events and focus on the total time spent before the transition to 'locked-clamp' conformation.

Data and software availability

All information for replication is included in the submission and data corresponding to each figure are provided as source data files. MATLAB software packages TwoTone and ebFRET are available on GitHub (<https://github.com/annawang692/TwoTone2018>, Wang, 2020 and <http://ebfret.github.io/>).

Acknowledgements

We thank Dr Anssi Malinen for discussions and early work on the development of real-time smFRET assays, and Dr Horst Steuer for the development of custom software. This work was supported by the Wellcome Trust [110164/Z/15/Z to ANK] and NIH [GM041376 to RHE].

The illustrations in Figures 1A, 2A, 3A, and 4 were created using Biorender (Biorender.com).

Additional information

Funding

Funder	Grant reference number	Author
Wellcome Trust	110164/Z/15/Z	Achillefs Kapanidis
NIH Office of the Director	GM041376	Richard H Ebright

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Abhishek Mazumder, Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – original draft, Investigation, Methodology; Richard H Ebright, Conceptualization, Resources, Writing – review and editing, Methodology; Achillefs N Kapanidis, Conceptualization, Funding acquisition, Resources, Writing – review and editing, Data curation, Investigation, Methodology

Author ORCIDs

Abhishek Mazumder  <http://orcid.org/0000-0002-9339-6256>

Richard H Ebright  <http://orcid.org/0000-0001-8915-7140>

Achillefs N Kapanidis  <http://orcid.org/0000-0001-6699-136X>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.70090.sa1>

Author response <https://doi.org/10.7554/eLife.70090.sa2>

Additional files

Supplementary files

- Transparent reporting form

Data availability

All information for replication is included in the submission and data corresponding to each figure are provided as source data files. MATLAB software packages TwoTone and ebFRET are available on Github (<https://github.com/annawang692/TwoTone2018> and <http://ebfret.github.io/>).

References

- Auner H, Buckle M, Deufel A, Kutateladze T, Lazarus L, Mavathur R, Muskhelishvili G, Pemberton I, Schneider R, Travers A. 2003. Mechanism of transcriptional activation by FIS: role of core promoter structure and DNA topology. *Journal of Molecular Biology* 331: 331–344. DOI: [https://doi.org/10.1016/S0022-2836\(03\)00727-7](https://doi.org/10.1016/S0022-2836(03)00727-7), PMID: 12888342

- Axelrod D**, Thompson NL, Burghardt TP. 1983. Total internal reflection fluorescent microscopy. *Journal of Microscopy* **129**: 19–28. DOI: <https://doi.org/10.1111/j.1365-2818.1983.tb04158.x>, PMID: 6827590
- Bae B**, Feklistov A, Lass-Napiorkowska A, Landick R, Darst SA. 2015. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife* **4**: e08504. DOI: <https://doi.org/10.7554/eLife.08504>, PMID: 26349032
- Belogurov G**, Vassilyeva MN, Svetlov V, Klyuyev S, Grishin NV, Vassilyev DG, Artsimovitch I. 2007. Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Molecular Cell* **26**: 117–129. DOI: <https://doi.org/10.1016/j.molcel.2007.02.021>
- Belogurov G.A**, Vassilyeva MN, Sevostyanova A, Appleman JR, Xiang AX, Lira R, Webber SE, Klyuyev S, Nudler E, Artsimovitch I, Vassilyev DG. 2009. Transcription inactivation through local refolding of the RNA polymerase structure. *Nature* **457**: 332–335. DOI: <https://doi.org/10.1038/nature07510>, PMID: 18946472
- Boyaci H**, Chen J, Jansen R, Darst SA, Campbell EA. 2019. Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature* **565**: 382–385. DOI: <https://doi.org/10.1038/s41586-018-0840-5>, PMID: 30626968
- Chakraborty A**, Wang D, Ebricht YW, Korlann Y, Kortkhonjia E, Kim T, Chowdhury S, Wigneshweraraj S, Irschik H, Jansen R, Nixon BT, Knight J, Weiss S, Ebricht RH. 2012. Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**: 591–595. DOI: <https://doi.org/10.1126/science.1218716>, PMID: 22859489
- Chen J**, Chiu C, Gopalkrishnan S, Chen AY, Olinares PDB, Saecker RM, Winkelman JT, Maloney MF, Chait BT, Ross W, Gourse RL, Campbell EA, Darst SA. 2020. Stepwise Promoter Melting by Bacterial RNA Polymerase. *Molecular Cell* **78**: 275–288. DOI: <https://doi.org/10.1016/j.molcel.2020.02.017>
- Cowing DW**, Mecas J, Record MT Jr, Gross CA. 1989. Intermediates in the formation of the open complex by RNA polymerase holoenzyme containing the sigma factor sigma 32 at the groE promoter. *Journal of Molecular Biology* **210**: 521–530. DOI: [https://doi.org/10.1016/0022-2836\(89\)90128-9](https://doi.org/10.1016/0022-2836(89)90128-9), PMID: 2693737
- Darst SA**, Opalka N, Chacon P, Polyakov A, Richter C, Zhang G, Wriggers W. 2002. Conformational flexibility of bacterial RNA polymerase. *PNAS* **99**: 4296–4301. DOI: <https://doi.org/10.1073/pnas.052054099>, PMID: 11904365
- Davis CA**, Bingman CA, Landick R, Record MT Jr, Saecker RM. 2007. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *PNAS* **104**: 7833–7838. DOI: <https://doi.org/10.1073/pnas.0609888104>, PMID: 17470797
- Duchi D**, Bauer DLV, Fernandez L, Evans G, Robb N, Hwang LC, Gryte K, Tomescu A, Zawadzki P, Morichaud Z, Brodolin K, Kapanidis AN. 2016. RNA Polymerase Pausing during Initial Transcription. *Molecular Cell* **63**: 939–950. DOI: <https://doi.org/10.1016/j.molcel.2016.08.011>, PMID: 27618490
- Duchi D**, Mazumder A, Malinen AM, Ebricht RH, Kapanidis AN. 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. DOI: <https://doi.org/10.1093/nar/gky482>, PMID: 29878276
- Dulin D**, Bauer DLV, Malinen AM, Bakermans JJW, Kaller M, Morichaud Z, Petushkov I, Depken M, Brodolin K, Kulbachinskiy A, Kapanidis AN. 2018. Pausing controls branching between productive and non-productive pathways during initial transcription in bacteria. *Nature Communications* **9**: e03902-9. DOI: <https://doi.org/10.1038/s41467-018-03902-9>
- Duval-Valentin G**, Ehrlich R. 1986. Interaction between *E. coli* RNA polymerase and the tetR promoter from pSC101: homologies and differences with other *E. coli* promoter systems from close contact point studies. *Nucleic Acids Research* **14**: 1967–1983. DOI: <https://doi.org/10.1093/nar/14.5.1967>, PMID: 3960716
- Feklistov A**, Darst SA. 2011. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**: 1257–1269. DOI: <https://doi.org/10.1016/j.cell.2011.10.041>, PMID: 22136875
- Feklistov A**, Bae B, Hauver J, Lass-Napiorkowska A, Kalesse M, Glaus F, Altmann KH, Heyduk T, Landick R, Darst SA. 2017. RNA polymerase motions during promoter melting. *Science* **356**: 863–866. DOI: <https://doi.org/10.1126/science.aam7858>, PMID: 28546214
- Floyd DL**, Harrison SC, van Oijen AM. 2010. Analysis of kinetic intermediates in single-particle dwell-time distributions. *Biophysical Journal* **99**: 360–366. DOI: <https://doi.org/10.1016/j.bpj.2010.04.049>, PMID: 20643053
- Glyde R**, Ye F, Jovanovic M, Kotta-Loizou I, Buck M, Zhang X. 2018. Structures of bacterial RNA Polymerase complexes reveal the mechanism of DNA loading and transcription initiation. *Molecular Cell* **70**: 1111–1120. DOI: <https://doi.org/10.1016/j.molcel.2018.05.021>, PMID: 29932903
- Helmann JD**, deHaseth PL. 1999. Protein–Nucleic Acid Interactions during Open Complex Formation Investigated by Systematic Alteration of the Protein and DNA Binding Partners. *Biochemistry* **38**: 5959–5967. DOI: <https://doi.org/10.1021/bi990206g>
- Holden SJ**, Uphoff S, Hohlbein J, Yadin D, Le Reste L, Britton OJ, Kapanidis AN. 2010. Defining the limits of single-molecule FRET resolution in TIRF microscopy. *Biophysical Journal* **99**: 3102–3111. DOI: <https://doi.org/10.1016/j.bpj.2010.09.005>, PMID: 21044609
- Hudson BP**, Quispe J, Lara-González S, Kim Y, Berman HM, Arnold E, Ebricht RH, Lawson CL. 2009. Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *PNAS* **106**: 19830–19835. DOI: <https://doi.org/10.1073/pnas.0908782106>, PMID: 19903881
- Hwang H**, Myong S. 2014. Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions. *Chemical Society Reviews* **43**: 1221–1229. DOI: <https://doi.org/10.1039/c3cs60201j>, PMID: 24056732

- Kapanidis AN**, Lee NK, Laurence TA, Doose S, Margeat E, Weiss S. 2004. Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *PNAS* **101**: 8936–8941. DOI: <https://doi.org/10.1073/pnas.0401690101>, PMID: 15175430
- Kapanidis AN**, Margeat E, Ho SO, Kortkhonja E, Weiss S, Ebright RH. 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**: 1144–1147. DOI: <https://doi.org/10.1126/science.1131399>, PMID: 17110578
- Ko J**, Heyduk T. 2014. Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse. *The Biochemical Journal* **463**: 135–144. DOI: <https://doi.org/10.1042/BJ20140179>, PMID: 24995916
- Koh HR**, Roy R, Sorokina M, Tang GQ, Nandakumar D, Patel SS, Ha T. 2018. Correlating Transcription Initiation and Conformational Changes by a Single-Subunit RNA Polymerase with Near Base-Pair Resolution. *Molecular Cell* **70**: 695–706. DOI: <https://doi.org/10.1016/j.molcel.2018.04.018>, PMID: 29775583
- Lee NK**, Kapanidis AN, Wang Y, Michalet X, Mukhopadhyay J, Ebright RH, Weiss S. 2005. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophysical Journal* **88**: 2939–2953. DOI: <https://doi.org/10.1529/biophysj.104.054114>, PMID: 15653725
- Lim HM**, Lee HJ, Roy S, Adhya S. 2001. A “master” in base unpairing during isomerization of a promoter upon RNA polymerase binding. *PNAS* **98**: 14849–14852. DOI: <https://doi.org/10.1073/pnas.261517398>, PMID: 11734629
- Lin W**, Das K, Degen D, Mazumder A, Duchi D, Wang D, Ebright YW, Ebright RY, Sineva E, Gigliotti M, Srivastava A, Mandal S, Jiang Y, Liu Y, Yin R, Zhang Z, Eng ET, Thomas D, Donadio S, Zhang H, et al. 2018. Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). *Molecular Cell* **70**: 60–71. DOI: <https://doi.org/10.1016/j.molcel.2018.02.026>
- Mazumder A**, Kapanidis AN. 2019. Recent Advances in Understanding σ 70-Dependent Transcription Initiation Mechanisms. *Journal of Molecular Biology* **431**: 3947–3959. DOI: <https://doi.org/10.1016/j.jmb.2019.04.046>, PMID: 31082441
- Mazumder A**, Wang A, Uhm H, Ebright RH, Kapanidis AN. 2021. RNA polymerase clamp conformational dynamics: long-lived states and modulation by crowding, cations, and nonspecific DNA binding. *Nucleic Acids Research* **49**: 2790–2802. DOI: <https://doi.org/10.1093/nar/gkab074>
- Mekler V**, Kortkhonja E, Mukhopadhyay J, Knight J, Revyakin A, Kapanidis AN, Niu W, Ebright YW, Levy R, Ebright RH. 2002. Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**: 599–614. DOI: [https://doi.org/10.1016/S0092-8674\(02\)00667-0](https://doi.org/10.1016/S0092-8674(02)00667-0), PMID: 11893332
- Mukhopadhyay J**, Mekler V, Kortkhonja E, Kapanidis AN, Ebright YW, Ebright RH. 2003. Fluorescence resonance energy transfer (FRET) in analysis of transcription-complex structure and function. *Methods Enzymol* **371**: 144–159. DOI: [https://doi.org/10.1016/S0076-6879\(03\)71010-6](https://doi.org/10.1016/S0076-6879(03)71010-6)
- Mukhopadhyay J**, Das K, Ismail S, Koppstein D, Jang M, Hudson B, Sarafianos S, Tuske S, Patel J, Jansen R, Irschik H, Arnold E, Ebright RH. 2008. The RNA polymerase “switch region” is a target for inhibitors. *Cell* **135**: 295–307. DOI: <https://doi.org/10.1016/j.cell.2008.09.033>, PMID: 18957204
- Murakami KS**, Masuda S, Campbell EA, Muzzin O, Darst SA. 2002a. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285–1290. DOI: <https://doi.org/10.1126/science.1069595>, PMID: 12016307
- Murakami KS**, Masuda S, Darst SA. 2002b. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296**: 1280–1284. DOI: <https://doi.org/10.1126/science.1069594>, PMID: 12016306
- Murakami KS**, Darst SA. 2003. Bacterial RNA polymerases: the whole story. *Current Opinion in Structural Biology* **13**: 31–39. DOI: [https://doi.org/10.1016/s0959-440x\(02\)00005-2](https://doi.org/10.1016/s0959-440x(02)00005-2), PMID: 12581657
- Niu W**, Kim Y, Tau G, Heyduk T, Ebright RH. 1996. Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell* **87**: 1123–1134. DOI: [https://doi.org/10.1016/s0092-8674\(00\)81806-1](https://doi.org/10.1016/s0092-8674(00)81806-1), PMID: 8978616
- Rogozina A**, Zaychikov E, Buckle M, Heumann H, 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. DOI: <https://doi.org/10.1093/nar/gkp560>, PMID: 19578065
- Ruff EF**, Record MT, Artsimovitch I. 2015. Initial events in bacterial transcription initiation. *Biomolecules* **5**: 1035–1062. DOI: <https://doi.org/10.3390/biom5021035>, PMID: 26023916
- Sambrook J**, Russell D. 2001. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory.
- Schickor P**, Metzger W, Werel W, Lederer H, Heumann H. 1990. Topography of intermediates in transcription initiation of *E. coli*. *The EMBO Journal* **9**: 2215–2220. PMID: 2192861.
- Spassky A**, Kirkegaard K, Buc H. 1985. Changes in the DNA structure of the lac UV5 promoter during formation of an open complex with *Escherichia coli* RNA polymerase. *Biochemistry* **24**: 2723–2731. DOI: <https://doi.org/10.1021/bi00332a019>, PMID: 3896305
- Suh WC**, Ross W, Record MT. 1993. Two open complexes and a requirement for Mg²⁺ to open the lambda PR transcription start site. *Science* **259**: 358–361. DOI: <https://doi.org/10.1126/science.8420002>, PMID: 8420002
- Unarta IC**, Cao S, Kubo S, Wang W, Cheung PP-H, Gao X, Takada S, Huang X. 2021. Role of bacterial RNA polymerase gate opening dynamics in DNA loading and antibiotics inhibition elucidated by quasi-Markov State Model. *PNAS* **118**: e2024324118. DOI: <https://doi.org/10.1073/pnas.2024324118>

- van de Meent J-W**, Bronson JE, Wiggins CH, Gonzalez RL Jr. 2014. Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments. *Biophysical Journal* **106**: 1327–1337. DOI: <https://doi.org/10.1016/j.bpj.2013.12.055>, PMID: 24655508
- Vassilyev DG**, Sekine S, Laptenko O, Lee J, Vassilyeva MN, Borukhov S, Yokoyama S. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**: 712–719. DOI: <https://doi.org/10.1038/nature752>, PMID: 12000971
- Wang A**. 2020. TwoTone2018. 9e8d7d1. Github. <https://github.com/annawang692/TwoTone2018>
- Winkelman JT**, Vvedenskaya IO, Zhang Y, Zhang Y, Bird JG, Taylor DM, Gourse RL, Ebricht RH, Nickels BE. 2016. Multiplexed protein-DNA cross-linking: Scrunching in transcription start site selection. *Science* **351**: 1090–1093. DOI: <https://doi.org/10.1126/science.aad6881>, PMID: 26941320
- Winkelman JT**, Pukhrambam C, Vvedenskaya IO, Zhang Y, Taylor DM, Shah P, Ebricht RH, Nickels BE. 2020. XACT-Seq Comprehensively Defines the Promoter-Position and Promoter-Sequence Determinants for Initial-Transcription Pausing. *Molecular Cell* **79**: 797–811. DOI: <https://doi.org/10.1016/j.molcel.2020.07.006>, PMID: 32750314
- Young BA**, Gruber TM, Gross CA. 2002. Views of transcription initiation. *Cell* **109**: 417–420. DOI: [https://doi.org/10.1016/S0092-8674\(02\)00752-3](https://doi.org/10.1016/S0092-8674(02)00752-3), PMID: 12086598
- Zhang Y**, Feng Y, Chatterjee S, Tuske S, Ho MX, Arnold E, Ebricht RH. 2012. Structural basis of transcription initiation. *Science* **338**: 1076–1080. DOI: <https://doi.org/10.1126/science.1227786>, PMID: 23086998
- Zuo Y**, Steitz TA. 2015. Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Molecular Cell* **58**: 534–540. DOI: <https://doi.org/10.1016/j.molcel.2015.03.010>, PMID: 25866247