CHEMICAL REVIEWS

Review

pubs.acs.org/CR

Molecular Mechanisms of Transcription through Single-Molecule Experiments

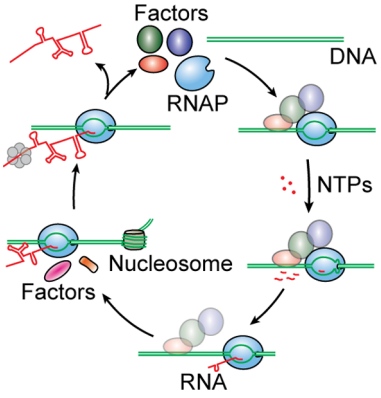
Manchuta Dangkulwanich,[†,‡] Toyotaka Ishibashi,[†,§,▽] Lacramioara Bintu,[†,∥,○] and Carlos Bustamante*[,†,‡,§,∥,⊥,#]

[†]Jason L. Choy Laboratory of Single-Molecule Biophysics, [‡]Department of Chemistry, [§]California Institute for Quantitative Biosciences, [∥]Department of Physics, and [⊥]Department of Molecular and Cell Biology, Howard Hughes Medical Institute, and Kavli Energy NanoSciences Institute, University of California, Berkeley, Berkeley, California 94720, United States

[#]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[▽]Division of Life Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR

[○]Department of Bioengineering, California Institute of Technology, Pasadena, California 91125, United States



CONTENTS

1. INTRODUCTION

Transcription represents the first step in gene expression. It is therefore not surprising that transcription is a highly regulated process and its control is essential to understand the flow and processing of information required by the cell to maintain its homeostasis. During transcription, a DNA molecule is copied into RNA molecules that are then used to translate the genetic information into proteins; this logical pattern has been conserved throughout all three kingdoms of life, from Archaea to Eukarya, making it an essential and fundamental cellular process. Even though some viruses that encode their genome in an RNA molecule use it as a template to make mRNA, others synthesize an intermediate DNA molecule from the RNA, a process known as reverse transcription, from which regular transcription of viral genes can then proceed in the host cells.[1]

Why has transcription evolved into such an essential cellular process? Why not directly express the information encoded in the DNA genome into proteins? There are several reasons to justify the evolution of transcription as an intermediate step for the synthesis of proteins. First, transcription expands the variety of gene products by allowing for splicing. Second, copying the information within DNA into many RNA molecules increases the rate of total protein synthesis in the cell and avoids the bottleneck that would result from expression of a gene directly from the DNA. Third, the number of RNA molecules available at any given time to synthesize proteins can be precisely regulated to give a burst of products. The signal amplification implicit in the transcription process increases the dynamic range of the expression, allowing the cell to control its RNA throughput with higher precision and in a gene-specific manner. This amplification also gives rise to stochasticity in gene

Special Issue: 2014 Single Molecule Mechanisms and Imaging

Received: December 23, 2013
Published: February 6, 2014

   dx.doi.org/10.1021/cr400730x | Chem. Rev. 2014, 114, 3203−3223

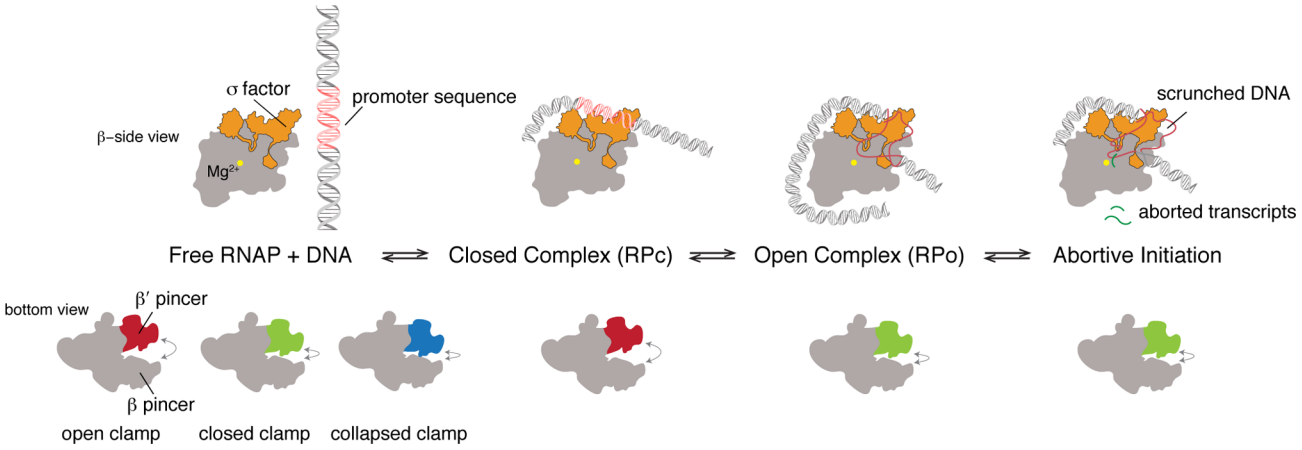


Figure 1. Summary of steps in transcription initiation for prokaryotic RNAP. The clamp of free RNAP exists in three different conformations: open (red), closed (green), and collapsed (blue) states. The RNA holoenzyme interacts with the promoter via the σ factor to form the closed promoter complex; only the open clamp state was observed in this state. In the open promoter complex, the RNAP wraps the upstream DNA around itself, closes the clamp, and unwinds the promoter. Next, the RNAP synthesizes and releases short transcripts in a process called abortive initiation before transiting into the elongation phase. The schematic of RNAP is inspired by Murakami and Darst.[117]

expression, making it possible to produce various outcomes from genetically identical cells.[2]

RNA synthesis in the cell is a complex process that requires a finite time for completion. Having the capability to follow the time course of transcription and its progression in real time is, therefore, essential to understand its regulation. In bulk, one can hope to follow, at most, the progression of transcription as an average of unsynchronized contributions from individual molecules within a population. This averaging obscures crucial information contained in the time-dependent behavior of individual molecules. Single-molecule methods overcome the limitations inherent to the ensemble averaging of bulk methods by allowing one to follow the trajectories of individual molecules in real time. The picture that emerges from single-molecule studies of transcription is that of a rich and complex process that provides many checkpoints for regulation throughout transcription.

Over the past two decades, various methods of single-molecule manipulation and detection have been employed to characterize all three stages of transcription. In the first stage of transcription initiation, RNA polymerase (RNAP) must locate specific promoter sites on the genome in the densely packed cellular environment. Single-molecule methods, such as atomic force microscopy (AFM) and fluorescence-based approaches, have provided insights into how RNAP locates its promoter and unwinds the DNA duplex. Because of the DNA helical structure, unwinding of the duplex is accompanied by changes in its twist. Through the use of magnetic tweezers, it has been possible to both apply torque and follow the torsional states of individual initiating RNAP complexes. During the second stage of elongation, RNAP operates as a molecular motor, converting difference between high-energy phosphoanhydride bonds and lower energy phosphodiester bonds into mechanical work, through the generation of force (in piconewton range) and displacement (in subnanometer scale). Methods of single-molecule manipulation, such as optical tweezers, are ideally suited to precisely measure forces and displacements on this scale; thus, optical tweezers are capable of providing unique insight on the mechanochemical conversion in the transcription process as well as the mechanisms by which transcription factors regulate the dynamics and the progress of the enzyme. When the RNAP finishes synthesizing the full-length transcript, it must stop at a specific location and release the transcript in a controlled manner. Single-molecule techniques make it possible to selectively apply loads on either the DNA template or the RNA transcript, and to dissect regulatory elements in the final stage of transcription, termination.

Here, we present a review of the various aspects of transcription that have been addressed using methods of single-molecule detection and manipulation. We have organized this Review along the three stages of transcription. In the initiation and termination stages, where the factors involved differ substantially between the prokaryotic and eukaryotic systems, we will describe first the results established in prokaryotes prior to detailing those obtained in eukaryotes.

2. TRANSCRIPTION INITIATION

Whereas single-subunit viral polymerases such as T7 and SP6 RNAP can start transcription at a promoter region without additional cofactors, multisubunit bacterial and eukaryotic RNAPs require transcription factors that aid the enzyme to recognize and bind to the promoter. Together they form an initiation complex that unwinds the DNA at the promoter and produces a nascent RNA transcript that stabilizes the complex and primes the enzyme for the processive synthesis of a full-length RNA message. Although similar events occur in prokaryotes and eukaryotes, different factors participate in the initiation process. Next, we address them in order.

2.1. Prokaryotic Transcription Initiation

In prokaryotes, core RNAP (consisting of five subunits, $\beta\beta'\alpha_2\omega$) can nonspecifically bind to DNA. In vitro, it can initiate transcription from DNA ends. However, to initiate transcription from a promoter, the core polymerase must assemble into the holoenzyme by binding to a σ factor, which recognizes specific sequences of promoter DNA, facilitates DNA unwinding, and influences the early phase of transcription elongation.[3] To begin specific transcription initiation, RNAP must first locate its promoter, unwind the DNA to form an open promoter complex, and begin transcribing the DNA until it releases the promoter and transitions into the elongation phase (Figure 1). As we will describe in detail in this section, various single-molecule approaches, such as single-molecule

dx.doi.org/10.1021/cr400730x | Chem. Rev. 2014, 114, 3203−3223

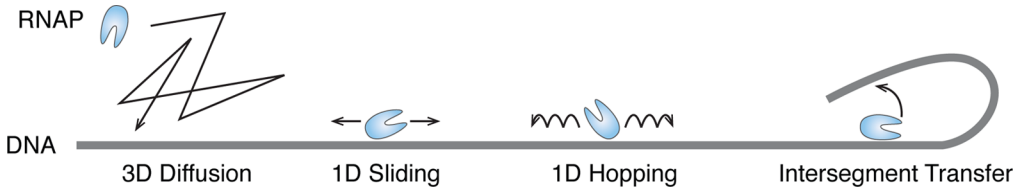


Figure 2. Schematic of various promoter search mechanisms by RNA polymerase. The blue clamp-like shape represents RNA polymerase, and the gray line represents the DNA. 1D sliding is the one-dimensional diffusion of the RNAP along the DNA segment. 1D hopping is a transient association and disassociation of the RNAP along the DNA. Intersegment transfer is the translocation of the RNAP from one point on the DNA to a distant point via a loop intermediate.

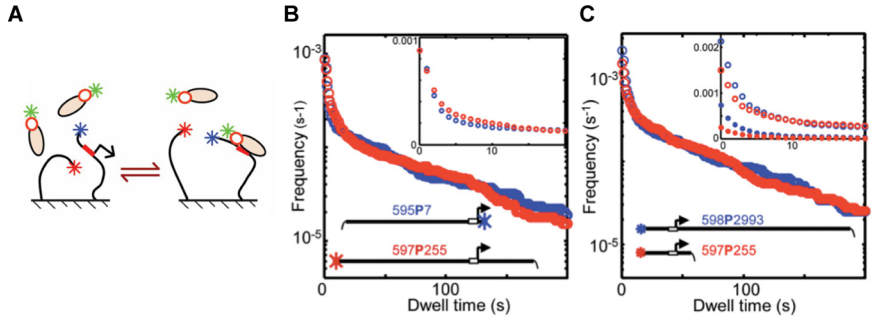


Figure 3. A CoSMos experiment to study the influence of 1D sliding to promoter search. (A) The DNA templates were labeled with different fluorescent dyes (red and blue), and immobilized on the glass surface. The σ^{54} holoenzyme was labeled with another fluorescent dye. Binding events were scored from colocalization of fluorescence signals. (B) Frequency of binding events with lifetimes greater than or equal to the indicated dwell time on DNAs with 7 bp downstream from the promoter (blue), or a control (red). Inset shows a magnified view of the short dwell time plotted in a linear scale. (C) Frequency of binding events with lifetimes greater than or equal to the indicated dwell time on DNAs with 2993 bp downstream from the promoter (blue), or a control (red). Inset shows a magnified view of the short dwell time plotted on a linear scale. Reprinted with permission from ref 13. Copyright 2013 National Academy of Sciences.

fluorescence-based methods, atomic force microscopy, and magnetic tweezers assays, have provided a detailed picture of transcription initiation.

2.1.1. Promoter Search. The investigation of how DNA-binding proteins search and find their targets was pioneered in the studies of *lac* repressor, which was shown to bind to the target operator much faster than the 3D diffusion limit.[4] Subsequent investigations have shown that the target search process of DNA-binding proteins can be accelerated by reducing the dimensionality of the search.[5,6] According to these thoughts, the protein finds the target much faster by binding nonspecifically on the DNA genome and subsequently sliding one-dimensionally along the double helix, hopping within a DNA segment, or transferring between contacting segments of the DNA molecule (Figure 2).

Single-molecule studies have directly observed the trajectories of RNAP molecules during the search process. Earlier experiments observed one-dimensional sliding of fluorescently labeled *E. coli* RNAP holoenzyme along the DNA; it was estimated that the 1D diffusion coefficient for the RNAP was 1×10^4 nm^2/s over a short distance of ~90 nm,[7] although a longer diffusion distance of over 10 μm was also observed in an independent study.[8] During this sliding process, RNAP tracks along the DNA groove, as was observed by rotation of fluorescent-bead labeled DNA when it was being dragged over immobilized RNAP holoenzyme.[9] Atomic force microscopy (AFM) has also been used to observe the 1D sliding of the σ^{70} RNAP on the DNA and reported the 1D diffusion coefficient of 1.1×10^1 nm^2/s.[10] The smaller diffusion coefficient reported in the AFM study may be due to the surface−DNA interactions that are expected to hinder the diffusion process. Events of hopping and intersegment transfer were also observed in AFM imaging.[11] A recently developed fast-scanning AFM, which has a temporal resolution of 1−2 frames per second, was used to re-examine the promoter search mechanism of holoenzyme RNAP and found that RNAP utilizes all of the mechanisms described for facilitated targeting, that is, 1D sliding, hopping, and intersegment transfer of RNAP to locate the promoter sequence.[12] Together, these studies showed that RNAP uses various mechanisms of facilitated targeting; however, they did not address the relative contribution of these mechanisms to the binding of RNAP to the promoter.

Two recent fluorescence-based in vitro studies have investigated the relative contributions of 1D sliding, hopping, and intersegment transfer to promoter search by RNAP. Friedman et al. have used colocalization single-molecule fluorescent spectroscopy (CoSMos) to assess the contribution of 1D sliding in target finding using the σ^{54} RNAP holoenzyme.[13] The σ^{54} RNAP holoenzyme is responsible for transcription initiation of genes required for survival under conditions of stress, such as heat shock and nitrogen depletion.[14] The widely studied σ^{70}-bound holoenzyme, on the other hand, is involved in the expression of housekeeping genes needed during the exponential growth phase. In the CoSMos study by Friedman et al., two DNA templates, with or without a promoter sequence, were labeled with different fluorescent dyes and immobilized on glass surface (Figure 3A). The positions of these templates were identified using fluorescence imaging of the DNA-specific dyes. The holoenzyme was labeled with a different fluorescent dye and introduced into the chamber. Binding events were scored by an appearance of the enzyme fluorescence in the same spot with the immobilized DNA template for 0.2 s or longer.[13]

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

To determine if short-lived events such as sliding or hopping play a role in promoter search, the authors varied the length of the DNA segment flanking a promoter site. The assumption here was that if nonspecific binding contributes to the promoter search, shortening or eliminating the flanking DNA should decrease the observed rates of promoter binding. However, shortening the flanking DNA to 7 base pair (bp) or increasing it to 3000 bp had no effect on the promoter binding rate of RNAP (Figure 3B and C). This observation is inconsistent with the prediction from the facilitated 1D diffusion model. Therefore, this study excluded the facilitated diffusion mechanism over distances between 7 and 3000 bp as a significant pathway for σ^{54} RNAP to reach the promoter, favoring instead 3D diffusion as the major promoter search mechanism.

Wang et al. followed quantum dot-labeled *E. coli* σ^{70} RNAPs as they searched for promoters on a nanofabricated array of stretched λ phage DNA, called DNA curtains (Figure 4A), and,

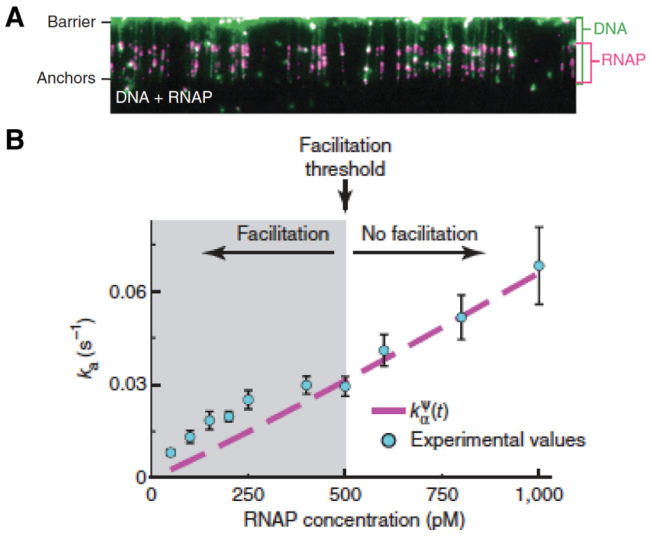


Figure 4. DNA curtain experiment for RNAP-promoter binding. (A) Quantum dot-labeled RNAP (magenta) bound to tethered DNA curtain (green). (B) Observed rate promoter association rates (k_a) as a function of RNAP concentration. The gray region shows the regime that the theoretical target association rate from 3D diffusion (magenta) is lower than the observed rate, which may reflect promoter-finding acceleration due to facilitated diffusion. Reprinted with permission from ref 15. Copyright 2013 Nature Publishing Group.

in contrast to prior studies, observed no 1D sliding of RNAP.[15] However, given the temporal (10 ms) and spatial (~40 nm) resolutions of the experiments, fast sliding events or events occurring on shorter length scales may not have been detected. The authors, therefore, derived a theoretical framework to estimate the contributions of hopping and 1D diffusion for promoter binding by RNAP that occur below the temporal and spatial resolutions of the experiments. They concluded that the most important contribution to the observed rate in these calculations (for stretched DNA) is the 3D diffusion term, which is dependent on the concentration of the RNAP. The authors then measured the promoter association rate to DNA curtains at various concentrations of RNAP; they observed that the rates of target association at concentrations below 500 pM of RNAP exceeded the expected rates from 3D diffusion solely (Figure 4B). The probability of finding the promoter target by direct collision with the curtains in 3D increases with protein abundance and will eventually surpass the probability of success target location through facilitated diffusion process in one dimension (Figure 4B). On the basis of these curtain experiments, these authors concluded that the higher RNAP concentrations present in vivo render 3D diffusion an efficient promoter search mechanism.

Thus, while the AFM studies have shown that RNAP can use all three target-search mechanisms, 1D sliding, hopping, and intersegment transfer,[11,12] the two recent fluorescence-based studies favored 3D diffusion and concluded that long-range 1D sliding does not play a significant role in promoter search.[13,15] However, these in vitro studies differ from the in vivo environment in two main aspects. First, a bacterial nucleoid exists in a coiled structure; therefore, the contributions from intersegment transfer to promoter search in vivo may be more significant than assumed. By design, DNA curtain experiments eliminate the possible contributions of intersegment transfer to promoter search. Second, the interior of the cell is crowded with other macromolecules; therefore, it is unrealistic that promoter search occurs strictly through 3D diffusion. Rather, it is possible that intersegment transfer events followed by short (few nanometers) 1D diffusion runs play a significant role during promoter search in vivo. Notwithstanding these considerations, the importance of intersegment transfer for promoter search in vivo remains to be established.

2.1.2. Closed-to-Open Complex Transition. The binding of RNAP to the promoter leads to the formation of the closed promoter complex (RPc, Figure 1). In this complex, the DNA is still in double-stranded form and has not been inserted into the RNAP active center cleft. Subsequently, the enzyme unwinds ~13 bp of the DNA duplex and isomerizes into the RNAP open promoter complex (RPo, Figure 1). Biochemical studies have characterized the sequence of events during the RPo complex formation at the λP_R promoter, which consists of at least three intermediate steps: DNA loading, DNA unwinding, and assembling of the polymerase clamp on the downstream DNA (see Saecker et al.[16] for a review). In one experiment, DNA footprinting assays revealed an increased protection of 10 bp on the upstream side of the RPc relative to the RPo (~100 bp in the RPc and ~90 bp in the RPo are protected).[17] These extended protections suggest that the DNA bends and wraps around the RNAP in these complexes.

Although footprinting assays are not able to directly detect DNA wrapping, RPo has been observed to wrap DNA by AFM imaging. Using the λP_R promoter, Rivetti et al. found that ~90 bp of DNA were unaccounted for in the images of RPo complexes (Figure 5A). These authors proposed that the missing DNA resulted from the wrapping of DNA around the enzyme by ~300° (Figure 5B).[18] This wrapping generates mechanical stress in the DNA; assuming a persistent length of 53 nm for DNA,[19] the bending energy stored in the 90 bp segment of DNA wrapped 300° around the enzyme is ~14 kcal/mol. This energy must be paid by the binding energy between the enzyme and the DNA. It has been estimated that the binding free energy of RNAP to the promoter to form the RPo at the λP_R promoter is −9 kcal/mol at 37 °C, corresponding to a dissociation constant of ~3.7×10^{-8} M.[20] Thus, without the energy cost of DNA bending, the free energy for RPo formation would be around −23 kcal/mol, which is equivalent to the free energy from hydrolysis of 2.5 mol of ATP, and would render the bound complex too stable, making promoter clearance more difficult. Subsequent studies have shown that the decrease in DNA contour length observed in

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

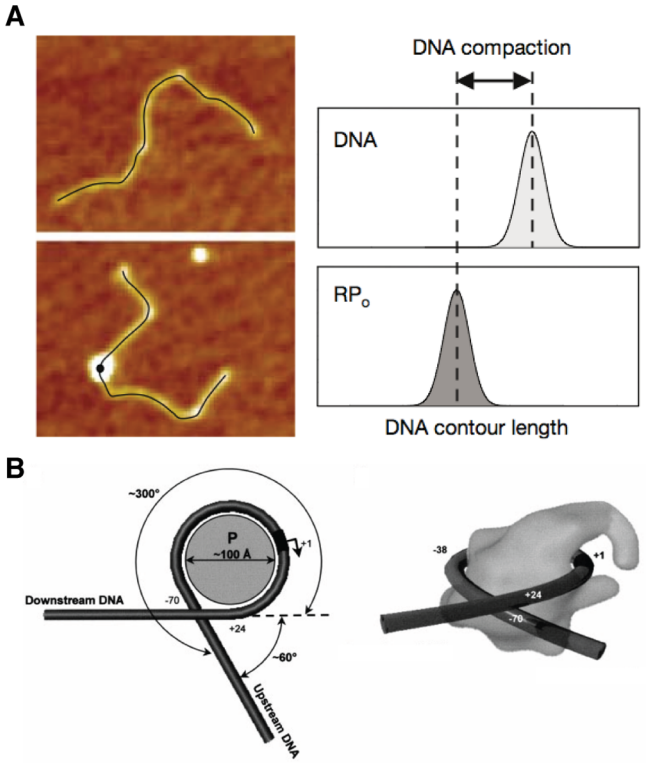


Figure 5. Wrapping of DNA around RNAP in the open promoter complex. (A) AFM images of DNA in the absence (top) and presence of RNAP (bottom). The contour length of DNA in RPo is shorter than that in the free DNA. (B) A schematic of σ^{70} *E. coli* RNAP wraps DNA around itself in the RPo. (A) Reprinted with permission from ref 21. Copyright 2007 Wiley-VCH Verlag GmbH & Co. KGaA. (B) Reprinted with permission from ref 18. Copyright 1999 Wiley-VCH Verlag GmbH & Co. KGaA.

the AFM images (from which wrapping can be inferred) varies among different promoter sequences ranging from ∼6 to ∼90 bp. Moreover, it was found that the extent of DNA wrapping in RPo decreases by 2−3-fold when the carboxy terminal domains of α-subunits are removed, indicating that this domain may mediate the interactions between RNAP and DNA in this process.[21]

During RPo formation, the polymerase unwinds the promoter, and the DNA outside the enzyme positively supercoils to compensate for the promoter unwinding. This process has been observed in real time at the single-molecule level using magnetic tweezers, which can apply torque to twist the DNA and allow it to form plectonemes (Figure 6A).[22] Using the *lacCONS* promoter, Revyakin et al. observed an extension increase of 50 ± 5 nm of negatively supercoiled DNA.[22] When the experiment was done with positively supercoiled DNA, its extension decreased by 80 ± 5 nm during RPo formation. These observations allowed the authors to separate the contributions from the unwinding and wrapping of the DNA around the enzyme during RPo formation to the DNA extensions. The authors determined that DNA unwinds 1.2 ± 0.1 turns or 13 ± 1 bp and wraps by 15 ± 5 nm or ∼44 bp,[22] in good agreement with the wrapping observed in an AFM study.[21] No DNA unwinding intermediates were observed in this study.

In addition to inducing bending and wrapping of the DNA, RNAP also closes its clamp upon RPo formation (Figure 1). This conformational change was detected by single-molecule FRET between the tips of the β' and β pincers labeled with fluorophores (Figure 1). Chakraborty et al. observed that the free σ^{70} holoenzyme and the core enzyme could exist in three noninterconverting FRET states, which they denoted open, closed, and collapsed clamp states, in a ratio of 2:1:1 approximately.[23] The authors estimated that the width between the pincers is ∼20 Å in the open state, sufficient to accommodate dsDNA. This width reduces to ∼12 Å in the closed state, capable of accommodating only ssDNA. In the collapsed state, the width further reduces to ∼8 Å, which is too small for ssDNA. Upon the addition of DNA, the σ^{70} RNAP quickly transits through the RPc and readily forms the RPo complex, and only the closed clamp conformation was observed. This closed clamp conformation was retained from this point onward throughout the initiation and elongation phases (Figure 1).[23] Because σ^{70} RNAP has a short lifetime in the RPc state, σ^{54} RNAP, which transits from the RPc to the RPo state only in the presence of AAA$^+$ ATPase activator NtrC, was studied.[23] In the RPc and various intermediates to the RPo of the σ^{54} RNAP, the majority of the molecules are found in the open clamp state, with a small population of the closed clamp state. Only when both ATP and NtrC1 are present, the RPo is formed and only the closed clamp state is observed.[23] This result indicates that the RNAP clamp only closes in the RPo

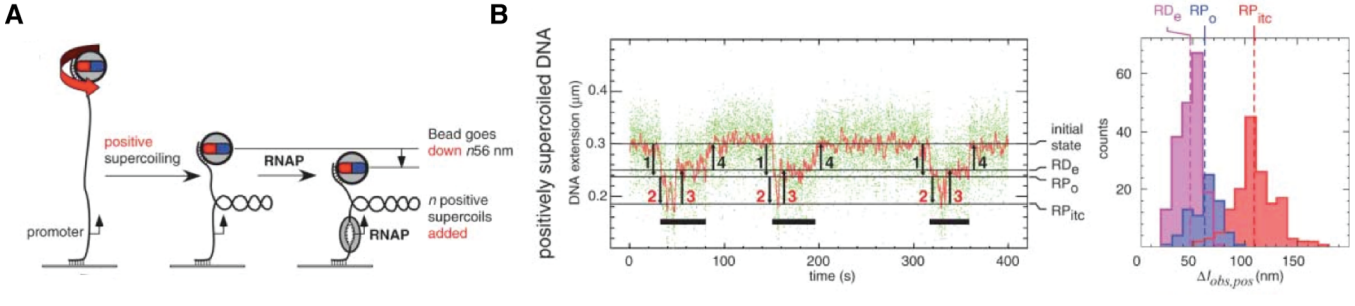


Figure 6. Promoter unwinding and DNA scrunching from single-molecule magnetic tweezers experiments. (A) A promoter DNA is tethered between a glass slide and a magnetic bead, and then positively (or negatively, not shown) supercoiled by rotating the bead, which results in shortening the end-to-end distance. Unwinding of *n* turns of DNA by RNAP results in the compensatory gain of *n* positive supercoils, which is detected by the movement of the bead. (B) An example trajectory of DNA extension over time showing transitions between the four states detected. Numbered arrows indicate unwinding and rewinding events. Numbers 2 and 3 mark events of scrunching and reversal of scrunching, respectively. The histogram shows the end-to-end distance change of the RPo, the initial transcribing complex, and the elongation complex from the initial state.[25] Reprinted with permission from ref 25. Copyright 2006 American Association for the Advancement of Science.

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

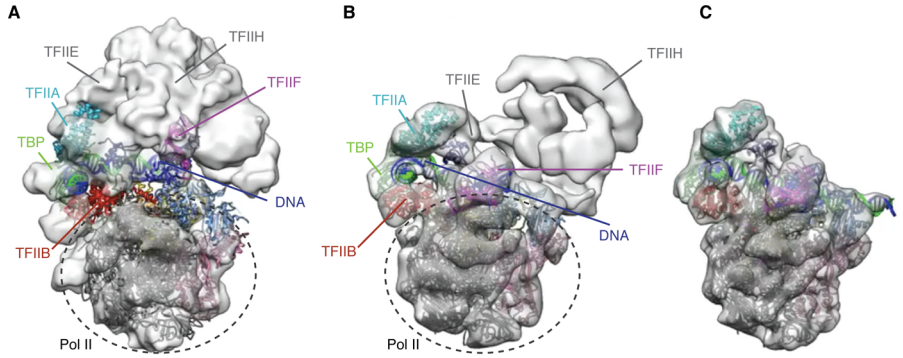


Figure 7. Eukaryotic preinitiation complex. (A) Cryo-EM reconstruction of purified yeast PIC with available crystal structures docked, from Murakami et al.[38] (B) Negative stain reconstruction of stepwise assembled human PIC with available crystal structures docked, from He et al.[39] (C) Cryo-EM reconstruction of a partial human PIC that lacks TFIIH with available crystal structures docked, from He et al.[39] Images courtesy of E. Nogales.

state upon loading of DNA into, and unwinding of DNA in the active-center cleft of the enzyme.

The σ^{54} RNAP has also been employed to study the kinetics of RPo formation using the previously discussed CoSMos study by Friedman and Gelles.[24] In this experiment, the DNA template was labeled with a fluorescent dye and immobilized on the surface. The σ^{54} RNAP was labeled with a different dye, and binding of the enzyme to the DNA was detected by colocalization of the two fluorescent dyes. From the lifetimes of the individual RNAP−DNA complexes, two distinct types of RPc were discovered: a short-lived RPc with a lifetime of 2.3 ± 0.5 s that is a precursor of a longer lived RPc, which has a lifetime of 79 ± 13 s. After adding the NtrC and ATP, the longer-lived RPc isomerizes into RPo with the rate of 1.9×10^{-3} s^{-1}, which is the rate-limiting step of the initiation process.[24] The mechanism by which NtrC uses the energy from ATP to catalyze the open complex formation is currently unknown.

2.1.3. Abortive Initiation and Promoter Clearance. After unwinding the DNA duplex, RNAP synthesizes and often releases small transcripts up to 11 nucleotides in a series of prematurely terminated initiation events termed "abortive initiation". This process represents the early stages of transcript synthesis before the enzyme commits itself to processive elongation. Both a single-molecule magnetic tweezers assay and a single-molecule FRET experiment have shown that abortive initiation occurs via a "scrunching" mechanism, wherein the polymerase remains stationary and reels the downstream DNA into the active site.[25,26] Kapanidis et al. used single-molecule FRET to determine the location of RNAP relative to the template.[26] These authors found that the leading and trailing edges of the enzyme remain stationary, while the downstream DNA moves closer to the active site during abortive initiation.[26] In a parallel study, Revyakin et al. used a magnetic tweezers to observe the DNA unwinding at different stages of transcription initiation at the *lac*CONS promoter (Figure 6A).[25] The authors observed four different DNA extensions, which were assigned to (1) the initial state, (2) RPo, (3) RNAP−promoter initial transcribing complex, and (4) RNAP−DNA elongation complex.[25] DNA scrunching was seen as an overshoot in DNA extension that followed RPo formation and preceded the formation of the elongation complex (Figure 6B). This scrunching signature was observed in 80% of the transcription traces. When the RNAP was allowed to synthesize only a 2-nt RNA transcript, scrunching was not observed. Scrunching occurred when the enzyme was allowed to synthesize a 4- or 8-nt RNA transcript, and the duplex in the RPo was seen to unwind by 2 or 6 additional bp, respectively.[25] These observations suggest that the polymerase can accommodate 2 nt of RNA transcript before it starts to scrunch the DNA template.[25] Scrunching should lead to unwinding and compaction of the DNA that would result in the generation of a stressed intermediate as previously hypothesized.[27] The enzyme can relieve this stress either by releasing the DNA to the downstream side, aborting transcript synthesis, and returning to the RPo, or by releasing the DNA to the upstream side, breaking the interactions of its trailing edge with the DNA, and escaping from the promoter. Indeed, each base pair of DNA scrunched stores on average a free energy of ∼2 kcal/mol from breakage of the base pair.[28] At a typical promoter, RNAP synthesizes ∼9−11 nt, corresponding to ∼7−9 bp of scrunched DNA, before it can proceed to elongation. It is possible that part of the ∼14−18 kcal/mol stored in this process is used to overcome the RNAP−promoter interactions, estimated to be ∼7−9 kcal/mol at 37 °C.[20]

2.1.4. The Fate of the σ Factor. The fate of the σ factor, once the polymerase transits from the initiation to the elongation phase, is still unclear. It may remain bound to the promoter, be released from the promoter, or be retained in the transcribing elongation complex. Recent studies observed retention of the σ^{70} factor in the elongation complex.[29−31] A single-molecule FRET study showed that ∼20−90% of early elongation complexes harboring 11- or 14-nt transcripts retain the σ^{70} factor with a half-life between 20 and 90 min.[29] The initial extent of σ^{70} retention may increase to ∼70−100%, depending on the DNA sequence.[29] In addition, about 50% of the late elongation-phase RNAPs harboring 50-nt RNA transcripts still retain the σ^{70} with a half-life of over 50 min.[29] These observations raise the possibility that σ^{70} may remain associated with the polymerase throughout most of the elongation phase; however, this hypothesis is yet to be tested for polymerases approaching the termination phase. The role of σ^{70} during the elongation phase is unknown.

Nonetheless, the fate of σ^{70} may not be shared by other σ factors. The σ^{54} factor was shown to be released from >90−95% of the complexes within 10 s after RNAP binding to the promoter, while RNA transcript was detected ∼20 s after σ^{54} release.[24] This result raises interesting possibilities that different σ factors might play different roles in transcription regulation. While the σ^{70} may remain bound to the core enzyme and

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

participate in transcription elongation, σ^{54} may only function in transcription initiation, and its release could trigger the transition of RNAP into the elongation phase.[24]

2.2. Eukaryotic Transcription Initiation

To date, most of our understanding of transcription initiation has come from studies in the prokaryotic system. However, biophysical methods have recently begun to shed light on the structure and dynamics of transcription initiation in eukaryotes. As a result, we are beginning to understand the mechanism of open complex formation, of abortive initiation, as well as the rates for these steps in the eukaryotic system.

While the core RNAP bound to the σ factor is sufficient to locate the promoter and to initiate transcription in prokaryotes, an additional set of general transcription factors that include TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH are involved in eukaryotic transcription initiation. In vitro biochemical reconstitution studies have revealed the order of assembly of these factors to form a functional preinitiation complex (PIC).[32] The first factor to locate and bind the promoter is TFIID, one of whose subunits is the TATA-box binding protein (TBP). TFIIA and TFIIB are recruited next and help stabilize the interactions of TBP with the complex. The bound TFIIB recruits RNA polymerase II (Pol II) and TFIIF. TFIIE binds next to the complex followed by TFIIH, whose ATPase and helicase activities are required to unwind the DNA duplex to form the transcription bubble. The binding of TFIIH completes the formation of the PIC. Recently, two different cryo-electron microscopy (cryo-EM) structures of eukaryotic PIC have been communicated and both structures are supported by cross-linking experiments.[33−37]

Murakami et al. published a yeast PIC structure consisting of two lobes.[38] The authors assigned one lobe to Pol II and the other lobe to the general transcription factors and the DNA (Figure 7A). In this structure, the DNA only interacts with the transcription factors, TBP, TFIIF, TFIIE, and TFIIH, but does not interact with the polymerase.

He et al. characterized the human PIC and studied the structural changes that accompanied the sequential addition of the various transcription factors.[39] In this structure, the DNA makes interactions with most of the transcription factors, as seen by Murakami et al., but also interacts directly with the polymerase in the presence of TFIIF (Figure 7B).[39] Binding of TFIIF to the TBP-TFIIA-TFIIB-Pol II bound promoter leads to an engagement of the DNA along the cleft of the polymerase. Binding of TFIIE further stabilizes the PIC and provides a platform for binding of TFIIH, which positions its XPB helicase domain directly on the downstream DNA. The authors hypothesized that the XPB domain could translocate and twist the DNA to generate stress, which would be relieved by duplex unwinding, favoring the formation of the open promoter complex.[39] One other significant difference is seen in the fitting quality resulting from the docking of the crystal structures of the various transcription factors to the two cryo-EM densities (Figure 7A and B). The crystal structures fit very well to He et al.'s cryo-EM densities, whereas, in the study by Murakami et al., parts of the crystal structures of Pol II and the transcription factors did not fit well to their cryo-EM densities, and parts of the density assigned to Pol II were not accounted for in the crystal structure (Figure 7).

Even though these structures are derived from different organisms, the differences observed between them are somewhat surprising. To explain these differences, Murakami et al. have suggested that the structure from He et al. represents an incomplete PIC, which lacks TFIIE and TFIIH. These authors also claimed that their cryo-EM densities cover a larger fraction of the transcription factors than those presented in the structure by He et al. However, it should be pointed out that He et al.'s cryo-EM densities fully accounted for the structured sections of these proteins observed in crystal structures. Thus, at the point of writing this Review, the controversy persists as to which structure more accurately represents the eukaryotic PIC. Perhaps, the two structures represent different stages of the PIC. The structure proposed by Murakami et al. could correspond to the promoter containing all of the transcription factors that has just recruited an additional polymerase after the first round of transcription, thus explaining the lack of interactions between the DNA and the enzyme. The structure proposed by He et al. could represent the sequential binding of transcription factors and Pol II to the promoter to initiate the first round of transcription. Future experiments should shed light on the origin of these differences.

Superresolution imaging is now being used to characterize the spatial distribution of initiation complexes of RNAP in live eukaryotic cells.[40] In parallel, single-molecule methods have recently begun to examine eukaryotic transcription initiation in vitro both structurally and dynamically. By fluorescently labeling various positions on the template DNA, nontemplate DNA, TBP, TFIIB, and the Rpb7 subunit of yeast Pol II, Treutlein et al. derived a model for a minimal open promoter complex structure.[41] Their model of the open complex suggests that the downstream DNA can adopt two different conformations: lying between the Pol II clamps, as appears in crystal structures of the elongation complex,[42] and on top of this cleft, closer to the Rpb7. The rates at which the DNA transits between the two conformations were determined to be $k_{\text{in clamp}} = 1.5 \pm 0.4\ s^{-1}$ and $k_{\text{out of lamp}} = 0.8 \pm 0.1\ s^{-1}$, respectively. However, it should be pointed out that this model was determined from a minimal open promoter complex that lacks TFIIA, TFIIE, TFIIH, and other subunits of TFIID, and an 11-nt mismatched DNA was employed to mimic a transcription bubble instead of a complementary nontemplate DNA.

To study the dynamics of eukaryotic transcription initiation, Revyakin et al. used fluorescence imaging to count the rounds of transcription that initiated from the same promoter by determining the number of photobleaching events of fluorescently labeled RNA probes at a given time. They found that the distribution of photobleaching events follows a Poisson distribution.[43] This observation suggests that transcription initiation and reinitiation are noncooperative and independent, which is somewhat inconsistent with a previous notion that the scaffold of transcription initiation factors remains bound to the promoter and reinitiates multiple rounds of transcription at the end of each abortive event.[44]

3. TRANSCRIPTION ELONGATION

The elongation phase of transcription starts once the polymerase has produced a long enough RNA chain and has cleared the promoter region. In this phase, the polymerase uses the energy of NTP incorporation into the growing RNA chain to advance on DNA. Single-molecule studies have provided important insights into the molecular mechanism of transcription by analyzing the temporal and spatial dynamics of elongating RNA polymerase molecules. These experiments are usually performed with optical tweezers, an instrument that

dx.doi.org/10.1021/cr400730x | Chem. Rev. 2014, 114, 3203−3223

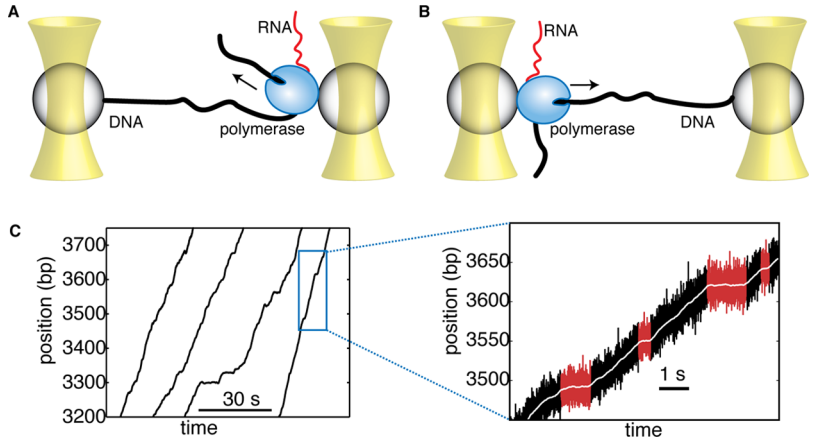


Figure 8. Dual-trap optical tweezers setup for transcription elongation assays. (A) Assisting force. (B) Opposing force. The arrow indicates the direction of RNA polymerase elongation. (C) Single-molecule elongation traces of Pol II during assisting force (left) with detail (right) showing pauses (red) and active elongation (black). Adapted with permission from ref 53. Copyright 2009 American Association for the Advancement of Science.

allows the manipulation of micrometer-sized beads with a focused laser beam. The elongating polymerase is attached to the surface of a bead held in a trap, while one of the ends of the DNA being transcribed is attached to another bead held in another trap (Figure 8). As elongation proceeds, one can infer the precise position of the polymerase on DNA by monitoring the movement of the two beads. In addition, this setup permits the application of force to the polymerase either in the same direction as transcription (assisting force, Figure 8A) or in the opposite direction (opposing force, Figure 8B). Studying the effect of force on the movement of molecular motors is an ideal tool to unravel how these motors couple chemical energy to mechanical motion.[45]

3.1. The Kinetic Cycle of Transcription Elongation

Initially, single-molecule transcription elongation experiments were performed with the *E. coli* RNAP.[46−49] These studies have shown that elongation has two phases: active translocation and pausing (Figure 8C). The same phases were also observed for the eukaryotic RNA polymerase II (Pol II),[50] and for the mitochondrial Rpo41.[51] During on-pathway active translocation, the polymerase incorporates nucleotides into the nascent RNA chain and advances along the DNA template. The paused states were shown to be off-pathway from the NTP incorporation cycle.[46] In the paused state, the polymerase can be stationary at one position on DNA, or it can diffuse backward along the DNA and then recover from these backtracks. Both on- and off-pathway phases of elongation are of interest, as different transcription factors can interact with the polymerase to modify either its active cycle or the probability of entering or remaining in the paused state. As we will describe in detail in this section, single-molecule experiments demonstrated that RNAP translocation occurs through a Brownian ratchet mechanism rectified into a forward movement by NTP binding, and that long pauses correspond to periods in which the enzyme backtracks on DNA, diffusing back and forth on the template until its active site re-engages with the 3′ end of the RNA transcript. There still remains some controversy about the exact details of the nucleotide incorporation cycle, and about the origin of short pauses.

3.1.1. The Nucleotide Incorporation Cycle. Single-molecule studies have confirmed structural and biochemical data indicating that RNA polymerase advances one base pair at a time as it incorporates one ribonucleotide during the synthesis of an RNA chain.[52] When pauses that are not considered part of the nucleotide incorporation cycle are removed, the resulting mean elongation velocities, pause-free velocities, at saturating concentrations of NTPs are: 15−23 bp/s for Pol II,[50,51,53] 10−25 bp/s for the bacterial RNAP,[52,54,55] and 20−24 bp/s for the mitochondrial Rpo41.[51] We provide a range of mean values here, as the pause-free velocities vary slightly depending on the DNA template, forces that aid or oppose the enzyme, and the algorithms that detect and remove pauses.

One early question that single-molecule experiments have addressed is whether the transcriptional translocation event is a power stroke, where the enzyme's translocation is directly coupled to a chemical step, or a Brownian ratchet, where thermally driven movement of the polymerase is rectified to one direction by its substrate. The answer to this question has implications about the possible modes of action of various inhibitors and transcription factors. The pause-free velocity is given by the distance the polymerase has to translocate during one cycle (d), divided by the total time it takes to complete the elongation cycle. This time is the sum of the times necessary to bind the incoming NTP (τ_{NTP}), complete the condensation reaction that incorporates the NTP to the RNA chain (τ_{cond}), and release the pyrophosphate before starting a new cycle (τ_{PPi}):

$$v = \frac{d}{\tau_{NTP} + \tau_{cond} + \tau_{PPi}}$$

If translocation is driven by thermal noise and biased forward by NTP binding followed by the irreversible condensation reaction (Brownian ratchet), τ_{NTP} should be sensitive to force, because that is the step associated with net movement on DNA. Instead, if pyrophosphate release induces, or coincides with, a change in conformation of the elongation complex, triggering translocation (power stroke), τ_{PPi} would be sensitive to force instead. At limiting NTP concentrations, the time it takes to bind NTPs (τ_{NTP}) becomes dominant over the time of pyrophosphate release (τ_{PPi}). In these conditions, if elongation follows a Brownian ratchet mechanism, the velocity should be sensitive to force, while if it follows a power stroke, the velocity should not depend on force. Single-molecule data have shown

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

that at low NTPs, the pause-free velocity does depend on force.[52] This result is thus inconsistent with the power stroke mechanism, and supports the Brownian ratchet mechanism for transcription elongation, originally determined from bulk studies (Figure 9).[56] To date, no motor has been shown to

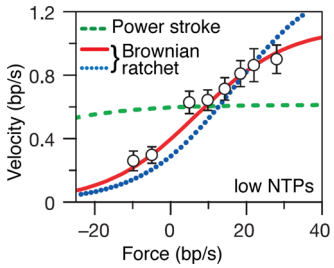


Figure 9. Pause-free velocity as a function of force at low NTPs concentrations. The data fit a Brownian ratchet model, but do not fit a power stroke model.[52] Reprinted with permission from ref 52. Copyright 2005 Nature Publishing Group.

change its mechanism of operation as a function of its substrate concentrations; thus, the ratchet is likely to be valid for RNAP at all NTP concentrations.

While it is accepted that RNA polymerases act as Brownian ratchets, some aspects of the nucleotide addition cycle are still under debate.[57] The nucleotide addition cycle has multiple steps: the movement of the polymerase from a pre-translocated to a post-translocated state, binding of the incoming NTP, and catalysis of a phosphodiester bond between the NTP and the nascent RNA chain. The simplest model includes these steps sequentially in a linear fashion (Figure 10A). Because it is difficult to directly measure individual rates involved in nucleotide incorporation, one measures the elongation velocity of the polymerase in different conditions and fits the rates of the model to the data. The data obtained from single-molecule experiments offer the advantage to define and focus selectively on pause-free velocity, which in principle reflects only the time the enzyme spent in the on-pathways reactions. Note, however, that the temporal and spatial resolution of the experiments and algorithms used to detect pauses typically limit their extraction to those longer than 1 s. Additionally, the force applied to the polymerase during transcription allows one to make quantitative predictions about the dependence of the pause-free velocity on this applied force, and, as such, it can be used to put any elongation model to the test.

When fitting the single-molecule data, most studies up to date have made the simplifying assumption that the rates of translocation and nucleotide binding are much faster than the rate of NTP incorporation. When using this simplified assumption and the linear model (Figure 10A), Abbondanzieri et al. and Larson et al. found that the velocity versus force data could not be fit well. This observation prompted them to propose a more complex, branched model of elongation that involves the existence of a secondary NTP binding site (Figure 10B).[52,54] While in the linear model the incoming NTP can only bind after the polymerase has translocated, in the branched model the NTP can bind to both the pre- and the post-translocated states.

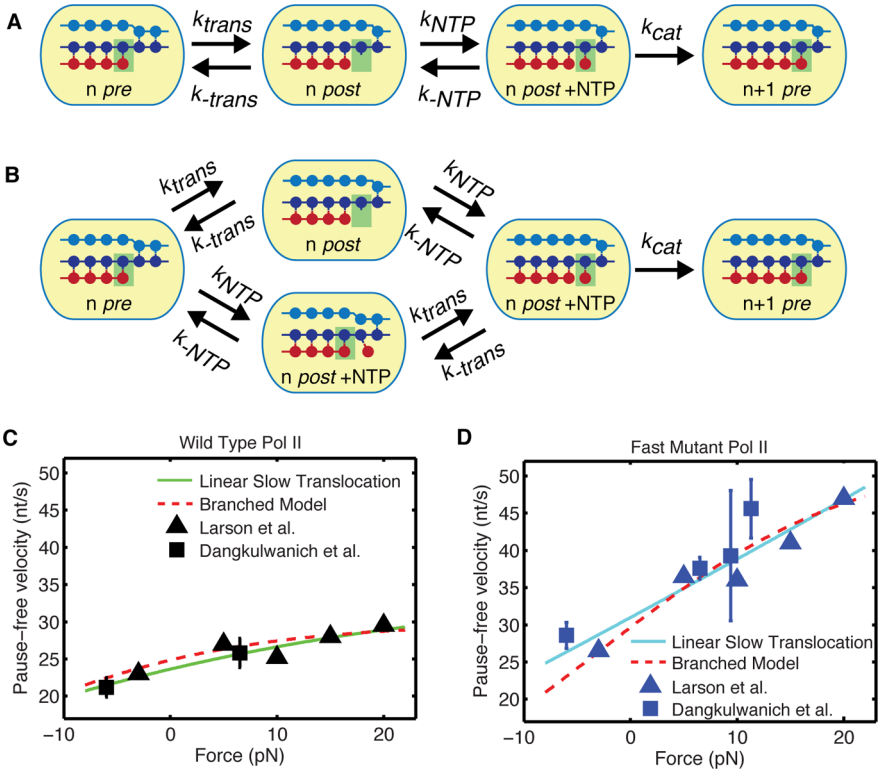


Figure 10. Two different kinetic models for the on-pathway nucleotide addition cycle. (A) Linear Brownian ratchet model of transcription elongation, which only allows NTP binding after translocation. (B) Branched Brownian ratchet model of transcription elongation that allows NTP binding either before or after translocation. (C) Fits of the linear model with slow translocation and the branched model to the force−velocity relationship of the wild-type Pol II.[58,54] (D) Fits of the linear model with slow translocation and the branched model to the force−velocity relationship of the fast mutant Pol II.[58,54]

dx.doi.org/10.1021/cr400730x | Chem. Rev. 2014, 114, 3203−3223

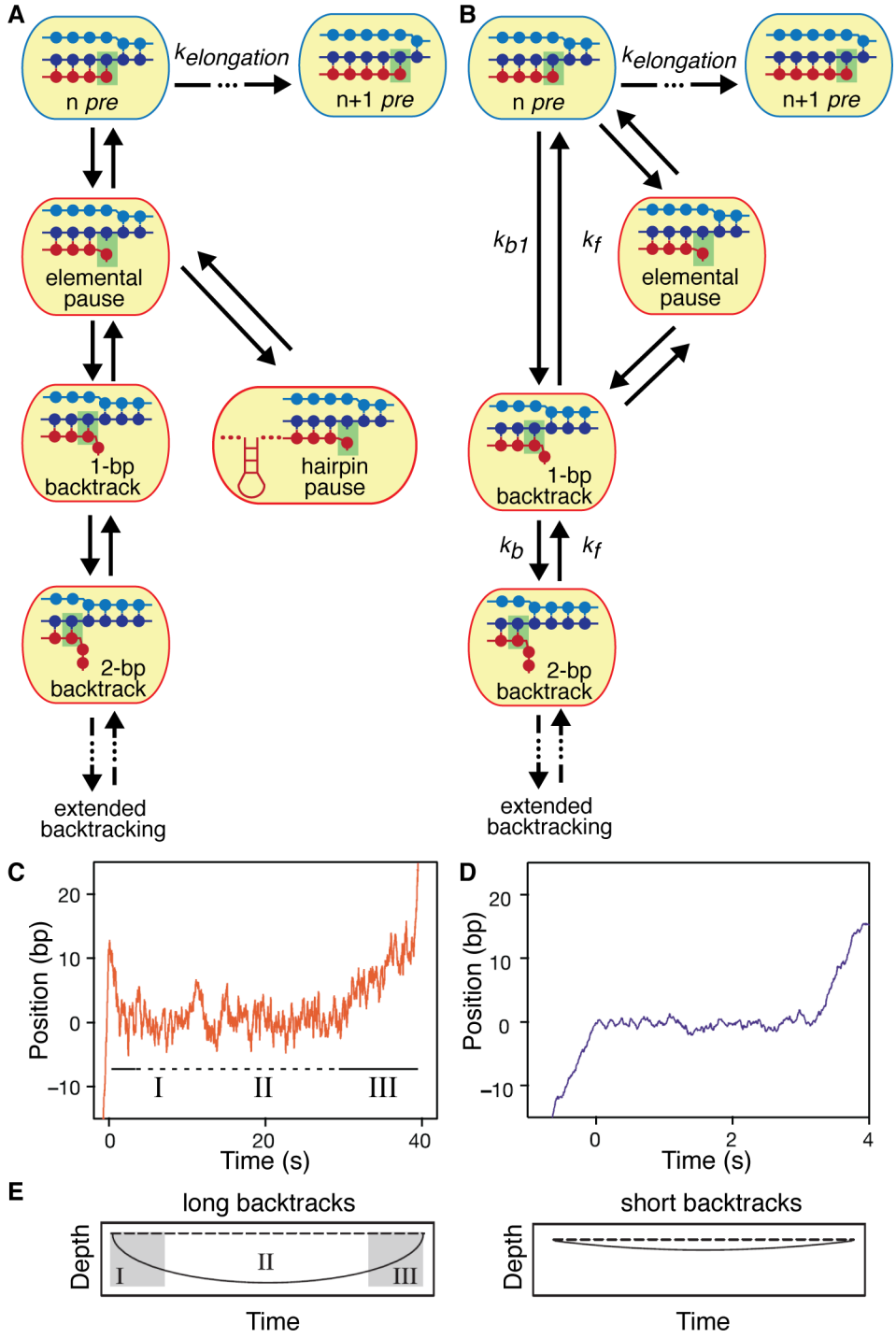


Figure 11. Mechanisms of transcriptional pauses. (A) Model of pausing including elemental/ubiquitous pauses that can lead to backtracking or hairpin stabilized pauses. (B) Backtracking model of pausing, which suggests that the elemental pause state is not obligatory. (C) Example single-molecule trace showing long pauses and backtracking. (D) Example trace showing a short pause. (E) Predictions of average trajectories for long and short pauses based on the backtracking model. (C,D) Reprinted with permission from ref 60. Copyright 2003 Nature Publishing Group. (E) Adapted with permission from ref 69. Copyright 2009 Elsevier.

However, Dangkulwanich et al. recently questioned the validity of the common assumption that the transition rates from the pre- to post-translocated states are much larger than those of other kinetic steps.[58] Whereas the application of external force affects both the forward and the reverse translocation rates of the enzyme, Dangkulwanich et al. recognized that a mechanical barrier such as a nucleosome only affects the forward rate, making it possible to separate it from the reverse rate. Using this approach, the authors found that the forward translocation rate is indeed comparable to the subsequent catalytic rates.[58] Thus, the fast translocation assumption previously used to fit the linear model to the data is not appropriate. Accordingly, the pause-free velocities as a function of force can be fit with the exact solution of the linear Brownian ratchet model just as well as to the branched model (Figure 10C). Therefore, it is not necessary to invoke a second

dx.doi.org/10.1021/cr400730x | Chem. Rev. 2014, 114, 3203−3223

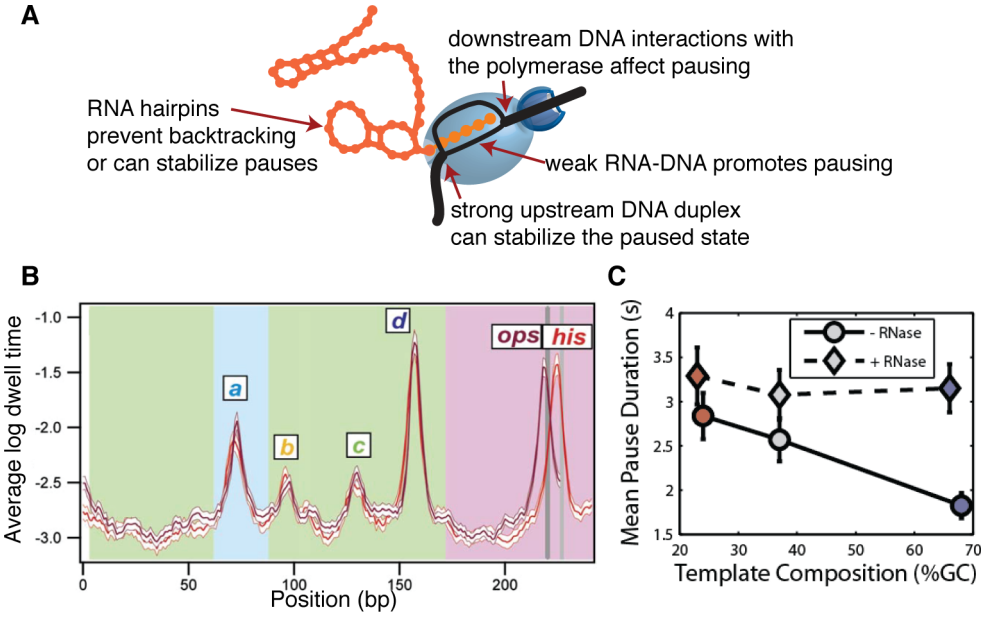


Figure 12. Regulations of transcriptional pauses. (A) Different elements that influence sequence-dependent pausing. (B) Dwell times from single-molecule data showing sequence-dependent pausing of RNAP.[64] (C) For Pol II, mean pause durations decrease as the GC content of the template increases.[51] The effect disappears as the nascent RNA is digested, suggesting the structure of the RNA aids pause recovery. (B) Adapted with permission from ref 64. Copyright 2006 Elsevier. (C) Adapted with permission from ref 51. Copyright 2012 National Academy of Sciences.

NTP binding site to explain the force−velocity data. This finding makes the simpler, linear Brownian ratchet model (Figure 10A) more appealing to explain the nucleotide incorporation cycle of the polymerase.

3.1.2. Pausing. Polymerase pauses can significantly reduce the overall transcription rate during elongation. For instance, at saturating concentrations of NTPs, the velocity of the eukaryotic Pol II including pauses is about one-half of its pause-free counterpart: 11.6 ± 2.5 bp/s[59] as compared to 22.9 ± 5.0 bp/s (same data as Bintu et al.,[59] unpublished result). This observation raises the possibility that various transcription factors can potentially regulate the length and the frequency of pauses to control the rate of transcription. Characterizing the nature of these pauses and how they are regulated is, therefore, essential for a comprehensive understanding of the regulation of gene expression.

It is accepted that a pause corresponds to a state that deviates from the main nucleotide incorporation pathway, as pausing has been shown to compete kinetically with elongation: the higher is the elongation rate, the fewer are the pauses.[46] However, the nature of these pauses is actively being studied. The duration of transcriptional pauses varies from under a second to minutes, and cannot be fit with a single exponential distribution.[50,52,55,59] This observation suggests that there is not a single paused state.

An initial analysis divided bacterial RNAP pauses into two categories: short pauses, under 20 s, and long pauses, above 20 s.[55,60] Short pause durations can be fit with two exponentials with time constants of 1.2 and 6 s. By contrast, long pauses are infrequent with durations that are broadly distributed. The long pauses were clearly shown to be associated with the backward movement of the polymerase on DNA averaging about 5 bp,[60] followed by the return of the polymerase to the original position. This phenomenon, initially described by ensemble footprinting studies, was termed "backtracking" (Figure 11C).[61,62] However, pauses defined as short in the single-molecule studies of the bacterial RNAP did not show backtracking (Figure 11D).[55,60] Moreover, the density (number of pauses per bp transcribed) and the durations of short pauses were reported to be insensitive to force.[55] Because backtracking involves movement on DNA, one would expect an assisting force to decrease pausing. Therefore, the force insensitivity of short pauses was interpreted as another piece of evidence that the short pauses are not associated with backtracking. These short pauses were called "ubiquitous" or "elemental", and were speculated to come from molecular rearrangements of the elongation complex that would render it elongation incompetent. In fact, a recent crystal structure of the elemental pause state of bacterial RNAP showed that the clamp is open, and the bridge helix is kinked and blocks the NTP binding site, whereas the RNA−DNA hybrid binding site along with the RNA exit channel are widened.[63] In one view of the mechanisms of transcriptional pausing, RNAP must first enter this elemental paused state.[64−66] These elemental pauses can be subsequently stabilized into longer-lived pauses by RNAP backtracking or by the formation of a hairpin structure in the nascent RNA transcript.[67,68]

While some of the short pauses may not be associated with backtracking (ubiquitous or elemental pauses), it is entirely possible that some of the pauses classified as ubiquitous are short backtracked pauses. An alternative view poses that the elemental pause state is not obligatory and attributes most pauses to backtracking. One expects short backtracking pauses to be associated with short backward excursions (under the resolution of these experiments ∼3 bp) and therefore to display an apparent force insensitivity.[69] In addition, the rate of entering the 1-bp backtracked state is faster than entering further backtracked states,[58] which predicts that short backtracks are on the same time scale (<1 s) as what has been called ubiquitous pauses.

In an analysis of eukaryotic Pol II, pauses longer than 1 s can be aptly described by a backtracking model in which a pause begins with the backward movement of Pol II by one base pair (Figure 11B).[50] As Pol II backtracks, the entire elongation

dx.doi.org/10.1021/cr400730x | Chem. Rev. 2014, 114, 3203−3223

bubble shifts, and the 3′ end of RNA loses its register with the active center of the polymerase and inhibits NTP incorporation. The polymerase performs a random walk back and forth on the DNA until the 3′ end of the transcript realigns with the active center to allow Pol II to resume elongation. According to this model, the diffusion of the enzyme along the DNA explains a wide distribution of pause durations: in some cases, the random walk finishes in a few steps, while in others the polymerase randomly backtracks many base pairs, so it takes much longer to recover. In fact, the distribution of pause durations can be derived exactly from this model by calculating the probability of observing a random trajectory with n steps, multiplying it by the distribution of times that it takes to perform those steps (which is given by a Gamma distribution), and summing over all allowed values of n.[53] The resulting pause duration distribution is given by:

$$\psi(t) = \sqrt{\frac{k_f}{k_b}}\,\frac{\exp[-(k_f + k_b)t]}{t} I_1(2t\sqrt{k_f k_b})$$

where I_1 is the modified Bessel function of the first kind, and k_f and k_b are the forward and backward stepping rates during backtracking (Figure 11B).[53,69] These rates depend on the applied force (F), the distance to the transition state for a step (d), and the back and forth stepping rate of the backtracked polymerase on DNA in the absence of applied force (k_0) as follows: $k_f = k_0 e^{F\cdot d/k_B T}$ and $k_b = k_0 e^{-F\cdot d/k_B T}$. Note that for short durations, on the order of the rate of backtracking, the distribution $\psi(t)$ behaves as an exponential, while for long durations and small forces, the probability distribution follows a $t^{-3/2}$ power law. In fact, the experimental distribution of pauses for Pol II was found to follow this power law.[50]

The fact that the backtracking model fits all pause durations very well, with just one parameter (k_0), makes it quite appealing.[50,53,69,70] However, the pausing mechanism in the eukaryotic Pol II could be different from that of the prokaryotic RNAP. In fact, Kireeva et al. showed the *E. coli* RNA polymerase paused at certain sequences without backtracking, while yeast Pol II does not recognize the same pause signal.[71] Therefore, the possibility that some of the short pauses are not associated with backtracks is likely.

3.2. DNA Sequence and Nascent RNA Effects on Elongation Dynamics

Ensemble biochemical studies revealed that RNA polymerases pause or arrest on certain DNA templates.[71] An important question is then how the transcribed sequence modulates pausing. Possible candidates are the local stability of the DNA−RNA hybrid, the DNA−DNA interactions upstream and downstream of the hybrid, the extent and stability of RNA secondary structures behind the polymerase, and allosteric interactions of the RNA with the enzyme (Figure 12A). Although some of these mechanisms have been shown to affect the pausing probability at a given sequence, the process by which the enzyme pauses in a sequence-dependent manner remains an area of active research.

Two kinetic models have explicitly calculated the energy differences between the pre-translocated state, the post-translocated state, and backtracked states at each position on DNA.[72,73] One of the models[72] posits that there is a high activation barrier to enter the backtracking state (on the order of 40 $k_B T$), and therefore predicts that most of the sequence specific pauses are trapped in the pre-translocated state. The alternative model[73] assumes that entering a backtrack is relatively easy, but that backtracks are on average limited to 9 bp by the folding of the nascent RNA. These models can correctly predict 60−80% of pauses observed experimentally by RNAP. However, in addition to missing pauses, both models predict pauses at positions that are not experimentally associated with pausing: 20−40% of predicted pauses are false positives.

Single-molecule data can test the predictions of these models by analyzing the pause dynamics and thus inferring the mechanism of pausing at various sequences. However, single-molecule studies on the sequence dependence of pausing are difficult, because even if changes in polymerase positions can be determined with base-pair resolution, the accuracy with which the absolute position of the polymerase on DNA can be determined is much lower. One study overcame this shortcoming by rescaling all traces so that they are perfectly aligned at the point where RNAP runs off the template.[74] Using this technique, the authors were able to achieve an accuracy of about 5 bp, and they showed that for a particular sequence (ΔtR2), the pause durations were sensitive to the applied force, suggesting that backtracking takes place at this pause site.

A different single-molecule study of pausing at multiple sites (Figure 12B) did not detect backtracking at sequences that induced RNAP pausing.[64] Note that the alignment of RNAP position was performed using the peak pause sites at the well-studied *his* sequence as a reference point. For the alignment to work correctly, pausing at the *his* site has to always take place without backtracking, which is believed to be the case. The lifetimes of the paused states at each pausing site could be fit with one exponential, and the time scales of these pauses varied between 1 and 6 s. By correlating the length of the RNA transcript from bulk transcription to the pause peaks in single-molecule traces, the authors proposed that these sequence-specific pauses, which occur with the polymerase stuck in the pre-translocated state,[75,76] are in fact the same as the ubiquitous pauses identified before, and that they are prerequisites for longer pauses stabilized by backtracking or RNA hairpins. By analyzing the DNA sequence at the pause sites, the authors noted that in all seven sequences, G or C bases were present immediately upstream of the RNA−DNA hybrid (at position −10 and −11). The mechanism by which these sequences induced RNAP pausing is still not clear. These bases could extend the RNA−DNA hybrid to the upstream side and generate stress in the enzyme, stabilize a state backtracked by 1 or 2 base pairs, or delay the elongation bubble from moving from the pre- to the post-translocated state. Because the authors did not observe backtracking at these pause sites, the mechanisms involving backtracked state stabilizing are not likely.

From these two experimental results, it seems that different sequences can induce different types of pausing. Moreover, it appears that the eukaryotic Pol II and the prokaryotic RNAP respond to the same pause sequences differently.[71] This result raises the question of whether interactions of the polymerase with the surrounding DNA or RNA are different in the prokaryotic and eukaryotic enzymes. The crystal structures of the prokaryotic and eukaryotic polymerases show them to be quite similar in the region that interacts with the elongation bubble. However, this might not be the case for interactions of the nucleic acids with the outer part of the enzymes. To address the question, one has to map the location of the DNA and RNA once they exit the polymerase. This type of experiment was performed in Pol II by mapping the RNA exiting Pol II

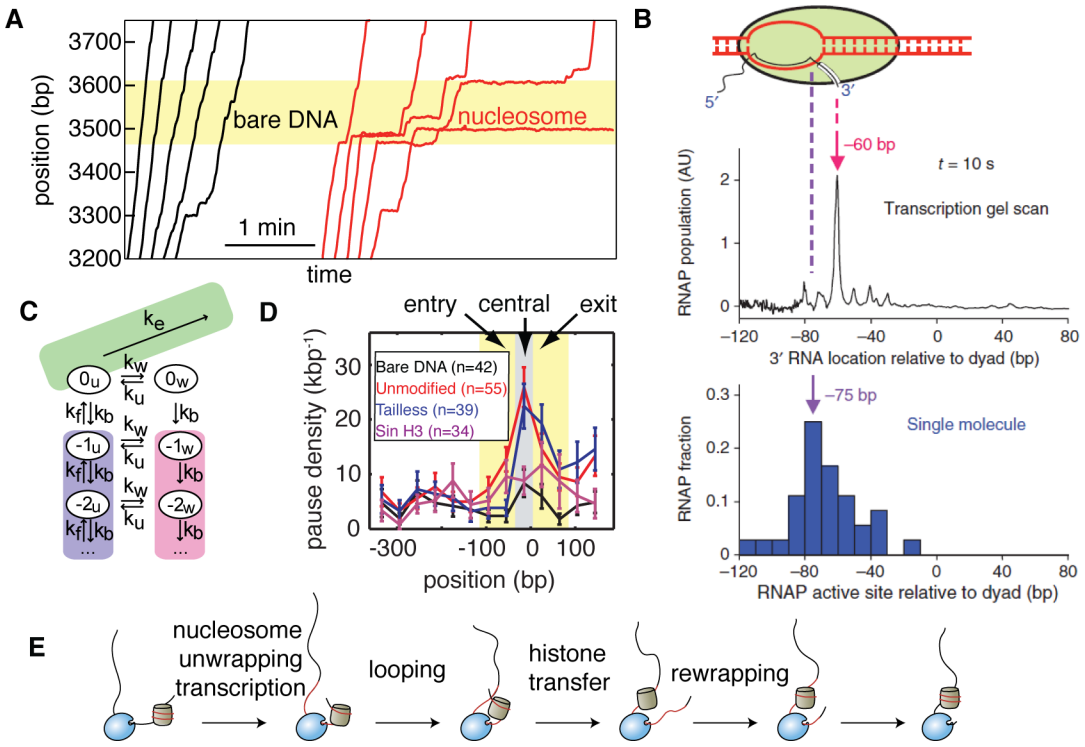


Figure 13. Dynamics of nucleosomal transcription. (A) Single-molecule traces showing increased pausing at the nucleosome as compared to bare DNA.[53] (B) Comparison of 3′ end of the RNA position to the RNAP position on the template reveals backtracking at the nucleosome.[80] (C) Kinetic model for transcription at the nucleosome.[53] (D) Pausing at the nucleosome for tailless and Sin-modified nucleosomes.[70] (E) Model of histone transfer through looping.[59] (B) Adapted with permission from ref 80. Copyright 2010 Nature Publishing Group. (C) Adapted with permission from ref 53. Copyright 2009 American Association for the Advancement of Science. (D) Adapted with permission from ref 70. Copyright 2012 Elsevier.

using single-molecule fluorescent techniques.[77] Their results showed that after leaving the exit channel, the 5′ end of the RNA is initially free, but then the RNA reassociates with the base of the Pol II dock domain when it reaches about 26−29 nt in length. This type of interaction could prevent backtracking or otherwise allosterically modulate pausing.

In addition, most of the sequence-dependent data analysis has focused on the region immediately surrounding the transcription bubble. However, sequences already transcribed can significantly alter the dynamics of pausing via the secondary structure of the nascent RNA. Recent results show that templates that are richer in GC base pairs lead to fewer and shorter pauses than AT-rich templates for both the nuclear yeast Pol II as well as the mitochondrial Rpo41 (Figure 12C).[51] The difference disappears when the RNA transcript harbored by the enzyme is digested with RNase A, suggesting that GC-rich RNAs form stronger secondary structures that prevent pausing, most likely by limiting backtracking.

3.3. Transcription through the Nucleosome

In eukaryotes, the DNA is wrapped in nucleosomes, which act as mechanical barriers to the advancing polymerase. It has been shown that chromatin structure and organization constitute an important mechanism of control of gene expression.[78,79] As such, it is of great interest to establish what happens when a transcribing enzyme encounters a nucleosome. Conversely, it is of interest to understand the fate of histones during and after transcription.

We can propose two ways Pol II can overcome the nucleosomal barrier. In one scenario, the enzyme actively unwraps the nucleosomal DNA from the surface of the histone octamer to advance, and, in the other, the polymerase simply stops and waits for a spontaneous unwrapping of the nucleosomal DNA before it can advance. These different modes of interactions between the polymerase and the nucleosome have different implications for the regulation of gene expression.

Single-molecule experiments in which a nucleosome was placed downstream of the transcribing Pol II have shown that advancement of the enzyme through the barrier depends on a fine interplay between the enzyme dynamics (including pausing and translocation), and the fast dynamics of nucleosomal fluctuations.[53] Upon encountering a nucleosome, the pause durations and pause densities of Pol II increased, while the pause-free velocity was seen to decrease considerably (Figure 13A). Increased pause densities at the nucleosome are consistent with a model where Pol II cannot actively unwrap the nucleosomal DNA even when it is in an elongation competent state.[53] Instead, Pol II has to wait for nucleosomal fluctuations to allow its access to the downstream DNA. These interpretations are consistent with the observation that Pol II can only generate a maximum of ∼7.5 pN of force (stall force) before entering irrecoverable backtracks on bare DNA.[50] While mechanical pulling from both ends of the DNA requires a force of ∼8 pN to peel the DNA off the octamer surface (at 300 mM KCl),[70] wedging of a transcribing polymerase through a nucleosomal barrier is a different process as the enzyme only exerts force through one end of the nucleosomal DNA. Even *E. coli* RNAP, which has a stall force 3 times higher than yeast Pol II,[48] cannot peel the DNA from the octamer surface.[80] Thus, disrupting the histones−DNA interactions via this wedging mechanism requires an enzyme that can generate much higher force. In the presence of the nucleosome, Pol II pauses, and the

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

distribution of pause durations was similar to that on bare DNA, except that it shifted toward longer pauses. Accordingly, this distribution was well described by the same backtracking model previously used for bare DNA with one modification: recovery from backtracks is only possible when the DNA downstream of the enzyme is unwrapped (Figure 13B).[50,53] Because nucleosome fluctuations are fast relative to the rate of elongation and the rate of diffusion of the enzyme during backtracking, the rate of recovery from backtracks (k_f) is reduced by the probability of finding the nucleosome locally unwrapped. This modification explains the observed increase in the extent and duration of backtracks when the nucleosome is present.

The increase in backtracking of a polymerase upon encountering a nucleosome was directly shown in another single-molecule study,[80] although using the prokaryotic RNAP as a substitute for Pol II. In this case, the authors unzipped the two strands of DNA to find the position of the polymerase after allowing it to transcribe through the nucleosome for various amounts of time. By comparing the positions of polymerases on DNA from these single-molecule experiments with the distribution of RNA lengths obtained by running the labeled product of transcription on a gel, the authors conclude that the polymerase backtracks by about 15 bp on average when it encounters the nucleosome (Figure 13C). Consistent with this finding, it has been shown that the presence of a second trailing enzyme helps the leading polymerase overcome the nucleosome by preventing it from backtracking.[80] The leading polymerase also helps the trailing one transcribe through the nucleosome by preventing the latter from rewrapping, and thus providing a clear path for the second enzyme. A theoretical analysis of the mechanism through which a pair of polymerases can influence each other activities has appeared recently.[81] This mechanism could be at work in genes with high rates of transcription, where multiple polymerases would cooperate to transcribe through the nucleosome in a synergistic manner.

In vivo, transcription elongation is regulated by specialized machineries that modify or remodel nucleosomes. However, in many cases, it is unclear whether these modifications directly affect polymerase elongation, or indirectly regulate transcription through recruitment of other factors. A recent single-molecule study of transcription through modified nucleosomes showed that removal or mock acetylation of the histone tails has a modest effect on overall transcription elongation.[70] The number and durations of pauses for transcription through tailless and acetylated nucleosomes decreased only in the entry region of the nucleosome (Figure 13D), but remained unchanged in the central region, which constitutes the main barrier to Pol II. This finding suggests that the histone tails control the gate into the nucleosomal region for Pol II and possibly for other remodeling factors. In contrast, single amino-acid mutations of residues in histones H3 or H4 that make direct contact with the DNA near the nucleosome dyad dramatically decreased Pol II, pausing in the central region of the nucleosome (Figure 13D). These mutations, while not natural, highlight the contribution of histone−DNA contacts to the magnitude of the barrier in the central region, and suggest that a nucleosome binding- or remodeling-factor that can disrupt even a single one of these contacts can dramatically enhance the efficiency of elongation through the nucleosome.

What is the fate of histones during and after transcription? To answer this question, Hodges et al. applied force between the upstream DNA and the polymerase (Figure 8B) after the latter had transcribed through a nucleosome. These authors found that ∼60% of the histone cores remained associated to the DNA after the passage of the enzyme. Interestingly, it was found that this fraction reduced to less than 10% when a force of only 3−5 pN was applied during the passage of the enzyme through the histone core.[53] This observation suggests that the partially unwrapped histones contact the upstream DNA forming a (tension sensitive) loop that allows their transfer upstream of the elongation complexes (Figure 13E). A subsequent study using AFM to image elongation complexes at different points during transcription has provided additional support to this looping model.[59] Significantly, this AFM study showed that the fate of the transcribed histones (detachment, partial dissociation, or upstream transfer) depends on the elongation rates of the transcribing enzyme, making it possible to rationalize the observations made by different laboratories when the polymerase encounters the nucleosome. In some studies, a histone octamer has been reported to move upstream of its initial nucleosome positioning sequence;[82−84] in other studies, a mixture of octamers and hexamers was observed,[85,86] and, in yet other studies, complete detachment of the nucleosome from the template was reported.[84] Bintu et al. suggest that the outcome of the process depends on a kinetic competition among the elongation rate of the transcribing polymerase, the rate of octamer transfer behind the polymerase, and the rate of H2A−H2B dissociation from the octamer.[59] According to their interpretations, as Pol II advances through the nucleosome, the DNA is being detached from the histones. The octamer being a collection of positively charged proteins is unstable at salt concentrations under 1 M. Unless the histones contact another piece of DNA that neutralizes their charges and stabilizes their association, the octamer may dissociate with partial loss of its components. Initially, as the nucleosome partially unwraps during Pol II advancement, enough of the histone core is exposed to allow contact with the upstream DNA through a temporary DNA loop, but not so much as to cause H2A/H2B dissociation. During slow transcription (100 μM NTPs), this partially exposed histone intermediate lasts long enough to allow transfer of the intact octamer onto the upstream DNA. However, if the rate of transcription is increased slightly, more of the nucleosome will unwrap, and, as enough of the histone core becomes exposed, dimer dissociation starts competing with octamer transfer to the upstream DNA. Under these conditions, representative for transcription at 200 and 1000 μM NTPs, both octamers and hexamers can be found as a result of transcription. Finally, when the rates of transcription are even higher, enough DNA is unwrapped from the surface of the histone core that the complete histone detachment from DNA greatly outcompetes the rates of histone transfer and histone−histone dissociation, thus leading to bare DNA formation.

3.4. Transcription Factors That Modulate Elongation

In addition to the effects of DNA sequence, nascent RNA, and nucleosomes, many transcription factors regulate the elongation phase in the cell. Certain factors stimulate elongation, while other factors hinder it. Some also regulate the fidelity of transcription. Optical tweezers-based single-molecule studies have investigated the underlying mechanisms of a few of these elongation factors, GreA,[60] GreB,[60] NusG,[87] and NusA from *E. coli*,[88] and TFIIS from *S. cerevisiae*.[50]

Prokaryotic transcription factors GreA, GreB, and their eukaryotic functional homologue, TFIIS, have been identified

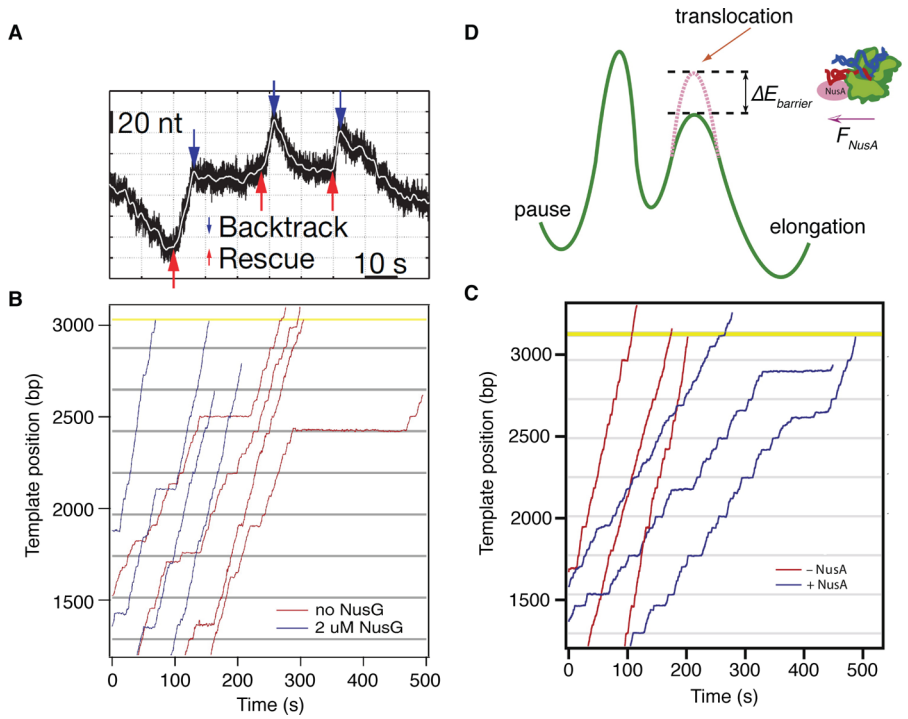


Figure 14. Effects of transcription factors on the elongation dynamics. (A) An example of cycles of backtracking and TFIIS rescue at 18 pN.[50] (B,C) Representative traces of transcription by RNAP along the ops repeat template in the absence of any transcription factors (red), in the presence of 2 μM NusG (blue in (B)),[87] or in the presence of 0.5 μM NusA (blue in (C)).[88] (D) Energy diagram illustrating the transcriptional modulation by NusA.[88] (A) Adapted with permission from ref 50. Copyright 2007 Nature Publishing Group. (B) Adapted with permission from ref 87. Copyright 2010 Elsevier. (C,D) Adapted with permission from ref 88. Copyright 2011 Elsevier.

to enhance transcriptional fidelity by stimulating cleavage of misincorporated nucleotides.[89−91] At the single-molecule level, both GreA and GreB were observed to decrease the frequency and duration of long transcriptional pauses. However, these two factors affect pauses differently. While GreB was found to decrease both the frequency and the duration of long pauses, GreA mainly decreased the frequency. In terms of backtracked distance, GreB prevents the rearward movement during a pause, whereas GreA does not.[60] These observations are consistent with the known activities of these two factors. GreA only stimulates the cleavage of 2 or 3 nt of backtracked RNA, thus rescuing the polymerase at the pause entry, but it is unable to completely eliminate pauses associated with longer backtracks.[92] In contrast, GreB stimulates the cleavage of larger fragments, thus completely eliminating the need of recovery from backtracks.[92] The presence of a nucleotide analogue ITP increases both the density and the durations of transcriptional pauses. GreA and GreB decrease the durations of the pauses substantially, while they slightly decrease the pause density.[60] These observations indicate that these factors aid the polymerase to exit from misincorporation-induced pauses without affecting pause entering, consistent with their known function in proofreading through the stimulation of the intrinsic cleavage activity of RNAP.

In eukaryotes, TFIIS like GreB stimulates the cleavage of the backtracked RNA transcript, and assists Pol II to recover faster from backtracks. In the absence of TFIIS, Pol II was only able to transcribe against an opposing load of 7.5 ± 2.0 pN, because its propensity to backtrack stalled the enzyme at high opposing loads. By rescuing the backtracked Pol II, TFIIS allowed the enzyme to transcribe up to a force of 16.9 ± 3.4 pN (Figure 14A).[50] Note that the stall force of the enzyme is highly dependent on the DNA template. This result is consistent with ensemble experiments that have shown that TFIIS can help Pol II transcribe against the nucleosome,[93,94] which, like an opposing force, also induces long backtracks of the enzyme.

Another *E. coli* transcription factor that is known to assist elongation is NusG, which associates with RNAP and increases the overall velocity of the process. Single-molecule assays of *E. coli* RNAP transcription elongation in the presence of NusG have shown that it increased the pause-free velocity of RNAP by 10−20% and simultaneously decreased its pause frequency (Figure 14B). As pausing kinetically competes with translocation, the observation that NusG decreased the pause frequency suggests that it shifts the equilibrium between the pre- and post-translocated state of the enzyme toward the latter. Interestingly, this result implies that translocation of the *E. coli* polymerase must be one of the rate-limiting steps of the kinetic cycle, just as has been recently shown for the eukaryotic enzyme.[58]

In contrast to the factors discussed above, optical tweezers experiment demonstrated that *E. coli* NusA slowed the pause-free velocity of RNAP and dramatically decreased the force where the pause-free velocity is half-maximal (Figure 14C) (without NusA 18 ± 2 pN of assisting load, with NusA 1 ± 4 pN of opposing load).[88] The effects of NusA are equivalent to exerting a hindering load of 19 ± 6 pN, which corresponds to an average energy barrier to forward translocation of 2.0 ± 0.4 k_BT, a magnitude equivalent to that of the intrinsic translocation barrier of Pol II.[58] NusA increases the pause frequency without affecting the duration of pauses. These observations are consistent with a model where NusA shifts the equilibrium of the translocation step toward the pre-translocated state (Figure 14D). In addition, the effects of NusA and NusG on pause

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

frequency vary for different engineered pause sequences. The mechanisms responsible for the sequence-dependent roles of NusA and NusG remain to be determined.

3.5. Transcription under Torsion

Another important regulator of transcription elongation in cells is DNA supercoiling. As RNAP transcribes the DNA template, it generates positively supercoiled DNA ahead of itself and negatively supercoiled DNA in its wake, which in turn affects the dynamics of the enzyme.[95] A recent study used a nanofabricated quartz cylinder held in an angular optical trap,[96] which can simultaneously control and measure rotation, torque, displacement, and force of the trapped cylinder to address the effect of DNA supercoiling on transcription elongation.[97] The experiments show that downstream torque (positive supercoil) decreases the pause-free velocity of the *E. coli* RNAP, while it increases the pause density and the duration of pauses. Elongation is halted when the enzyme transcribed against a downstream torque load of 11.0 ± 3.7 pN nm.[96] When the torsional stress is relaxed, the RNAP resumes transcription. Similarly, the enzyme could work against an average upstream torque (negative supercoil) of 10.6 ± 4.1 pN nm. The amount of torque required to melt DNA is ∼10 pN nm.[96,98] These results indicate that the *E. coli* RNAP can generate sufficient torque in its wake to melt the DNA. The torque so generated could regulate other processes, such as displacement of histones or other DNA-binding proteins that regulate transcription, or could influence transcription of other RNA polymerases on the same or opposite strand. In vivo, the interplay between topoisomerases and torque-generating processes, such as transcription, should ultimately regulate the torsional state of the DNA, which in turn should affect all DNA transactions.[99]

4. TRANSCRIPTION TERMINATION

Termination is the last stage of transcription wherein the elongation complex (EC) reaches the termination sequence, releases the nascent RNA, and the RNA polymerase disengages from the DNA template. It is known that bacterial and eukaryotic transcription termination mechanisms are quite different. Bacteria employ two termination pathways: intrinsic and Rho-dependent termination. During intrinsic termination (also called Rho-independent termination), a stem-loop hairpin encoded in the termination sequence causes the EC to dissociate. Formation of these RNA hairpin structures can be modulated by the binding of ligands such as adenine, thiamine pyrophosphate, or antiterminator RNA binding proteins. The other transcription termination mechanism involves Rho, a ring-shaped helicase that translocates along the mRNA and presumably pulls it out of the RNA−DNA hybrid inside the EC or induces a conformational change in the EC. Recently, both mechanisms of bacterial transcription termination have been studied using single-molecule techniques. Eukaryotic transcription termination mechanisms are more complex and less understood as compared to the bacterial counterpart. Therefore, we will mainly discuss bacterial transcription termination mechanisms in this section.

4.1. Rho-Dependent Termination

The Rho-dependent mechanism constitutes 20−50% of all bacterial RNA synthesis termination. Rho is a ∼50 kDa ring-shaped hexameric ATPase motor protein that translocates along the RNA and subsequently disrupts the RNA−DNA hybrid inside the EC to release the transcript. Rho contains two domains: the N-terminal RNA binding domain and the C-terminal RecA-like ATPase domain.[100] Rho-dependent termination starts with the binding of Rho to the RNA at a Rho utilization (rut) site, which is a C-rich region of 85−97 nt. Rho binds to the rut site with high affinity: $K_a \approx 10^{10}$ M^{-1}.[101,102] Once the RNA associates with the primary binding site of Rho (RNA binding domain), the motor translocates RNA through its center. The affinity of RNA to the secondary site (RecA-like ATPase domain) is relatively weak: $K_a \approx 4 \times 10^6$ M^{-1}.[101,103] After Rho reaches a paused RNAP, other transcription factors such as NusA and NusG also participate in the transcript release process.[104] NusA decreases the efficiency of Rho-dependent transcription termination at the *tiZ1* and *tiZ2* intragenic terminators, while NusG increases it.[104] Surprisingly, the roles of these factors in transcription termination are opposite from what one would expect from their effects in the elongation phase. Because NusA decreases the overall elongation velocity, we would expect it to enhance termination. Conversely, NusG increases efficiency of elongation; hence, we would expect it to reduce termination efficiency. It is unclear what mechanisms allow these factors to affect termination and elongation in such different manners.

Recently, the binding of Rho was examined using optical tweezers.[105] These experiments revealed that Rho binds 57 ± 2 nt of RNA in the primary site (9−10 nt per monomeric protein) and 28 ± 2 nt of RNA in the secondary site (Figure 15A). In addition, the data suggest that Rho translocation occurs by the "tethered tracking mechanism", in which Rho remains bound to the rut site while the motor threads the downstream RNA sequence in the 5′ to 3′ direction through the secondary binding site (Figure 15B).[102]

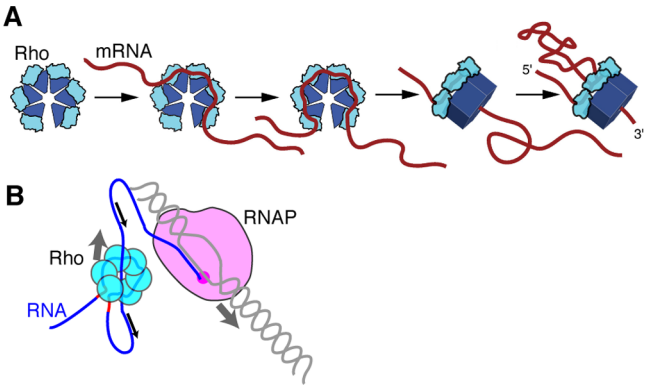


Figure 15. A model for Rho-dependent termination. (A) A general model of Rho binding to mRNA and translocation through the pore.[105] The N-terminal domain of Rho (cyan) associates with the RNA transcript, which is then passed into the center of the hexamer, allowing Rho to translocate downstream toward RNAP with a loop. (B) Proposed model for Rho-dependent termination. (A) Adapted with permission from ref 105. Copyright 2012 Elsevier.

However, many mechanistic details about Rho-dependent termination remain unknown. For instance, how fast does Rho translocate? How many bases does Rho translocate per ATP-hydrolysis? What are the roles of NusA and NusG during transcription termination? What is the role of Rho during EC dissociation? Future single-molecule studies of Rho will help elucidate these mechanistic details.

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

4.2. Intrinsic Termination

Intrinsic termination involves a nascent RNA transcript, which forms a stem loop structure followed by 7−9 nt of a U-rich RNA−DNA hybrid (Figure 16).[101,102] Larson et al. employed a

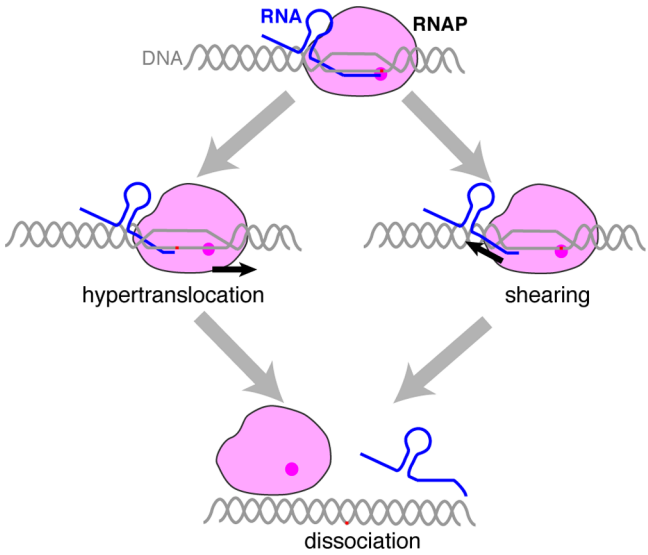


Figure 16. Intrinsic transcription termination. RNAP synthesizes the GC-rich RNA hairpin (blue loop), followed by U-rich segment, which forms a weak hybrid with DNA. RNAP can translocate forward along the DNA template without NTP incorporation in a process called hypertranslocation. Alternatively, the hairpin can destabilize the DNA−RNA hybrid either by shearing or by allosteric interactions. These destabilized structures induce the dissociation of the elongation complex.

single-molecule optical tweezers assay to study the mechanism of intrinsic transcription termination.[106] By quantifying the termination efficiency at different applied forces on either the DNA or the RNA, the authors characterized mechanisms of transcription in three representative bacterial intrinsic terminators (*his*, *t500*, and *tR2*). Transcription termination efficiencies at these sequences vary from 30% to 98% in bulk.[106] The termination efficiency for *his* and *tR2* terminators did not change with assisting or hindering loads applied on the DNA. This observation indicates that entering termination at these locations does not require translocation of the RNAP along the DNA. On the other hand, the *t500* terminator, which contains a U−A pair instead of a G−C pair at the base of the hairpin and two interruptions in the U tract, shows a force dependency in both the kinetics and the efficiency of termination in a manner corresponding to a forward translocation of the RNAP by ∼2.9 bp (Figure 16, hypertranslocation).

The integrity of the terminator RNA hairpin is required to induce the changes in the EC involved in termination. Results from force applied to the nascent RNA are consistent with this conclusion. The authors proposed a model in which the terminator hairpin closure induces shearing of the RNA−DNA hybrid to dissociate the elongation complex (Figure 16, shearing), but alternative mechanisms involving dissociation through an allosteric mechanism cannot be ruled out.

4.3. Effects of RNA Structure Dynamics on Intrinsic Transcription Termination

As we described above, the secondary structure of the nascent RNA can determine the fate of transcription termination. The control of transcription termination by RNA-binding proteins that modulate RNA structure is an important regulatory mechanism of gene expression in bacteria. Several antitermination proteins, including HutP, GlcT, and LicT, directly bind single-stranded RNA, and stabilize a competitive alternative secondary structure.[107−109] Using single-molecule FRET, the RNA hairpin structure fluctuations induced by binding of LicT and SacY antiterminator proteins from *Bacillus subtillis* were monitored.[110] These proteins prevent transcription termination by forming a shorter RNA hairpin (antiterminator RNA hairpin) that precludes the folding of the intrinsic RNA hairpin terminator (Figure 16). The strength of the antiterminator depends on the stability of the protein−RNA interactions. The length of the stem is the main determinant of the stability of these RNA hairpins.[110] These experiments were done on purified RNA strands; it will be interesting to study how these antiterminator proteins affect termination as the nascent RNA emerges from the polymerase one nucleotide at a time and folds cotranscriptionally. The effect of cotranscriptional folding of nascent RNA in termination was addressed in an independent study of adenine-sensitive riboswitch.

Riboswitches are regulatory segments of mRNA that bind a small molecule, and are involved in transcription termination, inhibition of translation, and splicing.[110,111] A riboswitch consists of two parts: an aptamer that binds a small molecule and an expression platform that regulates gene expression (Figure 17). A recent optical tweezers study monitored the

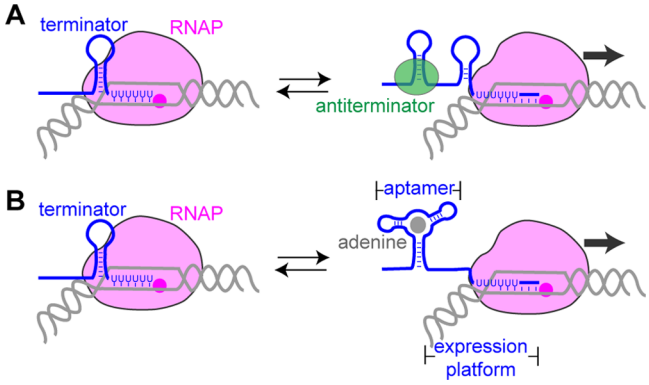


Figure 17. Regulation mechanisms of transcription termination by (A) antiterminator or (B) riboswitch. In these cases, the binding of either antiterminator protein or adenine inhibits the formation of the terminator hairpin and prevents transcription termination.

cotranscriptional folding of nascent RNA of an adenine-sensitive riboswitch.[112] In the absence of adenine, when the nascent transcript is held at a constant force below 7 pN, a small hairpin is seen to fold first, sequestering 21 ± 1 nt. This event is followed by the folding of the terminator hairpin, which sequesters 50 ± 1 nt, and releases the transcript. In the presence of adenine, a higher fraction of the polymerase transcribes through the termination sequence due to the formation of an alternative aptamer structure instead of the terminator hairpin. At forces higher than ∼13 pN, the unfolding force of this terminator hairpin, transcription termination efficiency decreases.[112] These experiments highlight the

dx.doi.org/10.1021/cr400730x | *Chem. Rev.* 2014, 114, 3203−3223

importance of competing, alternative secondary structures in the nascent RNA and their folding dynamics, as a basic mechanism through which small molecule effectors and transcription factors can control gene expression.

4.4. Eukaryotic Transcription Termination

As compared to transcription initiation and elongation, little is known about the mechanism of transcription termination in eukaryotes. Termination mechanisms differ among various RNA polymerases. RNA polymerase I (Pol I) terminates at a poly (T) stretch by binding of the NTS1 family of silencing proteins, Nsi 1[113] (see more details in the review by Németh et al.[114]). Pol III termination is induced by a poly (T) stretch alone. More details about the mechanism of Pol III termination can be found in work by Arimbasseri et al.[115] The transcription termination mechanism in Pol II is different from that of other polymerases. It involves phosphorylation of the C-terminal domain of Pol II and association of several transcription factors (more detailed information can be found in work by Mischo and Proudfoot[116]). Future studies, including single-molecule approaches, should shed light on the detailed mechanisms of transcription termination in eukaryotes.

5. CONCLUDING REMARKS

Transcription is a highly regulated process made up of distinct, strictly regulated steps. The emergence of single-molecule studies over the last two decades has provided many mechanistic insights into this process from the mechanisms by which polymerases locate and bind to promoters to molecular events that lead to the release of the RNA transcript from the complex. These studies have also begun to provide a coherent picture of the energy flow during the transcription cycle. In particular, the processes through which the energy released in the binding and hydrolysis of NTPs is converted into mechanical movement of the enzyme along the template and the generation of force and torque are starting to appear in sharper focus. As a result, scientists are now adding rich dynamic information derived from carefully designed single-molecule experiments to the increasing number of crystal structures depicting snapshots of RNA polymerases in different states of their kinetic cycle.

Although initially much information was derived from studies on bacterial transcription, more recent investigations have paid increasing attention to the behavior of the more complex eukaryotic process. Comparison between these two systems will surely provide important additional insights on the molecular mechanisms underlying the regulation of the three phases of transcription, and their evolution.

Single-molecule methods are maturing at a fast pace, making it possible to design and execute ever more complex experiments involving many different molecular actors. These advances should help scientists to obtain a more realistic picture of the complex dynamics and control of transcription resembling more closely those operating in vivo, without sacrificing the precision and quantitative description afforded by in vitro studies. Likewise, through the combination of single-molecule manipulation and single-molecule fluorescence methods in the same experiment, it should be possible to follow, for example, the internal dynamics of the polymerase or the binding of a regulatory factor and simultaneously monitor the mechanical variables of position, force, and torque. The result of these efforts will be a multidimensional picture of transcription that will provide crucial information about the relative timing of various molecular events and therefore reveal their causal connection.

At the end of the last century, a large gap existed between our knowledge of the structures responsible for the readout of the genetic information and our understanding of the functional and structural transitions of those structures. The difficulty to synchronize the individual trajectories of large numbers of molecules undergoing complex transitions in an attempt to arrive at a timed-ordered sequence of events was largely responsible for that gap. The advent of single-molecule methods has begun to fill this knowledge gap, and scientists can now confidently expect that future research using these methods will eventually yield a detailed "moving picture" of the molecular processes underlying transcription.

AUTHOR INFORMATION

Corresponding Author

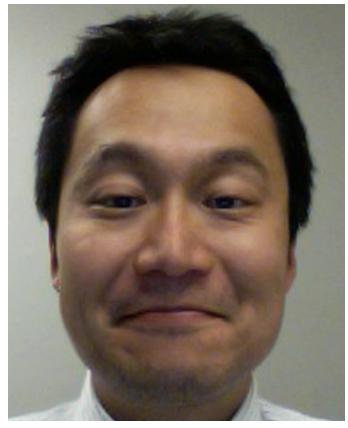
*E-mail: carlosb@berkeley.edu.

Notes

The authors declare no competing financial interest.

Biographies

Manchuta Dangkulwanich obtained her B.S. in Chemistry from Duke University, NC, in 2009. Through her undergraduate works on crystallography with Prof. Christian Raetz and biomolecule mass spectrometry with Prof. Michael C. Fitzgerald, she became interested in biophysical characterization and the inner work of the cell. She joined the laboratory of Prof. Carlos Bustamante at the University of California, Berkeley, where she characterized the kinetics of yeast RNA polymerase II transcription using optical tweezers.



Toyotaka Ishibashi received his B.S., M.S., and Ph.D. degrees from the department of applied biological sciences at Tokyo University of Sciences, Japan. He was a postdoctoral research fellow in the

laboratories of Prof. Juan Ausio at the University of Victoria, Canada, and Prof. Carlos Bustamante at the University of California, Berkeley. In 2013, he became an assistant professor at the Life Science Division at the Hong Kong University of Science and Technology. His research focuses on biophysical characterization of the nucleosomes including histone post-translational modifications and histone variants using single-molecule techniques.

Lacramioara Bintu obtained her B.S. in physics from Brandeis University, MA. During her undergraduate study, Prof. Jané Kondev and Prof. Rob Phillips introduced her to gene regulation and the exciting world of physical biology. She followed this philosophy of taking biology, turning it into numbers, and then back into biological insight as a graduate student in Prof. Carlos Bustamante's lab, at the University of California, Berkeley. There, she followed RNA polymerases in real time to tease out how they transcribe the DNA wrapped into nucleosomes. Currently, she is a postdoctoral research fellow in the laboratory of Prof. Michael Elowitz at the California Institute of Technology. Her research involves measuring the dynamics of decision-making in individual cells and determining the role that external factors, intrinsic noise, and the network architecture have on the fate of the cells.

Carlos J. Bustamante received his Ph.D. in Biophysics from his work with Prof. Ignacio Tinoco. After a short postdoctoral training at the Lawrence Berkeley National Laboratory, he started his own laboratory in 1982 at the University of New Mexico, where he became interested in the physical forces at work in biological processes. In 1991, he moved to the University of Oregon, where he continued to develop methods for single-molecule detection. He moved to the University of California at Berkeley in 1998, where he is currently a Raymond and Beverly Sackler Chair of Biophysics. His laboratory has applied novel methods of single-molecule detection, such as scanning force microscopy, and manipulation, such as optical and magnetic tweezers, to study the structure and function of nucleoprotein assemblies, and characterize the mechanochemical coupling of molecular motors. The biological processes of interest in his laboratory include DNA translocation, transcription, translation, protein folding, and protein degradation.

ACKNOWLEDGMENTS

We thank Drs. Shixin Liu, Troy A. Lionberger, and Liang Meng Wee for critical readings of the manuscript, and Prof. Eva Nogales for providing Figure 7 and helpful discussions. This work was supported in part by grants from the National Institutes of Health (R01-GM032543) to C.B. and the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Contract No. DE-AC02- 05CH11231 to C.B.

REFERENCES

(1) Temin, H. M.; Mizutani, S. *Nature* **1970**, *226*, 1211.
(2) Kærn, M.; Elston, T. C.; Blake, W. J.; Collins, J. J. *Nat. Rev. Genet.* **2005**, *6*, 451.
(3) Gruber, T. M.; Gross, C. A. *Annu. Rev. Microbiol.* **2003**, *57*, 441.
(4) Riggs, A. D.; Bourgeois, S.; Cohn, M. *J. Mol. Biol.* **1970**, *53*, 401.
(5) Berg, O. G.; Winter, R. B.; von Hippel, P. H. *Biochemistry (Moscow)* **1981**, *20*, 6929.
(6) Von Hippel, P. H.; Berg, O. G. *J. Biol. Chem.* **1989**, *264*, 675.
(7) Harada, Y.; Funatsu, T.; Murakami, K.; Nonoyama, Y.; Ishihama, A.; Yanagida, T. *Biophys. J.* **1999**, *76*, 709.
(8) Kabata, H.; Kurosawa, O.; Arai, I.; Washizu, M.; Margarson, S. A.; Glass, R. E.; Shimamoto, N. *Science* **1993**, *262*, 1561.
(9) Sakata-Sogawa, K.; Shimamoto, N. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14731.
(10) Guthold, M.; Zhu, X.; Rivetti, C.; Yang, G.; Thomson, N. H.; Kasas, S.; Hansma, H. G.; Smith, B.; Hansma, P. K.; Bustamante, C. *Biophys. J.* **1999**, *77*, 2284.
(11) Bustamante, C.; Guthold, M.; Zhu, X.; Yang, G. *J. Biol. Chem.* **1999**, *274*, 16665.
(12) Suzuki, Y.; Shin, M.; Yoshida, A.; Yoshimura, S. H.; Takeyasu, K. *FEBS Lett.* **2012**, *586*, 3187.
(13) Friedman, L. J.; Mumm, J. P.; Gelles, J. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 9740.
(14) Reitzer, L.; Schneider, B. L. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 422.
(15) Wang, F.; Redding, S.; Finkelstein, I. J.; Gorman, J.; Reichman, D. R.; Greene, E. C. *Nat. Struct. Mol. Biol.* **2013**, *20*, 174.
(16) Saecker, R. M.; Record, M. T., Jr.; Dehaseth, P. L. *J. Mol. Biol.* **2011**, *412*, 754.
(17) Davis, C. A.; Bingman, C. A.; Landick, R.; Record, M. T., Jr.; Saecker, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7833.
(18) Rivetti, C.; Guthold, M.; Bustamante, C. *EMBO J.* **1999**, *18*, 4464.
(19) Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S. *Science* **1994**, *265*, 1599.
(20) Roe, J. H.; Burgess, R. R.; Record, M. T., Jr. *J. Mol. Biol.* **1985**, *184*, 441.
(21) Cellai, S.; Mangiarotti, L.; Vannini, N.; Naryshkin, N.; Kortkhonjia, E.; Ebright, R. H.; Rivetti, C. *EMBO Rep.* **2007**, *8*, 271.
(22) Revyakin, A.; Ebright, R. H.; Strick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4776.
(23) Chakraborty, A.; Wang, D.; Ebright, Y. W.; Korlann, Y.; Kortkhonjia, E.; Kim, T.; Chowdhury, S.; Wigneshweraraj, S.; Irschik, H.; Jansen, R.; Nixon, B. T.; Knight, J.; Weiss, S.; Ebright, R. H. *Science* **2012**, *337*, 591.
(24) Friedman, L. J.; Gelles, J. *Cell* **2012**, *148*, 679.
(25) Revyakin, A.; Liu, C.; Ebright, R. H.; Strick, T. R. *Science* **2006**, *314*, 1139.
(26) Kapanidis, A. N.; Margeat, E.; Ho, S. O.; Kortkhonjia, E.; Weiss, S.; Ebright, R. H. *Science* **2006**, *314*, 1144.
(27) Straney, D. C.; Crothers, D. M. *J. Mol. Biol.* **1987**, *193*, 267.

(28) Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3746.
(29) Kapanidis, A. N.; Margeat, E.; Laurence, T. A.; Doose, S.; Ho, S. O.; Mukhopadhyay, J.; Kortkhonjia, E.; Mekler, V.; Ebright, R. H.; Weiss, S. *Mol. Cell* **2005**, *20*, 347.
(30) Margeat, E.; Kapanidis, A. N.; Tinnefeld, P.; Wang, Y.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S. *Biophys. J.* **2006**, *90*, 1419.
(31) Bar-Nahum, G.; Nudler, E. *Cell* **2001**, *106*, 443.
(32) Roeder, R. G. *Trends Biochem. Sci.* **1996**, *21*, 327.
(33) Chen, H.-T.; Hahn, S. *Cell* **2004**, *119*, 169.
(34) Miller, G.; Hahn, S. *Nat. Struct. Mol. Biol.* **2006**, *13*, 603.
(35) Chen, H.-T.; Warfield, L.; Hahn, S. *Nat. Struct. Mol. Biol.* **2007**, *14*, 696.
(36) Eichner, J.; Chen, H.-T.; Warfield, L.; Hahn, S. *EMBO J.* **2010**, *29*, 706.
(37) Fishburn, J.; Hahn, S. *Mol. Cell. Biol.* **2012**, *32*, 12.
(38) Murakami, K.; Elmlund, H.; Kalisman, N.; Bushnell, D. A.; Adams, C. M.; Azubel, M.; Elmlund, D.; Levi-Kalisman, Y.; Liu, X.; Gibbons, B. J.; Levitt, M.; Kornberg, R. D. *Science* **2013**, *342*, 1238724.
(39) He, Y.; Fang, J.; Taatjes, D. J.; Nogales, E. *Nature* **2013**, *495*, 481.
(40) Cisse, I. I.; Izeddin, I.; Causse, S. Z.; Boudarene, L.; Senecal, A.; Muresan, L.; Dugast-Darzacq, C.; Hajj, B.; Dahan, M.; Darzacq, X. *Science* **2013**, *341*, 664.
(41) Treutlein, B.; Muschielok, A.; Andrecka, J.; Jawhari, A.; Buchen, C.; Kostrewa, D.; Hög, F.; Cramer, P.; Michaelis, J. *Mol. Cell* **2012**, *46*, 136.
(42) Gnatt, A. L.; Cramer, P.; Fu, J.; Bushnell, D. A.; Kornberg, R. D. *Science* **2001**, *292*, 1876.
(43) Revyakin, A.; Zhang, Z.; Coleman, R. A.; Li, Y.; Inouye, C.; Lucas, J. K.; Park, S.-R.; Chu, S.; Tjian, R. *Genes Dev.* **2012**, *26*, 1691.
(44) Yudkovsky, N.; Ranish, J. A.; Hahn, S. *Nature* **2000**, *408*, 225.
(45) Mehta, A. D.; Rief, M.; Spudich, J. A.; Smith, D. A.; Simmons, R. M. *Science* **1999**, *283*, 1689.
(46) Davenport, R. J.; Wuite, G. J.; Landick, R.; Bustamante, C. *Science* **2000**, *287*, 2497.
(47) Forde, N. R.; Izhaky, D.; Woodcock, G. R.; Wuite, G. J.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11682.
(48) Wang, M. D.; Schnitzer, M. J.; Yin, H.; Landick, R.; Gelles, J.; Block, S. M. *Science* **1998**, *282*, 902.
(49) Yin, H.; Wang, M. D.; Svoboda, K.; Landick, R.; Block, S. M.; Gelles, J. *Science* **1995**, *270*, 1653.
(50) Galburt, E. A.; Grill, S. W.; Wiedmann, A.; Lubkowska, L.; Choy, J.; Nogales, E.; Kashlev, M.; Bustamante, C. *Nature* **2007**, *446*, 820.
(51) Zamft, B.; Bintu, L.; Ishibashi, T.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 8948.
(52) Abbondanzieri, E. A.; Greenleaf, W. J.; Shaevitz, J. W.; Landick, R.; Block, S. M. *Nature* **2005**, *438*, 460.
(53) Hodges, C.; Bintu, L.; Lubkowska, L.; Kashlev, M.; Bustamante, C. *Science* **2009**, *325*, 626.
(54) Larson, M. H.; Zhou, J.; Kaplan, C. D.; Palangat, M.; Kornberg, R. D.; Landick, R.; Block, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 6555.
(55) Neuman, K. C.; Abbondanzieri, E. A.; Landick, R.; Gelles, J.; Block, S. M. *Cell* **2003**, *115*, 437.
(56) Guajardo, R.; Sousa, R. *J. Mol. Biol.* **1997**, *265*, 8.
(57) Kireeva, M.; Kashlev, M.; Burton, Z. F. *Biochim. Biophys. Acta* **2010**, *1799*, 389.
(58) Dangkulwanich, M.; Ishibashi, T.; Liu, S.; Kireeva, M. L.; Lubkowska, L.; Kashlev, M.; Bustamante, C. J. *eLife* **2013**, *2*, e00971.
(59) Bintu, L.; Kopaczynska, M.; Hodges, C.; Lubkowska, L.; Kashlev, M.; Bustamante, C. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1394.
(60) Shaevitz, J. W.; Abbondanzieri, E. A.; Landick, R.; Block, S. M. *Nature* **2003**, *426*, 684.
(61) Komissarova, N.; Kashlev, M. *J. Biol. Chem.* **1997**, *272*, 15329.
(62) Komissarova, N.; Kashlev, M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1755.
(63) Weixlbaumer, A.; Leon, K.; Landick, R.; Darst, S. A. *Cell* **2013**, *152*, 431.
(64) Herbert, K. M.; La Porta, A.; Wong, B. J.; Mooney, R. A.; Neuman, K. C.; Landick, R.; Block, S. M. *Cell* **2006**, *125*, 1083.
(65) Sydow, J. F.; Brueckner, F.; Cheung, A. C.; Damsma, G. E.; Dengl, S.; Lehmann, E.; Vassylyev, D.; Cramer, P. *Mol. Cell* **2009**, *34*, 710.
(66) Toulokhonov, I.; Zhang, J.; Palangat, M.; Landick, R. *Mol. Cell* **2007**, *27*, 406.
(67) Artsimovitch, I.; Landick, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7090.
(68) Herbert, K. M.; Greenleaf, W. J.; Block, S. M. *Annu. Rev. Biochem.* **2008**, *77*, 149.
(69) Depken, M.; Galburt, E. A.; Grill, S. W. *Biophys. J.* **2009**, *96*, 2189.
(70) Bintu, L.; Ishibashi, T.; Dangkulwanich, M.; Wu, Y. Y.; Lubkowska, L.; Kashlev, M.; Bustamante, C. *Cell* **2012**, *151*, 738.
(71) Kireeva, M. L.; Kashlev, M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 8900.
(72) Bai, L.; Shundrovsky, A.; Wang, M. D. *J. Mol. Biol.* **2004**, *344*, 335.
(73) Tadigotla, V. R.; Maoileidigh, D. O.; Sengupta, A. M.; Epshtein, V.; Ebright, R. H.; Nudler, E.; Ruckenstein, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4439.
(74) Shundrovsky, A.; Santangelo, T. J.; Roberts, J. W.; Wang, M. D. *Biophys. J.* **2004**, *87*, 3945.
(75) Toulokhonov, I.; Landick, R. *Mol. Cell* **2003**, *12*, 1125.
(76) Toulokhonov, I.; Artsimovitch, I.; Landick, R. *Science* **2001**, *292*, 730.
(77) Andrecka, J.; Lewis, R.; Brückner, F.; Lehmann, E.; Cramer, P.; Michaelis, J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 135.
(78) Hartzog, G. A.; Speer, J. L.; Lindstrom, D. L. *Biochim. Biophys. Acta* **2002**, *1577*, 276.
(79) Kornberg, R. D.; Lorch, Y. *Cell* **1999**, *98*, 285.
(80) Jin, J.; Bai, L.; Johnson, D. S.; Fulbright, R. M.; Kireeva, M. L.; Kashlev, M.; Wang, M. D. *Nat. Struct. Mol. Biol.* **2010**, *17*, 745.
(81) Galburt, E. A.; Parrondo, J. M. R.; Grill, S. W. *Biophys. Chem.* **2011**, *157*, 43.
(82) Clark, D. J.; Felsenfeld, G. *Cell* **1992**, *71*, 11.
(83) Studitsky, V. M.; Clark, D. J.; Felsenfeld, G. *Cell* **1994**, *76*, 371.
(84) Studitsky, V. M.; Kassavetis, G. A.; Geiduschek, E. P.; Felsenfeld, G. *Science* **1997**, *278*, 1960.
(85) Kireeva, M. L.; Walter, W.; Tchernajenko, V.; Bondarenko, V.; Kashlev, M.; Studitsky, V. M. *Mol. Cell* **2002**, *9*, 541.
(86) Kulaeva, O. I.; Hsieh, F.-K.; Studitsky, V. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11325.
(87) Herbert, K. M.; Zhou, J.; Mooney, R. A.; Porta, A. L.; Landick, R.; Block, S. M. *J. Mol. Biol.* **2010**, *399*, 17.
(88) Zhou, J.; Ha, K. S.; La Porta, A.; Landick, R.; Block, S. M. *Mol. Cell* **2011**, *44*, 635.
(89) Erie, D. A.; Hajiseyedjavadi, O.; Young, M. C.; von Hippel, P. H. *Science* **1993**, *262*, 867.
(90) Jeon, C.; Agarwal, K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13677.
(91) Thomas, M. J.; Platas, A. A.; Hawley, D. K. *Cell* **1998**, *93*, 627.
(92) Borukhov, S.; Sagitov, V.; Goldfarb, A. *Cell* **1993**, *72*, 459.
(93) Kireeva, M. L.; Hancock, B.; Cremonaqq, G. H.; Walter, W.; Studitsky, V. M.; Kashlev, M. *Mol. Cell* **2005**, *18*, 97.
(94) Luse, D. S.; Spangler, L. C.; Újvári, A. *J. Biol. Chem.* **2011**, *286*, 6040.
(95) Liu, L. F.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7024.
(96) Ma, J.; Bai, L.; Wang, M. D. *Science* **2013**, *340*, 1580.
(97) Deufel, C.; Forth, S.; Simmons, C. R.; Dejgosha, S.; Wang, M. D. *Nat. Methods* **2007**, *4*, 223.
(98) Bryant, Z.; Stone, M. D.; Gore, J.; Smith, S. B.; Cozzarelli, N. R.; Bustamante, C. *Nature* **2003**, *424*, 338.
(99) Champoux, J. J. *Annu. Rev. Biochem.* **2001**, *70*, 369.
(100) Skordalakes, E.; Berger, J. M. *Cell* **2003**, *114*, 135.

(101) Kim, D. E.; Patel, S. S. *J. Biol. Chem.* **2001**, *276*, 13902.
(102) Richardson, J. P. *J. Biol. Chem.* **1982**, *257*, 5760.
(103) Richardson, J. P. *Biochim. Biophys. Acta* **2002**, *1577*, 251.
(104) Burns, C. M.; Richardson, L. V.; Richardson, J. P. *J. Mol. Biol.* **1998**, *278*, 307.
(105) Koslover, D. J.; Fazal, F. M.; Mooney, R. A.; Landick, R.; Block, S. M. *J. Mol. Biol.* **2012**, *423*, 664.
(106) Larson, M. H.; Greenleaf, W. J.; Landick, R.; Block, S. M. *Cell* **2008**, *132*, 971.
(107) Gopinath, S. C. B.; Balasundaresan, D.; Kumarevel, T.; Misono, T. S.; Mizuno, H.; Kumar, P. K. R. *Nucleic Acids Res.* **2008**, *36*, 3463.
(108) Langbein, I.; Bachem, S.; Stülke, J. *J. Mol. Biol.* **1999**, *293*, 795.
(109) Schnetz, K.; Stülke, J.; Gertz, S.; Krüger, S.; Krieg, M.; Hecker, M.; Rak, B. *J. Bacteriol.* **1996**, *178*, 1971.
(110) Clerte, C.; Declerck, N.; Margeat, E. *Nucleic Acids Res.* **2013**, *41*, 2632.
(111) Serganov, A.; Patel, D. J. *Annu. Rev. Biophys* **2012**, *41*, 343.
(112) Frieda, K. L.; Block, S. M. *Science* **2012**, *338*, 397.
(113) Lang, W. H.; Reeder, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9781.
(114) Németh, A.; Perez-Fernandez, J.; Merkl, P.; Hamperl, S.; Gerber, J.; Griesenbeck, J.; Tschochner, H. *Biochim. Biophys. Acta* **2013**, *1829*, 306.
(115) Arimbasseri, A. G.; Rijal, K.; Maraia, R. J. *Biochim. Biophys. Acta* **2013**, *1829*, 318.
(116) Mischo, H. E.; Proudfoot, N. J. *Biochim. Biophys. Acta* **2013**, *1829*, 174.
(117) Murakami, K. S.; Darst, S. A. *Curr. Opin. Struct. Biol.* **2003**, *13*, 31.