SINGLE-MOLECULE FLUORESCENCE ANALYSIS OF OPENING AND CLOSING OF THE RNA POLYMERASE CLAMP

by

ANIRBAN CHAKRABORTY

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Biochemistry

written under the direction of

Prof. Richard H. Ebright

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2013

ABSTRACT OF THE DISSERTATION

Single-molecule fluorescence analysis of opening and closing of the RNA polymerase clamp

By ANIRBAN CHAKRABORTY

Dissertation Director:
Richard H. Ebright

Crystal structures of RNA polymerase (RNAP) indicate that the RNAP β' pincer ("clamp") can exist in conformational states, ranging from a fully open conformation that permits entry and exit of DNA, to a fully closed conformation that prevents entry and exit of DNA. It has been hypothesized that the clamp also adopts multiple conformational states in solution and conformational changes in the clamp are important for function.

In this work, a single-molecule fluorescence resonance energy transfer (smFRET) approach was developed that enables determination of RNAPclamp conformation in solution. smFRET was measured between a probe at the tip of the RNAP clamp and a probe at a fixed reference point in RNAP. A computational framework was then employed to interpret measured FRET efficiencies in terms of structural changes.

Using this approach, RNAP clamp conformation was defined in each step of σ^{70}-dependent transcription initiation and elongation and in each step in σ^{54}-dependent transcription initiation. Additionally, effects of four RNAP inhibitors, myxopyronin, corallopyronin, ripostatin and Gp2 on RNAP clamp conformation were assessed.

It was observed that the clamp is predominantly open in free RNAP and in all steps leading up to the formation of a catalytically-competent-transcription-initiation complex. Upon formation of a catalytically-competent-transcription-initiation complex, the clamp closes, and continues to remain closed during transcription elongation. It was further observed that myxopyronin, corallopyronin, ripostatin and Gp2, prevent opening of the RNAP clamp.

The results lead to the proposal that, the open clamp state is important for entry of DNA into, and unwinding of DNA in, the RNAP active center cleft during formation of a catalytically-competent-transcription initiation complex. The results lead to the proposal that, after entry of DNA into the RNAP active-center cleft upon formation of the catalytically competent transcription initiation complex, electrostatic interactions between the negatively charged DNA and the positively charged inner facet of the clamp, induce and/or stabilize clamp closure. The results are in agreement with the proposal that, clamp closure is important for stability of the catalytically competent transcription initiation complex and for stability and processivity of the transcription elongation complex.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to Prof. Richard Ebright for his insights and continuous support. Roadblocks encountered in this work were many, and spanned over multiple disciplines. But I always found Dr. Ebright steering me in the correct direction with his exemplary guidance. His undivided enthusiasm and passion for science is a trait I hope to acquire.

I would like thank my thesis committee members, Prof. Edward Arnold, Prof. Charalampos Kalodimos and Prof. Bryce Nickels for careful reading of my thesis.

I would like to extend my gratitude to Prof. Shimon Weiss for providing me the opportunity to visit his lab and learn the nuances of single-molecule fluorescence spectroscopy.

I would like to thank Dr. Jennifer knight for developing the molecular-modelling framework, which was used in this work.

I would like to thank Dr. Yon Ebright her cooperation during the past five years. I would like to thank her, in particular, for the help in development and execution of organic synthesis procedures.

I would like to thank each of the past and present members of the Ebright laboratory, I have the interacted with, for their help and advice.

I will be forever indebted to my wife, Kajori Chakraborty, for all the sacrifices she has made to support my efforts. Without her presence, nothing would have been possible.

I would like to thank my parents, Amit Chakraborty and Anita Chakraborty, and my late grandfather, Amal Das, for their support and belief in me.

Results presented in this dissertation have been published in two journal articles – Chakraborty et al., *Science* (2012) and Chakraborty et al., *Methods in enzymology* (2010).

DEDICATION

Dedicated to my wife Kajori

Table of Contents

Lists of tables

List of illustrations

1. Introduction

The process of DNA directed RNA synthesis is called transcription. Transcription is the first step in gene expression. Transcription is also the most regulated step in gene expression. The enzyme responsible for transcription is called RNA polymerase (RNAP). Multi-subunit RNA polymerases are related in sequence, structure and function, in all three domains of life – bacteria, archea and eukarya. All multi-subunit RNAPs contain, at minimum, five conserved subunits — β', β, α^{I}, α^{II}, and ω in bacteria; A, B, D, L, and K in archaea; RPB1, RPB2, RPB3, RPB11, and RPB6 in eukarya. Bacterial RNAP consists of only the five conserved subunits. Archeal and eukaryotic enzymes additionally contain 8 and 7 more subunits respectively [1-7]. Bacterial RNAP is the simplest and the most studied form of multi-subunit RNAP, and thus provides the framework for understanding more evolutionarily complex systems in archea and eukarya.

1.1 Bacterial RNAP core

1.1.1 Overview

Bacterial RNAP core contains five subunits -- β', β, α^{I}, α^{II}, and ω (Fig.1). Bacterial RNAP core can carry out all the steps of transcription -- it can initiate transcription, elongate the RNA chain and terminate RNA synthesis. However, it can do this with no selectivity based on sequence determinants in DNA [8, 9].

High resolution crystal structures of *Thermus thermophilus* RNAP and *Thermus aquaticus* RNAP have been solved [10, 11]. A low resolution cryo-electron-microscopy structure of *E.coli* RNAP has been solved [12]. The structures reveal that RNAP core has dimensions of ~150 Å × ~100 Å × ~100 Å with molecular mass of ~380 kDa. It has a shape reminiscent of a crab claw with two pincers (Fig.1). One of the pincers, formed by the largest subunit of RNAP, β', is called the clamp [4, 12-14]. The other pincer is formed by the second largest subunit of RNAP, β. The two pincers define a central cleft which serves as the active center cleft of the enzyme (Fig.1). The active center is marked by the presence of a Mg^{2+} ion which is chelated in position by three aspartate residues.

RNAP employs a two- metal-ion reaction mechanism for catalytic nucleotide addition by phosphodiester bond formation. In this mechanism, a first metal ion lowers the affinity of the 3'-OH for hydrogen, enabling a 3'-O^- nucleophilic attack on the α-phosphate and the second metal ion assists in leaving of the released pyrophosphate. The active site Mg^{2+} plays a pivotal role in this process by serving as the first metal ion. The second metal ion is brought in by the incoming NTP [15]. In addition to phosphodiester bond formation during RNA synthesis, the active center of the enzyme is also weakly capable of hydrolysis of phosphodiester bonds. During hydrolysis of phosphodiester bonds, the affinity of the active center for the incoming Mg^{2+} is reduced. The RNAP active center modulates its affinity for the incoming Mg^{2+} by an unknown mechanism to switch from RNA synthesis mode to editing mode in order to recover from backtracked states or repair misincorporated NTPs [16, 17].

In addition to the main active center cleft, RNA polymerase contains two other channels which connect the active center of the enzyme to the external milieu –the NTP-entrance channel and the RNA-exit channel. The NTP-entrance channel, also known as the secondary channel serves as a passage for inlet of substrate NTPs and inhibitory molecules into the RNAP active center (Fig.1B). Protein molecules such as Gre factors are also known to utilize the secondary channel for exerting their regulatory effects [18-20]. The RNA-exit channel serves as a passage for outlet of the growing RNA chain from the active center cleft. The active center of RNA can accommodate a DNA/RNA hybrid of 9 nt length. Any further addition of nucleotide to the growing RNA transcript requires entry of the RNA into the RNA exit channel. The RNA exit can accommodate and protect 5 nt of RNA before the RNA emerges out of the enzyme into solution [21, 22].

RNAP core contains three putative mobile modules – the β' clamp, the β' jaw, β lobe – which can potentially undergo rigid body rotations and occupy different regions in space relative to each other (Fig.1). The β' clamp has been proposed to be involved in modulating the dimensions of the active center cleft, permitting entry DNA into the active center, containment of DNA in the active center and exit of DNA from the active center [23, 24]. The β' clamp has also been proposed to be a hotspot of regulatory activities.

The β' jaw has been suggested to stabilize RNAP-promoter-open complex by establishing interactions with the far downstream region of the promoter (+10 to +20) [25, 26].

β lobe has been suggested to participate in DNA unwinding, during formation of RPo, by directing the downstream DNA towards β' clamp [27].

1.1.2. RNAP subunits

The β' subunit is the largest subunit of RNAP with a mass of ~155 kDa (Fig.1). It forms the RNAP clamp and makes crucial interactions with DNA and RNA. It makes up a large part of the active center and consequently contains most of the determinants for the catalysis.

β is the second largest subunit of RNAP (Fig.1). Together with the β' subunit it comprises the bulk mass of RNAP. It forms the β pincer and contains sequence determinants for interactions with DNA and RNA. It also partly makes up the active center of RNAP.

β' and β subunits make extensive interactions with each other. They participate in the formation of the active center of the enzyme which is located at the bottom of the cleft. The β' subunit forms a helical structure at the bottom of the cleft called the bridge helix which spans the active center of the enzyme. The β' subunit also forms a flexible loop in the vicinity of the active center Mg^{2+}, which plays a crucial role in catalysis during processive RNA synthesis [21, 28].

β' and β' subunits participate in the formation of the 'switch region', located at the base of the β' pincer. Switch region consists of 5conserved switches – switch 1, switch 2 and

switch 5 belong to the β' subunit, whereas switch 3 and switch 4 belong to the β subunit [29, 30].

α is the third largest subunit of RNAP (Fig.1). α is present in two copies – α^{I} and α^{II}, which are identical in sequence and structure. α^{I} and α^{II} make specific interactions with the rest of the RNAP -- α^{I} interacts with the β subunit and α^{II} interacts with the β' subunit. α is involved in interaction with upstream DNA sequences and interaction with a vast majority of regulatory factors. Each α subunit consists of a N-terminal domain (NTD; amino acids 1-235), an unstructured flexible linker and a C-terminal domain (CTD) (amino acids 245-329) [10, 11, 31]. α NTD contains structural determinants for dimerization of two α subunits. The α^{I}NTD-α^{II}NTD dimer acts as a scaffold for initiating the assembly of the RNAP core. α-CTD is separated from α-NTD by an unstructured linker. The unstructured linker confers a certain degree of flexibility to α-CTD to occupy various positions on the promoters. α-CTD performs two important regulatory functions: i) it is involved in interaction with DNA, particularly at sites upstream of the -35 element, called UP-elements. ii) it is involved in interaction with transcription regulators which bind to an upstream site on the DNA. Additionally, αCTD contains weak structural determinants for dimerization [32, 33].

ω is the smallest of all the subunits of RNAP and associates primarily with the β' subunit (Fig.1). ω is non-essential for function of RNAP. It is also non-essential for growth under normal conditions [34]. However it performs the critical function of acting as a molecular chaperone. ω has a structure that resembles a latch and consequently, it fastens together

the N-terminus and C-terminus of the unfolded β' subunit, reducing the configurational entropy of β'. This low entropic state facilitates the folding of β'[35]. There have been also some studies implying a role of ω in stringent response [36].

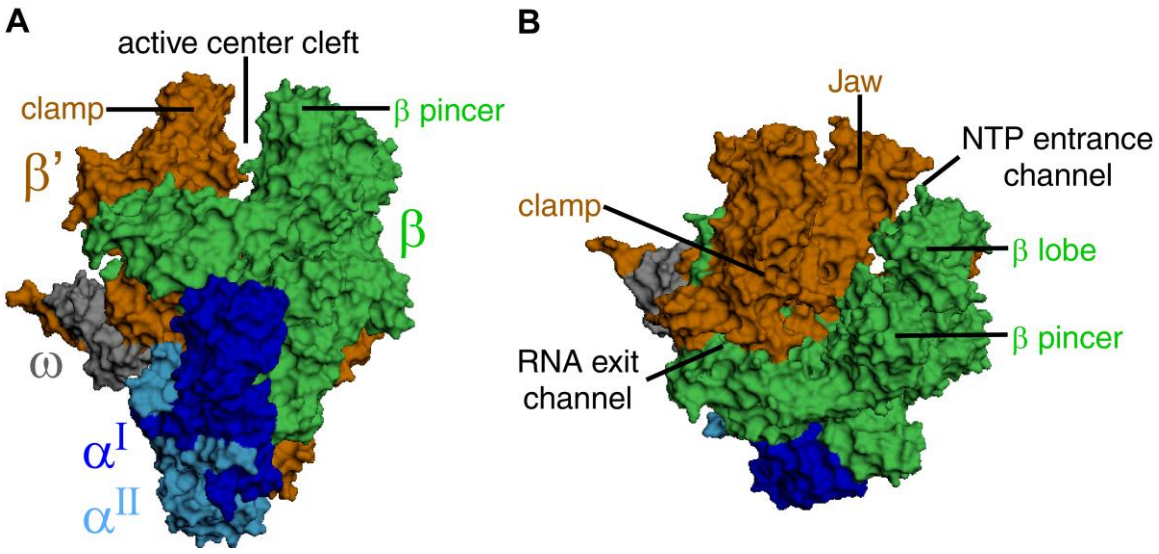


Fig. 1. RNAP core.
β' is shown in orange; β is hown in green; α^{I} is shown in light blue; α^{II} is shown in dark blue; ω is shown in gray.

(A) Upstream view.

(B) Topographic view.

1.2. Bacterial RNAP holoenzyme

In order to efficiently initiate transcription from specific promoter DNA sequences, RNAP core binds to a dissociable initiation factor, σ, forming RNAP holoenzyme (Fig.2A, B) [13, 14, 21, 37]. Apart from their roles in recognizing promoter sequences, σ factors play critical roles in promoter-unwinding, promoter-escape and transcriptional pausing [38, 39]. σ factors are also targets for various transcription factors [40-42].

Most bacterial species contain multiple σ factors. However, only one of the σ factors is responsible for transcription from a vast majority of promoters and thus performs the housekeeping function. In case of *E.coli,* the primary σ factor is a 70 kDa protein called σ^{70}. σ^{70} makes extensive interactions with RNAP core, by spreading across the upstream face of RNAP core (Fig.2A, B) [10, 41, 43-47]. RNAP-σ^{70} holoenzyme binds to promoter DNA containing two sequence elements centered around positions -35 (TATAAT) and -10 (TTGACA) upstream of the +1 transcription start site. σ^{70} contains four distinct domains: σ-region-1.1 (σR1.1), σ-region-2 (σR2), σ-region-3 (σR3) and σ-region-4 (σR4). Determinants in σR1.1, σR2 and σR3 domains recognize distinct promoter elements [14, 43, 48-50].

σR1.1 is an unstructured, flexible, N-terminal region of σ^{70} comprising of residues 1 through ~100. In free σ^{70}, σR1.1 occludes the DNA binding activity of the σR2 and σR4 [44, 51]. Since σ^{70} is present in about 1.6 molar excess over RNAP *in vivo,* auto-

inhibition of DNA binding in σ^{70} ensures that no promoter is pre-occupied by σ^{70} and transcription can proceed only upon ninding of RNAP –σ^{70} holoenzyme.

In addition to the crystal structure of RNAP-σ^{70} holoenzyme, crystal structures of *E.coli* and *T.acquaticus* σR2 have also been solved [10, 51, 52]. The structures reveal that σR2 is composed of three α-helices. One of the α-helices interacts with the β' subunit while another α-helix is involved in DNA melting and -10 element promoter recognition [10]. σR3 is also composed of three helices, one of which is involved in interactions with upstream DNA.

σR4 is composed of 2 pairs of α- helices. The C-terminal helix contacts promoter DNA between -30 to -38 regions. σR4 is also a hotspot for regulator binding [40, 41].

σR3 and σR4 are connected by an unstructured, highly negatively charged, flexible linker [51].

σR1.1 and σR4 serve as molecular mimics of DNA. In free RNAP-σ^{70} holoenzyme σR1.1 occupies the active center cleft and the σR3/4 linker occupies the RNA exit channel as placeholders for DNA, preventing nonspecific interactions. Prior to entry of DNA in the active center cleft and entry of RNA into the exit channel, σR1.1 and σR3/4 linker are ejected to permit melting of promoter DNA and RNA synthesis respectively [44].

In most bacterial species, in addition to primary sigma factors, a repertoire of alternate σ factors is available. Alternate σ factors can replace the primary σ factor, resulting in changed promoter preferences for a proportion of RNA polymerases present in the cell [49, 53]. Alternate σ factors are typically expressed under specific environmental stress

conditions and are utilized to drive expression from genes responsible for counteracting the stress. Variability of environmental conditions encountered by the bacteria appears to directly correlate with the diversity of σ factors expressed by the bacteria. For example, intracellular pathogen *Mycoplasma genitalium* contains only one σ factor, butgut bacteria *E.coli* contains 6 alternate σ factors and soil bacteria *Streptomyces coelicolor* contains 63 alternate σ factors. Alternate σ factors usually have short half-lives and are often quickly degraded upon removal of the environmental stress [49, 53, 54].

The best studied alternate σ factor, present in ~60% of bacterial species is a 54 kDa protein called σ^{54} [55]. σ^{54}, first identified for its role in Nitrogen fixation, has now been shown to be involved in various cellular responses such as utilization of carbon from alternative sources, cell motility and production of extracellular aligante [50, 56-59]. σ^{54} shares little sequence homology and promoter specificity with σ^{70} [49, 60]. In contrast to σ^{70}, σ^{54}containing holoenzyme recognizes and bind to sequence elements at positions -24 (GG) and -12 (TGC) and in a manner analogous to eukaryotic RNAP, requires assistance from an additional ATP dependent activator protein to perform unwinding of DNA [55, 61-63]. Biochemical and cryo-electron microscopic studies reveal that σ^{54} can be divided into three functional regions – region I, region II and region III (Fig.2C). Region I (residues 1-56 in *E.coli*) occupies the entrance to the active center cleft of RNAP, occluding the entry of DNA into, and unwinding of DNA in, the active cleft [60, 62-66]. Region I is also involved in interactions with the activator proteins and recognition of the -12 promoter DNA element. Region II (residues 57-107 in *E.coli*), which links Region I and Region III is variable in length and sometimes can even be absent. Region III

(residues 108-477 in *E.coli*) contains determinants for interaction with RNAP core and -24 promoter element. [64].

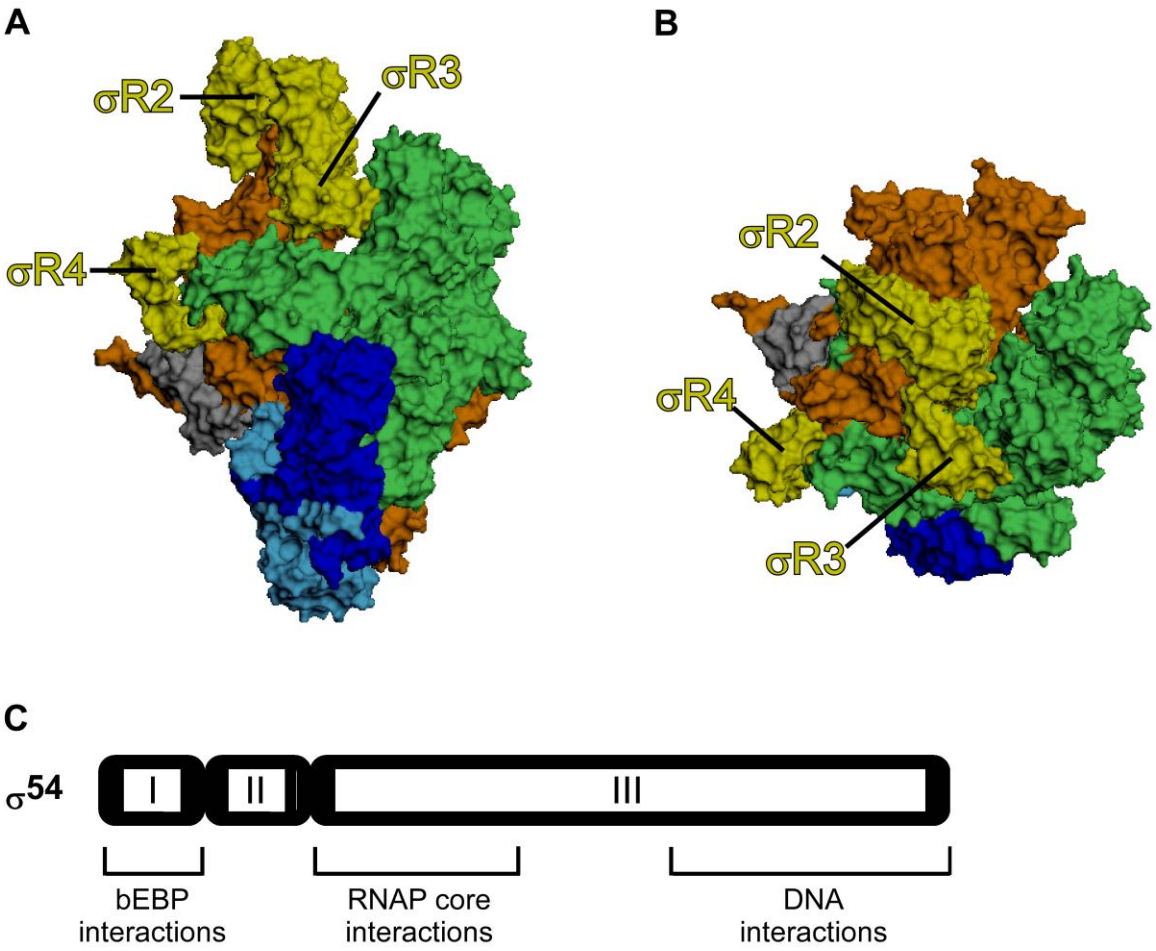


Fig. 2. RNAP holoenzyme:domain architechture of σ^{70} and functional regions of σ^{54}.
β' is shown in orange; β is shown in green; α^{I} is shown in light blue; α^{II} is shown in dark blue; ω is shown in gray; σ^{70} is shown in yellow.
(A) RNAP-σ70 holoenzyme: Upstream view.
(B) RNAP-σ70 holoenzyme: Topographic view.
(C) Functional regions of σ^{54}.

1.3. σ^{70}-dependent transcription initiation

σ^{70}-dependent transcription initiation is a multi-step process (Fig. 3) [41, 67]:

(i) RNAP binds to promoter DNA, yielding RNAP-promoter closed complex, in which dsDNA makes initial interactions with the enzyme and the DNA remains outside the RNAP active-center cleft (RPc; where "closed" refers to the fact that the DNA is double-stranded). At most promoters, RPc is unstable and it either rapidly dissociates or isomerizes to more stable complexes via various intermediates.

(ii) RNAP loads DNA into, and unwinds DNA in (from -12 to +2 positions, with respect to transcription start site, +1) the RNAP active-center cleft, yielding an RNAP-promoter open complex (RPo; where "open" refers to the fact that the DNA is partly single-stranded).

(iii) RNAP synthesizes the first ~10 nt of RNA using a "scrunching" mechanism, in which RNAP remains stationary relative to promoter DNA and pulls downstream DNA into itself (*15*,*22*), yielding an RNAP-promoter initial transcribing complex (RPitc).

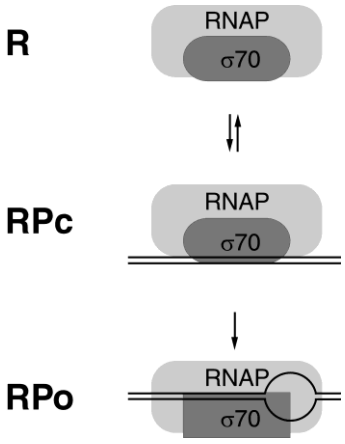


Fig. 3. Intermediates in σ^{70}-dependent transcription initiation.

1.4. σ^{54}-dependent transcription initiation

σ^{54}-dependent transcription initiation in bacteria is mechanistically different from the σ^{70}-dependent transcription initiation pathway [41, 62].

(i) σ^{54}-RNAP holoenzyme binds to a promoter containing key -12 (GG) and -24 (TGC) promoter elements, yielding stable RNAP-promoter closed complex, RPc. Unlike σ^{70} dependent RPc, σ54-dependent RPc is stable and cannot spontaneously isomerize to a RNAP-promoter open complex, RPo.

(ii) An inactive enhancer binding protein (bEBP; such as NtrC1 or PspF) dimer binds to an upstream enhancer binding site on DNA.

(iii) Binding of the bEBP dimer leads to the recruitment of additional subunits resulting in the formation of an active hexamer.

(iv) DNA looping (which can be optionally aided by DNA bending proteins) permits the bEBP to interact with RPc.

(v) Binding of ATP to the active site of bEBP causes conformational change in loop L1 of bEBP, which adopts a raised conformation allowing the bEBP to interact with RPc, yielding RNAP-promoter-intermediate-complex-1 (RPi1).

(vi) ATP hydrolysis causes remodeling of σ^{54} domains, resulting in downstream movement of promoter DNA and relocation of σ^{54} region I to partially block entry of DNA into the active center cleft of RNAP, yielding RNAP-promoter-intermediate-complex-2 (RPi2).

(vii) Completion of ATP hydrolysis results in lowering of the L1 loop and relocation of σ^{54}, allowing entry of DNA into the active center cleft and unwinding of DNA in the active center cleft yielding RPo.

(viii) ADP is released from the bEBP, resulting in dissociation of the bEBP from DNA.

(ix) Steps following formation of RPo are similar in σ^{54}-dependent initiation pathway and σ^{70}-dependent initiation pathway (see step iii of the preceding section).

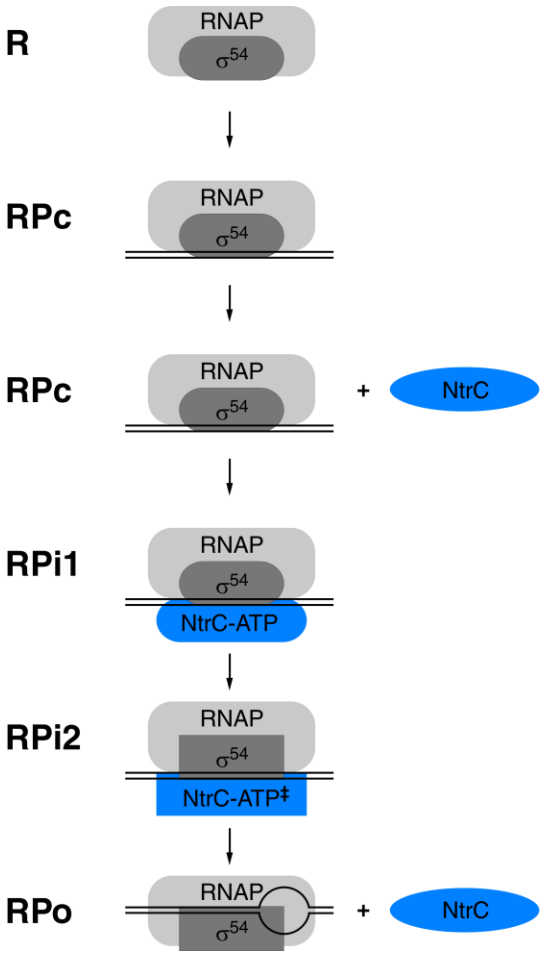


Fig. 4. Intermediates in σ^{54}-dependent transcription initiation.

1.5. Transcription elongation

RNAP initial transcribing complex is involved in abortive cycle of RNA production. Upon reaching a threshold of about 8-9 nt RNA product, RNA polymerase breaks free of its interactions with the promoter and begins processive RNA synthesis. At this stage structural and mechanistic rearrangements occur in RNAP leading to the formation of RNAP-DNA elongation complex (RDe). In RDe, the clamp undergoes a rotation of ~11˚ with respect to the RNAP-σ^{70}holoenzyme, to adopt a closed conformation. In RDe, interactions with σ^{70} are sufficiently weakened to allow optional release of σ^{70} during early stages of elongation (RNA product of up to 12 bases). σ^{70} is obligatorily released in mature elongation complexes (RNA product greater than 12 bases) [68-70].

Crystal structures have revealed that RNAP continues to maintain the crab claw shape in RDe [21]. The downstream DNA enters the active center cleft and forms the transcription bubble. The growing RNA is attached to the catalytic site with its 3'-end and forms an 8-9 bp hybrid duplex with the DNA strand. The upstream DNA enters the RNA exit channel.

Elongation complexes are marked by great processivity and stability. Single molecule force measurements have revealed that RNAP uses a stepping mechanism, in which it translocates 1 bp at a time relative to DNA [71].

In RDe, key structural dynamics have been proposed to occur in the vicinity of RNAP active center – involving the trigger loop. Crystallographic structures have indicated that the trigger loop cycles between the open and closed conformations [21, 28]. In the unfolded open conformation, catalytic residues on the trigger loop are far away for the

active site Mg^{2+}, and water molecules can freely access the active center of the enzyme. During catalysis, the trigger loop adopts a folded closed conformation which brings the catalytic residues on the loop close to the active center Mg^{2+} allowing synthesis of phosphodiester bonds and exclusion of unwanted water molecules.

1.6. *in crystallo* conformational states of RNAP clamp

Over a decade back, Roger Kornberg's group reported crystal structures of two structurally distinct forms of eukaryotic RNA polymerase II [23]. A striking difference between the two forms was a massive conformational change in the RNAP clamp. Since then, crystal structures of RNAP across different species, across various crystal contexts have captured the RNAP clamp in conformational states ranging from fully open clamp state to a fully closed (Fig. 5) [10-12, 21, 23, 24, 28, 45, 72, 73]. The open clamp state has dimensions which permit entry of dsDNA into the active center cleft and the closed state has dimension which do not permit entry or exit of DNA from the active center cleft. It has been proposed that clamp closure involves a swinging motion of the β' pincer, with rotation by ~30º about a hinge region at the base of β' pincer, and with displacement by ~30Å of residues at the tip of the β' pincer. The rotation about the hinge region is proposed to be mediated by a 'switch region' comprising of five flexible switches located at the base of the β' pincer (Fig.5) [23, 24, 29]. Upon formation of a catalytically competent transcription complex All five switches undergo conformational changes. Switches 1, 2 and 3 make direct interactions with the unwound DNA in the RNAP active center cleft, making it plausible that interactions involving ssDNA and the RNAP switch region may induce clamp closure.

Contradictory to the above proposals, an alternative model of transcription initiation exists. It has been proposed, that unwinding of DNA to form the transcription bubble occurs outside of the active center cleft, during which the clamp does not fully open and the strands of DNA are inserted in the active center cleft, one at a time [28]. This proposal, at least partly, stems from the fact that a fully open conformational state of the RNAP clamp has not been observed in bacterial RNAP holoenzyme crystal structures. Therefore, it remai0ns a possibility that the putative conformational states observed in RNAP crystal structures do not represent clamp states in solution and are results of crystal lattice artifacts.

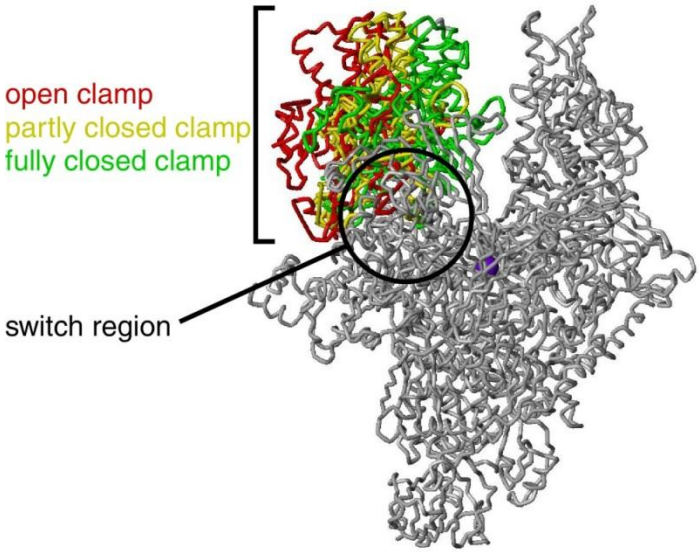


Fig. 5. RNAP clamp conformational states as observed in crystal structures.
Open, partly closed and closed states of the RNAP clamp (PDB 1I3Q, 1HQM, and 1I6H). σ and the β' non-conserved domain are omitted for clarity in this and subsequent figures.

1.7. Switch-region-target-inhibitors

Recent research has identified a RNAP switch region as a novel target for a class of small molecule inhibitors [29, 74, 75]. The switch region target inhibitors have been shown to bind to the switch region of RNAP -- a region proposed to mediate opening and closing of the RNAP clamp by serving as a hinge at the base of the β' pincer [5, 23, 29, 72]. The switch region is highly conserved across gram positive and gram negative bacterial species and does not overlap with the binding pocket of rifampicin, a popular drug for treatment of tuberculosis [29, 76].

Myxopyronin (Myx), corallopyronin (Cor), and ripostatin (Rip) belong to the class of switch-region-target inhibitors (Fig. 6). Myx, Cor and Rip have been proposed to block a step in the isomerization of RPc to RPo. It has been proposed that switch-region-target inhibitors function by inducing a conformational restrain on the movement of the RNAP clamp [29].

Myx is an α-pyrone antibiotic (MW = 417 Da for Myx A and 431 Da for Myx B) produced by the *Myxococcus fulvus* Mf50 (Fig. 6) [77, 78]. Myx potently inhibits both Gram-positive bacterial RNAP and Gram-negative bacterial RNAP, but does not inhibit eukaryotic RNAP [30]. Myxopyronin interacts with a significant section of the switch region, including switch 2 and amino acids on β' and β subunits adjacent to switch 2. Additionally it has been shown that Myx interferes with interactions between RNAP and

the promoter region of promoter (-11 to +15) that is typically unwound to form the transcription bubble [29, 30].

Cor is a α-pyrone antibiotic (MW = 542 Da) structurally related to Myx, differing by possession of a seven-carbon side-chain extension (Fig. 6). Cor is produced by the Myxobacterium *Corallococcus coralloides [79]*. Cor interacts with exactly the same amino acid residues as Myx, along with an addition residue 1326 on β subunit. Analogous to Myx, Cor interferes interaction of RNAP with promoter position -11 to +15 [29, 30].

Rip is a 14-membered macrocylic-lactone antibiotic (MW = 494 Da) structurally unrelated to Myx and Cor (Fig. 6). Rip is produced by the *Sorangium cellulosum* Soce377 [80]. The inhibitory properties of Rip are similar to Myx. In spite of the structural dissimilarities, Rip interacts with the exact same residues as Cor and also interferes with RNAP binding to promoter region spanning -11 to +15 [29, 30].

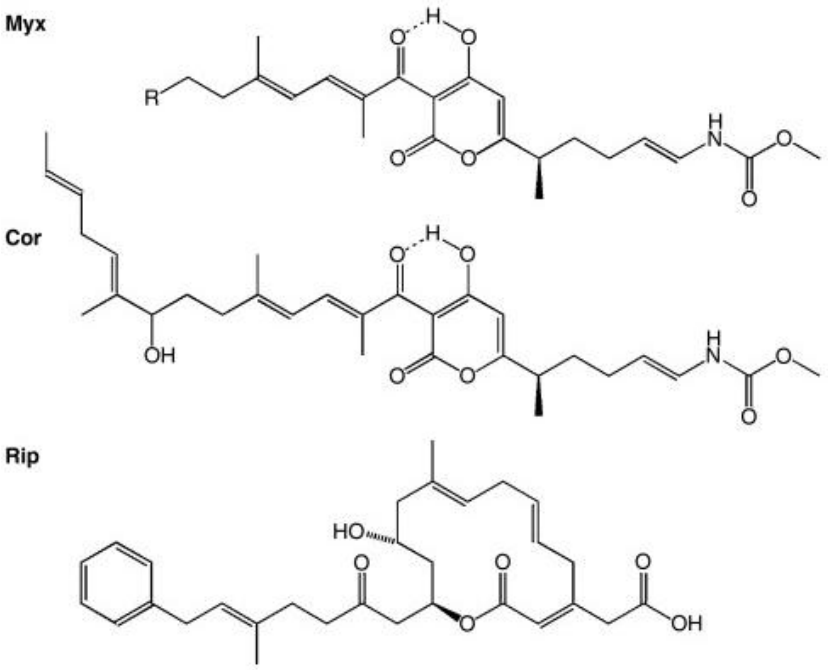


Fig. 6. Structures of switch-region-target inhibitors: Myxopyronin (Myx), Corallopyronin (Cor) and Ripostatin (Rip).

1.8. Gp2

The transcription of bacteriophage T7 genome proceeds in three stages: (i) early genes are transcribed by the host RNAP. (ii) middle genes are transcribed by the viral RNAP. (iii) late genes are transcribed by the viral RNAP.

Gp2 (7 kDa) is a non-DNA-binding transcription factor encoded by the middle genes of the T7 bacteriophage. Recent crystallographic and Biochemical studies indicate that Gp2 binds to the β' jaw domain of RNAP [81, 82].

Gp2 inhibits the function of host RNAP by blocking a late step in the isomerization of RPc to RPo [82, 83]. It has been proposed that Gp2 may interfere with the binding of downstream DNA to the β' jaw. However, it has been shown that Gp2 is capable of inhibiting RNAP in complex with DNA lacking DNA downstream of the +6 position. Thus, the mechanism of inhibition of RNAP by Gp2 may be a complex process involving allosteric effects in, or near, the active center of the RNAP, including effects on RNAP clamp and/or σR1.1[81, 83].

1.9. Single molecule Fluorescence resonance energy transfer

In a series of seminal papers from 1946 to1955, Theodor Foster outlined the theoretical basis of transfer of non-radiative electronic excitation energy between two molecules in solution, a process termed Fluorescence resonance energy transfer (FRET) [84-86]. Theodor Foster's work was built upon the work of Jean Perrin and Francis Perrin who observed that 2 chromophores in solution could interact without collision.

Quantum mechanical processes underlying FRET are outlined below (???):

1. A donor fluorophore absorbs energy to reach an excited singlet state, S1, from a ground state, S0.

2. Several excited states are available to the donor fluorophore, but rapid vibrational relaxation to S1 by internal conversion ensures majority of emission occurs from this state.

3. If a suitable acceptor is nearby, then non-radiative energy transfer can occur between the donor and the acceptor. The transfer involves resonance between the singlet-singlet electronic transitions of the two fluorophores generated by coupling of emission transition dipole moment of the donor and absorption transition dipole moment of the acceptor. Thus the efficiency of FRET and the range of distances over which it can occur are determined by spectral properties if a given donor-acceptor pair.

FRET can be measured either by decrease of quantum yield of the donor in the presence of an acceptor or by the enhancement of fluorescence emission of the acceptor in the presence of the donor. Quantum mechanics dictates that the efficiency of energy transfer is proportional to the inverse sixth power of the distance separating the fluorophores, R, given by the relation [86-88]:

$$E = 1 / [1+(R/R_0)^6]$$

Where R_0 is the donor-acceptor distance at which the energy transfer efficiency is half maximal. In other words, $R=R_0$ when FRET efficiency is 50%. R_0 in Å distance can be determined using the following relation:

$$R_0^6 = 8.8\times10^{-28}n^{-4}\kappa^2 Q_D J$$

where n is the refractive index of the medium (1.4; *46*); κ^2 is the orientation factor relating donor emission dipole and acceptor excitation dipole; Q_D is the quantum yield of the donor in the absence of acceptor; and J is the overlap integral of donor emission spectrum and acceptor excitation spectrum.

FRET is routinely used as a sensitive reporter of molecular distance in the range of ~20 to ~100 Å. Once limited to ensemble biophysical measurements, advances in microscopy techniques have enable determination of FRET at the level of single-molecules. Many complex biological processes are inherently heterogenous and/or stochastic in nature. In contrast to measurements which usually report a mean value representing an ensemble, single-molecule FRET (smFRET) measurements report population distributions of observed FRET species. smFRET observables permit the study processes such as conformational changes in biological macromolecules and multi-step reaction mechanisms, allowing identification of distinct and/or even rare states.

In-solution single-pair FRET (spFRET) measurements typically employ setups where a donor laser is focused tightly into a femtoliter scale excitation volume. Molecules present in an extremely dilutes solution, by virtue of Brownian diffusion freely enter the femtoliter-scale excitation volume and illuminated by the donor laser which generates a characteristic burst of photons from each of the single fluorophores [89-93]. If an acceptor fluorophore is present in close proximity (~20-~100 Å) [94], energy is transferred from donor to acceptor, and the acceptor also emits a burst of photons. The photons from donor and acceptor are counted in their respective donor and acceptor detector channels. The FRET efficiency for each donor-acceptor pair is determined by the ratio of photons detected in the acceptor channel and the number of photons detected in both donor and acceptor channels. Complications due to chemically and photophysically induced artifactual low FRET species limit the application of spFRET to only high FRET measurements. (greater than 50 Å inter-dye distance). Moreover, sub-stoichiometric complexes cannot be identified by spFRET measurements.

Alternating laser excitation (ALEX) based smFRET measurements can overcome the issues with spFRET measurements [68, 95, 96]. ALEX confocal microscopy employs two lasers, tightly focused into a femtoliter-scale excitation volume, alternating at donor excitation and acceptor excitation wavelengths (Fig. 7A). ALEX with micro-second time-scale alternation period enables illumination of freely diffusing single-molecules as they transit the femtoliter-scale excitation volume -- permitting recovery of distinct donor and acceptor emission signatures. The emission signatures consist of small bursts of photons, detected in the donor and/or acceptor channels, per single-molecule observed. Since, in

ALEX, emission from acceptor can be measured independent of donor excitation, it allows computation of a stoichiometry parameter, S, which reports on the relative stoichiometry of donor and acceptor probes for each molecule observed. ALEX also permits computation of FRET efficiency parameter, E, for each molecule observed. This allows the construction of an E/S plot, in which the S, with values ranging between 0 and 1, can be used to virtually sort species that contain both donor and acceptor dyes (clustered around S values of ~0.5) from donor-only (clustered around S values of ~1.0) species and acceptor only species (clustered around S values of ~0) (Fig. 7B). The desired doubly labelled species be selected and the corresponding E distributions can be determined. The results define the equilibrium population distributions of E values. The distributions can be fitted to Gaussians, in which the number of Gaussians correlates with the number of sub-populations and the mean of the Gaussian defines the mean E value of the sub-population.

ALEX measurements can be performed both in-solution and on surface immobilized complexes. In-solution ALEX measurements, used in this work have certain advantages over surface immobilized complexes [68, 95-98]:

(i) it does not require development of a surface immobilization chemistry

(ii) it can avoid structural and functional perturbations that can be induced by surface immobilization of protein molecules.

(iii) there are no photophysical issues related to close proximity to the surface (higher background, non-radiative transfer etc.)

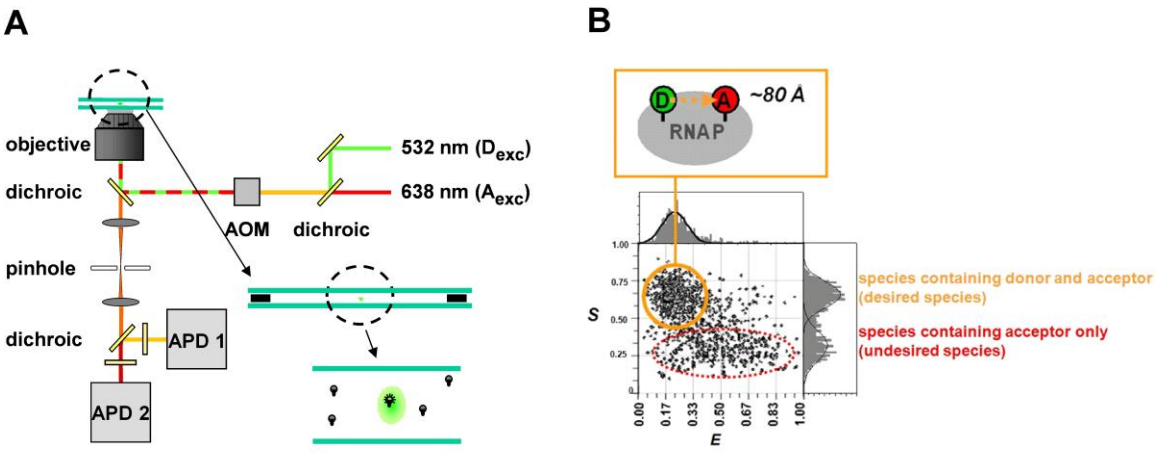


Fig. 7. Measurement of smFRET: two-color alternating laser excitation (ALEX).
(A) Experimental setup. Confocal optical microscopy with alternating laser excitation (ALEX) is used to generate a femtoliter-scale excitation and detection volume in a sample chamber (inset at right center) and to detect and quantify fluorescent signals from fluorescently labelled single-molecules freely diffusing in solution as they transit the femtoliter-scale excitation and detection volume (inset at lower right). The experimental setup contains two lasers: one providing excitation at the donor excitation wavelength (532 nm) and the other providing excitation at the acceptor excitation wavelength (638 nm). The lasers are alternated on the μs time scale using an acousto-optical modulator (AOM). Emission is monitored in the donor channel and the acceptor channel using avalanche photodiodes (APD 1 and APD 2).
(B) Two-color stoichiometry-based molecular sorting. Two-color ALEX yields two parameters for each molecule detected: a donor-acceptor stoichiometry parameter (*S*) and a donor-acceptor smFRET efficiency (*E*). The figure shows an *S/E* plot (at bottom), in which each dot represents a single observation--i.e., the observed value of *S* and the observed value of *E* for a single molecule transiting the femtoliter-scale excitation and detection volume. The distribution of observations on the *S*-axis (histogram at right) enables distinction between species containing both the donor and the acceptor (desired species) and species containing only the donor or only the acceptor (undesired species, comprising incompletely labeled complexes, incompletely assembled complexes, or complexes in photophysical dark states). Considering only observations from species containing both the donor and the acceptor, the distribution of *E* values on the *E*-axis (histogram at top) defines mean *E* and permits calculation of mean donor-acceptor distance, *R* (image at top). The results provide an equilibrium population distribution of *E*. The number of peaks in the distribution defines the number of distinguishable subpopulations, and, for each distinguishable subpopulation, defines mean *E* and mean *R*.

2. Experimental strategy

This work aims at defining the conformational states of RNAP clamp, in solution, at each step in transcription by use of single-molecule fluorescence resonance energy transfer (Fig 8A). The strategy involves:

(i) synthesis of phosphine derivatives of fluorescent donor and acceptor probes suitable for single-molecule fluorescence spectroscopy.

(ii) Use of unnatural amino acid mutagenesis and Staudinger ligation for site-specific incorporation of a first donor fluorescent probe at the tip of the β pincer serving as a fixed reference point (Fig 8B).

(iii) Use of unnatural amino acid mutagenesis and Staudinger ligation site-specific incorporation of a second donor fluorescent probe at the tip of the clamp serving as a mobile reporter of distance (Fig 8B).

(iv) *in vitro* reconstitution of RNAP using fluorescently labelled β' and β subunits.

(v) Development of a computational framework to interpret FRET efficiencies in terms of structural changes in RNAP (Fig 8B).

(vi)Development of a single-molecule FRET based assay to monitor clamp conformation at each step in transcription initiation and transcription elongation.

In addition to defining the role of RNAP clamp in each stage in transcription, this work also provides insight into possible mechanism of inhibition of transcription by small molecule inhibitors and a transcriptional regulator via conformational changes in the RNAP clamp.

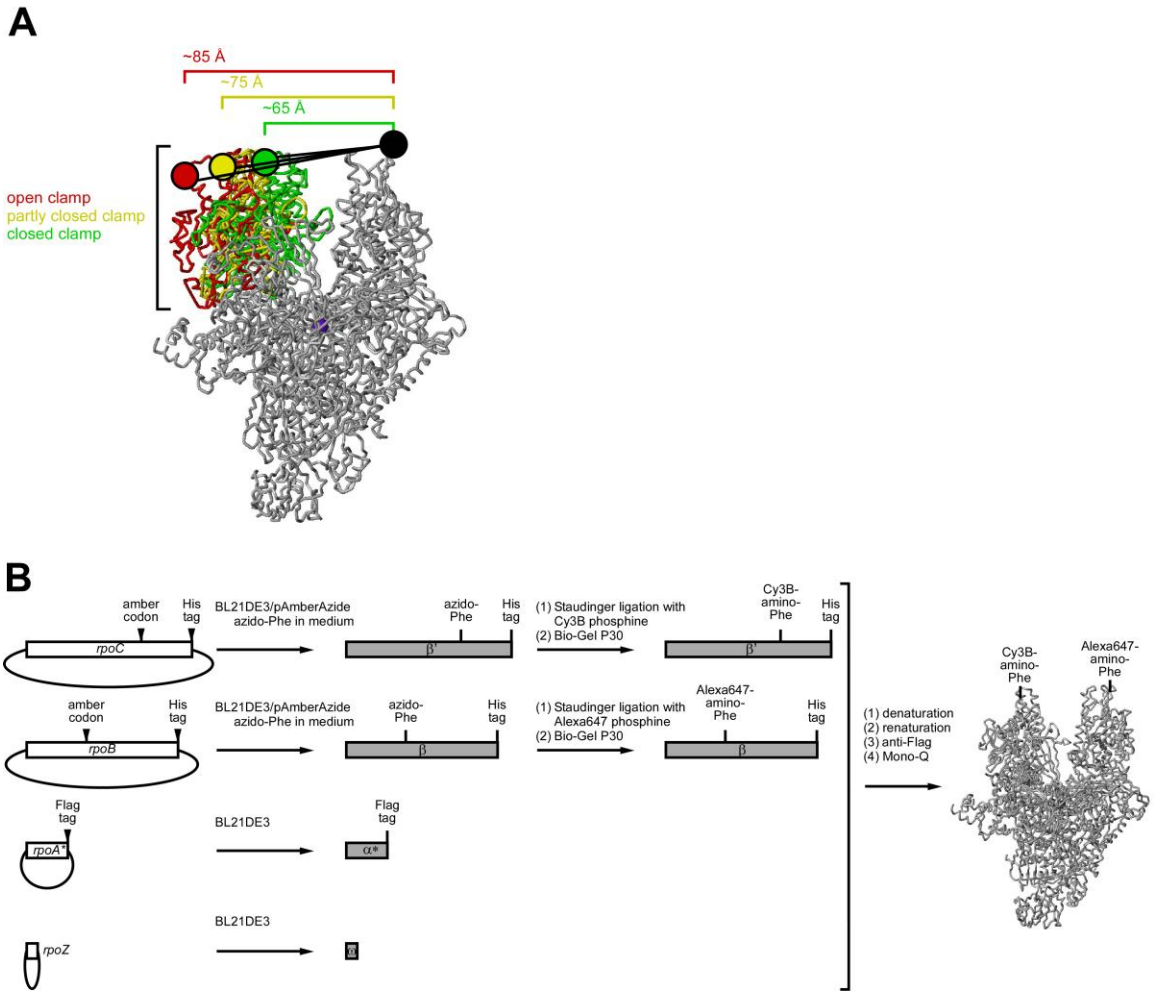


Fig. 8. Determination of RNAP clamp conformation in solution.
(A) Measurement of smFRET between fluorescent probes incorporated at the tips of the RNAP β' pincer (clamp) and the RNAP β pincer. Open (red), partly closed (yellow), and closed (green) RNAP clamp conformational states are as observed in crystal structures (PDB 1I3Q, 1HQM, and 1I6H).
(B) Incorporation of fluorescent probes at the tips of the RNAP β' pincer (clamp) and the RNAP β pincer, by unnatural amino acid mutagenesis [99] to incorporate 4-azidophenylalanine at sites of interest in β' and β, followed by Staudinger ligation [100, 101] to incorporate fluorescent probes at 4-azidophenylalanines in β' and β, followed by in vitro reconstitution of RNAP from labelled β' and β and unlabelled α* (covalently linked α-N-terminal-domain dimer) and ω(see Materials and Methods) [102, 103]. Plasmids, genes, and proteins are shown as ovals, open bars, and closed bars, respectively.

3. Materials and Methods

3.1. Plasmids.

Plasmid pET21a-rpoC-CH6, encoding C-terminally hexahistidine-tagged *Escherichia coli* RNAP β' subunit, was constructed by replacing the BsmI-HindIII segment of pET21a-rpoC by the BsmI-HindIII *rpoC* segment of plasmid pRL663 [104]. Plasmid pET21a-rpoC284TAG-CH6 was constructed by use of site-directed mutagenesis (QuikChange II; Stratagene, Inc.) to replace *rpoC* codon 284 by an amber codon (TAG).

Plasmid pET21d-rpoB-CH6, encoding C-terminally hexahistidine-tagged *E. coli* RNAP β subunit, was constructed by replacing the NcoI-HindIII segment of plasmid pET21d (Novagen, Inc.) by the NcoI-HindIII *rpoB* segment of plasmid pRL706 [105], followed by deleting 44 bp between the NcoI site and the *rpoB* start codon by use of site-directed mutagenesis. Plasmid pET21d-rpoB106TAG-CH6 was constructed by use of site-directed mutagenesis to replace *rpoB* codon 106 by an amber codon (TAG).

Plasmid pET21a-NFα, encoding N-terminally FLAG-tagged *E. coli* RNAP α subunit, was constructed by replacing the Xbal-BamHI segment of plasmid pET21a (Novagen, Inc.) by the Xbal-BamHI *rpoA* segment of plasmid pHTf1Bam-NFα [106].

Plasmid pET28a-NF-αNTD^{I}-αNTD^{II}, encoding an N-terminally Flag-tagged first *E. coli* RNAP α subunit N-terminal domain (α residues 1-235; αNTD^{I}), followed by GlySerGlyGlySerGly, followed by a second *E. coli* RNAP α subunit N-terminal domain (α residues 1-235; αNTD^{II}) (Flag-αNTD^{I}-GSGGSG-αNTD^{II}; α* in Fig. 8B), was constructed by replacing the NcoI-NotI segment of plasmid pET28A (Novagen, Inc.) by an add-on-PCR-generated NcoI-BamHI DNA segment encoding N-terminally FLAG-

tagged αNTD^{I} [prepared using plasmid pET21a-NFα as template and 5'-GGGAATTCCATGGACTACAAGGACG ACGATGACAAGG-3' and 5'-TCGCGGATCCACGTAAGTCAACGAAAGCTTCCAGTTGTTCAGC-3' as primers] and an add-on-PCR-generated BamHI-NotI DNA segment encoding GlySerGlyGlySerGly-αNTD^{II} [prepared using plasmid pREII-NHα [107] as template and 5'-TCGCGGATCCGGTGGCAGC GGGATGCAGGGTTCTGTGACAGAGTTTCTAAAACCG-3' and 5'-TGCGGCGGCCGCTTAACGT AAGTCAACGAAAGCTTCCAGTTGTTCAGC-3' as primers].

3.2. DNA fragments.

DNA fragments *lacCONS-14*(-107/+56), *lacCONS(A+2G)-14*(-107/+56) and [+56-Alexa488]-*lacCONS-14*(-107/+56) (sequences in Fig. 16) were prepared according to procedures in *35*. DNA fragment *nifH*(-60/+28) was prepared as in [108].

3.3. Fluorescent probes.

Alexa488 C5-maleimide was purchased from Invitrogen, Inc.

Fluorescent donor probe Cy3B-carboyl-ethylenediaminyl-phosphine (Cy3B-phosphine) was synthesized in three steps as follows [100, 101] (Fig. 9):

(i) Cy3B-carboyl-ethylenediaminyl-trityl (II). Mono-trityl-ethelenediamine (acetic acid salt; 23.5 mg; 65 μmol), **I** (Cy3B-NHS; 5.0 mg; 6.5 μmol), and TEA (60 μl; 430 umol) were added, in turn, to 200 μl anhydrous DMF, and the reaction mixture was incubated 1

h at room temperature. The product was purified by reversed-phase HPLC (solvent A: water; solvent B: 90% acetonitrile, 10% water; gradient 30 to 80% B in 30 min at 2 ml/min) and lyophilized. MS (MALDI): calculated, *m/z* 845.6 (MH^+); found, 845.6.

(ii) Cy3B-carboyl-ethylenediamine (III). TFA (50 μl; 0.65 mmol) was added to **II** (4.2 mg; 5.0 μmol) in 200 μl choloroform, and the reaction mixture was incubated at 1 h at room temperature. The product was purified by reversed-phase HPLC (solvent A: 0.1% TFA in water; solvent B: 100% acetonitrile; gradient: 20 to 80% B in 30 min at 2 ml/min) and lyophilized. MS (MALDI): calculated, *m/z* 603.3 (MH^+); found, 603.3.

(iii) Cy3B-carboyl-ethylenediaminyl-phosphine (Cy3B-phosphine; IV). EDAC (12.5 mg; 65 umol) in 50 μl DMF, NHSS (8.8 mg; 65 μmol) in 50 μl DMF, and MDPT (24 mg; 60 μmol) in 50 μl DMF were combined. **III** (2.4 mg; 4.0 μmol) in 50 μl DMF was added, followed by DIPEA (23 μl; 130 μmol), and the reaction mixture was incubated 3 h at 37˚C. The product was purified by reversed-phase HPLC (solvent A: 0.1% TFA in water; solvent B: 100% acetonitrile; gradient: 30 to 100% B in 30 min at 2 ml/min) and lyophilized. MS (MALDI): calculated, *m/z* 948.5 (MH^+); found, 948.5.

Fluorescent acceptor probe Alexa647-phosphine$^{20\ Å}$ was synthesized in three steps as follows [100, 101] (Fig. 9):

(i) Alexa647-pentanoyl-ethylenediaminyl-trityl (VI). N-trityl-1,2-ethanediamine (hydrobromide salt) (23 mg; 60 μmol) was added, to **V** (Alexa Fluor 647 NHS ester; 5.0 mg; 5.0 μmol) in 1 ml DMF. TEA (10.0 μl; 71 μmol) was added and the reaction

mixture was incubated 30 min at room temperature. The reaction mixture was dried under vacuum, re-dissolved in 0.5 ml ethanol and 20 μl ammonium hydroxide. The product was isolated by flash chromatography and dried under vacuum.

(ii) Alexa647-pentanoyl-ethylenediamine (VII). TFA (100 μl; 1.3 mmol) was added to **VI** (5.0 mg; 4.4 μmol) in 200 μl choloroform, and the reaction mixture was incubated at 30 min at room temperature. The reaction mixture was dried under vacuum, and the product was purified using flash chromatography. MS (MALDI): calculated. *m/z* 901 (MH^+); found, 901.

(iii) Alexa647-pentanoyl-ethylenediaminyl-phosphine (Alexa647-phosphine$^{20\ Å}$; VIII). EDAC (21 mg; 110 μmol) in 250 μl degassed water, NHSS (21 mg; 78 μmol) in 250 μl degassed water, **VII (**5.0 mg; 5.5 μmol) in 200 μl DMF and 50 μl degassed water, and MDPT (30 mg; 75 μmol) in 250 μl DMF were combined. A precipitate was observed. DMF (~700 μl) was added, resulting in dissolution of the precipitate. DIPEA (28 μl; 160 μmol) was added, and the mixture was incubated 3 h at 37˚C. The product was purified by reversed-phase HPLC (solvent A: 0.1% TFA in water; solvent B: 100% acetonitrile; gradient: 30 to 100% B in 30 min at 2 ml/min) and was dried under vacuum. MS (MALDI): calculated, *m/z* 1248.4 (MH^+); found, 1248.4.

3.4. β’ and β derivatives.

Inclusion bodies containing *E. coli* RNAP β' and β subunits were prepared and stored as in *[102]*.

4-azidophenylalanine-labelled β' and β derivatives were prepared as follows (see [99]: *E. coli* strain BL21(DE3) (Stratagene, Inc) was transformed with plasmid pDule-azide *[99]*; kindly provided by Peter Schultz, The Scripps Research Institute, La Jolla CA and Ryan Mehl, Franklin & Marshall College, Lancaster PA) and subsequently was transformed with plasmid pET21a-rpoC284TAG-CH6 or plasmid pET21d-rpoB106TAG-CH6. Single colonies of the resulting transformants were inoculated into 5 ml LB broth *[109]* containing 100 μg/ml ampicillin and 25 μg/ml tetracycline, and cultures were incubated 10 h at 37°C with shaking. Cells were harvested by centrifugation (4000xg; 10 min at 4˚C), resuspended in 5 ml M9+ broth [M9 broth (*38*) containing 0.4% glucose, 2 mM $MgCl_2$, 0.1 M $CaCl_2$, 3 nM $(NH4)6Mo_7O_{24}$, 400 nM H_3BO_3, 30 nM $CoCl_2$, 10 nM $CuSO_4$, 80 nM $MnCl_2$, 20 nM $ZnSO_4$, 2 mg/ml thiamine, 0.4 mg/ml choline chloride, 0.5 mg/ml folic acid, 0.5 mg/ml nicotinamide, 1 mg/ml myoinositol, 1 mg/ml pyridoxal-HCl, 0.05 mg/ml riboflavin, and 1 mg/ml biotin], used to inoculate 50 ml M9+ broth containing 100 μg/ml ampicillin and 25 μg/ml tetracycline, and cultures were incubated 16 h at 37°C. Aliquots (10 ml) were used to inoculate 1 L M9+ containing 1mM 4-azido-L-phenylalanine (Chem-Impex International, Inc.), 100 μg/ml ampicillin, and 25 μg/ml tetracycline, cultures were incubated at 37°C with shaking until OD_{600} = 0.6, were induced by addition of isopropyl-beta-D-thiogalactopyranoside (IPTG) to 1 mM, and were incubated for an additional 4 h at 37°C with shaking. Cells were harvested by centrifugation (4000xg; 20 min at 4˚C), were lysed, and inclusion bodies containing 4-azidophenylalanine-labelled β' and β derivatives were prepared and stored according to procedures in [102].

Fluorescent-probe-labelled β' and β derivatives were prepared as follows (see [100]: Reaction mixtures (3 ml) contained 20 μM 4-azidophenylalanine-labelled β' or β derivative (solubilized immediately before use) and 200 μM fluorescent-probe-phosphine [Cy3B-phosphine or Alexa647-phosphine$^{20\ Å}$ (*18*)] in 50 mM Tris-HCl, pH 7.9, 6 M guanidine-HCl, and 5% glycerol. Reaction mixtures were incubated 15 h at 37˚C. Reaction mixtures were applied to 10 ml columns of Bio-Gel P30 (BioRad, Inc.) pre-equilibrated in 50 mM Tris-HCl, pH 7.9, 6 M guanidine-HCl, and 5% glycerol; columns were washed with 3 ml of the same buffer; and fluorescent-probe-labelled β' and β derivatives were eluted in 3 ml of the same buffer.

Efficiencies of incorporation of fluorescent probes were determined from UV/Vis-absorbance measurements and were calculated as follows [100]:

$$\text{concentration of product} = [A_{280} - \in_{F,\,280}(A_{max\,/} \in_{F,\,max})]/\in_{P,\,280}$$

$$\text{labelling efficiency} = 100\%[(A_{max} /\in_{F,\,max})/(\text{concentration of product})]$$

where A_{280} is the measured absorbance at 280 nm, A_{max} is the measured absorbance at the long-wavelength absorbance maximum of fluorescent probe *F* (559 nm and 652 nm for Cy3B and Alexa647, respectively), $\in_{P,280}$ is the molar extinction coefficient of protein *P* at 280 nm (calculated as in *39*), $\in_{F,280}$ is the molar extinction coefficient of fluorescent probe *F* at 280 nm (10,400 M^{-1} cm^{-1} and 7,350 M^{-1} cm^{-1} for Cy3B and Alexa647, respectively), and $\in_{F,max}$ is the extinction coefficient of fluorescent probe *F* at its long-

wavelength absorbance maximum (130,000 M^{-1} cm^{-1} and 245,000 M^{-1} cm^{-1} for Cy3B and Alexa647, respectively).

Specificities of incorporation of fluorescent probes were measured from the observed efficiencies of labelling (see preceding paragraph) of (i) the product of the labelling reaction with the 4-azidophenylalanine-labelled β' or β derivative and (ii) the product of a parallel labelling reaction with unlabelled β' or β. The specificity of labelling was calculated as [100]:

labelling specificity = 100%[1 - [(labelling efficiency with *P*) / (labelling efficiency with *P*-azide)]

where *P*-azide is the 4-azidophenylalanine-labelled protein and *P* is the corresponding unlabelled protein.

Typical yields of fluorescent-probe-labelled β' and β derivatives were 15 mg/L and 18 mg/L, respectively. Typical labelling efficiencies were ~90% (Fig. 10C). Typical labelling specificities were >90% (Fig. 10C).

3.5. Flag-αNTD^{I}-GSGGSG-αNTD^{II}.

A fusion protein comprising an N-terminally Flag-tagged first *E. coli* RNAP α subunit N-terminal domain (α residues 1-235; αNTD^{I}), followed by GlySerGlyGlySerGly, followed by a second *E. coli* RNAP α subunit N-terminal domain (α residues 1-235; αNTD^{II}) (Flag-αNTD^{I}-GSGGSG-αNTD^{II}; α* in Fig. 8B) was prepared as follows: *E. coli* strain

BL21(DE3) (Invitrogen, Inc.) was transformed with plasmid pET28a-NF-αNTD^{I}-αNTD^{II}. Single colonies of the resulting transformants were used to inoculate10 ml LB broth containing 40 μg/ml kanamycin, and cultures were incubated 16 h at 37°C with shaking. Aliquots (5 ml) were used to inoculate 1 L LB broth containing 40 μg/ml kanamycin, cultures were incubated at 37°C with shaking until OD_{600} = 0.6, cultures were induced by addition of IPTG to 1 mM, and cultures were incubated an additional 3 h at 37°C. Cells were harvested by centrifugation (5,000xg; 20 min at 4˚C), resuspended in 25 ml of lysis buffer [20 mM Tris-HCl pH 7.9, 500 mM NaCl, 10 mM EDTA, and one protease inhibitor cocktail tablet (Roche, Inc)] and lysed using an Avestin EmulsiFlex-C5 cell disrupter (Avestin, Inc.). Lysates were centrifuged (20,000xg; 20 min at 4°C), supernatants were collected, and protein was precipitated by addition of ammonium sulfate (35 g per 100 ml supernatant) followed by centrifugation (20,000xg; 10 min at 4°C). Pellets were dissolved in 10 ml TBS (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 5% glycerol) and loaded onto two 5 ml columns packed with ANTI-FLAG M2 affinity gel (Sigma, Inc.) pre-equilibrated in TBS. Columns were washed with 50 ml TBS and were eluted with twenty-five 1 ml fractions of TBS containing 0.1 mg/ml FLAG peptide (Sigma, Inc.). Fractions containing αNTD^{I}-αNTD^{II} were pooled, concentrated, and stored in 50 mM Tris-HCl pH 7.9, 200 mM KCl, 10 mM $MgCl_2$, 1 mM EDTA, 5 mM β-mercaptoethanol, and 50% glycerol at -80°C. Typical yields of Flag-αNTD^{I}-GSGGSG-αNTD^{II} were 20 mg/L.

3.6. ω.

Inclusion bodies containing *E. coli* RNAP ω subunit were prepared and stored as in [102].

3.7. RNAP derivatives.

Unlabelled *E. coli* RNAP core and RNAP-σ^{70} holoenzyme were prepared as in [102].

Fluorescent-probe-labelled *E. coli* RNAP core and RNAP-σ^{70} holoenzyme were reconstituted from fluorescent-probe-labelled and unlabelled subunits as follows: Reaction mixtures (60 ml) contained: 7.8 mg (50 nmol) fluorescent-probe-labelled or unlabelled β' derivative, 3.0 mg (20 nmol) fluorescent-probe-labelled or unlabelled β derivative, 4.2 mg (80 nmol) Flag-αNTD^{I}-GSGGSG-αNTD^{II}, and 2.0 mg (200 nmol) ω, in denaturation buffer (6 M guanidine-HCl, 50 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM EDTA, 10 mM DTT, and 10% glycerol). Reaction mixtures were dialyzed 12 h at 4°C against 2 L reconstitution buffer (50 mM Tris-HCl, pH 7.9, 200 mM KCl, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM EDTA, 5 mM β-mercaptoethanol, and 20% glycerol), further dialyzed 10 h at 4°C against 2 L of the same buffer, and further dialyzed 12 h against 2 L of the same buffer minus β-mercaptoethanol. Following dialysis, insoluble material was removed by centrifugation (20,000xg; 30 min at 4°C), and supernatants (for preparation of RNAP core) or supernatants supplemented with 0.7 mg (10 nmol) σ^{70} (for preparation of RNAP-σ^{70} holoenzyme) were incubated 45 min at 30°C. Following incubation, insoluble material was removed by centrifugation (12000xrpm; 20 min at 4°C)., and supernatants were applied to 3 ml columns packed with ANTI-FLAG M2 affinity gel (Sigma, Inc) pre-equilibrated in TBS. Columns were washed with 30 ml TBS and eluted with fifteen 1 ml fractions of TBS containing 0.1

mg/ml FLAG peptide (Sigma, Inc). Fractions containing fluorescent-probe-labelled RNAP core and RNAP-σ^{70} holoenzyme derivatives were pooled and were dialyzed against TGEB (30 mM Tris-HCl, pH 7.9, 200 mM NaCl, 0.1 mM EDTA, 1 mM β-mercaptoethanol and 5% glycerol). Fluorescent-probe-labelled RNAP core and RNAP-σ^{70} holoenzyme derivatives were further purified by anion-exchange chromatography on Mono-Q according to procedures in *35* and were stored in 25 mM Tris-HCl, pH 7.9, 0.1 M NaCl, 0.1 mM EDTA, 1 mM β-mercaptoethanol, and 50% glycerol at -20°C.

SDS-PAGE, analytical size exclusion chromatography, and transcription assays confirm that RNAP-σ^{70} holoenzyme and core have correct subunit stoichiometries, correct monomeric aggregation states, and high specific activities (Fig. 11).

The use in this procedure of Flag-αNTD^{I}-GSGGSG-αNTD^{II} (80 nmol) instead of N-terminally Flag-tagged wild-type RNAP α subunit (160 nmol) results in higher yields and equal or higher specific activities (unpublished data). RNAP derivatives containing Flag-αNTD^{I}-GSGGSG-αNTD^{II} behave indistinguishably from RNAP derivatives containing wild-type α subunit in σ^{70}-dependent transcription initiation and elongation at *lacCONS-14* and in σ^{54}-dependent transcription initiation at *nifH*(-35/-1) (unpublished data).

Typical yields of fluorescent-probe-labelled RNAP core and RNAP-σ^{70} holoenzyme derivatives were 0.2 mg per 60 ml reconstitution reaction mixture. Typical purities were >95% (Fig. 11A,B). Typical specific activities were 70-80% those of unlabelled RNAP core and RNAP-σ^{70} holoenzyme (Fig. 11C).

3.8. σ^{70}

Unlabelled *E. coli* σ^{70} was prepared as in *[108]*.

Alexa488-labelled *E. coli* σ^{70} was prepared from the single-cysteine σ^{70} derivative [Ser132;Ser291;Ser295;Cys596]-σ^{70} (*35*) and Alexa488 C5-maleimide (Invitrogen, Inc.), according to procedures in [108].

3.9. σ^{54}.

σ^{54} was prepared as in [110].

3.10. NtrC1.

NtrC1(121-387) was prepared as in [111].

3.11. Inhibitors.

Myxopyronin A, corallopyronin A, and ripostatin A were prepared as in *24*. Rifampin, streptolydigin, and BR703 were purchased from Sigma, Inc., Sourcon-Padena GmbH, and Ryan Scientific, Inc., respectively. Bacteriophage T7 Gp2 and [Ala56]Gp2, a Gp2 mutant defective in interaction with RNAP (Gp2* in Fig. 25), were prepared as in [82].

3.12. Analytical size exclusion chromatography: RNAP core and RNAP-σ^{70} holoenzyme.

Aggregation states of fluorescent-probe-labelled RNAP core and RNAP-σ^{70} holoenzyme derivatives were analyzed by analytical size exclusion chromatography. RNAP derivatives (50 μl of 50 nM in TB) were applied to a Superdex 200 HR 10/30 column (GE Healthcare, Inc.) pre-equilibrated in TB, were eluted with transcription buffer at a flow rate of 0.4 ml/min in 46 x 0.5 ml or 23 x 1 ml fractions, and were detected by measurement of fluorescence emission of fractions (GENios Pro fluorescence microplate reader; TECAN, Inc.; excitation wavelength = 614 nm; emission wavelength = 670 nm). Columns were calibrated by performing runs with high-molecular-mass markers (HMW calibration Kit; GE Healthcare, Inc.), using the same elution buffer and flow rate, and detecting markers by UV-absorbance scanning at 280 nm.

3.13. Transcription assays: RNAP-σ^{70} holoenzyme.

Radiochemical transcription assays were performed according to procedures in [18]. Reaction mixtures (18 μl) contained: 100 nM RNAP-σ^{70} holoenzyme derivative and 20 nM DNA fragment *lacCONS-14*(-107/+56) in TB (50 mM Tris-HCl, pH 7.9, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 100 μg/ml bovine serum albumin, and 5% glycerol). Reaction mixtures were incubated 15 min at 37˚C, 0.5 μl 1 mg/ml heparin (Sigma, Inc.) was added, and reaction mixtures were further incubated 2 min at 37°C. Transcription was started by addition of 0.5 μl 7 μM [α-^{32}P]UTP (0.6 Bq/fmol) and 0.5 μl 5 mM each of UTP, ATP, GTP, CTP, was allowed to proceed 5 min at 37˚C, and was stopped by addition of 10 μl 80% formamide, 10 mM EDTA, 0.04% bromophenol blue, and 0.04% xylene cyanol and incubation 5 min at 90˚C. Products were analyzed by PAGE on 10% TBE-urea gels (BioRad, Inc.) followed by storage-phosphor imaging.

3.14.Transcription assays: RNAP-σ^{54} holoenzyme.

Radiochemical transcription assays were performed according to procedures in [112]. RNAP core derivative (50 nM) and σ^{54} (100 nM) were incubated 10 min at 37˚C in 18μl TB; 1 μl 400 nM DNA fragment *nifH*(-60/+28) of *36* was added, and reaction mixtures were further incubated 5 min at 37˚C; 0.5 μl 4 μM NtrC1(121-387) heptamer and 1 μl 40 mM ATP were added, and the reaction mixture was further incubated 10 min at 37C. Transcription was started by addition of 0.5 μl 7 μM [α-^{32}P]UTP (0.6 Bq/fmol) and 0.5 μl 5 mM each of UTP, ATP, GTP, CTP, was allowed to proceed 5 min at 37˚C, and was stopped by addition of 5 μl 80% formamide, 10 mM EDTA, 0.04% bromophenol blue, and 0.04% xylene cyanol and incubation 5 min at 90˚C. Products were analyzed by PAGE on 10% TBE-urea gels (BioRad, Inc.) followed by storage-phosphor imaging.

3.15. Small-primed-RNA-synthesis assay: RNAP-σ^{54}.

Small-primed-RNA (spRNA) synthesis assays were performed essentially as in [113]. RNAP core derivative (100 nM) and σ^{54} (400 nM) in 20 μl TB were incubated 10 min at 30°C and then were supplemented with either (i) 0.5 μl 800 nM DNA fragment "pre-opened" *nifH* (-10-1/WT) of *43* in TB (RPc) or (ii) 0.5 μl 800 nM DNA fragment "pre-opened" *nifH*(-10-1/WT), 200 nM NtrC1(121-387) heptamer, 200 μM ADP, 5 mM NaF, and 200 μM $AlCl_3$ in TB (RPi2). Samples were further incubated 10 min at 37°C, 0.5 μl 4 mg/ml heparin was added, and spRNA (UpGpGpG) synthesis was started by addition of 1 μl 10mM UpG and 4 μCi [α-^{32}P]UTP. Reactions were allowed to proceed

20 min at 37˚C, and were stopped by addition of 5 μl 80% formamide, 10 mM EDTA, 0.04% bromophenol blue, and 0.04% xylene cyanol and 5 min incubation at 90˚C. Products were analyzed by PAGE on 20% TBE-urea gels [109] followed by storage-phosphor imaging.

The σ^{54}-dependent transcription initiation intermediate RPc-NtrC (formed in reactions in the absence of ADP-AlF_x) does not support spRNA synthesis ([113]; Fig. 23). The σ^{54}-dependent transcription initiation intermediate RPi2 (formed in reactions in the presence of ADP-AlF_x) supports spRNA synthesis ([113]; Fig. 23).

3.16. Electrophoretic mobility shift assays: RNAP-σ^{54}.

Electrophoretic mobility shift assays were performed according to procedures in *36*. RNAP core derivative (0 or 50 nM) and σ^{54} (0 or 100 nM) in 20 μl TB were incubated 10 min at 37°C; 0 or 20 μM myxopyronin was added, reaction mixtures were further incubated 10 min at 37°C, 0.2 μl 1.5 μM ^{32}P-5'-end-labelled DNA fragment *nifH*(-35/+1) of [112] (0.6 Bq/fmol; ^{32}P-labelled using T4 polynucleotide kinase and [γ^{32}P]-ATP) was added, and reaction mixtures were further incubated 15 min at 37°C. Products were analyzed by PAGE on 5% TBE gels (BioRad, Inc) followed by storage-phosphor imaging.

The σ^{54}-dependent transcription initiation intermediate RPc (formed in reactions containing RNAP σ^{54} holoenzyme and DNA) is stable to electrophoresis and migrates as a high-molecular-mass species ([109, 112]; Fig. 24C).

3.17. smFRET sample preparation: σ^{70}-dependent transcription.

For analysis of RNAP core and RNAP-σ^{70} holoenzyme, samples containing 30 nM fluorescent-probe-labelled RNAP core derivative or fluorescent-probe-labelled RNAP-σ^{70} holoenzyme derivative in 30 μl TB were incubated 15 min at 37˚C, and aliquots (0.25 μl) were transferred to pre-warmed tubes containing 70 μl KG7 (40 mM HEPES-NaOH, pH 7.0, 100 mM potassium glutamate, 10 mM $MgCl_2$, 1 mM DTT, 100 μg/ml bovine serum albumin, 2 mM Trolox, and 10% glycerol). Following further incubation for 5 min at 37°C, smFRET data collection was performed.

For analysis of RPo, RPitc,≤2, RPitc,≤4, RPitc,≤7, RDe,14, and RDe,15, reaction mixtures containing 30 nM fluorescent-probe-labelled RNAP-σ^{70} holoenzyme derivative and 100 nM DNA fragment (*lacCONS-14*(-107/+56), *lacCONS(A+2G)*-14(-107/+56), or [+56-Alexa488]-*lacCONS-14*(-107/+56) in 30 μl TB were incubated 15 min at 37˚C. Heparin-Sepharose (GE Healthcare, Inc.) was added to 1 mg/ml, reaction mixtures were further incubated 1 min at 37°C, reaction mixtures were centrifuged (2,100xg; 30 s), and supernatants were transferred to new pre-warmed tubes. Aliquots (0.25 μl) were transferred to pre-warmed tubes containing 70 μl KG7 and the following additional components: (i) none (RPo); (ii) 500 μM initiating dinucleotide [ApA for *lacCONS-14*(-107/+56) and [+56-Alexa488]-*lacCONS-14*(-107/+56); ApG for *lacCONS*(A+2G)-14(-107/+56)] (RPitc,2); (iii) 500 μM initiating dinucleotide and 50 μM UTP (RPitc4); (iv) 500 μM initiating dinucleotide and 50 μM each of UTP and GTP (RPitc,7); (v) 500 μM initiating dinucleotide and 50 μM each of UTP, GTP, and ATP, (RDe,14); and (vi) 500 μM initiating dinucleotide and 50 μM each of UTP, GTP, ATP, and 3'-O-MeCTP

(Ribomed, Inc.) (RDe,15). Following further incubation for 10 min at 37°C, smFRET data collection was performed.

Transcription experiments verify the formation of RPo and verify the synthesis of 7, 14, and 15 nt RNA products in RPitc,7, RDe,14 and RDe,15 (methods as in [70]).

For analysis of complexes of RNAP-σ^{70} holoenzyme with inhibitors, reaction mixtures contained 30 nM fluorescent-probe-labelled RNAP-σ^{70} holoenzyme derivative and 0 or 20 uM inhibitor (50 μM for corallopyronin; 120 nM for Gp2 and [Ala56]Gp2) in 30 μl transcription buffer. Following incubation 10 min at 37˚C, aliquots (0.25 μl) were transferred to pre-warmed tubes with 70 μl KG7 containing the same concentration of inhibitor, and smFRET data collection was performed.

3.18. smFRET sample preparation: σ^{54}-dependent transcription.

For analysis of RNAP-holoenzyme-σ^{54}, RPc, RPc+NtrC1, RPi1, RPi2, and RPo, reaction mixtures containing 30 nM fluorescent-probe-labelled RNAP core derivative and 90 nM σ^{54} in 30 μl TB were incubated 10 min at 30°C and then were supplemented with the following additional components at the following final concentrations: (i) none (RNAP-σ^{54} holoenzyme); (ii) 180 nM DNA fragment "early melt" *nifH*(-35/+1) of *36*, (RPc); (iii) 180 nM DNA fragment "early melt" *nifH*(-35/+-1) and 100 nM NtrC1(121-387) heptamer (RPc+NtrC1); (iv) 180 nM DNA fragment "early melt" *nifH*(-35+/-1), 100 nM NtrC1(121-387) heptamer, 200 μM ADP, 5 mM NaF, and 200 μM $BeCl_2$ (RPi1); and (iv) 180 nM DNA fragment "early melt" *nifH*(-35/+1), 100 nM NtrC1(121-387) heptamer, 200 μM ADP, 5mM NaF, and 200 μM $AlCl_3$ (RPi2); and (v) 180 nM DNA fragment

"early melt" *nifH*(-35/+1), 100 nM NtrC1(121-387) heptamer, and 2 mM ATP (RPo). Following incubation 10 min at 37°C, aliquots (0.25 µl) were transferred to pre-warmed tubes holding 70 µl KG7 containing the same additional components, samples were further incubated 10 min at 37°C, and smFRET data collection was performed.

Electrophoretic-mobility-shift assays verify the formation of RPc (methods as in [112]). spRNA-synthesis assays verify the formation of RPi2 (methods as in [113]). Transcription experiments verify the formation of RPo (methods in [112]).

For analysis of complexes of RNAP-σ^{54} holoenzyme with inhibitors, reaction mixtures contained 30 nM RNAP-core, 90 nM σ^{54}, and 0 or 20 µM inhibitor in 30 µl TB. Following incubation 10 min at 37˚C, aliquots (0.25 µl) were transferred to pre-warmed tubes with 70 µl KG7 containing the same concentration of inhibitor, and smFRET data collection was performed.

3.19. smFRET sample preparation: transcription pausing.

For Nucleic acid scaffold reconstitution, annealing mixtures contained 1.2 µl of 10 µM RNA, 1.0 µl of 10 µM template DNA strand, and 1.5 µl of 10 µM of non-template DNA strand, in 10 µl R buffer (20 mM Tris-HCl pH 7.9, 20 mM NaCl, 0.1 mM EDTA). The mixture was heated to 70˚C, rapidly cooled to 45˚C and then slowly (1˚/min) cooled to 25˚C.

For analysis of clamp conformation in RDe and hairpin-containing RDe, reaction mixtures contained 100 nM RNAP, 150 nM scaffold, 1X glycerol/BSA mix (5% glycerol, 10 mM MgCl2, 2 mM BME, 50 ug/ml BSA) in 20 µl R buffer.

Following incubation 10 min at 25˚C, samples were diluted to 50 nM in R buffer, aliquots (0.2 μl) were transferred to pre-warmed tubes with 100 μl KG7, and smFRET data collection was performed.

3.20. smFRET data collection and data analysis.

Alternating laser excitation microscopy (ALEX), data acquisition, and data analysis were as in [68, 95-98].

For two-color ALEX experiments, a green laser (532 nm; Compass 215M-20; Coherent, Inc.) was used for direct excitation of the donor, and a red laser (638 nm; Radius 635-25; Coherent, Inc.) was used for direct excitation of the acceptor (Fig. 7A; *11-14*). Lasers were operated at continuous-wave excitation intensities of 100 μW at 532 nm and 80 μW at 638 nm and were alternated at 25 μs intervals using an acousto-optical modulator (Neos Technologies, Inc.). Fiber-coupled collimated beams were directed to an Olympus IX71 inverted microscope (Olympus America, Inc.), reflected by a beamsplitter, and focused into the sample through a 60x oil-immersion objective. Fluorescence emission from the sample was collected through the objective, filtered through a 100 μm pinhole, spectrally split by a dichroic mirror, and focussed onto two avalanche photodiode detectors (APD; SPCM-AQR-15; Perkin-Elmer, Inc.). Data acquisition times ranged from 25-40 min, depending on sample concentration.

Photons detected at the donor emission channel (D_{em}) and the acceptor emission channel (A_{em}) were assigned to donor excitation (D_{exc}) or acceptor excitation (A_{exc}) based on photon arrival times. The stoichiometry parameter (S) was calculated for each above-threshold photon burst, as follows [68, 95-98]:

$$S = \left(F_{D_{exc}}^{A_{em}} + \gamma_{D2\text{-}A2} F_{D_{exc}}^{D_{em}}\right) \Big/ \left(F_{D_{exc}}^{A_{em}} + \gamma_{D2\text{-}A2} F_{D_{exc}}^{D_{em}} + F_{A_{exc}}^{A_{em}}\right)$$

where $F_{D_{exc}}^{A_{em}}$, $F_{D_{exc}}^{D_{em}}$, and $F_{A_{exc}}^{D_{em}}$ are fluorescence emission intensities assigned to donor excitation and acceptor emission, donor excitation and donor emission, and acceptor excitation and donor emission, where $F_{D_{exc}}^{A_{em}}$ is corrected for direct excitation and leakage, and where $\gamma_{D2\text{-}A2}$ is a detection factor calculated as follows (1 in this work; [68, 95-98]:

$$\gamma_{D2\text{-}A2} = \gamma_{D1-A1}\left(\phi_{A2}/\phi_{A1}\right)\left(\phi_{D1}/\phi_{D2}\right)$$

The donor-acceptor smFRET efficiency (*E*) for each above-threshold photon burst was calculated as follows:

$$E = \left(F_{D_{exc}}^{A_{em}}\right) \Big/ \left(F_{D_{exc}}^{A_{em}} + F_{D_{exc}}^{D_{em}}\right)$$

Two-dimensional *E*-*S* plots were used to distinguish species containing donor only (D-only), acceptor only (A-only), and both donor and acceptor (D-A) (Fig. 7B; *11-14*). For species containing both donor and acceptor (D-A), one-dimensional *E* histograms were plotted and were fitted with Gaussian curves (Fig. 7B; *11-14*). The resulting histograms

provide equilibrium population distributions of *E*, define numbers of subpopulations with distinguishable *E*, and, for each subpopulation, define mean *E*.

For three-color ALEX experiments, a blue laser (488 nm; Ar-ion laser 532-AP; Melles-Griot, Inc) was added to the two-color ALEX experimental setup (Fig. 18A; [114]. Light from the blue, green, and red lasers was spatially coupled using an optical fiber and was collimated and reflected using a triple-band beamsplitter. Fluorescence emission from the sample was collected through the objective, filtered through a 100 μm pinhole, spectrally split by two dichroic mirrors, and focussed onto three APDs. Data acquisition times ranged from 20-30 min, depending on sample concentration.

Blue (B), green (G) and red (R) excitation sources were assigned to detected photon streams and stoichiometry parameters were determined as follows [114]:

.

$$S_{BG} = F_{B_{exc}} / \left(F_{B_{exc}} + F_{G_{exc}} \right)$$

$$S_{GR} = F_{G_{exc}} / \left(F_{G_{exc}} + F_{R_{exc}} \right)$$

$$S_{BR} = F_{B_{exc}} / \left(F_{B_{exc}} + F_{R_{exc}} \right)$$

Species containing all three probes were identified based on stoichiometry parameters (Fig. 18B,C; *15*), and, for each such species, the G-R smFRET efficiency (E_{GR}) was determined as follows [114]:

$$E_{GR} = F_{G_{exc}}^{R_{em}} / \left(F_{G_{exc}}^{G_{em}} + F_{G_{exc}}^{R_{em}} \right)$$

Values of E_{GR} were plotted as histograms, and fitted with Gaussian curves (Fig. 18C; [114]. The resulting histograms provide equilibrium population distributions of E_{GR}, define numbers of subpopulations with distinguishable E_{GR}, and, for each subpopulation, define mean E_{GR}.

Mean donor-acceptor distances (R) were calculated from mean E values obtained from two-color ALEX experiments or mean E_{GR} values obtained from three-color ALEX experiments, as follows [68, 95-97, 114]:

$$R = R_0\left[\left(1/E\right)-1\right]^{1/6}$$

where R_0 is the Főrster parameter, calculated as follows [88]:

$$R_0 = 9780(n^{-4}\kappa^2 Q_D J)^{1/6}\ \text{Å}$$

where n is the refractive index of the medium (1.4; [88]); κ^2 is the orientation factor relating donor emission dipole and acceptor excitation dipole [approximated as 2/3--justified by fluorescence anisotropy measurements (Fig. 12B; performed essentially as in [44]) indicating donors and acceptors reorient on the timescale of donor excited-state life time [88] and also by the fact that, in all cases, mean E < 0.5 [115]; Q_D is the quantum yield of the donor in the absence of acceptor (Fig. 12A); and J is the overlap integral of

donor emission spectrum and acceptor excitation spectrum (Fig. 12A; determined using corrected spectra for donor-only and acceptor-only controls; [88].

Experiments were performed in triplicate. For two-color smFRET experiments, standard error of the mean (SEM) for mean *E* and mean *R* from triplicate experiments was ≤8% *E* (mean 2%; range 0-8%) and ≤2% *R* (mean 0.5%; range 0-2%), respectively (Table S1). For three-color smFRET experiments, SEM for mean *E* and mean *R* from triplicate experiments was ≤15% *E* (mean 11%; range 0-15%) and ≤4% *R* (mean 2%; range 0-4%), respectively (Table S2).

3.21. Molecular modelling.

Starting from atomic coordinates from a crystal structure of *T. thermophilus* RNAP-σ^A holoenzyme (PDB code 1IW7; [10]), a set of eighteen structural models having RNAP clamp conformations ranging from fully open to fully closed in 2˚ increments were generated. Models were generated by deletion of the β' non-conserved domain (β' residues 158-453; residues numbered as in *T. thermophilus* RNAP), followed by rigid-body rotation of σ^A region 2 and the clamp about the clamp rotation axis defined based on comparison of crystal structures in Fig. 13 (i.e., rigid-body rotation of σ^A residues 74-256, β' residues 3-157, 453-621, and 1441-1455, and β residues 1080-1116 in 2^o increments about a line connecting C□ atoms of β' residues 621 and 1398; see *[44, 116]*.

Probes and linkers were incorporated into models essentially as in [44] and [116], using the CHARMM molecular modelling package [117] and MMTSB tool set [118], and generating force-field-parameter and topology files using CHARMM Generalized Force

Field parameters [119] and MATCH [120]. Probes and linkers were incorporated at β' residue 559 and β residue 104 in *T. thermophilus* models. At each probe site, sterically allowed conformations of probes and linkers were identified [sampling all linker torsion angles in 60° increments, accepting conformations with van der Waals energy <5,000 kcal/mol in the CHARMM force field, and randomly selecting 1,000 accepted conformations], and, for each, a probe pseudoatom, corresponding to the center of the probe chromophore, was defined.

For each structural model, 1,000,000 pairwise donor-acceptor distances were calculated ($R_{calc(Di\text{-}Ai)}$), 1,000,000 donor-acceptor FRET efficiencies were calculated ($E_{calc(Di\text{-}Ai)}$), and the mean donor-acceptor FRET efficiency was calculated (mean $E_{calc(Di\text{-}Ai)}$).

To assess error in $E_{calc(Di\text{-}Ai)}$, for each structural model, $E_{calc(Di\text{-}Ai)}$ was calculated ten times using ten different sets of randomly accepted probe and linker conformations. SEM of $E_{calc(Di\text{-}Ai)}$ was 0.3% to 0.6% of $E_{calc(Di\text{-}Ai)}$.

$E_{calc(Di\text{-}Ai)}$ was calculated in parallel for two limiting cases [121] concerning the relative timescales of probe rotational averaging, linker and macromolecule conformational dynamics, and donor excited-state lifetime. In the first limiting case--expected to be more probable for this system in view of the sizes and structures of the linker and clamp--the timescale for probe rotational averaging (τ_c) is small relative to the donor excited-state lifetime, and the timescale for linker and macromolecule conformational dynamics (τ_p) is large relative to the donor excited-state lifetime [121]. For this case, $E_{calc(Di\text{-}Ai)}$ is calculated as 6follows [121]:

$$\langle E \rangle = \int \left(1 + \left(r/R_o\right)^6\right)^{-1} P(r) dr$$

$$\langle E \rangle = \frac{1}{N_j N_k} \sum_{j}^{N_j} \sum_{k}^{N_k} \left(1 + \left(r_{jk}/R_o\right)^6\right)$$

where *P(r)* is the normalized donor-acceptor distance distribution, and where bin width is infinitely small, such that each pair of donor and acceptor, *j* and *k*, are in a separate bin.

In the second limiting case--expected to be less probable for this system in view of the sizes and structures of the linker and clamp--the timescales for both probe rotational averaging (τ_c) and linker and macromolecule conformational dynamics (τ_p) are small relative to the donor excited-state lifetime [121]. For this limiting case, $E_{calc(Di\text{-}Ai)}$ is calculated as follows [121]:

$$\langle E \rangle = \frac{\int \left(R_o/r\right)^6 P(r) dr}{1 + \int \left(R_o/r\right)^6 P(r) dr}$$

Modelling results presented in the main text, supplemental tables, and supplemental figures are from calculations according to the first limiting case. Modelling results from calculations according to the second limiting case are qualitatively identical (defining open, closed, and collapsed states in free RNAP, and closed states in RPo, RPitc, RDe, RNAP-Myx, RNAP-Cor, RNAP-Rip, and RNAP-Gp2) but quantitatively yield clamp

states that are slightly more open (defining an open state more open by ~2^{o}, a closed state more open by ~4^{o}, and a collapsed state more open by ~8^{o}) (data not shown).

Modelling results presented are from calculations using structural models based on the crystal structure of *T. thermophilus* RNAP-σ^{A} holoenzyme.

4. Results

4.1. Synthesis of phosphine derivatives of fluorescent probes, suitable for smFRET applications.

Single-molecule fluorescence spectroscopy requires fluorescent probes that have exceptionally high brightness and exceptionally high photostability [122]. smFRET experiments require a pair of fluorescent probes capable of serving as donor and acceptor

, where fluorescence emission spectrum of the donor probe overlaps with the fluorescence excitation spectrum of the acceptor probe. Furthermore, smFRET results can be affected by length and flexibility of biomolecule-probe linkers. Therefore, a choice of fluorophores with different linker lengths can be desirable.

In this work, Cy3B-phosphine , Alexa 647-phosphine, Alexa 488-phosphine, Dylight 549-phosphine and Dylight 649-phosphine [123-125], capable of reacting with azide moieties on biomolecules, were tested for suitability in serving as donor-acceptor pair in single-molecule FRET measurements involving RNAP clamp conformation.

In order to select an appropriate biomolecule-acceptor-probe-linker length for smFRET measurements involving RNAP clamp conformation, procedures were developed for synthesis of Alexa647-phosphine$^{20\ Å}$ (Fig. 9), which yields a moderate-length biomolecule probe linker and Alexa647-phosphine$^{24\ Å}$, which yields a longer biomolecule-probe linker [100]. Both versions of A647-phosphine were tested in RNAP clamp conformation assays.

Although all the probe pairs tested yielded qualitatively similar results, use of Cy3B/Alexa647-phosphine$^{20\ Å}$ pair produced strongest bursts of photons and yielded best smFRET efficiency distribution profiles. All smFRET experiments described in this work, unless otherwise stated, were performed with Cy3B/Alexa647$^{20\ Å}$ as donor/acceptor pair.

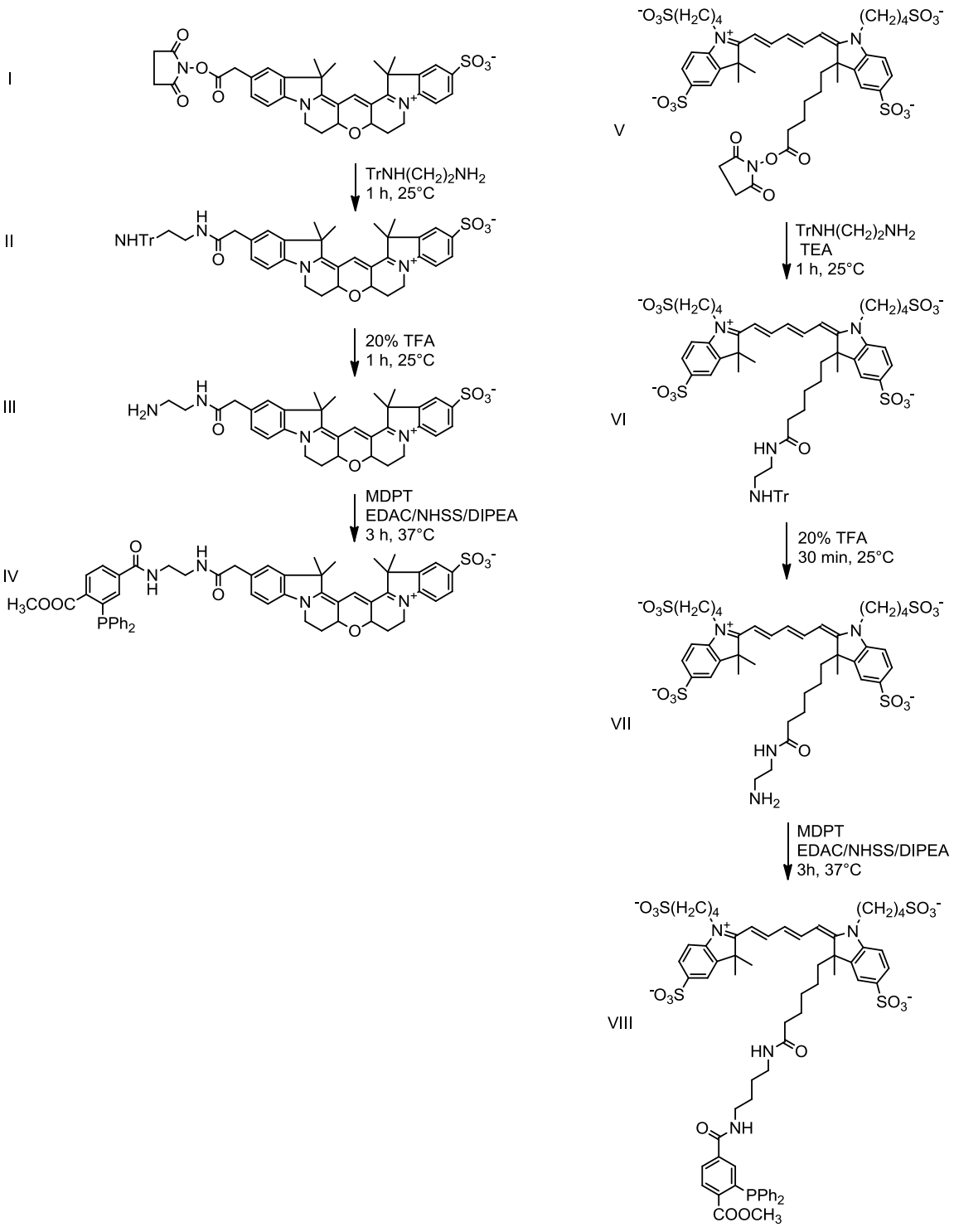
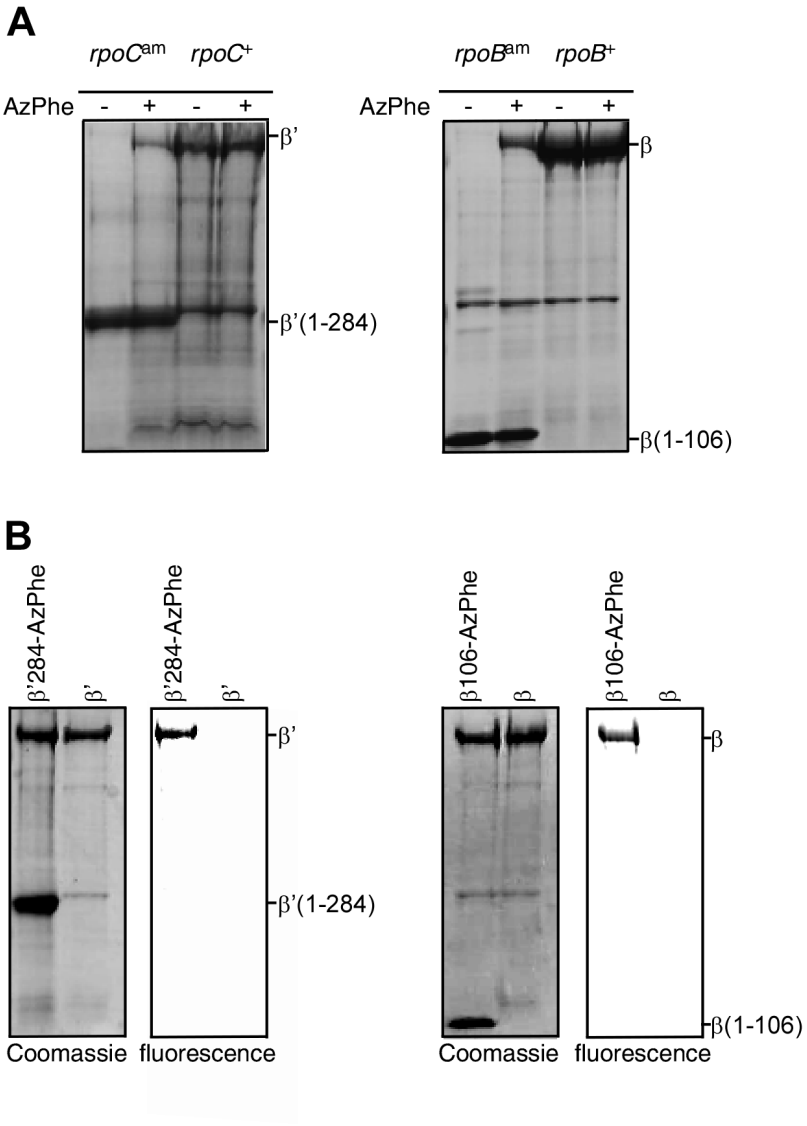


Fig. 9. Synthesis of CyB-phosphine and A647[20 Å]-phosphine

4.2. Site-specific incorporation of fluorescent probes into β’ and β subunits

In order to incorporate probes site-specifically at the tips of the two pincers, an approach comprising of unnatural amino acid mutagenesis and Staudinger ligation was used. The procedure involved (Figs. ???, 10A, B): (i) site-directed mutagenesis of genes expressing β’ and β subunits to incorporate nonsense amber (TAG) codons at the position of interest; (ii) preparation of β’ and β subunits containing 4-azidophenylalanine at the position of interest, accomplished by growing cells which contain an engineered suppressor-transfer RNA (tRNA)/aminoacyl-tRNA-synthetase pair, in a medium supplemented with 4-azidopheylalanine; (iii) incorporation of fluorescent probes Cy3B (donor) and Alexa 647 (acceptor) into β’ and β subunits by azide-specific chemical modification, accomplished by Staudinger ligation using phosphine derivatives of fluorescent probes.

The resulting labelled proteins showed high labelling efficiencies and high labelling specificities (Fig. 10C).



C

RNAP subunit	labelling efficiency (%)	labelling specificity (%)
β'284-Cy3B	88	94
β106-A647	90	92

Fig. 10. Incorporation of fluorescent probes into RNAP β' and β subunits.
(A) Use of unnatural amino acid mutagenesis to incorporate 4-azidophenylalanine(AzPhe) at specified positions in RNAP β' and β subunits. *Left subpanel:* Incorporation of AzPhe at position 284 of β' subunit (*rpoC*^{am}, *rpoC* gene with amber codon at position 284; *rpoC*$^{+}$, wild-type *rpoC* gene; β'(1-284), amber fragment

comprising β’ residues 1-284; β', full-length β’). In experiments with *rpoC*^{am}, full-length β’ was observed only in the presence of added AzPhe. *Right subpanel:* Incorporation of AzPhe at position 106 of β subunit (*rpoB*^{am}, *rpoB* gene with amber codon at position 106; *rpoB*$^{+}$, wild-type *rpoB* gene, ; β(1-106), amber fragment comprising β residues 1-106; β, full-length β). In experiments with *rpoB*^{am}, full-length β was observed only in the presence of added AzPhe.

(B) Use of Staudinger ligation to incorporate fluorescent probes at AzPhe-containing positions in RNAP β’ and β subunits. *Left subpanel:* Products of labelling reactions of β’284-AzPhe (Fig. 10A) and β’ with Cy3B-phosphine, as detected by Coomassie staining (left) and fluorescence scanning with 532nm excitation and 580 nm bandpass emission filters (right). Fluorescent labelling is observed only in reactions with β’284-AzPhe and only for full-length products. *Right subpanel:* Products of labelling reactions of β106-AzPhe (Fig 10A) and β with Alexa647-phosphine, as detected by Coomassie staining (left) and fluorescence scanning with 633 nm excitation and 670 nm bandpass emission filters (right). Fluorescent labelling is observed only in reactions with β106-AzPhe and only for full-length products.

(C) Labelling efficiencies and labelling specificities (see Materials and Methods: β' and β derivatives).

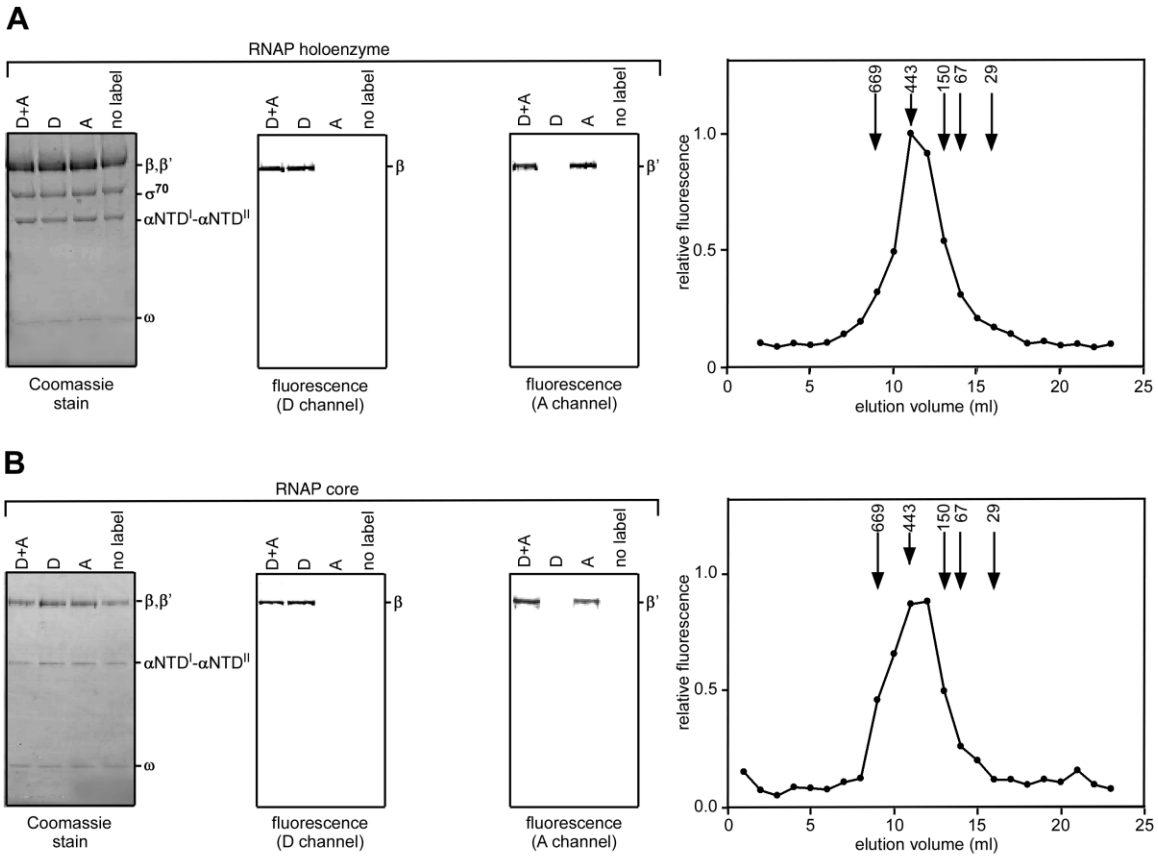
4.3. Incorporation of fluorescent probes into functional RNAP

In order to incorporate fluorescent probes into RNAP, *in vitro* reconstitutions of RNAP core and RNAP holoenzyme were performed. Unlabelled and fluorescently labelled subunits of RNAP core ($\beta'/\beta/\alpha^{I}/\alpha^{II}/\omega$) were combined, denatured, renatured and purified to prepare intact RNAP. Fluorescently labelled RNAP holoenzyme ($\beta'/\beta/\alpha^{I}/\alpha^{II}/\omega/\sigma^{70}$) was prepared using procedures analogous to the preparation of RNAP core, but comprising of an additional step involving σ^{70} (see materials and methods) .

Subunit stoichiometries and the presence of fluorescent probes were confirmed by denaturating polyacrylamide gel electrophoresis (PAGE) of donor-acceptor labelled (D+A) RNAP, donor (D) labelled RNAP, acceptor (A) labelled RNAP and unlabelled RNAP, followed by Coomassie staining and fluorescence scanning (Fig. 11A, B). The prepared RNAP derivatives exhibited correct subunit stoichiometries and correct fluorescent emissions.

To determine the aggregation states of the RNAP core and RNAP holoenzyme, analytical size exclusion chromatography was performed. The results showed both RNAP core and RNAP holoenzyme have the expected monomeric aggregation state (Fig. 11A, B; *right subpanel*).

D+A labelled RNAP, D labelled RNA and A labelled RNAP showed high specific activities in transcription assays (Fig. 11C).



C

RNAP	specific activity (%)
[β106-A647; β'284-Cy3B]-RNAP (D+A)	70
[β'284-Cy3B]-RNAP (D)	80
[β106-A647]-RNAP (A)	70
RNAP (no label)	100

Fig. 11. Incorporation of fluorescent probes into intact RNAP.
(A) *In vitro* reconstitution of RNAP holoenzyme from fluorescently labelled β', fluorescently labelled β, Flag-αNTD^{I}-GSGGSG-αNTD^{II}, ω, and σ^{70}. *Left subpanel*: Preparation of donor+acceptor-labelled (D+A), donor-labelled (D), acceptor-labelled (A), and unlabelled RNAP-σ^{70} holoenzyme, as detected by Coomassie staining fluorescent scanning in the D channel (532 nm excitation and 580 nm bandpass emission filters), and fluorescent scanning in the A channel (633 nm excitation and 670 nm bandpass emission filters). Fluorescence is observed in the D channel for D+A and D, and is observed in the A channel for D+A labelled and A. *Right subpanel*: Analytical size exclusion chromatography of donor+acceptor-labelled (D+A) RNAP-σ^{70} holoenzyme.

(B) *In vitro* reconstitution of RNAP core from fluorescently labelled β', fluorescently labelled β, Flag-αNTD^{I}-GSGGSG-αNTD^{II}, and ω. *Left subpanel*: Preparation of donor+acceptor-labelled (D+A), donor-labelled (D), acceptor-labelled (A), and unlabelled RNAP core, as detected by Coomassie staining, fluorescent scanning in the D channel (532 nm excitation and 580 nm band pass emission filters), and fluorescent scanning in the A channel (633 nm excitation and 670 nm bandpass emission filters). Fluorescence is observed in the D channel for D+A and D. and is observed in the A channel for D+A and A. *Right subpanel*: Analytical size exclusion chromatography of donor+acceptor-labelled (D+A) RNAP core.
(C) Transcriptional activities of RNAP-σ^{70} holoenzyme derivatives (D+A, D, A, and unlabelled).

4.4. Spectral properties of fluorescent probes in RNAP holoenzyme and RPo

To determine the spectral properties of labelled fluorophores, Ro was determined, in parallel, in samples of RNAP holoenzyme and RNAP-promoter open complex (RPo). Calculation of Ro involves determination of donor quantum yield (Q_D) and donor emission-acceptor excitation spectral overlap integral (J). The results showed that all Q_D, J and Ro remains essentially unchanged in RNAP holoenzyme and RPo (Fig. 12A).

Measurement of steady-state anisotropies of labelled RNAP was also performed in parallel in RNAP holoenzyme and RPo. The results suggest that anisotropies remain relatively similar in RNAP holoenzyme and RPo (Fig. 12B).

Comparison of Ro and anisotropy values in RNAP holoenzyme and RPo revealed that the spectral properties of fluorescently labelled RNAP holoenzyme are not affected by binding of DNA – suggesting any changes in smFRET efficiencies can be attributed to changes in D-A distance only.

A

RNAP	RNAP holo			RPo		
	Q_D	J (M^{-1} cm^3)	**R_0** (Å)	Q_D	J (M^{-1} cm^3)	**R_0** (Å)
[β106-A647; β'284-Cy3B]-RNAP	0.55	8.9×10^{-13}	**60.1**	0.55	9.1×10^{-13}	**60.4**
[β106-Cy3B; β'284-A647]-RNAP	0.55	9.2×10^{-13}	**60.5**	0.55	9.1×10^{-13}	**60.7**

B

RNAP	anisotropy	
	RNAP holo	RPo
[β106-A647]-RNAP	0.27	0.28
[β106-Cy3B]-RNAP	0.27	0.26
[β'284-A647]-RNAP	0.25	0.26
[β'284-Cy3B]-RNAP	0.24	0.23

Fig. 12. Measurement of spectroscopic properties.
(A) Spectroscopic properties of fluorescent probes, measured in the context of RNAP-σ^{70} holoenzyme (center) and in the context of RPo (right). Q_D, quantum yield of donor; J, spectral overlap integral; R_o, Förster radius.
(B) Steady-state anisotropies of fluorescent probes, measured in the context of RNAP-σ^{70} holoenzyme (center) and in the context of RPo (right).

4.5. Computational and structural framework to relate smFRET efficiencies (*E*) to RNAP clamp conformational states: *T. thermophilus*-σ^A holoenzyme.

In order to infer changes in smFRET efficiencies with respect to structural changes in RNAP, a computational framework was developed (Fig. 13). Starting with the crystal coordinates of *T.thermophilus* holoenzyme, A series of 18 structural models, ranging from a fully open clamp state to a fully closed clamp, were created by rotating the clamp in 2˚ increments about a hinge located at the bottom of the β' pincer. The fluorescent probes were modeled at their respective positions. Each torsion angle on the probes is rotated in 60˚ increments to generate ensembles of sterically acceptable probe conformations. 1000 random donor positions were selected and 1000 acceptor positions were selected. 1000000 pairwise distances were determined, followed by calculation of corresponding 1000000 FRET efficiency (*E*) values. Mean of the 1000000 E values represents the mean E of the model.

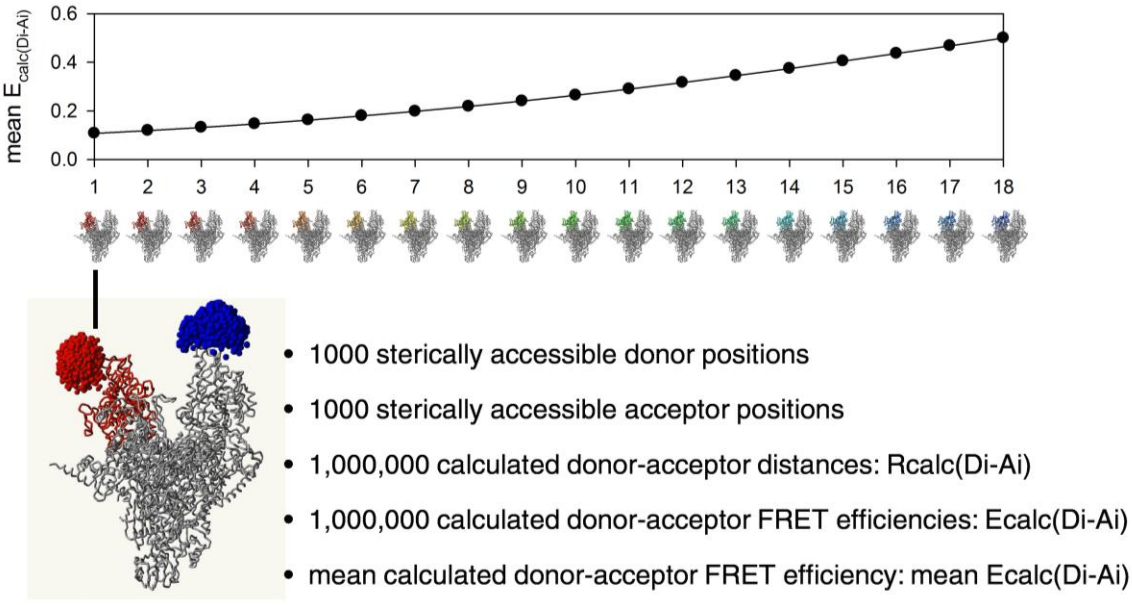


Fig. 13. Computational and structural framework to relate smFRET efficiencies (*E*) to RNAP clamp conformational states: *T. thermophilus* RNAP-σ^A holoenzyme. Starting from atomic coordinates from a crystal structure of *T. thermophilus* RNAP-σ^A holoenzyme (PDB code 1IW7), 18 structural models having RNAP clamp conformations ranging from fully open to fully closed in 2˚ increments were generated (structures immediately beneath graph; σ omitted for clarity; see Materials and Methods: Molecular modelling). For each of the 18 structural models, 1,000 sterically accessible donor positions were defined (red spheres on structure at lower left), 1,000 sterically accessible acceptor positions were defined (blue spheres on structure at lower left), 1,000,000 pairwise donor-acceptor distances ($R_{calc(Di\text{-}Ai)}$) were calculated, 1,000,000 donor-acceptor FRET efficiencies ($E_{calc(Di\text{-}Ai)}$) were calculated, and the mean donor-acceptor FRET efficiency (mean $E_{calc(Di\text{-}Ai)}$) was calculated. Results were plotted (graph at top). The resulting graph and set of structures allow a smFRET efficiency (*E*) to be input and a structure to be output.

4.6. Clamp conformation in RNAP-σ^{70} holoenzyme and RNAP core

To determine RNAP clamp conformation in solution, confocal optical microscopy with alternating laser excitation (ALEX) was employed to measure the smFRET efficiencies between the probes incorporated at the tips of the clamp and the β pincer. Clamp conformation was first analyzed in free RNAP holoenzyme. It was observed that that RNAP-σ^{70} holoenzyme exhibited a broad multimodal distribution of smFRET efficiencies (Fig. 15; *upper panel*). The distribution could be fitted with three Gaussian distributions, corresponding to three populations, in which the mean of the Gasussians define the mean smFRET efficiency of the population:

(i) a major subpopulation with mean E = 0.15 and mean R = 81 Å, corresponding to an open clamp state.

(ii) a minor subpopulation with mean E = 0.28 and mean E = 69 Å, corresponding to a closed clamp state in which the clamp is rotated inwards by ~14˚.

(iii) a minor subpopulation with mean E = 0.40 and mean E = 64 Å, corresponding to a collapsed clamp state in which the clamp is rotated inwards by ~22˚.

The open state has solvent accessible active center widths of 20 Å, sufficient to accommodate dsDNA. The closed state has solvent accessible active center widths of ~12 Å, sufficient to accommodate ssDNA, but insufficient to accommodate dsDNA. The

collapsed state has solvent accessible active center widths of 8 Å, insufficient to accommodate either ssDNA or dsDNA.

The results demonstrate that the RNAP-σ^{70} holoenzyme can sample multiple conformational states in solution ranging from an open state to a collapsed state. The open clamp state, which can permit unimpeded entry of dsDNA in the active center cleft is the predominant state in solution.

Identical results were obtained with RNAP core, indicating the RNAP clamp conformational states are an intrinsic property of RNAP core (Fig. 15; *lower panel*).

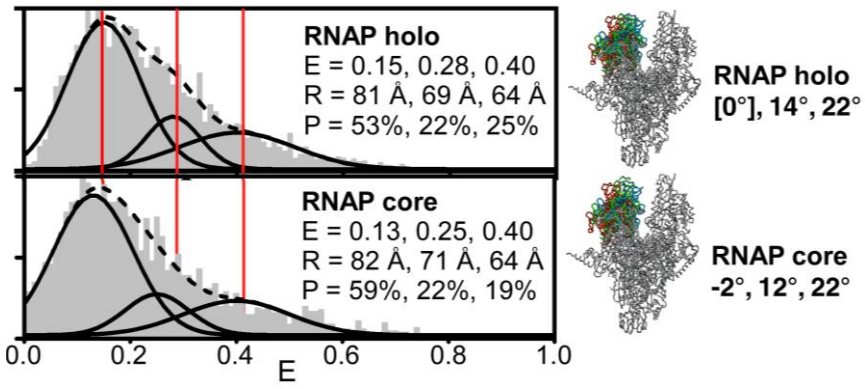


Fig. 15. Clamp conformation in RNAP-σ^{70} holoenzyme and RNAP core.
Panels show histograms and Gaussian fits of observed donor-acceptor smFRET efficiencies, *E* (at left); mean *E*, mean *R*, and percentage, *P*, for each subpopulation (in inset at left); inferred structural states of the RNAP clamp (at right; colored as in Fig. 13); and inferred extents of closure of the RNAP clamp (at far right; in degrees relative to the open state). The vertical red lines denote mean *E* of the open, closed, and collapsed states in RNAP-σ^{70} holoenzyme.

4.7. RNAP clamp conformation in σ^{70}-dependent transcription initiation and elongation

Clamp conformation was analyzed in each step in σ^{70}-dependent transcription initiation and elongation. Starting with a RNAP-σ^{70} holoenzyme, upon addition of promoter DNA and formation of RPo, the distribution of smFRET efficiencies became narrow and unimodal (Fig. 17, *second panel;* Fig. 16A; Table 1). Only the subpopulation corresponding to the closed clamp state exclusively remained, with the clamp rotated inward by ~16˚. The results suggest that after loading of DNA into, and unwinding of DNA in, the RNAP active center cleft to form RPo, the clamp closes. It is plausible that direct electrostatic interactions between the negatively charged ssDNA in active center cleft induces or stabilizes clamp closure.

Initial transcribing complex, RPitc,2 was formed by the addition of an initiating dinucleotide, ApA (Fig. 17, *third panel*; Fig. 16A; Table 1). Upon formation of RPitc2, no change in the distribution of smFRET efficiencies was observed. The distribution remained narrow and unimodal with the clamp rotated inwards ~16˚.Upon further additions of nucleotide subsets in sequence to form RPitc,4 (Fig. 17, *fourth panel*; Fig. 16A ; Table 1) and RPitc,7 (Fig. 17, *fifth panel*; Fig. 16A; Table 1) respectively, no change in distribution of smFRET efficiencies was observed. In RPitc4 and RPitc7, in order to synthesize abortive RNA products of up to 4 nt and 7 nt, RNAP employs a scrunching mechanism, in which the enzyme remains stationary relative to DNA, unwinds 2 and 5 additional bp of downstream DNA and pulls the downstream DNA within itself. Thus, upon transition from a catalytically-competent initiation complex in

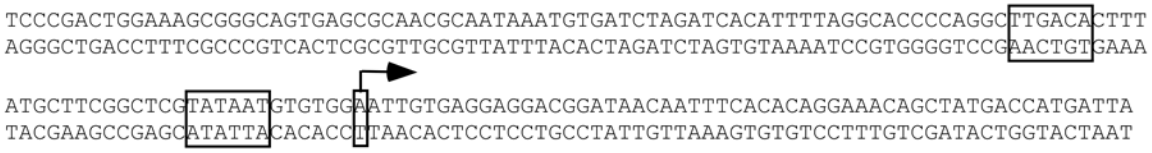
RPo to a mechanistically distinct catalytically-active intial transcribing complex in RPitc, the clamp continues to remain closed.

Upon further addition of a NTP subset, yielding RDe14, in which RNA product of up to 14 nt in length is present and the RNA exit channel is completely occupied with RNA, no change in smFRET efficiencies was observed. – the distribution remained narrow and unimodal (Fig. 17, *sixth panel*; Fig. 16A; Table 1). Thus upon transition from a scrunching mechanism of RNA synthesis in RPitc to a stepping mechanism of RNA synthesis, in which the RNAP translocates relative to DNA 1 bp at a time, the clamp continues to remain closed.

Upon futher adiition of a NTP substrate, yielding RDe15, in which up to 15 nt of RNA is synthesized and the RNA extends beyond the RNA-exit channel and emerges into solution, the clamp continues to remain closed (Fig. 17, *seventh panel*; Fig. 16A; Table 1).

A

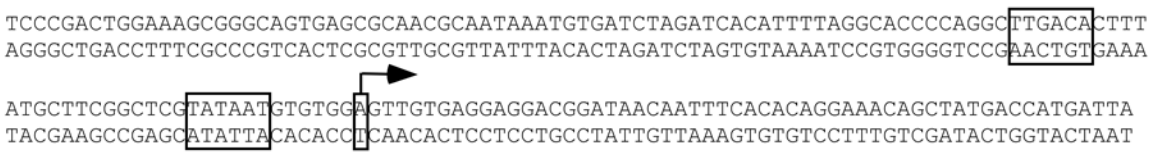
***lacCONS-14*(-107/+56)**



NTP subset	RNA product	complex
RPo		RPo
RPo + ApA	AA	RPitc,2
RPo + ApA/UTP	AAUU	RPitc,4
RPo + ApA/UTP/GTP	AAUUGUG	RPitc,7
RPo + ApA/UTP/GTP/ATP	AAUUGUGAGGAGGA	RDe,14
RPo + ApA/UTP/GTP/ATP/3'-OMeCTP	AAUUGUGAGGAGGAC*	RDe,15

B

***lacCONS(A+2G)-14*(-107/+56)**



NTP subset	RNA product	complex
RPo		RPo
RPo + ApG	AG	RPitc,2
RPo + ApG/UTP	AGUU	RPitc,4
RPo + ApG/UTP/GTP	AGUUGUG	RPitc,7
RPo + ApG/UTP/GTP/ATP	AGUUGUGAGGAGGA	RDe,14
RPo + ApG/UTP/GTP/ATP/3'-OMeCTP	AGUUGUGAGGAGGAC*	RDe,15

Fig. 16. DNA fragments analyzed.
In each panel, the top subpanel shows the DNA fragment sequence (promoter -35 element, -10 element, and transcription start site boxed), and the bottom subpanel shows the expected maximum-length RNA product and the expected transcription complex for each for each reaction condition analyzed.
(A) DNA fragment *lacCONS-14*(-107/+56) [contains *lacCONS-14* promoter (*45*); yields transcription elongation complexes that contain σ^{70} [68, 70, 126]; used in experiments in Figs. 17, 19, 20].

(B) DNA fragment *lacCONS(A+2G)-14*(-107/+56) [contains *lacCONS(A+2G)* promoter (*24*); yields transcription elongation complexes that do not contain σ^{70} [68, 126]; used in experiments in Fig. 21].

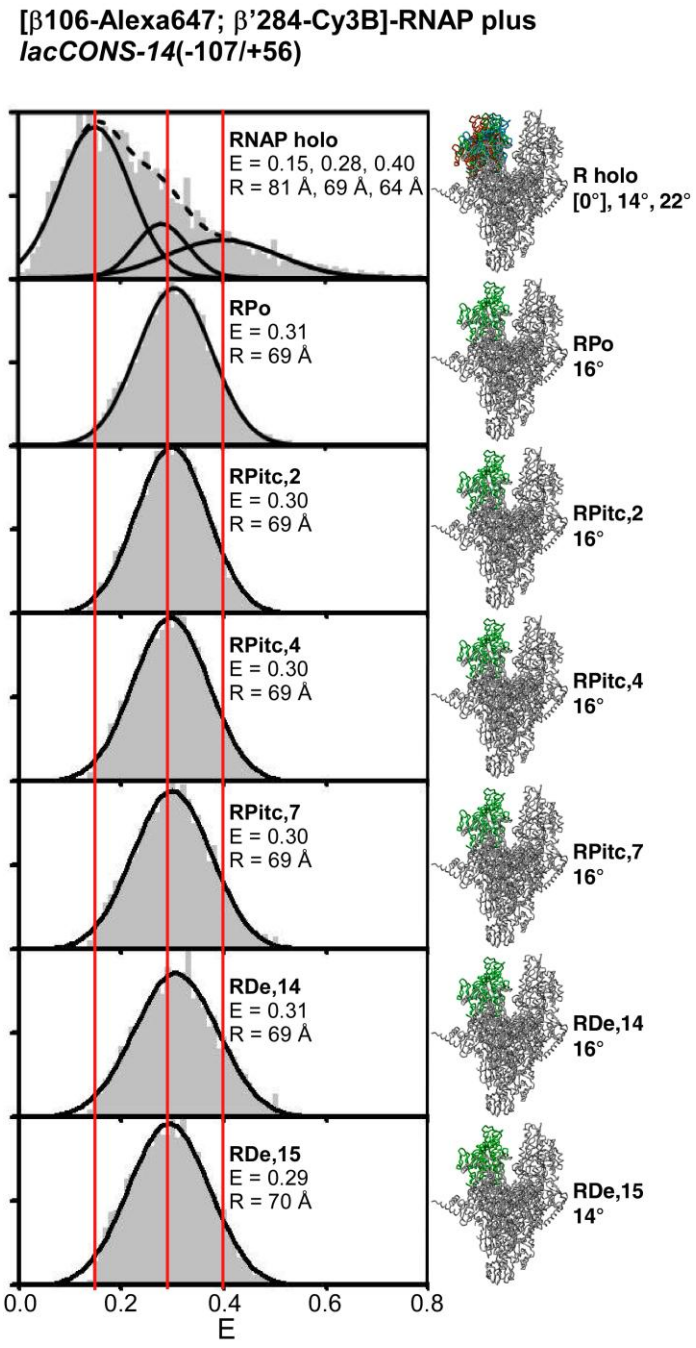


Fig. 17. RNAP clamp conformation in σ^{70}-dependent initiation and elongation. RNAP clamp conformation in RNAP-σ^{70} holoenzyme, RPo, RPitc,2 (2 nt RNA), RPitc,4 (4 nt RNA), RPitc,7 (7 nt RNA), RDe,14 (14 nt RNA), and RDe,15 (15 nt RNA). Panels show histograms and Gaussian fits of observed donor-acceptor smFRET efficiencies, *E* (at left); mean *E*, mean *R*, and percentage, *P*, for each subpopulation (in inset at left); inferred structural states of the RNAP clamp (at right; σ omitted for clarity; colored as in Fig. 13); and inferred extents of closure of the RNAP clamp (at far right; in degrees relative to the open state). The vertical red lines denote mean *E* of the open, closed, and collapsed states in RNAP-σ^{70} holoenzyme.

Table 1: Summary of smFRET efficiencies (*E*), distances (*R*), and clamp rotation states: two-color FRET experiments

complex	subpopulation	figure	*E* mean (±SEM)	*R* mean (±SEM) (Å)	clamp rotation mean (±SEM) (degrees)
RNAP-σ^{70} holoenzyme	open (53±3%)	17	0.15(±0.00)	81(±0.3)	[0]
	closed (22±2%)	17	0.28(±0.00)	70(±0.2)	14(±2)
	collapsed (25±3%)	17	0.40(±0.00)	81(±0.1)	22(±2)
RPo	closed (100%)	17	0.31(±0.00)	69(±0.2)	16(±2)
RPitc	closed (100%)	17	0.30(±0.00)	69(±0.0)	16(±2)
RDe	closed (100%)	17	0.31(±0.00)	69(±0.2)	16(±2)
RNAP core	open (59±3%)	15	0.13(±0.00)	82(±0.1)	-2(±2)
	closed (22±2%)	15	0.25(±0.01)	72(±0.4)	12(±4)
	collapsed (19±3%)	15	0.40(±0.00)	64(±0.2)	22(±2)
RNAP-σ^{54} holoenzyme	open (48±2%)	22	0.15(±0.00)	81(±0.3)	0(±2)
	closed (30±2%)	22	0.28(±0.00)	70(±0.0)	14(±2)
	collapsed (22±3%)	21	0.45(±0.00)	62(±0.1)	24(±2)
RPc	open (72±3%)	22	0.15(±0.01)	81(±0.7)	0(±2)
	closed (28±3%)	22	0.30(±0.01)	69(±0.7)	14(±4)
RPc+NtrC	open (70±2%)	22	0.17(±0.01)	78(±0.5)	2(±4)
	closed (30±2%)	22	0.30(±0.01)	69(±0.7)	14(±2)
RPi1	open (65±2%)	22	0.16(±0.01)	79(±0.6)	2(±2)
	closed (35±2%)	22	0.30(±0.01)	69(±0.0)	14(±2)
RPi2	open (66±2%)	22	0.15(±0.01)	81(±1.3)	0(±6)
	closed (34±2%)	22	0.30(±0.01)	69(±0.4)	14(±4)
RPo	closed (100%)	22	0.26(±0.01)	72(±0.8)	12(±2)

4.8. RNAP clamp conformation in σ^{70}-dependent transcription initiation: three-color alternating laser excitation: control experiment

To validate the loss of σ^{70} did not complicate the interpretations in Figs. 15 and 17, three color smFRET experiments were performed (see Fig. 18 for experimental setup and data analysis), in which, in addition to the probes on the RNAP clamp and β pincer, a probe of a third color (Alexa 488) was incorporated on σ^{70}. RNAP clamp conformation was monitored by measuring smFRET efficiencies between the probe at the tip of the clamp and the probe at the tip of the β pincer. Post collection, data was filtered to include only complexes that contained all the three probes, i.e. complexes that contained σ^{70}. The results obtained matched previous results, suggesting the previous results were not influenced by the loss of σ^{70} (Fig. 19; Table 2).

Analogous three color experiments were performed to ensure the previous results were not complicated by loss of DNA. A probe of a third color was incorporated in DNA. Clamp conformation was monitored and results were filtered to include only complexes that contained all three probes. The results obtained again matched previous results suggesting the previous results were not complicated by loss of DNA (Fig. 19; Table 2).

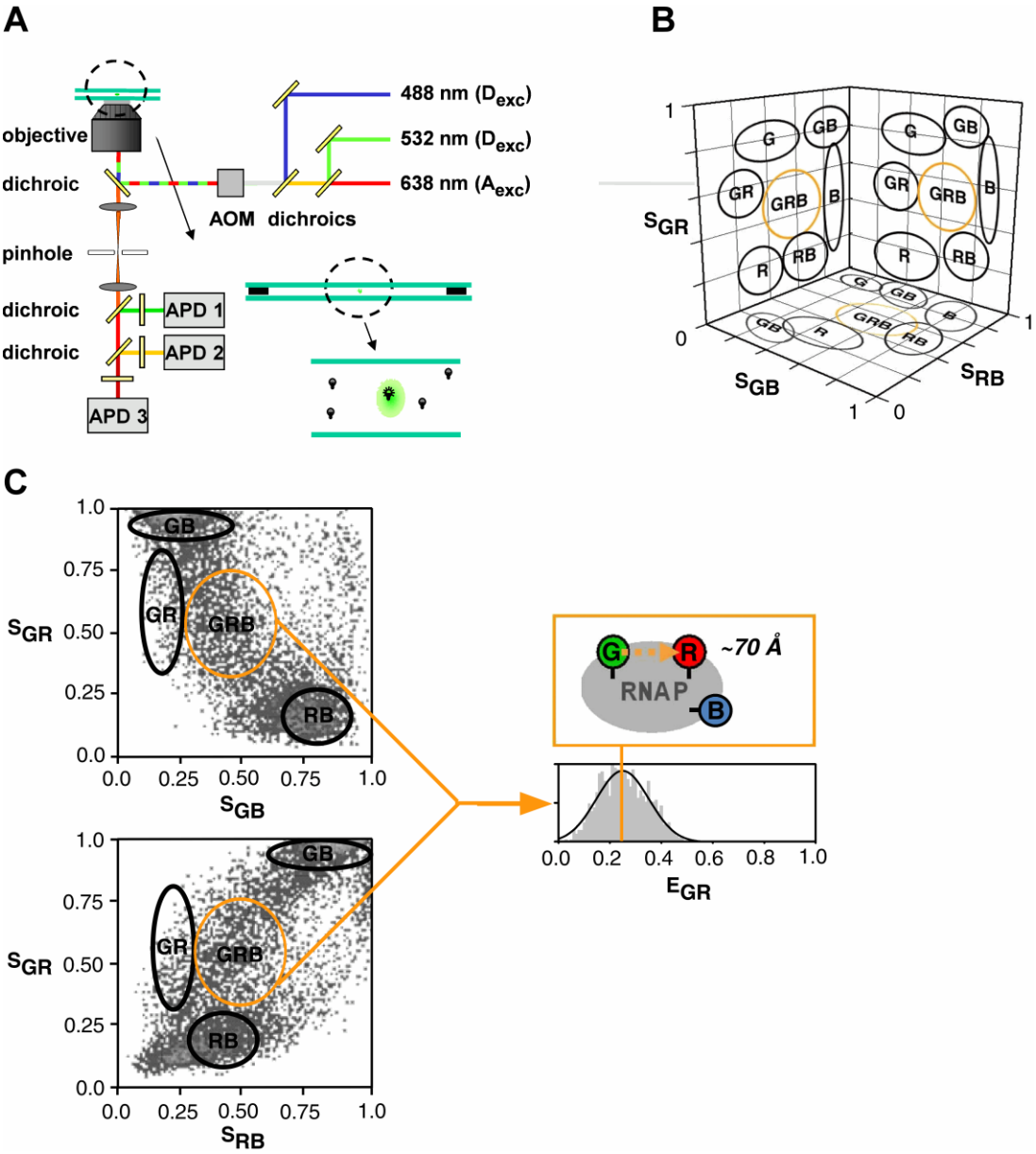


Fig. 18. Measurement of smFRET: three-color alternating laser excitation (ALEX). **(A)** Experimental setup. Confocal optical microscopy with alternating laser excitation (ALEX) was used to generate a femtoliter-scale excitation and detection volume in a sample chamber (inset at right center) and to detect and quantify fluorescent signals from single molecules of fluorescently labelled transcription complexes freely diffusing in solution as they transit the femtoliter-scale excitation and detection volume (inset at lower right). The experimental setup contains three lasers: one providing excitation at the excitation wavelength of fluorescent probe "G" (532 nm), one providing excitation at the excitation wavelength of fluorescent probe "R" (638 nm), and one providing excitation at the excitation wavelength of fluorescent probe "B" (488 nm). The lasers are alternated on the μs time scale using an acousto-optical modulator (AOM). Emission is monitored in the G, R, and B channels using avalanche photodiodes (APD 1, APD 2, and APD3).

(B) Three-color stoichiometry based molecule sorting: concept (*15*). The figure shows a three-dimensional $S_{GR}/S_{GB}/S_{RB}$ plot with predicted locations of observations of species containing one probe (G, R, and B), species containing two probes (GR, GB, and RB), and species containing all three probes (GRB; orange ovals). Observations of species containing only three probes (GRB; orange ovals)--the species of interest--are centered at 0.5/0.5 on the on S_{GB}/S_{RB} plane, 0.5/0.5 on the S_{GB}-S_{GR} plane, and 0.5/0.5 on the S_{RB}/S_{GR} plane.

(C) Three-color stoichiometry-based molecule sorting: representative data. The left half of the figure shows representative S_{GB}/S_{GR} and S_{RB}/S_{GR} plots. Observations for species that contain all three probes (GRB; orange oval in central part of each plot)--the species of interest--readily can be distinguished from observations for species that do not contain all three probes (black ovals in peripheral parts of each plot). The right half of the figure shows a representative histogram of G-R smFRET efficiencies (E_{GR}) for species that contain all three probes. The histogram defines the equilibrium population distribution of E_{GR} for species that contain all three probes, defines the number of subpopulations with distinguishable E_{GR} (one in this case), and defines mean E_{GR} and mean R_{GR}.

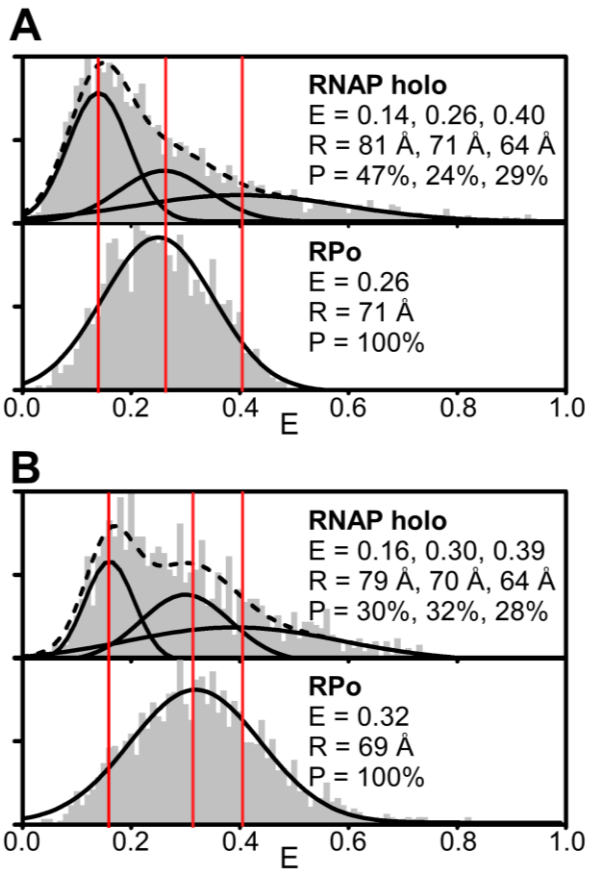


Fig. 19. Three-color alternating laser excitation: RNAP-σ^{70} holoenzyme and RPo.
(A) Control three-color FRET experiments with third probe on σ^{70} (data filtered to include only molecules containing σ^{70}).
(B) Control three-color FRET experiments with third probe on DNA (data for RPo filtered to include only molecules containing DNA).

Table 2: Summary of smFRET efficiencies (*E*), distances (*R*), and clamp rotation states: three-color FRET experiments

complex	subpopulation	figure	*E* mean (±SEM)	*R* mean (±SEM) (Å)	clamp rotation mean (±SEM) (degrees)
RNAP-σ^{70} holoenzyme	open (47±2%)	18A	0.14(±0.02)	81(±1.7)	[0]
	closed (24±1%)	18A	0.26(±0.04)	71(±1.9)	12(±8)
	collapsed (29±2%)	18A	0.40(±0.06)	64(±2.7)	22(±6)
RPo	closed (100%)	18A	0.26(±0.02)	71(±1.3)	12(±4)
RNAP-σ^{70} holoenzyme	open (30±6%)	18B	0.16(±0.01)	79(±0.6)	[0]
	closed (32±3%)	18B	0.30(±0.02)	70(±0.8)	16(±4)
	collapsed (28±2%)	18B	0.39(±0.05)	64(±2.2)	22(±8)
RPo	closed (100%)	18B	0.32(±0.02)	69(±1.0)	16(±6)

4.9. RNAP clamp conformation in σ^{70}-dependent transcription initiation and elongation: Control experiment: reversal of donor and acceptor probe positions

Straightforward correlations of FRET efficiencies with distances can sometimes be complicated due to orientation of fluorophore transition dipole. One potential uncertainty in extraction of distance information from FRET efficiencies lies in the orientation factor κ^2, which depends upon the orientation of donor emission dipole, orientation of acceptor absorbance dipole and the line connecting the center of donor and acceptor fluorophores. The values of κ^2 can range from 0 to 4. If the fluorophores undergo isotropic reorientation on timescale which is much shorter than the excited state lifetime of the donor, then κ^2 = 2/3. In most FRET measurements, this approximation holds true for fluorophores which have a single point of attachment and are capable of freely rotating. Accurate estimation of κ^2 is not possible in many cases.

To rule out complications arising because of potential lack of rotational flexibility of the fluorophores, due to the steric hindrances, 3 control experiments were performed:

(i) RNAP clamp conformation was monitored in samples with reversed donor and acceptor probe labelling positions, resulting in similar donor-acceptor distances, but different fluorophore environment (Fig. 20); (ii) RNAP clamp conformation was monitored in samples labelled at alternate positions (data not shown); (iii) RNAP clamp conformation was monitored using alternate dye pairs (data not shown).

The results from all the above control experiments matched previous results, indicating the original results were not complicated by RNAP-probe interactions.

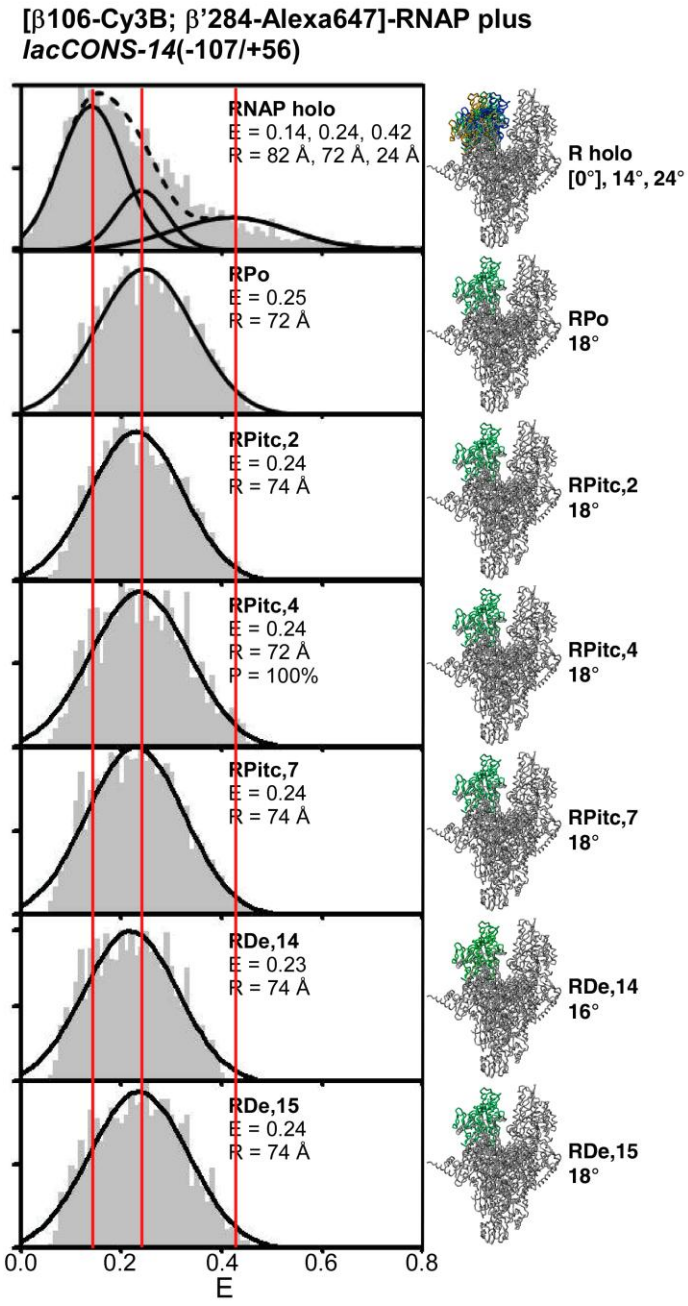


Fig. 20. Fig. RNAP clamp conformation in σ^{70}-dependent initiation and elongation: control experiment: reversal of positions of donor and acceptor probes.
The figure presents results of experiments identical to those in Fig. 17, but using an RNAP derivative that has the donor probe at β position 106 and the acceptor probe at β' position 284 ([β106-Cy3B; β'284-Alexa647]-RNAP-σ^{70} holoenzyme).

4.10. RNAP clamp conformation in σ^{70}-dependent transcription initiation and elongation: analysis of promoter that yields a σ^{70}-free transcription elongation complexes.

Previous ensemble and single-molecule experiments have shown, at the LacCONS promoter, in early elongation complexes (RDe14), σ^{70} remains associated with RNAP. Subsequently, upon formation of mature elongation complexes (RDe14), σ^{70} is obligatorily released. Previous work has further shown that LacCONS promoter contains a "-10-like" sequence determinant which increases the half-life and extent of σ^{70} retention in early elongation complexes. A variant of the LaCONS (A+2G) promoter lacks the specific sequence determinants resulting in loss of σ70 in early elongation complexes.

To assess the effects of loss of σ^{70} in early elongation complexes, and to mimic the subunit composition of mature elongation complexes, RNAP clamp conformation was monitored in complexes formed with LacCONS (A+2G) promoter (Fig. 21). Results were similar to those obtained with LacCONS promoter -- the clamp remains closed to the same extent in σ^{70}-free elongation complexes.

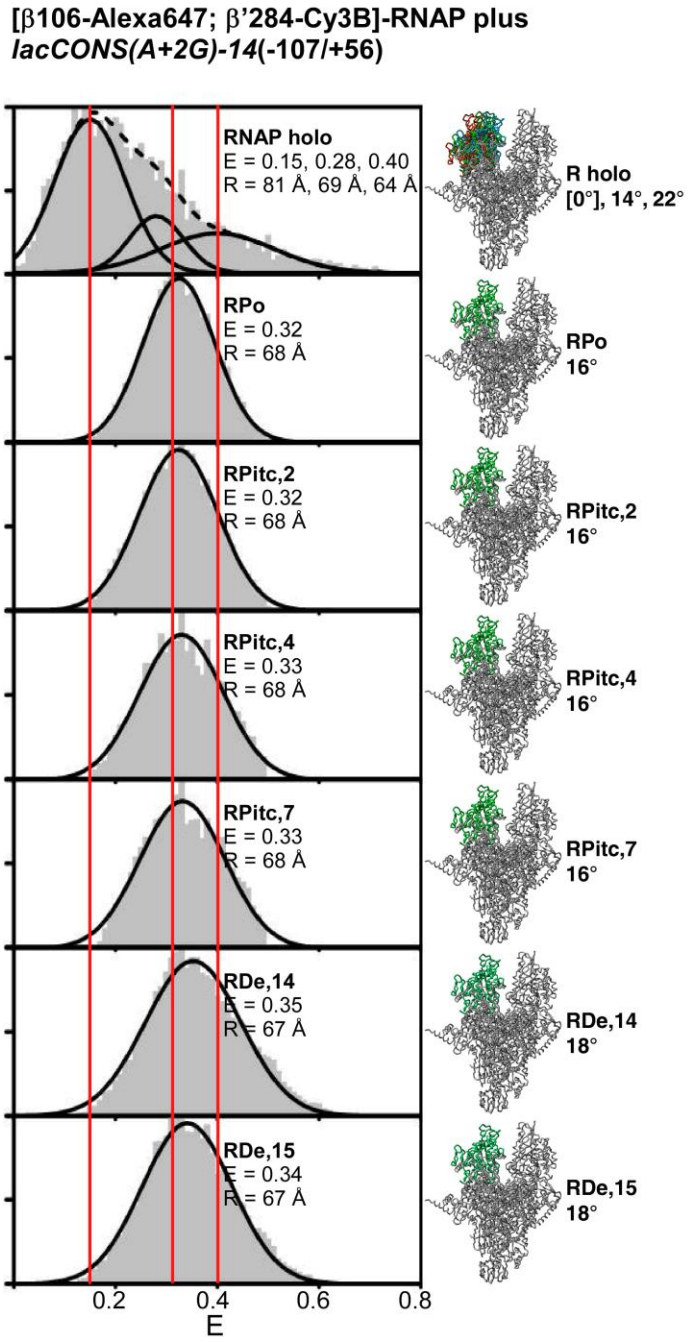


Fig. 21. RNAP clamp conformation in σ^{70}-dependent initiation and elongation: additional experiment: analysis of promoter that yields σ^{70}-free transcription elongation complexes.

The figure present results of experiments identical to those in Fig. 17, but using a promoter that yields transcription elongation complexes that do not retain σ^{70} [*lacCONS(A+2G)-14*(-107/+56); Fig. 16B].

4.11. RNAP clamp conformation in σ^{54}-dependent transcription initiation

To assess RNAP clamp conformation in the intermediates leading up to the formation of RPo, σ^{54}-dependent transcription initiation was analyzed (Fig. 22; Table 1). As mentioned earlier, in contrast to σ^{70} dependent transcription initiation, RNAP-σ^{54} holoenzyme forms stable RPc. σ54-RPc isomerizes to RPo only in the presence of an activator protein (in this case, NtrC1) and ATP hydrolysis. In addition to analysis of RPc and RPo, the σ^{54}-dependent initiation pathway permits trapping of complexes at different stages of ATP hydrolysis by using ATP-analogs; the ground state of ATP hydrolysis (RPi1) can be mimicked by using ADP-BeF_x, and the transition state of ATP hydrolysis (RPi2) can be mimicked by using ADP-AlF_x.

Clamp conformation was analyzed in RNAP-σ^{54} holoenzyme. RNAP-σ^{54} holoenzyme exhibited the same distribution of smFRET efficiencies as RNAP-σ^{70} holoenzyme: a major subpopulation with open clamp state; a minor subpopulation with closed clamp state (clamp rotated inward by ~14˚); a minor subpopulation with collapsed clamp state (clamp rotated inward by ~24˚) (Fig. 22, *first panel*; Table 1). The results suggest that σ^{54} does not alter the clamp conformation in RNAP holoenzyme. The results are consistent with the proposal that conformational states of the RNAP clamp in free holoenzyme are an intrinsic property of core and are not affected by the presence and in this case the nature of σ factor.

Upon formation of RPc by addition of DNA, the open clamp state continues to be the predominant state (Fig. 22, *second panel*; Table 1). The collapsed clamp state disappears but a minor subpopulation corresponding to the closed clamp state remains.

Upon formation of RPc+NtrC1 (by addition of NtrC1 to RPc) (Fig. 22, *third panel*; Table 1), RPi1 (by addition of ADP-BeF_x to RPc+NtrC1) (Fig. 22, *fourth panel*; Table 1) and RPi2 (by addition of ADP-AlF_x to RPc+NtrC1) (Fig. 22, *fifth panel*; Table 1), the clamp continues to remain predominantly open with a minor subpopulation corresponding to the closed clamp state.

Upon formation of RPo (by addition of Ntrc1 and ATP to RPc) the distribution of FRET efficiencies becomes narrow and unimodal (Fig. 22, *sixth panel*; Table 1). The open clamp state disappears and only the population corresponding to the closed clamp state remains, with the clamp rotated inward by 12˚.

The results suggest that the clamp is predominantly open in RNAP-σ^{54} holoenzyme and continues to be predominantly open in all steps leading up to formation of RPo. Only open formation of RPo which is accompanied by loading of DNA in the active center cleft and/or unwinding of DNA in the active center cleft, the clamp closes.

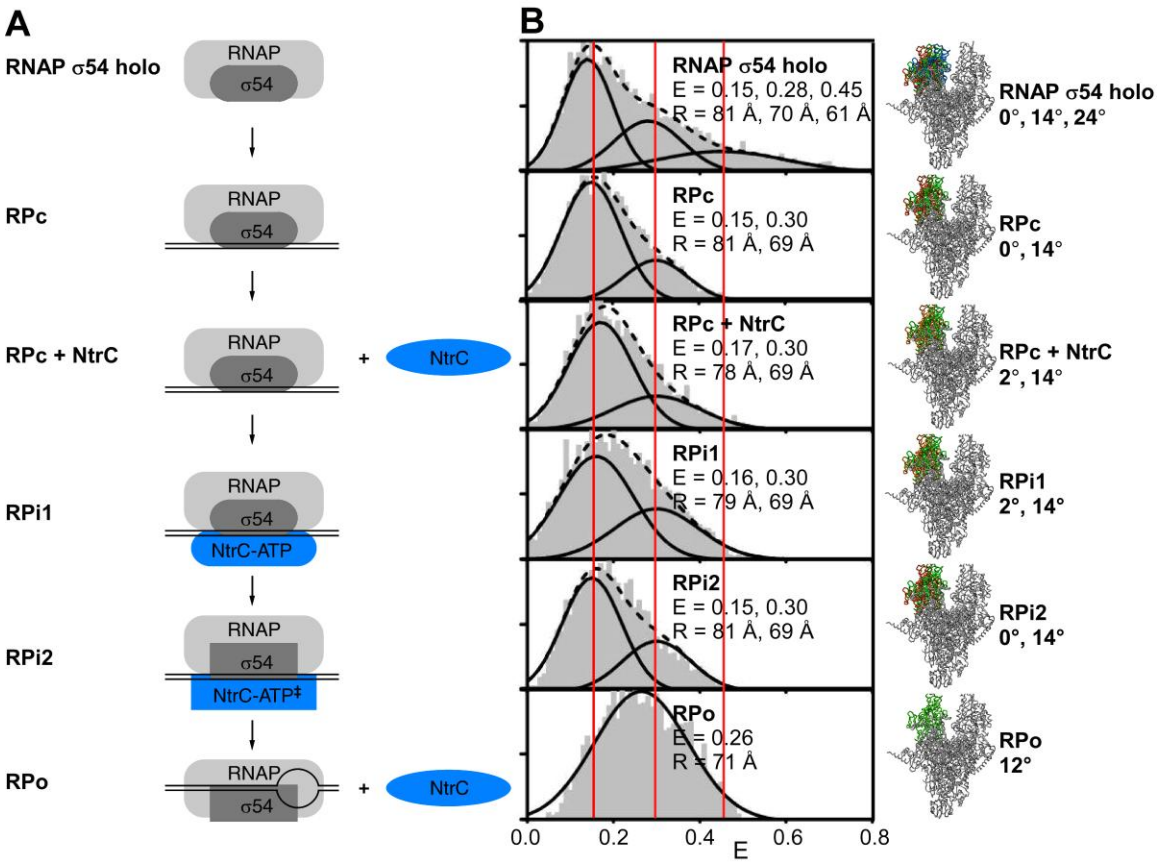


Fig. 22. RNAP clamp conformation in σ^{54}-dependent transcription initiation.
(A) Intermediates in σ^{54}-dependent transcription initiation.

(B) RNAP clamp conformation in RNAP-σ^{54} holoenzyme, RPc, RPc+NtrC1, RPi1, RPi2, and RPo.

4.12. Validation of formation of transcription σ54-dependent transcription intermediate: RPi2

Recent research has shown that σ^{54}-dependent RPi2 formed by addition of ADP-AlF_x (see Fig. 22A for $^{\sigma 54}$-dependent transcription initiation pathway), can engage in synthesis of short RNA products [113]. In a "small-primed-RNA synthesis" (spRNA) assay, addition of an initiating dinucleotide, UpG, and GTP, to RPi2, allows synthesis of 4 nt RNA product. In this work, spRNA assay was performed to compare unlabelled RPc+NtrC1 and and unlabelled RPi2 with fluorescently labelled RPc+NtrC1 and and fluorescently labelled RPi2 (Fig. 23). The results show, in agreement with the literature, that short RNA synthesis occurred only in RPi2 and not in RPc+NtrC. The results further show that fluorescent labelling of RNAP did not interfere with the formation of RPi2.

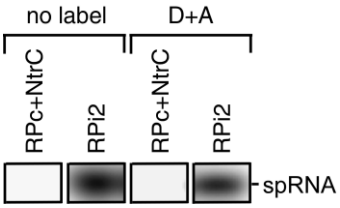


Fig. 23. Formation of transcription initiation intermediates: σ^{54}-dependent transcription initiation.

The figure presents results of small-primed-RNA (spRNA) synthesis assays of samples containing σ^{54}-dependent transcription initiation intermediates RPc+NtrC (does not support spRNA synthesis) and RPi 2 (supports spRNA synthesis). *Left subpanel:* Data for unlabelled RNAP core (no label). *Right subpanel:* Data for donor+acceptor-labelled RNAP core (D+A).

4.13. Effect of inhibitors on RNAP clamp conformation: σ^{70} dependent transcription initiation

To assess the effects of switch-region-target-inhibitors on RNAP clamp conformation in solution, myxopyronin (Myx), corallopyronin (Cor) and ripostatin (Rip) were tested in the single-molecule clamp conformation assay. It has been hypothesized that Myx, Cor and Rip bind to the RNAP switch region, the hinge that mediates opening and closing of the clamp.

Myx, Cor and Rip were found to depopulate the clamp open state, leaving only closed and collapsed clamp states, in which the clamp is rotated inward by ~8˚ to ~26˚ (Fig. 24, *left panels*).

Analogous experiments with non-switch-region-target inhibitors rifampicin (Rif), sterptolydigin (Stl) and CBR703 (CBR) which inhibit steps in transcription subsequent to isomerization of RPc to RPo (Fig. 24, *right panels*).

The results are consistent with the proposal that the switch-region-target-inhibitors lock the clamp in a closed conformation, preventing the isomerization of RPc to RPo, by inhibiting loading of DNA into and unwinding of DNA in, the RNAP active center cleft.

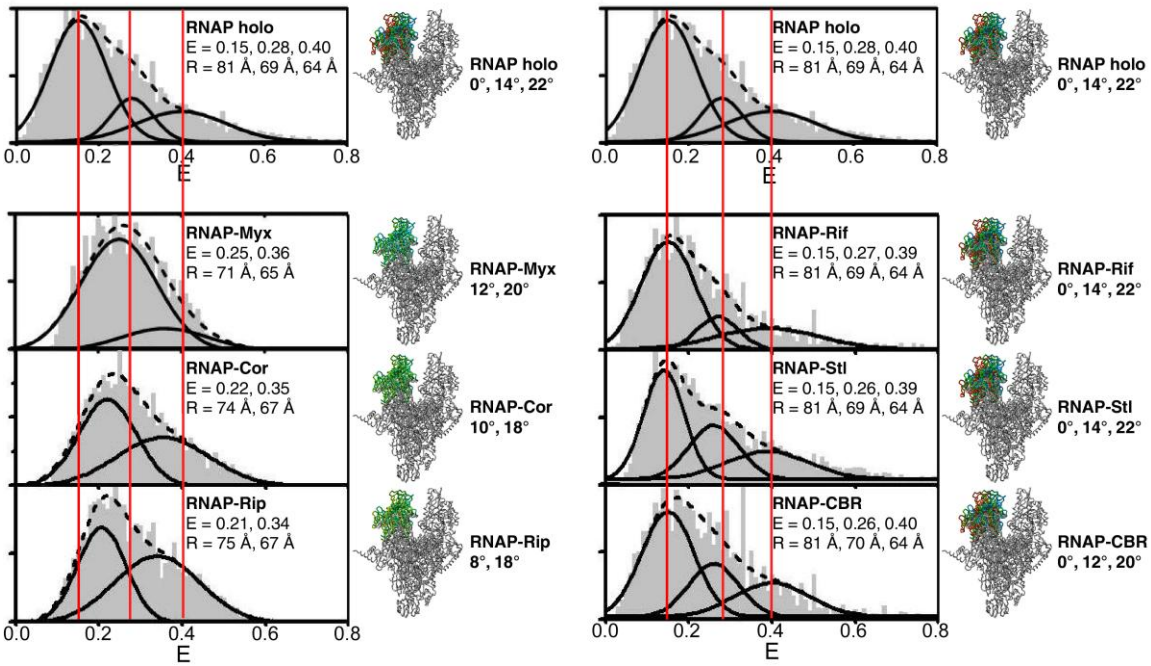


Fig. 24. Effects of inhibitors on RNAP clamp conformation.
(A) Effects of inhibitors that block the isomerization of RPc to RPo: Myx, Cor, Rip, and Gp2.

(B) Effects of inhibitors that block steps subsequent to the isomerization of RPc to RPo: Rif, Stl, and CBR.

4.14. Effect of inhibitors on RNAP clamp conformation: σ^{54} dependent transcription initiation

To assess the effect of the effect of Myx on σ^{54}-dependent intermediates to the formation of RPo, clamp conformation was monitored in RNAP-σ^{54} holoenzyme in presence of Myx and in RPc in presence of Myx (Fig. 25). It was observed that Myx induced closure of the RNAP clamp (clamp rotated inward by 12˚ to 22˚) in both RNAP-σ^{54} holoenzyme. (Fig. 25A). It was observed that in RPc formed byaddition of RNAP-σ^{54}holoenzyme, followed by Myx, followed by DNA, the clamp remained closed (clamp rotated inward by 12˚ to 22˚).

Radiochemical transcription assays showed that Myx inhibits σ^{54}-dependent transcription initiation (Fig. 25B).

Gel mobility shift assays performed using radioactively labelled DNA, show that Myx did not inhibit the formation of RPc (Fig. 25C).

Taken together the results presented in Fig. 26 suggest that Myx inhibits transcription in RNAP-σ^{54} holoenzyme, at least in part by locking the RNAP clamp in a closed conformation. Addition of Myx prior to the addition of DNA does not inhibit the formation of RPc and the clamp – this provides further support to the hypothesis that Myx inhibits a step in transcription following formation of RPc, but prior to the formation of RNAP holoenzyme [29].

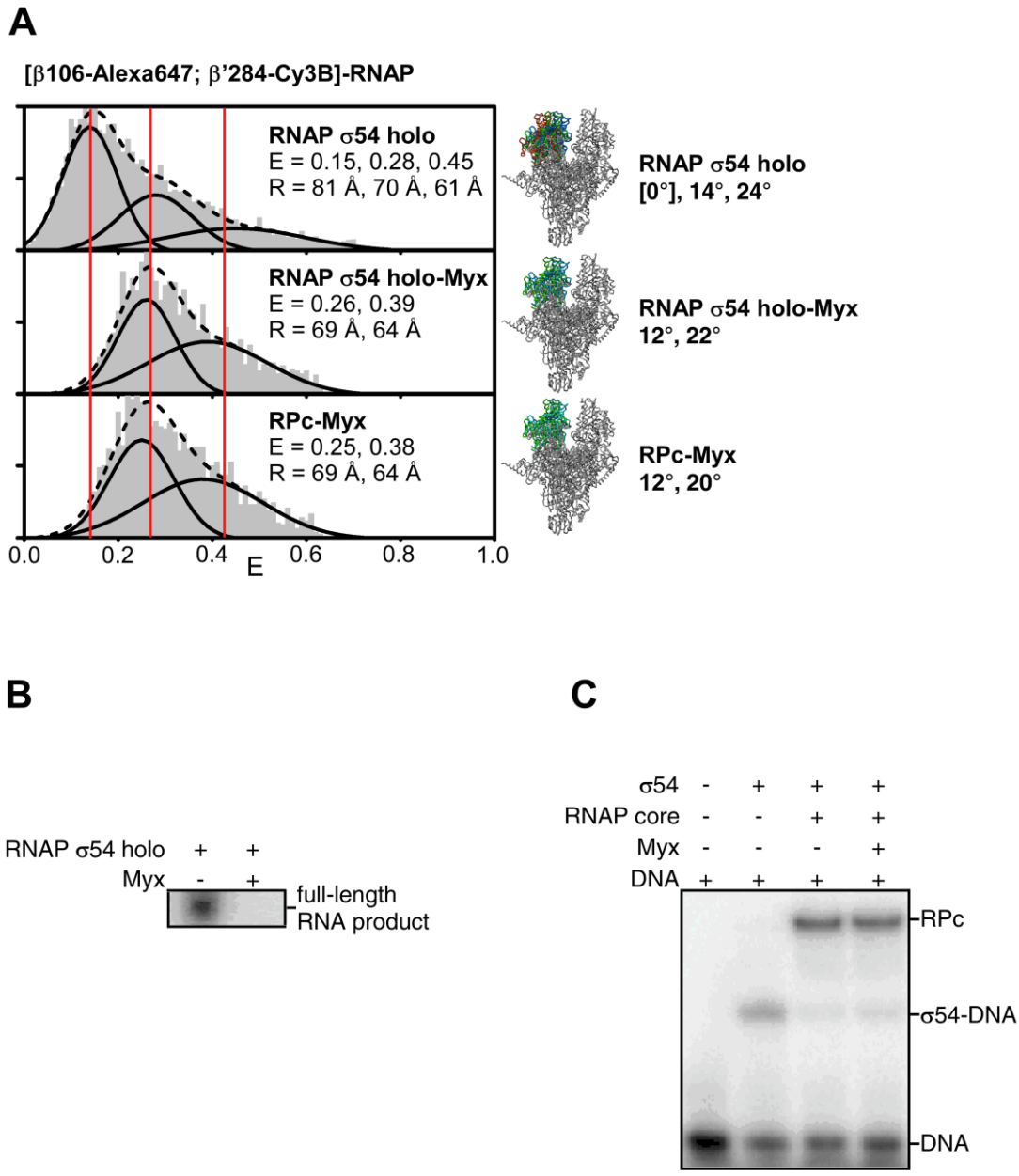


Fig. 25. Effects of inhibitors on RNAP clamp conformation: σ^{54}-dependent transcription initiation.
(A) Myxopyronin (Myx) depopulates the open clamp state in RNAP-σ^{54} holoenzyme and in RPc.
(B) Myx inhibits σ^{54}-dependent transcription.
(C) Myx does not inhibit formation of RPc.

4.15. Effect of transcriptional regulator Gp2 on RNAP clamp conformation.

To assess the effect of T7 Gp2 on RNAP clamp conformation, wild-type T7 Gp2 and a mutant Gp2 (Gp2*) incapable of binding to RNAP were tested (Fig. 26). It was observed that in presence of Gp2, the open clamp state is depopulated and the clamp adopts conformations ranging from closed to collapsed states (clamp rotated inward from 18˚ to 26˚). Analogous experiments with a mutant Gp2 incapable of binding RNAP ([Ala56]Gp2),Gp2*, did not show this effect [82].

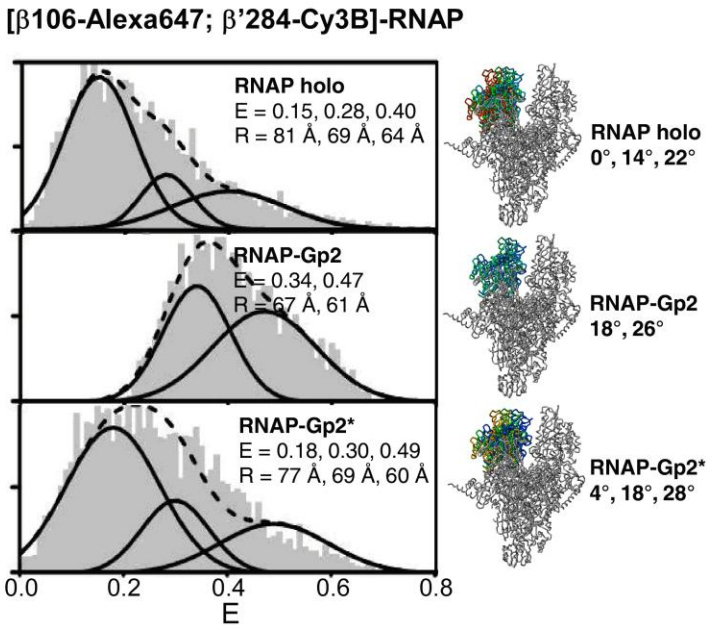


Fig. 26. Effects of inhibitor Gp2 on RNAP clamp conformation.
RNAP clamp conformation in RNAP-σ70 holoenzyme, RNAP-σ^{70}-Gp2 and RNAP-σ^{70}-Gp2*, where Gp2* is mutant defective in binding RNAP ([Ala56]Gp2; Gp2*).

4.16. Effect of hairpin-dependent transcription pause on RNAP clamp conformation.

In order to assess RNAP clamp conformation during transcription pause, clamp conformation was analyzed in RDe and in *his*-hairping-containing RDe. Elongation complexes were formed using RNAP fluorescently labelled at the tips of the two pincers and nucleic acid scaffolds. Nucleic acid scaffolds contained a probe of the third color (Alexa 488) incorporated at the 5' end of the non-template strand of DNA, to avoid complications arising due to unbound RNAP molecules. Three-color ALEX was performed. Data was filtered to consider only populations containing all three fluorescent probes. E distributions for the clamp conformation were plotted. The results show that the mean E in his-hairpin-containing RDe shifted to lower energy transfer efficiency in comparison to RDe, indicating the RNAP adopts a more open conformation (rotated outward by 8˚) during hairpin-dependent his pause.

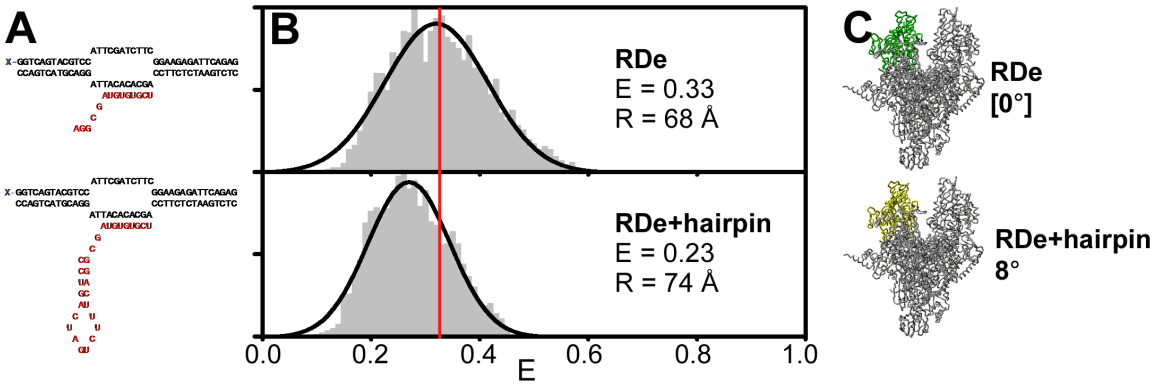


Fig. 27. Effects of inhibitor Gp2 on RNAP clamp conformation.
RNAP clamp conformation in RDe without pause (top panel) and with hairpin-dependent his pause (bottom panel).
(A) Nucleic Acid scaffolds.
(B) E histogram showing clamp conformation.
(C) Interpretation of structural changes of RNAP shown in degrees of rotation.

5. Discussion

Crystal structures have in different crystal contexts indicate that the RNAP clamp can adopt conformational states ranging from an open state to a closed state. Clamp closure involves a swinging motion of the β' pincer, with rotation by ~20º, about a hinge region ("switch region") at the base of β' pincer, and with displacement by ~20Å, of residues at the tip of the β' pincer. It has been hypothesized that different conformational states of the clamp are also present in solution and are important for function.

In this work, a smFRET assay was developed which permits determination of clamp conformational states in solution. The results from this work define the conformational states of the RNAP clamp in solution, and extend the previously reported observations in crystal structures of RNAP.

5.1 RNAP clamp has a predominantly open conformational state in holoenzyme and core

In this work, smFRET analysis revealed that, in freely diffusing molecule of RNAP holoenzyme, the clamp can adopt multiple conformational states. The conformational states can be fitted to three Gaussian distributions corresponding to three states of the RNAP clamp – an open state, a closed state and a collapsed state. The open state is the predominant conformational state in free RNAP holoenzyme. In the open clamp state, the RNAP active center cleft has solvent accessible width of ~20 Å, the dimensions of which

permits unimpeded entry of dsDNA into the active center cleft. The closed state and the collapsed RNAP clamp states exist as minor subpopulations, which do not permit entry of dsDNA into the RNAP active center cleft.

The clamp conformation in RNAP core is essentially identical to RNAP-σ^{70} holoenzyme -- the clamp is predominantly open, with minor subpopulations of closed and collapsed clamp states. This suggests that the clamp conformational dynamics are intrinsic to RNAP core. Although the one or more conformational states of the RNAP clamp may be stabilized by σ^{70}, the core is capable of sampling all the conformational states by itself. This conclusion is further supported by the observation that clamp conformations are essentially identical in RNAP-σ^{70} holoenzyme and RNAP-σ^{54} holoenzyme – undermining the influence of σ factor in clamp conformational dynamics..

5.2. RNAP clamp closes upon formation of a catalytically competent transcription initiation complex and remains closed in subsequent steps in transcription

Previous work done over the past several decades have defined the steps involved in σ^{70}-dependent transcription initiation and elongation, but the role of RNAP clamp in solution at each of these steps in transcription had not been directly ascertained. In this work, smFRET was used to directly measure the distance between the tip of β' clamp and the tip of β pincer at each stage in transcription – namely RPo, RPitc and RDe.

RPo is the first stable catalytically competent σ^{70}-dependent initiation complex that is formed by addition of DNA to RNAP-σ^{70} holoenzyme. In RPo, promoter DNA is loaded into the RNAP active center cleft and ~13 bp of DNA is unwound to form the transcription bubble. In this work, smFRET measurements revealed that the conformational state of RNAP clamp in holoenzyme and RPo are different –in constrast to the the predominantly open clamp state in RNAP holoenzyme, the clamp adopts an exclusively closed conformation in RPo. Based on this observation, it is attractive to propose that, the negatively charged unwound DNA in RPo and positively charged inner facet of the clamp may induce or stabilize the clamp closure.

Upon transition to initial transcribing complexes (RPitc) synthesizing RNA of 2, 4 and 7 nt of length, the clamp continues to remain closed. Initial transcribing complexes mechanistically different from RPo. At this stage, RNAP employs a scrunching mechanism of RNA synthesis, where RNAP remains stationary relative to DNA and pulls downstream DNA within itself to synthesize abortive RNA products. Observation of closed clamp state in RPitc is consistent with the proposal that, during abortive RNA synthesis, there is insufficient space in the RNAP active center cleft to accommodate the additional bp of scrunched DNA and some of the additional bp DNA have to emerge out into solution.

Upon transition to transcription elongation complex (RDe), in which RNAP translocates relative to DNA using a stepping mechanism, no change in RNAP clamp conformation is observed – the clamp continues to remain closed. Equivalent results were obtained upon

analysis of RDe14, in which RNA of 14 nt is produced which completely fills the RNA exit channel and RDe15 in which, RNA of 15 nt is produced which extends beyond the RNAP exit channel into solution. Equivalent results were also obtained upon analysis of early elongation complexes containing σ^{70} and early elongation complexes lacking σ^{70}, indicating σ^{70} does not influence clamp conformation in early elongation complexes. Furthermore, the early elongation complexes lacking σ^{70} are possible structural mimics of mature σ^{70}-free elongation complexes, suggesting mature elongation complexes probably function with the clamp in the closed state.

Taken together, observations of closed clamp state in RRo, RPitc and RDe are in agreement with the proposal that clamp closure accounts for the high stability of transcription initiation complexes and high stability and processivity of transcription elongation complexes.

5.3. RNAP clamp remains open in all steps of initiation prior to formation of the catalytically competent initiation complex

Previous work has shown, in contrast to RNAPσ^{70} holoenzyme which forms unstable RPc, RNAP-σ^{54} holoenzyme forms stable RPc upon addition of promoter DNA. σ^{54}-dependent RPc isomerizes to RPo only in the presence of an AAA+ ATPase such as NtrC1, and ATP hydrolysis. In this work, clamp conformation was defined in intermediates to the formation of RPo. It was observed that starting with RPc, all intermediates leading up to the formation of RPo had predominantly open clamp states

and only upon formation of RPo, the clamp closed. This observation further supports the proposal that upon loading of DNA into the RNAP, and/or unwinding of DNA in, the RNAP active center cleft, the clamp closes. These observations further support the proposal that electrostatic interactions between the negatively charge DNA in the RNAP active center cleft and positively charged inner facet of the clamp induce or stabilize clamp closure.

5.4. Switch-region-target-inhibitors depopulate the clamp open state

Previous work has shown that RNAP switch region is a target for inhibitors such as myxopyronin (Myx), corallopyronin (Cor), and ripostatin (Rip). Myx, Cor and Rip have been shown to inhibit a late step in isomerization of RPc to RPo. It has been proposed that these switch-region-target-inhibitors function by a "hinge-jamming" mechanism in which they lock the clamp in a conformation that may not permit entry of DNA in the active center cleft.

In this work, the hinge jamming hypothesis was directly tested. Consistent with the hinge-jamming hypothesis, smFRET measurements showed that, in presence of Myx, Cor and Rip, the clamp adopts conformational states ranging from closed to collapsed – the dimensions of which do not permit entry of dsDNA into the active center cleft. The results further indicate that, compounds which depopulate the open clamp state also inhibit transcription -- suggesting the open clamp state is an important state and perhaps an obligatory state to the formation of RPo. Additionally, results from this work show

that Myx can depopulate the open clamp state, not only in σ^{70}-containing RNAP holoenzyme, but also in σ^{54}-containing holoenzyme.

5.4. RNAP clamp re-opens in response to his-hairpin-dependent pause.

Based on results obtained in this work, it can be suggested that clamp closure is important for stability and processivity and stability of transcription elongation complexes. Ensuing this proposal, it is plausible that the clamp may re-open in states marked by low processivity such as transcription pausing and transcription termination.

Hairpin-dependent pauses are characteristic of class I pauses and transcription termination. Previous work has shown that RNAP leader region of his biosynthetic operon contains a hairpin dependent pause site [127]. It has been suggested that pausing of RNAP at this region permits ribosomes to synchronize translation with transcription. This synchronization is important for regulation of his operon through the attenuation mechanism [128].

Results in Fig. 27 show that the clamp indeed opens during hairpin-dependent-transcription pausing. The results are consistent with the proposal that his-pause hairpin may partly penetrate the exit channel and sterically clash with the switch region to open the clamp [24, 128-130].

References

1. Sweetser, D., M. Nonet, and R.A. Young, *Prokaryotic and eukaryotic RNA polymerases have homologous core subunits.* Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(5): p. 1192-6.

2. Ebright, R.H., *RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II.* Journal of molecular biology, 2000. **304**(5): p. 687-98.

3. Werner, F. and D. Grohmann, *Evolution of multisubunit RNA polymerases in the three domains of life.* Nature reviews. Microbiology, 2011. **9**(2): p. 85-98.

4. Darst, S.A., *Bacterial RNA polymerase.* Current opinion in structural biology, 2001. **11**(2): p. 155-62.

5. Cramer, P., *Multisubunit RNA polymerases.* Current opinion in structural biology, 2002. **12**(1): p. 89-97.

6. Cramer, P., K.J. Armache, S. Baumli, S. Benkert, F. Brueckner, C. Buchen, G.E. Damsma, S. Dengl, S.R. Geiger, A.J. Jasiak, A. Jawhari, S. Jennebach, T. Kamenski, H. Kettenberger, C.D. Kuhn, E. Lehmann, K. Leike, J.F. Sydow, and A. Vannini, *Structure of eukaryotic RNA polymerases.* Annual review of biophysics, 2008. **37**: p. 337-52.

7. Wild, T. and P. Cramer, *Biogenesis of multisubunit RNA polymerases.* Trends in biochemical sciences, 2012. **37**(3): p. 99-105.

8. Burgess, R.R., A.A. Travers, J.J. Dunn, and E.K. Bautz, *Factor stimulating transcription by RNA polymerase.* Nature, 1969. **221**(5175): p. 43-6.

9. Travers, A.A. and Burgessrr, *Cyclic re-use of the RNA polymerase sigma factor.* Nature, 1969. **222**(5193): p. 537-40.

10. Vassylyev, D.G., S. Sekine, O. Laptenko, J. Lee, M.N. Vassylyeva, S. Borukhov, and S. Yokoyama, *Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution.* Nature, 2002. **417**(6890): p. 712-9.

11. Zhang, G., E.A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S.A. Darst, *Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution.* Cell, 1999. **98**(6): p. 811-24.

12. Darst, S.A., N. Opalka, P. Chacon, A. Polyakov, C. Richter, G. Zhang, and W. Wriggers, *Conformational flexibility of bacterial RNA polymerase.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4296-301.

13. Borukhov, S. and E. Nudler, *RNA polymerase holoenzyme: structure, function and biological implications.* Current opinion in microbiology, 2003. **6**(2): p. 93-100.

14. Murakami, K.S. and S.A. Darst, *Bacterial RNA polymerases: the wholo story.* Current opinion in structural biology, 2003. **13**(1): p. 31-9.

15. Steitz, T.A., *A mechanism for all polymerases.* Nature, 1998. **391**(6664): p. 231-2.

16. Zenkin, N., Y. Yuzenkova, and K. Severinov, *Transcript-assisted transcriptional proofreading.* Science, 2006. **313**(5786): p. 518-20.

17. Borukhov, S., V. Sagitov, and A. Goldfarb, *Transcript cleavage factors from E. coli.* Cell, 1993. **72**(3): p. 459-66.

18. Mukhopadhyay, J., E. Sineva, J. Knight, R.M. Levy, and R.H. Ebright, *Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel.* Mol Cell, 2004. **14**(6): p. 739-51.

19. Kettenberger, H., K.J. Armache, and P. Cramer, *Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS.* Molecular cell, 2004. **16**(6): p. 955-65.

20. Nickels, B.E. and A. Hochschild, *Regulation of RNA polymerase through the secondary channel.* Cell, 2004. **118**(3): p. 281-4.

21. Vassylyev, D.G., M.N. Vassylyeva, A. Perederina, T.H. Tahirov, and I. Artsimovitch, *Structural basis for transcription elongation by bacterial RNA polymerase.* Nature, 2007. **448**(7150): p. 157-62.

22. Komissarova, N. and M. Kashlev, *Functional topography of nascent RNA in elongation intermediates of RNA polymerase.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(25): p. 14699-704.

23. Cramer, P., D.A. Bushnell, and R.D. Kornberg, *Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution.* Science, 2001. **292**(5523): p. 1863-76.

24. Landick, R., *RNA polymerase clamps down.* Cell, 2001. **105**(5): p. 567-70.

25. Korzheva, N., A. Mustaev, M. Kozlov, A. Malhotra, V. Nikiforov, A. Goldfarb, and S.A. Darst, *A structural model of transcription elongation.* Science, 2000. **289**(5479): p. 619-25.

26. Mekler, V., L. Minakhin, and K. Severinov, *A critical role of downstream RNA polymerase-promoter interactions in the formation of initiation complex.* The Journal of biological chemistry, 2011. **286**(25): p. 22600-8.

27. Brodolin, K., N. Zenkin, and K. Severinov, *Remodeling of the sigma70 subunit non-template DNA strand contacts during the final step of transcription initiation.* Journal of molecular biology, 2005. **350**(5): p. 930-7.

28. Vassylyev, D.G., M.N. Vassylyeva, J. Zhang, M. Palangat, I. Artsimovitch, and R. Landick, *Structural basis for substrate loading in bacterial RNA polymerase.* Nature, 2007. **448**(7150): p. 163-8.

29. Mukhopadhyay, J., K. Das, S. Ismail, D. Koppstein, M. Jang, B. Hudson, S. Sarafianos, S. Tuske, J. Patel, R. Jansen, H. Irschik, E. Arnold, and R.H. Ebright, *The RNA polymerase "switch region" is a target for inhibitors.* Cell, 2008. **135**(2): p. 295-307.

30. Srivastava, A., M. Talaue, S. Liu, D. Degen, R.Y. Ebright, E. Sineva, A. Chakraborty, S.Y. Druzhinin, S. Chatterjee, J. Mukhopadhyay, Y.W. Ebright, A. Zozula, J. Shen, S. Sengupta, R.R. Niedfeldt, C. Xin, T. Kaneko, H. Irschik, R. Jansen, S. Donadio, N. Connell, and R.H. Ebright, *New target for inhibition of bacterial RNA polymerase: 'switch region'.* Current opinion in microbiology, 2011. **14**(5): p. 532-43.

31. Lara-Gonzalez, S., J.J. Birktoft, and C.L. Lawson, *Structure of the Escherichia coli RNA polymerase alpha subunit C-terminal domain.* Acta crystallographica. Section D, Biological crystallography, 2010. **66**(Pt 7): p. 806-12.

32. Ebright, R.H. and S. Busby, *The Escherichia coli RNA polymerase alpha subunit: structure and function.* Current opinion in genetics & development, 1995. **5**(2): p. 197-203.

33. Lawson, C.L., D. Swigon, K.S. Murakami, S.A. Darst, H.M. Berman, and R.H. Ebright, *Catabolite activator protein: DNA binding and transcription activation.* Current opinion in structural biology, 2004. **14**(1): p. 10-20.

34. Mathew, R. and D. Chatterji, *The evolving story of the omega subunit of bacterial RNA polymerase.* Trends in microbiology, 2006. **14**(10): p. 450-5.

35. Minakhin, L., S. Bhagat, A. Brunning, E.A. Campbell, S.A. Darst, R.H. Ebright, and K. Severinov, *Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(3): p. 892-7.

36. Vrentas, C.E., T. Gaal, W. Ross, R.H. Ebright, and R.L. Gourse, *Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA.* Genes & development, 2005. **19**(19): p. 2378-87.

37. Borukhov, S. and K. Severinov, *Role of the RNA polymerase sigma subunit in transcription initiation.* Research in microbiology, 2002. **153**(9): p. 557-62.

38. Hsu, L.M., *Promoter clearance and escape in prokaryotes.* Biochimica et biophysica acta, 2002. **1577**(2): p. 191-207.

39. Perdue, S.A. and J.W. Roberts, *Sigma(70)-dependent transcription pausing in Escherichia coli.* Journal of molecular biology, 2011. **412**(5): p. 782-92.

40. Jain, D., B.E. Nickels, L. Sun, A. Hochschild, and S.A. Darst, *Structure of a ternary transcription activation complex.* Molecular cell, 2004. **13**(1): p. 45-53.

41. Lee, D.J., S.D. Minchin, and S.J. Busby, *Activating transcription in bacteria.* Annual review of microbiology, 2012. **66**: p. 125-52.

42. Campbell, E.A., L.F. Westblade, and S.A. Darst, *Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective.* Current opinion in microbiology, 2008. **11**(2): p. 121-7.

43. Young, B.A., T.M. Gruber, and C.A. Gross, *Views of transcription initiation.* Cell, 2002. **109**(4): p. 417-20.

44. Mekler, V., E. Kortkhonjia, J. Mukhopadhyay, J. Knight, A. Revyakin, A.N. Kapanidis, W. Niu, Y.W. Ebright, R. Levy, and R.H. Ebright, *Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex.* Cell, 2002. **108**(5): p. 599-614.

45. Murakami, K.S., S. Masuda, E.A. Campbell, O. Muzzin, and S.A. Darst, *Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex.* Science, 2002. **296**(5571): p. 1285-90.

46. Murakami, K.S., S. Masuda, and S.A. Darst, *Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution.* Science, 2002. **296**(5571): p. 1280-4.

47. Sharp, M.M., C.L. Chan, C.Z. Lu, M.T. Marr, S. Nechaev, E.W. Merritt, K. Severinov, J.W. Roberts, and C.A. Gross, *The interface of sigma with core RNA polymerase is extensive, conserved, and functionally specialized.* Genes & development, 1999. **13**(22): p. 3015-26.

48. Gross, C.A., C. Chan, A. Dombroski, T. Gruber, M. Sharp, J. Tupy, and B. Young, *The functional and regulatory roles of sigma factors in transcription.* Cold Spring Harbor symposia on quantitative biology, 1998. **63**: p. 141-55.

49. Paget, M.S. and J.D. Helmann, *The sigma70 family of sigma factors.* Genome Biol, 2003. **4**(1): p. 203.

50. Lonetto, M., M. Gribskov, and C.A. Gross, *The sigma 70 family: sequence conservation and evolutionary relationships.* Journal of bacteriology, 1992. **174**(12): p. 3843-9.

51. Campbell, E.A., O. Muzzin, M. Chlenov, J.L. Sun, C.A. Olson, O. Weinman, M.L. Trester-Zedlitz, and S.A. Darst, *Structure of the bacterial RNA polymerase promoter specificity sigma subunit.* Molecular cell, 2002. **9**(3): p. 527-39.

52. Malhotra, A., E. Severinova, and S.A. Darst, *Crystal structure of a sigma 70 subunit fragment from E. coli RNA polymerase.* Cell, 1996. **87**(1): p. 127-36.

53. Gruber, T.M. and C.A. Gross, *Multiple sigma subunits and the partitioning of bacterial transcription space.* Annual review of microbiology, 2003. **57**: p. 441-66.

54. Maeda, H., N. Fujita, and A. Ishihama, *Competition among seven Escherichia coli sigma subunits: relative binding affinities to the core RNA polymerase.* Nucleic acids research, 2000. **28**(18): p. 3497-503.

55. Wigneshweraraj, S., D. Bose, P.C. Burrows, N. Joly, J. Schumacher, M. Rappas, T. Pape, X. Zhang, P. Stockley, K. Severinov, and M. Buck, *Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor.* Molecular microbiology, 2008. **68**(3): p. 538-46.

56. Kim, Y., L.S. Watrud, and A. Matin, *A carbon starvation survival gene of Pseudomonas putida is regulated by sigma 54.* Journal of bacteriology, 1995. **177**(7): p. 1850-9.

57. Morett, E. and M. Buck, *In vivo studies on the interaction of RNA polymerase-sigma 54 with the Klebsiella pneumoniae and Rhizobium meliloti nifH promoters. The role of NifA in the formation of an open promoter complex.* Journal of molecular biology, 1989. **210**(1): p. 65-77.

58. Kimbara, K. and A.M. Chakrabarty, *Control of alginate synthesis in Pseudomonas aeruginosa: regulation of the algR1 gene.* Biochemical and biophysical research communications, 1989. **164**(2): p. 601-8.

59. Alm, R.A., P. Guerry, and T.J. Trust, *The Campylobacter sigma 54 flaB flagellin promoter is subject to environmental regulation.* Journal of bacteriology, 1993. **175**(14): p. 4448-55.

60. Buck, M., M.T. Gallegos, D.J. Studholme, Y. Guo, and J.D. Gralla, *The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor.* Journal of bacteriology, 2000. **182**(15): p. 4129-36.

61. Cannon, W.V., M.T. Gallegos, and M. Buck, *Isomerization of a binary sigma-promoter DNA complex by transcription activators.* Nat Struct Biol, 2000. **7**(7): p. 594-601.

62. Ghosh, T., D. Bose, and X. Zhang, *Mechanisms for activating bacterial RNA polymerase.* FEMS Microbiol Rev, 2010. **34**(5): p. 611-27.

63. Osterberg, S., T. del Peso-Santos, and V. Shingler, *Regulation of alternative sigma factor use.* Annual review of microbiology, 2011. **65**: p. 37-55.

64. Bose, D., T. Pape, P.C. Burrows, M. Rappas, S.R. Wigneshweraraj, M. Buck, and X. Zhang, *Organization of an activator-bound RNA polymerase holoenzyme.* Molecular cell, 2008. **32**(3): p. 337-46.

65. Bordes, P., S.R. Wigneshweraraj, J. Schumacher, X. Zhang, M. Chaney, and M. Buck, *The ATP hydrolyzing transcription activator phage shock protein F of Escherichia coli: identifying a surface that binds sigma 54.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(5): p. 2278-83.

66. Cannon, W.V., M.K. Chaney, X. Wang, and M. Buck, *Two domains within sigmaN (sigma54) cooperate for DNA binding.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(10): p. 5006-11.

67. Saecker, R.M., M.T. Record, Jr., and P.L. Dehaseth, *Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis.* Journal of molecular biology, 2011. **412**(5): p. 754-71.

68. Kapanidis, A.N., E. Margeat, T.A. Laurence, S. Doose, S.O. Ho, J. Mukhopadhyay, E. Kortkhonjia, V. Mekler, R.H. Ebright, and S. Weiss, *Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis.* Mol Cell, 2005. **20**(3): p. 347-56.

69. Mooney, R.A., S.A. Darst, and R. Landick, *Sigma and RNA polymerase: an on-again, off-again relationship?* Molecular cell, 2005. **20**(3): p. 335-45.

70. Mukhopadhyay, J., A.N. Kapanidis, V. Mekler, E. Kortkhonjia, Y.W. Ebright, and R.H. Ebright, *Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA.* Cell, 2001. **106**(4): p. 453-63.

71. Abbondanzieri, E.A., W.J. Greenleaf, J.W. Shaevitz, R. Landick, and S.M. Block, *Direct observation of base-pair stepping by RNA polymerase.* Nature, 2005. **438**(7067): p. 460-5.

72. Gnatt, A.L., P. Cramer, J. Fu, D.A. Bushnell, and R.D. Kornberg, *Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution.* Science, 2001. **292**(5523): p. 1876-82.

73. Tagami, S., S. Sekine, T. Kumarevel, N. Hino, Y. Murayama, S. Kamegamori, M. Yamamoto, K. Sakamoto, and S. Yokoyama, *Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein.* Nature, 2010. **468**(7326): p. 978-82.

74. Belogurov, G.A., M.N. Vassylyeva, A. Sevostyanova, J.R. Appleman, A.X. Xiang, R. Lira, S.E. Webber, S. Klyuyev, E. Nudler, I. Artsimovitch, and D.G. Vassylyev, *Transcription inactivation through local refolding of the RNA polymerase structure.* Nature, 2009. **457**(7227): p. 332-5.

75. Tupin, A., M. Gualtieri, J.P. Leonetti, and K. Brodolin, *The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site.* The EMBO journal, 2010. **29**(15): p. 2527-37.

76. Haebich, D. and F. von Nussbaum, *Lost in transcription--inhibition of RNA polymerase.* Angewandte Chemie, 2009. **48**(19): p. 3397-400.

77. Kohl, W., H. Irschik, H. Reichenbach, and G. Hofle, *Antibiotics from Gliding Bacteria .22. The Biosynthesis of Myxopyronin-a, an Antibiotic from Myxococcus-Fulvus, Strain Mx-F50.* Liebigs Annalen Der Chemie, 1984(6): p. 1088-1093.

78. Irschik, H., K. Gerth, G. Hofle, W. Kohl, and H. Reichenbach, *The myxopyronins, new inhibitors of bacterial RNA synthesis from Myxococcus fulvus (Myxobacterales).* The Journal of antibiotics, 1983. **36**(12): p. 1651-8.

79. Irschik, H., R. Jansen, G. Hofle, K. Gerth, and H. Reichenbach, *The corallopyronins, new inhibitors of bacterial RNA synthesis from Myxobacteria.* The Journal of antibiotics, 1985. **38**(2): p. 145-52.

80. Irschik, H., H. Augustiniak, K. Gerth, G. Hofle, and H. Reichenbach, *The ripostatins, novel inhibitors of eubacterial RNA polymerase isolated from myxobacteria.* The Journal of antibiotics, 1995. **48**(8): p. 787-92.

81. James, E., M. Liu, C. Sheppard, V. Mekler, B. Camara, B. Liu, P. Simpson, E. Cota, K. Severinov, S. Matthews, and S. Wigneshweraraj, *Structural and mechanistic basis for the inhibition of Escherichia coli RNA polymerase by T7 Gp2.* Molecular cell, 2012. **47**(5): p. 755-66.

82. Camara, B., M. Liu, J. Reynolds, A. Shadrin, B. Liu, K. Kwok, P. Simpson, R. Weinzierl, K. Severinov, E. Cota, S. Matthews, and S.R. Wigneshweraraj, *T7 phage protein Gp2 inhibits the Escherichia coli RNA polymerase by antagonizing stable DNA strand separation near the transcription start site.* Proc Natl Acad Sci U S A, 2010. **107**(5): p. 2247-52.

83. Mekler, V., L. Minakhin, C. Sheppard, S. Wigneshweraraj, and K. Severinov, *Molecular Mechanism of Transcription Inhibition by Phage T7 gp2 Protein.* Journal of molecular biology, 2011. **413**(5): p. 1016-27.

84. Forster, T., *Energiewanderung und Fluoreszenz.* Naturwissenschaften, 1946. **6**: p. 166-75.

85. Forster, T., *Fluoreszenzversuche an FarbstoVmischungen.* Angew Chem, 1947. **59**: p. 181-7.

86. Forster, T., *Zwischenmolekulare Energiewanderung und Fluoreszenz.* Ann. Phys., 1948. **2**(55-75).

87. Selvin, P.R., *The renaissance of fluorescence resonance energy transfer.* Nature structural biology, 2000. **7**(9): p. 730-4.

88. Clegg, R.M., *Fluorescence resonance energy transfer and nucleic acids.* Methods in enzymology, 1992. **211**: p. 353-88.

89. Deniz, A.A., M. Dahan, J.R. Grunwell, T. Ha, A.E. Faulhaber, D.S. Chemla, S. Weiss, and P.G. Schultz, *Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Forster distance dependence and subpopulations.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(7): p. 3670-5.

90. Best, R.B., K.A. Merchant, I.V. Gopich, B. Schuler, A. Bax, and W.A. Eaton, *Effect of flexibility and cis residues in single-molecule FRET studies of polyproline.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(48): p. 18964-9.

91. Merchant, K.A., R.B. Best, J.M. Louis, I.V. Gopich, and W.A. Eaton, *Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(5): p. 1528-33.

92. Schuler, B. and W.A. Eaton, *Protein folding studied by single-molecule FRET.* Current opinion in structural biology, 2008. **18**(1): p. 16-26.

93. Dahan, M., A.A. Deniz, T.J. Ha, D.S. Chemla, P.G. Schultz, and S. Weiss, *Ratiometric measurement and identification of single diffusing molecules.* Chemical Physics, 1999. **247**(1): p. 85-106.

94. Stryer, L. and R.P. Haugland, *Energy transfer: a spectroscopic ruler.* Proceedings of the National Academy of Sciences of the United States of America, 1967. **58**(2): p. 719-26.

95. Kapanidis, A.N., N.K. Lee, T.A. Laurence, S. Doose, E. Margeat, and S. Weiss, *Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules.* Proc Natl Acad Sci U S A, 2004. **101**(24): p. 8936-41.

96. Kapanidis, A.N., E. Margeat, S.O. Ho, E. Kortkhonjia, S. Weiss, and R.H. Ebright, *Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism.* Science, 2006. **314**(5802): p. 1144-7.

97. Lee, N.K., A.N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R.H. Ebright, and S. Weiss, *Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation.* Biophys J, 2005. **88**(4): p. 2939-53.

98. Kapanidis, A.N., T.A. Laurence, N.K. Lee, E. Margeat, X. Kong, and S. Weiss, *Alternating-laser excitation of single molecules.* Accounts of chemical research, 2005. **38**(7): p. 523-33.

99. Chin, J.W., S.W. Santoro, A.B. Martin, D.S. King, L. Wang, and P.G. Schultz, *Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli.* Journal of the American Chemical Society, 2002. **124**(31): p. 9026-7.

100. Chakraborty, A., D. Wang, Y.W. Ebright, and R.H. Ebright, *Azide-specific labeling of biomolecules by Staudinger-Bertozzi ligation phosphine derivatives of fluorescent probes suitable for single-molecule fluorescence spectroscopy.* Methods Enzymol, 2010. **472**: p. 19-30.

101. Saxon, E. and C.R. Bertozzi, *Cell surface engineering by a modified Staudinger reaction.* Science, 2000. **287**(5460): p. 2007-10.

102. Naryshkin, N., Y. Kim, Q. Dong, and R.H. Ebright, *Site-specific protein-DNA photocrosslinking. Analysis of bacterial transcription initiation complexes.* Methods Mol Biol, 2001. **148**: p. 337-61.

103. Tang, H., Y. Kim, K. Severinov, A. Goldfarb, and R.H. Ebright, *Escherichia coli RNA polymerase holoenzyme: rapid reconstitution from recombinant alpha, beta, beta', and sigma subunits.* Methods in enzymology, 1996. **273**: p. 130-4.

104. Wang, D., T.I. Meier, C.L. Chan, G. Feng, D.N. Lee, and R. Landick, *Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex.* Cell, 1995. **81**(3): p. 341-50.

105. Severinov, K., R. Mooney, S.A. Darst, and R. Landick, *Tethering of the large subunits of Escherichia coli RNA polymerase.* J Biol Chem, 1997. **272**(39): p. 24137-40.

106. Niu, W., *Identification and characterization of interactions between a transcription activator and the transcription machinery*, 1999, Rutgers University New Brunswick, NJ.

107. Niu, W., Y. Kim, G. Tau, T. Heyduk, and R.H. Ebright, *Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase.* Cell, 1996. **87**(6): p. 1123-34.

108. Mukhopadhyay, J., V. Mekler, E. Kortkhonjia, A.N. Kapanidis, Y.W. Ebright, and R.H. Ebright, *Fluorescence resonance energy transfer (FRET) in analysis of transcription-complex structure and function.* Methods Enzymol, 2003. **371**: p. 144-59.

109. Sambrook, J. and D.W. Russell, *Molecular cloning : a laboratory manual*. 3rd ed2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

110. Rappas, M., J. Schumacher, F. Beuron, H. Niwa, P. Bordes, S. Wigneshweraraj, C.A. Keetch, C.V. Robinson, M. Buck, and X. Zhang, *Structural insights into the activity of enhancer-binding proteins.* Science, 2005. **307**(5717): p. 1972-5.

111. Chen, B., T.A. Sysoeva, S. Chowdhury, L. Guo, S. De Carlo, J.A. Hanson, H. Yang, and B.T. Nixon, *Engagement of arginine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase.* Structure, 2010. **18**(11): p. 1420-30.

112. Wigneshweraraj, S.R., S. Nechaev, P. Bordes, S. Jones, W. Cannon, K. Severinov, and M. Buck, *Enhancer-dependent transcription by bacterial RNA polymerase: the beta subunit downstream lobe is used by sigma 54 during open promoter complex formation.* Methods Enzymol, 2003. **370**: p. 646-57.

113. Burrows, P.C., N. Joly, and M. Buck, *A prehydrolysis state of an AAA+ ATPase supports transcription activation of an enhancer-dependent RNA polymerase.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(20): p. 9376-81.

114. Lee, N.K., A.N. Kapanidis, H.R. Koh, Y. Korlann, S.O. Ho, Y. Kim, N. Gassman, S.K. Kim, and S. Weiss, *Three-color alternating-laser excitation of single*

molecules: monitoring multiple interactions and distances. Biophys J, 2007. **92**(1): p. 303-12.

115. Wu, P. and L. Brand, *Orientation factor in steady-state and time-resolved resonance energy transfer measurements.* Biochemistry, 1992. **31**(34): p. 7939-47.

116. Knight, J.L., V. Mekler, J. Mukhopadhyay, R.H. Ebright, and R.M. Levy, *Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability.* Biophys J, 2005. **88**(2): p. 925-38.

117. Brooks, B.R., C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus, *CHARMM: The Biomolecular Simulation Program.* J Comput Chem, 2009. **30**(10): p. 1545-1614.

118. Feig, M., J. Karanicolas, and C. Brooks, *MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology*. J Mol Graph Model, 2004. **22**(5): p. 377-395.

119. Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A.D. Mackerell, *CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields.* J Comput Chem, 2010. **31**(4): p. 671-690.

120. Yesselman, J.D., D.J. Price, J.L. Knight, and C.L. Brooks, 3rd, *MATCH: An atom-typing toolset for molecular mechanics force fields.* Journal of computational chemistry, 2012. **33**(2): p. 189-202.

121. Schuler, B., E.A. Lipman, P.J. Steinbach, M. Kumke, and W.A. Eaton, *Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence.* Proc Natl Acad Sci U S A, 2005. **102**(8): p. 2754-9.

122. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nature methods, 2008. **5**(6): p. 507-16.

123. Cooper, M., A. Ebner, M. Briggs, M. Burrows, N. Gardner, R. Richardson, and R. West, *Cy3B: improving the performance of cyanine dyes.* J Fluoresc, 2004. **14**(2): p. 145-50.

124. Panchuk-Voloshina, N., R.P. Haugland, J. Bishop-Stewart, M.K. Bhalgat, P.J. Millard, F. Mao, W.Y. Leung, and R.P. Haugland, *Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates.* J Histochem Cytochem, 1999. **47**(9): p. 1179-88.

125. Leung, W.Y., Cheung, C.Y., Yue,S., *Modified carbocyanine dyes and their conjugates*, U. patent, Editor 2005. p. 305.

126. Nickels, B.E., J. Mukhopadhyay, S.J. Garrity, R.H. Ebright, and A. Hochschild, *The sigma 70 subunit of RNA polymerase mediates a promoter-proximal pause at the lac promoter.* Nat Struct Mol Biol, 2004. **11**(6): p. 544-50.

127. Chan, C.L. and R. Landick, *Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin.* Journal of molecular biology, 1993. **233**(1): p. 25-42.

128. Toulokhonov, I., I. Artsimovitch, and R. Landick, *Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins.* Science, 2001. **292**(5517): p. 730-3.

129. Toulokhonov, I. and R. Landick, *The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination.* Molecular cell, 2003. **12**(5): p. 1125-36.

130. Toulokhonov, I., J. Zhang, M. Palangat, and R. Landick, *A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing.* Molecular cell, 2007. **27**(3): p. 406-19.