doi:10.1016/j.jmb.2011.01.024 J. Mol. Biol. (2011) 412, 814–831

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

REVIEW

Single-Molecule Studies of Transcription: From One RNA Polymerase at a Time to the Gene Expression Profile of a Cell

Feng Wang[1] and Eric C. Greene[1,2*]

[1]*Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA*
[2]*Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA*

Received 1 December 2010; received in revised form 5 January 2011; accepted 8 January 2011
Available online 19 January 2011

Edited by R. Ebright

Keywords:
single molecule;
transcription;
RNA polymerase

Single-molecule techniques have emerged as powerful tools for deciphering mechanistic details of transcription and have yielded discoveries that would otherwise have been impossible to make through the use of more traditional biochemical and/or biophysical techniques. Here, we provide a brief overview of single-molecule techniques most commonly used for studying RNA polymerase and transcription. We then present specific examples of single-molecule studies that have contributed to our understanding of key mechanistic details for each different stage of the transcription cycle. Finally, we discuss emerging single-molecule approaches and future directions, including efforts to study transcription at the single-molecule level in living cells.

Introduction

Transcription is the key step of gene expression and regulation in which the information encoded in genomic DNA is transcribed into RNA. The products of transcription can be messenger RNA (mRNA), ribosomal RNA (rRNA) or transfer RNA (tRNA), or other types of RNA molecules such as ribozymes. In addition, many of these different RNA molecules are processed to generate an increasingly diverse array of small RNA species which themselves can influence gene expression.[1–3] Numerous cis- and trans-acting factors work together to establish a complex network of regulatory features, allowing precise control over the expression of any given gene.[4–6] This regulation is achieved through the combined effects of promoter DNA sequences that dictate the sites of transcript initiation, along with the effects of a multitude of transcription factors and other regulatory elements that can influence the efficiency of transcript initiation, elongation, and/or termination. In eukaryotes, transcription regulation is further complicated by the higher-order organization of chromatin structure, with the positioning of nucleosomes and establishment of repressive chromatin structures greatly influencing the organization and regulation of gene expression.[6–9] At the heart of this regulatory network lies RNA polymerase (RNAP), which is the protein machinery directly responsible for RNA synthesis. The simplest RNAPs come from bacteriophages and consist of single polypeptides capable of carrying out all the basic steps of transcription.[10] In prokaryotes, a single multisubunit RNAP is responsible for all RNA production,

*Corresponding author. E-mail address: ecg2108@columbia.edu.

Abbreviations used: RNAP, RNA polymerase; Pol I, polymerase I; rNTP, ribonucleotide triphosphate; spFRET, single-pair fluorescence resonance energy transfer; FRET-ALEX, FRET with alternating-laser excitation; GFP, green fluorescent protein; FISH, fluorescence *in situ* hybridization; YFP, yellow fluorescent protein; TPM, tethered particle motion; AFM, atomic force microscopy; EC, elongation complex.

whereas in eukaryotes many different types of RNAP [e.g., polymerase I (Pol I), II, III] divide the labor required for production of different classes of RNA molecules.[11–13]

Before transcription can begin, RNAP must locate the promoter sequence that lies upstream from the gene that is to be expressed in a process that can be referred to as the promoter search. Following the promoter search, transcription itself can be divided into three stages: initiation, elongation, and termination; and all RNAPs, regardless of their origin, must complete the same basic set of reactions in order to generate RNA transcripts. The transcriptional machinery first unwinds the template DNA before beginning RNA synthesis, and this process is referred to as open complex formation. Bacteriophage and prokaryotic RNAP holoenzymes can conduct all of these steps with no need for accessory factors, whereas eukaryotes and archaea require a large preinitiation complex composed of RNAP and several additional transcription factors.[14–16] After open complex formation, RNAP undergoes abortive initiation, wherein it synthesizes many short transcripts ~9–11 nt in length, until it finally escapes the promoter and begins elongating the RNA.[17–19] During elongation, RNAP translocates along the template DNA while catalyzing successive addition of ribonucleotides [ribonucleotide triphosphates (rNTPs)] to the growing RNA chain. RNAP is highly processive during elongation, but its forward motion is not monotonic; rather, it exhibits frequent pauses and backtracking,[20,21] which are often coupled to proofreading mechanisms necessary to ensure fidelity.[21,22] Once a gene is transcribed, transcription must be terminated. *Escherichia coli* RNAP terminates transcription through two distinct mechanisms, either intrinsic termination or rho-dependent termination.[23,24] During intrinsic termination, a hairpin formed in the nascent RNA destabilizes the elongation complex (EC). In rho-dependent termination, rho disrupts transcribing RNAP through a process coupled to ATP-dependent translocation along the nascent transcript. Termination in eukaryotes is less well understood and is coupled to 3′ end processing of the transcript; however, the underlying processes may share some mechanistic similarities with *E. coli* RNAP. It has been proposed that either a polyadenylation sequence in the RNA changes the factors associated with the polymerase, making it less processive, or some "rho-like" helicase may bind the 5′ RNA end generated by the cleavage at the polyadenylation sequence.[25]

RNAP is the most important component of the transcription apparatus, and much of the existing knowledge regarding the mechanisms by which it functions have been garnered from single-molecule studies. The power of these studies lies in their ability to observe and measure individual molecules in real time, thereby eliminating the need for ensemble averaging and allowing direct detection of rare or transient intermediates within heterogeneous populations of molecules. In addition, some single-molecule techniques can physically manipulate individual molecules of RNAP, offering the ability to study the response of RNAP to externally applied forces as well as measure the forces that RNAP is able to exert. In this review, we highlight a number of critical single-molecule transcription studies, with emphasis placed on the latest progress in the field as well as future avenues of research that will push forward our understanding of transcription.

Single-Molecule Techniques for Studying Transcription

There are several categories of techniques used in the single-molecule studies of transcription. Here we present a brief overview of the most prevalent single-molecule techniques that have been -used to study transcription.

Fluorescence-based assays

Fluorescence-based assays are among the most widely used single-molecule techniques. Examples include single-pair fluorescence resonance energy transfer (spFRET) experiments for measuring distances and conformational rearrangements of RNAP, wide-field microscopy to watch single fluorescent molecules of RNAP as they bind to and move along individual strands of DNA, as well as more recent developments allowing *in vivo* imaging of fluorescently tagged transcripts or transcription factors.

spFRET is a powerful tool that allows the relative distance between a fluorescence donor and a fluorescence acceptor to be determined in real time (Fig. 1a).[31–33] spFRET experiments require the covalent attachment of organic donor and acceptor fluorophore pairs to the molecules of interest. By exciting the donor and measuring the energy transfer efficiency to the acceptor, spFRET can effectively report intra- and intermolecular distances ranging between 20 and 100 Å. spFRET has been used to study the conformation and dynamics of both *E. coli* and T7 RNAP.[31,33] A recent modification of this technique is FRET with alternating-laser excitation (FRET-ALEX). By illuminating the sample with alternating donor and acceptor excitation, FRET-ALEX provides direct information on both the presence of the acceptor dye and the distance between the donor and the acceptor pairs.[34]

Direct-labeling of individual RNAP molecules with fluorescent tags and observing their movement in a wide field is also an effective way to study transcription (Fig. 1b), which allows many individual RNAP molecules to be tracked simultaneously with millisecond temporal resolution and nanometer

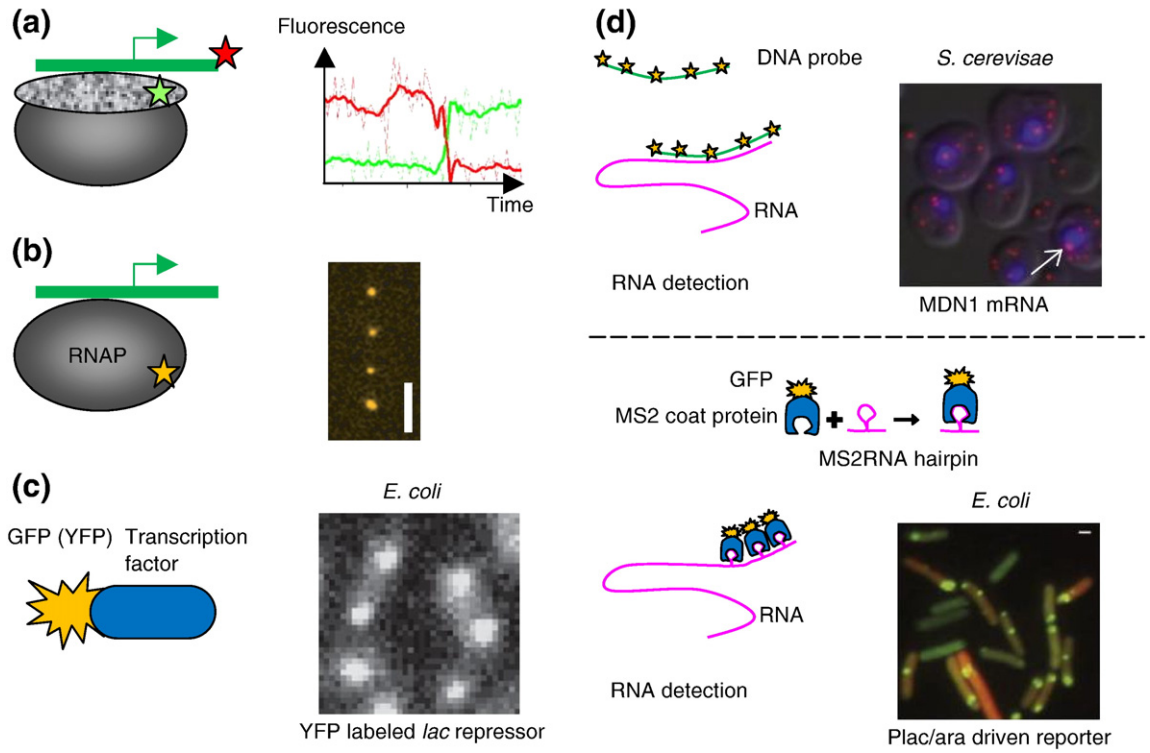

Fig. 1. Fluorescence-based single-molecule imaging. (a) FRET can monitor the distance change between a fluorescence donor (green star) and an acceptor (red star). A typical fluorescence intensity trace of the donor and acceptor is shown on the right panel (reprinted with permission from Ref. 26). (b) Wide-field fluorescence can track many individual fluorescently labeled transcription complexes. An image of rhodamine-labeled T7 RNAP molecules bound along a DNA template is shown on the right panel; the scale bar represents 5 μm (reprinted with permission from Ref. 27). (c) Fluorescent protein fusion technique is used to track transcription factor(s) of interest *in vivo*. An image of YFP-lac fusion protein in *E. coli* cells is shown on the right panel (with permission from Ref. 28). (d) FISH in fixed cells (top) and GFP-tagged RNA-binding protein (bottom) are used to detect single RNA molecules *in vivo*. Images of MDN1 mRNA molecules in *S. cerevisae* detected by FISH (top right) and mRNA molecules of a reporter gene in *E. coli* detected by GFP-MS2 system (bottom right) are shown (with permission from Refs. 29 and 30).

spatial resolution. For example, some of the earliest reported single-molecule studies of transcription involved the use of wide-field fluorescence microscopy to visualize fluorescently tagged RNAP as it interacted with "belts" of DNA.[35] More recent studies have observed fluorescent molecules of T7 RNAP as it slides along DNA,[27] and our group has recently reported the use of DNA curtains to visualize promoter binding by quantum-dot-tagged *E. coli* RNAP.[36] These approaches also offer the potential for multicolor detection, which could permit concurrent observation of fluorescently tagged RNAP as well as the template DNA, transcript RNA, and/or associated transcription factors.

Most single-molecule approaches are confined to *in vitro* measurements of highly purified molecules. However, a few laboratories are pushing the boundaries of these technologies and have begun developing methods to visualize transcription at the single-molecule level *in vivo*. These include using green fluorescent protein (GFP), or one of its variants, fused to the protein(s) of interest (Fig. 1c), fluorescence *in situ* hybridization (FISH) for transcript detection in fixed cells, and labeling transcripts with a GFP-tagged RNA-binding protein (Fig. 1d). For example, Elf *et al.* fused yellow fluorescent protein (YFP) to *lac* repressor (LacI) and observed its dynamics in a living *E. coli* cell.[28] In FISH, fluorescently tagged oligonucleotides are hybridized to transcripts of interest, and multiple positions along the mRNA can be labeled to increase signal.[29,37,38] Using probes labeled with different dyes, FISH can also be used to measure the expression levels of multiple genes within the same cell. However, FISH only provides a snapshot of mRNA abundance, and the cells must be fixed prior to probe hybridization. For visualizing mRNA in living cells, transcripts can be labeled with RNA-binding protein MS2. In the MS2 system, multiple MS2 RNA hairpins are inserted into the mRNA of interest and labeled with a coexpressed GFP-MS2 fusion protein.[30] The GFP-MS2 system can detect single mRNA molecules in living cells and provides a powerful tool for studying *in vivo* transcription.

Bead-based assays

Most bead-based assays fall into one of two general categories, those in which the bead is used only as a visual marker for monitoring a reaction, such as tethered particle motion (TPM), and those in which the bead is used as a handle to physically manipulate the transcription apparatus, such as optical traps.

TPM was the first experimental setup used to observe transcription with single-molecule sensitivity (Fig. 2a).[39] For TPM, a small bead is generally tethered to a DNA template that is held by an RNAP molecule immobilized on a solid supporting surface. The bead exhibits confined Brownian motion that can be monitored by optical microscopy, and the magnitude of the Brownian fluctuations provides a direct readout of the DNA length between the bead and the immobilized RNAP. The motion of the tethered bead is used to monitor the length change in the DNA due to movement of RNAP as it begins transcribing the template.[39,41,42] TPM has also been combined with optical or magnetic traps (see below) to improve the spatial resolution, and smaller fluorescent particles have also been tethered to the bead in order to detect its rotation, which in turn reports on the rotational movement of RNAP as it follows the twist of the DNA helix.[42,43]

Optical traps are another category of bead-based assays that have yielded a tremendous amount of information regarding the chemomechanical behaviors of RNAP (Fig. 2b). In a typical optical trap, a laser is focused onto a bead, causing the bead to become trapped at the focal point of the beam. Optical traps can be used to apply force by moving the bead, and the force exerted by transcribing RNAP can be measured by monitoring the bead displacement from the trapping center. In a single-

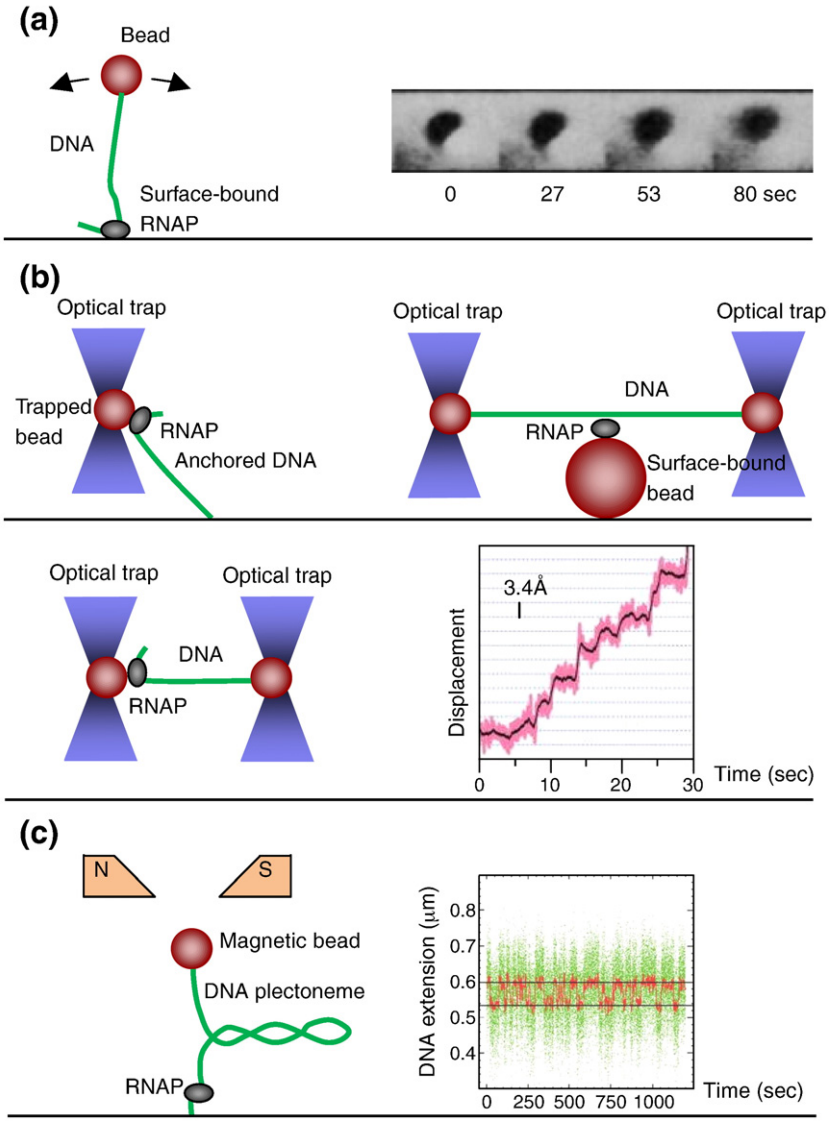

Fig. 2. Bead-based single-molecule techniques. (a) TPM assay. The Brownian motion of the tethered bead changes (the image of the bead blurs from 0 to 80 s, right panel; with permission from Ref. 39) when the RNAP translocates along the DNA increasing the length of DNA between the bead and RNAP. (b) Optical trap. Optical trapping is used to track the motion of the trapped bead or apply force to the bead. In a one-bead optical trap, the RNAP is bound to a bead and the DNA is anchored to a surface or a micropipette. In a two-bead optical trap, the transcription complex is lifted from the surface by two optical traps to reduce any vibration from the surface. In a three-bead optical trap, the RNAP is bound to a surface-attached bead. An RNAP translocation trace with angstrom precision obtained using a two-bead optical trap is shown (reprinted with permission from Ref. 40). (c) Magnetic traps. A magnetic bead is attached to one end of the DNA template through multiple bonds. The topology change (plectoneme formation or elimination) of the DNA due to unwinding by RNAP is monitored by the height of the bead. A typical trace of the magnetic bead in a magnetic trap is shown on the right panel (reprinted with permission from Ref. 32).

bead setup, DNA is tethered to the surface of a slide or a micropipette at one end, and RNAP is attached to the bead, allowing the movement of RNAP during transcription to be tracked with nanometer precision (Fig. 2b).[44,45] In a two-bead setup, the second end of the DNA is tethered to another bead in a second optical trap, which greatly reduces experimental noise from the reference surface, enabling single-base-pair spatial resolution (Fig. 2b).[40,46–50] In a three-bead setup, an RNAP-bound bead is anchored on a surface, and a DNA template with a bead at either end is brought to the anchored RNAP (Fig. 2b).[51] During transcription, the downstream bead moves toward RNAP, enabling accurate measurement of RNAP movement. An even more recent breakthrough in this field is the development of an angular optical trap, which can apply and detect both torque and force.[52] As DNA is normally supercoiled *in vivo*, this remarkable new technique will prove invaluable for studying transcription on supercoiled DNA molecules.

Magnetic traps are another type of bead-based assay, which can be used to exert and measure torque and are ideally suited for the study of DNA topology (Fig. 2c). The DNA template is usually attached to the chamber surface by one end and a magnetic bead by the other end through multiple bonds. This rotationally constrained attachment allows the DNA topology to be manipulated simply by rotating magnets located above the sample. Some supercoiling is usually introduced into the DNA to form plectonemes, and the introduction of plectonemes serves to amplify small local conformation changes in the DNA, such as those induced by the binding of a protein, by converting them into large changes in DNA length due to alterations in supercoiling. For example, RNAP unwinding a promoter has been observed by a magnetic trap via the supercoiling in the rest of the DNA template.[53]

Atomic force microscopy

Atomic force microscopy (AFM) uses a cantilever bearing a sharp tip (probe) at its end to scan the three-dimensional (3D) topology of a specimen surface. When the tip is brought into proximity of a sample surface, deflection of the cantilever is detected by a laser beam reflected from the cantilever to a quadrant photodiode (Fig. 3). AFM does not require the sample to be labeled with a bead or fluorophore and can work in liquid under physiological conditions. AFM has been used to directly observe movement of RNAP along DNA and promoter binding.[54,55] One limitation of AFM is the relatively slow speed (five to six frames per second) at which it can be used to probe a sample. However, recent technical developments in this field have led to high-speed AFM, which offers the potential for probing 3D topologies of reactions at acquisition rates of up to 30 frames per second.[56] Nevertheless, AFM studies are limited by the need to adsorb the molecules under investigation to a solid supporting surface (typically mica), and there is also the potential for disrupting the molecules being studied as a consequence of the physical interactions between the AFM probe and the sample.

Single-Molecule Studies of RNAP throughout the Transcription Cycle

Below we describe events throughout the transcription cycle, including initiation, elongation, and termination, and highlight examples of single-molecule studies that have contributed to our understanding of each of the events. Much of this section focuses on *E. coli* RNAP, which remains the most intensively studied RNAP.

The promoter search

Before a transcript can be made, RNAP must find appropriate promoter sequences. This promoter search problem is the molecular equivalent of "trying to find a needle in a haystack." For example, *E. coli* has on the order of ~3000 promoters, each ~35 bp in length.[57,58] Thus *E. coli* RNAP must locate a subset of sequences comprising <2% of the *E. coli* genome, and this promoter search problem may be one of the least understood aspects of transcription.

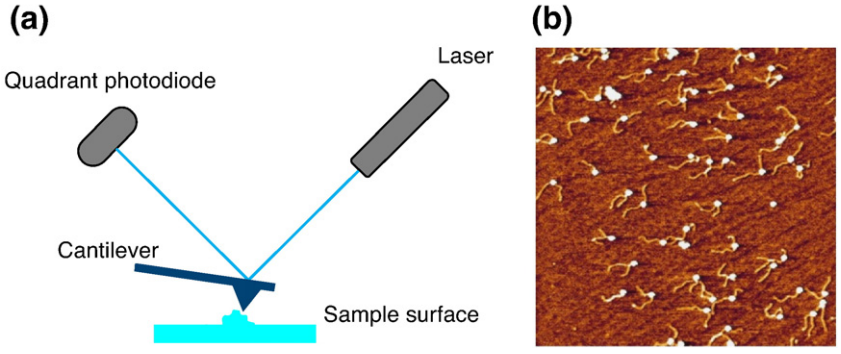

Fig. 3. Atomic force microscopy. (a) Diagram of AFM setup. (b) AFM image of open complex formed with *E. coli* RNAP and a DNA template containing the P_R–P_{RM} promoter of lambda phage (reprinted with permission from Ref. 54).

There are four potential mechanisms that might facilitate promoter search, and these four mechanisms are not mutually exclusive: (i) 1D hopping, where the protein moves along the same molecule of DNA via a series of microscopic dissociation and rebinding events; (ii) 1D sliding, which involves a random walk along the DNA without dissociation; (iii) intersegment transfer, involving movement from one site to another via a looped intermediate; and (iv) 3D diffusion (or jumping), where the protein moves over longer distances via dissociation and then rebinding at a distal location (Fig. 4).[59,60] Theoretical studies have shown that short-range 1D hopping or sliding can speed up target search by a DNA-binding protein beyond the limit imposed by simple 3D diffusion.[61] Bulk studies suggested that RNAP can move along DNA by 1D sliding;[62,63] Kabata *et. al.* observed the first single-molecule evidence of RNAP sliding using fluorescently labeled *E. coli* RNAP and reported that RNAP could slide several micrometers along DNA in the presence of a buffer flow.[35] Long-distance diffusion was not detected in a later study by Harada *et. al.*, where only 2.6% of the observed RNAP molecules (10 of 381) exhibited 1D diffusion detectable above instrument resolution limits ($\geq$0.2 μm).[64] These authors concluded that the RNAP used a 1D diffusion-based mechanism to search for promoters and reported a 1D diffusion coefficient (D_1) of 10^{-2} μm^2 s^{-1} with mean sliding distance (L_{sl}) of 300 bp. 1D diffusion has also been observed for fluorescently tagged T7 RNAP with D_1 of $\sim 10^{-1}$ μm^2 s^{-1} and L_{sl} in the micrometer range.[27] Another study by Guthold *et al.* used AFM to image RNAP bound to nonspecific DNA adsorbed onto a mica surface and showed that RNAP was able to slide along DNA while undergoing a random 1D walk.[55] On the basis of these AFM measurements, the authors reported D_1 of 10^{-5} μm^2 s^{-1} and a lifetime of 600 s. Interestingly, the 1D diffusion coefficients determined from single measurements are all 1–3 orders of magnitude lower than the values inferred in the earlier bulk studies.[62,65] On the basis of these cumulative studies, it had been largely accepted that RNAP searches for promoters by means of a 1D scanning mechanism. However, this conclusion has been challenged in the literature, and no promoter association rate has ever been reported that is higher than the limit that would be imposed by 3D diffusion ($\sim 10^8$–10^9 M^{-1} s^{-1}), suggesting that the

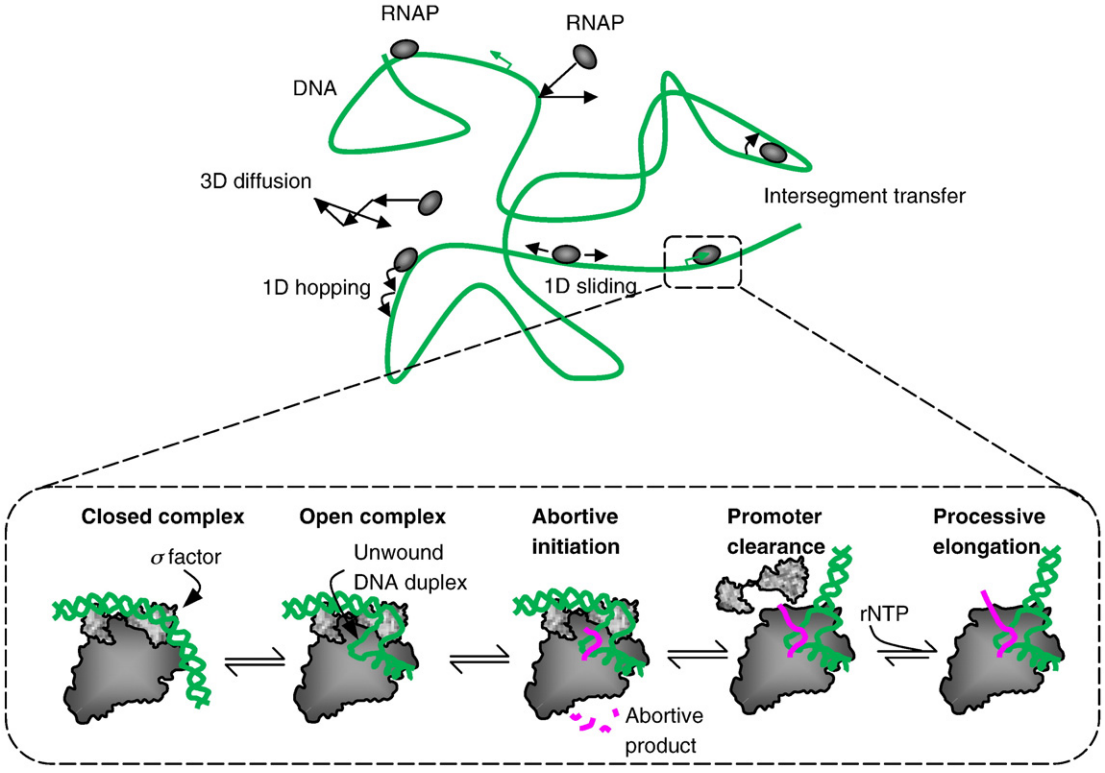

Fig. 4. Transcript initiation and elongation. Promoter search by RNAP may involve several mechanisms: 1D hopping, 1D sliding, intersegment transfer, and 3D diffusion. After promoter search, the RNAP holoenzyme binds tightly to the promoter and bends the DNA to form the closed complex. With the help of the σ factor, RNAP unwinds the DNA around the promoter and forms the open complex. RNAP repeatedly synthesizes short, abortive RNA products until it escapes the promoter region and enters the processive elongation phase.

search mechanism can be accounted for by a simple 3D collision.[63,66] The role and contribution of 3D diffusion in the promoter search mechanism is worthy of further investigation.

Initiation

Once a promoter has been located, RNAP forms an open complex (RPo), in which DNA around the transcription start site is unwound in preparation for RNA synthesis (Fig. 4). The transition from a closed to an open complex has been investigated by single-molecule studies in an effort to answer the following questions: Is DNA wrapped around RNAP in the closed and open complexes? What conformational changes take place during the transition from closed to open complex, and *vice versa*? What are the dynamics of these changes and how are they regulated by transcription factors? Using AFM, Rivetti *et al.* and Mangiarotti *et al.* showed that the lambda pR promoter underwent a 30-nm compaction upon open complex formation with *E. coli* RNAP.[54,67] On the basis of these findings, they inferred that DNA wrapped around RNAP for ~300° and that the interaction spanned 120 bp of template DNA, which is slightly larger than the value reported (90 bp) from ensemble footprinting experiments.[68] They further showed that DNA wrapping was mediated by the C-terminal domain of the RNAP α-subunit and that extensive wrapping seemed to be promoter-specific.[69] Open complex formation has also been observed in real time with single-molecule magnetic traps.[53] Unwinding and compaction of DNA as a result of open complex formation was distinguished by monitoring the change in DNA extension with either positively or negatively supercoiled DNA substrates. Compaction by RNAP reduced the end-to-end extension of the DNA, while unwinding caused either extension or reduction in length depending on whether the DNA is positively or negatively supercoiled. On the basis of this concept, Revyakin *et al.* demonstrated that DNA in the open complex was unwound by 1.2 ± 0.1 turns and compacted by 15 ± 5 nm relative to the closed complex.[53] Interestingly, they did not observe any difference in unwinding or compaction for a strong and a weak promoter, even though the lifetime of open complexes at the two promoters differed by 600-fold. They further showed that the initiating nucleotide and transcription effector ppGpp could stabilize or destabilize the open complex, depending on the extent of supercoiling. Because DNA may be positively or negatively supercoiled by other DNA interacting proteins locally, they proposed that supercoiling is a possible mechanism of transcription regulation *in vivo*.

RNAP begins transcript synthesis with abortive cycles of initiation, which generates short RNA products ~9–11 nt in length (Fig. 4). The position of the RNAP active center relative to the DNA template during abortive initiation has been controversial.[18,70–73] It was not clear whether the entire RNAP complex translocated back and forth along the DNA template while undergoing abortive initiation, whether some flexible element within RNAP allowed the active center to move forward while the remainder of the complex remained stationary, or whether RNAP pulled in the downstream DNA template while otherwise remaining stationary at the promoter (referred to as "DNA scrunching"). Single-molecule experiments finally provided a convincing answer for this controversy.[32,74] A study using spFRET monitored four different distances within the RNAP–DNA complex: (i) the position of the RNAP leading edge relative to the DNA template; (ii) the position of the RNAP trailing edge relative to the DNA; (iii) the expansion and contraction of RNAP during initiation; and (iv) the expansion and contraction of DNA during initiation.[32] By comparing these different distances, Kapanidis *et. al.* demonstrated that abortive initiation proceeds through a DNA scrunching mechanism in which RNAP remains stationary while reeling in the downstream DNA as it makes abortive RNA products. In a parallel study, using magnetic traps, Revyakin *et. al.* quantified RNAP-dependent DNA unwinding and revealed an increase in DNA unwinding when RNAP proceeds from the open to the initial transcribing complex and showed that the extent of this increase was dependent on the length of the RNA product.[74] Detailed inspection of population distributions and single-molecule time traces of spFRET data revealed that synthesis of the abortive product and forward translocation of the RNAP active center were faster than dissociation of the abortive transcript and reverse translocation of the RNAP active center.[75] These findings were all consistent with the DNA scrunching model.

The fate of the transcription factor σ^{70} after initiation has also been the subject of considerable debate.[76] σ^{70} is responsible for sequence-specific interactions with promoter DNA, facilitates promoter unwinding, and mediates interactions with regulators of transcription initiation. It has been proposed that σ^{70} is released from RNAP at the transition from initiation to elongation, and because of this potential difference in subunit composition the initiation and elongation complex (EC) may respond differently to DNA sequences and other potential transcription regulators. However, an alternative model proposes that σ^{70} is not released, but instead remains bound to RNAP during transcription elongation.[77–79] Single-molecule experiments have now given a much more quantitative description of RNAP–σ^{70} interaction.[31,75] Using FRET-ALEX, Kapanidis *et. al.* simultaneously monitored the position and the σ^{70} content of ECs and found that most early ECs (~70–90%) bearing

an 11-nt RNA product retained σ^{70}.[31] They further established that 50–60% of ECs bearing a 50-nt RNA product retained σ^{70}. Moreover, the half-life of σ^{70} retention was long relative to the rate of elongation, confirming that σ^{70} can remain bound to core RNAP during elongation.

Elongation

After promoter clearance, RNAP proceeds to elongation by processively adding nucleotides to the growing RNA chain. Elongation has been studied by means of many single-molecule techniques, especially TPM and optical traps. These studies typically begin with a stalled EC in which transcription is initiated with a subset of rNTPs (e.g., rA, rC, and rG) and stalls at the missing rNTP (e.g., rU). Stalled ECs are highly stable and can be introduced into an observation unit (usually a flow cell) of a single-molecule experimental setup, and then elongation can be resumed by addition of all four rNTPs.

The velocity of RNAP during elongation places inherent limitations on the rate of gene transcription and as such plays an important role in gene regulation. Early single-molecule studies using TPM revealed an average RNAP velocity of 6.2–20 bp/s at saturating rNTPs.[39,41] These studies also found that velocity could vary from molecule to molecule, but that the velocity of individual molecules did not vary significantly.[39,41] Detailed analysis of this heterogeneity was initially difficult because of pauses present in individual traces, but in later studies with higher-resolution optical traps, Herbert *et al.* and Neuman *et al.* could deconvolve active transcription and pauses from individual reaction trajectories.[47,80] In doing this, they found that the velocity distribution for a single RNAP was well described by a Gaussian function, but the distribution for a population of single RNAPs was too broad to be consistent with a homogeneous population. The underlying reason for this heterogeneity remains unclear.

The step size of RNAP has long been postulated to be 1 bp, as dictated by the sequential growth of the nascent transcript, but measurements of this sub-nanometer-scale movement was technically inaccessible. However, the development of an ultrastable optical trap by the Block laboratory has now allowed actively transcribing RNAP to be monitored with angstrom resolution.[40] Using low rNTP concentrations to slow the enzyme and a moderate load to reduce Brownian noise, Abbondanzieri *et al.* were able to observe RNAP, taking single steps of 3.7 ± 0.6 Å, which agrees well with the expectation that RNAP moves in single-nucleotide increments.[40] This high-resolution optical trap paves the way for future experiments where transcription can be studied in unprecedented detail.[47,81]

The movement of RNAP relative to the DNA template can be described by two general models, either as a "Brownian ratchet" or a "power stroke." In the Brownian ratchet model, the motion of the motor is driven by thermal fluctuations and rectified by rNTP addition. In the power stroke model, the free energy released during rNTP addition directly drives motion. These two models can be differentiated by the relationship of velocity *versus* applied force at low NTP concentrations. In the simple "Brownian ratchet" model, rNTP binding lowers the energy of the post-translocation position ($i+1$), while the applied hindering force increases the energy of the post-translocation position relative to the pretranslocation position (i). Therefore, force is expected to act as a competitive inhibitor to rNTP addition. In the "power stroke" model, rNTP binding is separated from translocation by several other irreversible chemical transitions; therefore, velocity should be largely insensitive to applied force. Using a high spatial resolution optical trap, Abbondanzieri *et al.* measured the force–velocity curves over a range of rNTP concentrations (~10 μM–1 mM) and showed that velocity was more sensitive to the applied force at low rNTP concentrations. They then generated global fits of these data to three different models for translocation, including a power stroke, a simple Brownian ratchet, and a more complex Brownian ratchet mechanism that incorporated a secondary rNTP-binding site that would allow RNAP to bind rNTPs in the pretranslocated state.[40] The results of this analysis were inconsistent with a power stroke mechanism but were consistent with both Brownian ratchet models, and the best fits were obtained from the Brownian ratchet model that included a secondary rNTP-binding site.[40]

Transcriptional pausing is a fundamental property of RNAP that can impact elongation rates and underlies many gene-regulation schemes.[22,82] There are two types of pauses: short-lived elemental (ubiquitous) pauses and longer-lived stabilized pauses (Fig. 5). Pausing is difficult to study using traditional gel-based biochemical assays, which are restricted to short DNA templates and cannot be used to accurately probe short duration pauses. In contrast, single-molecule studies can be used to study pauses with different lifetimes on long DNA templates. By aligning individual transcription traces to the DNA template with 1-bp resolution, Herbert *et al.* determined the distribution of pauses over a 2000-bp template.[47] This work demonstrated that elemental pauses did not occur randomly; rather, they were induced by specific sequences, although their exact origin remains unknown. On one hand, elemental pauses by *E. coli* RNAP do not show a force-dependent lifetime and frequency, which implies that backtracking is not involved.[80,83] Furthermore, elemental pauses still exist even when

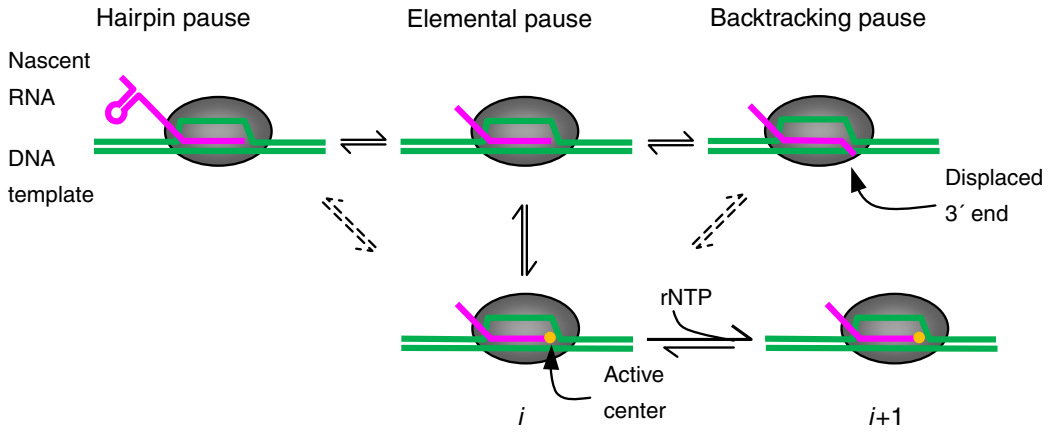

Fig. 5. Transcription pauses. RNAP makes frequent pauses during processive elongation. An elemental pause is believed to be the precursor of more stabilized pauses, such as hairpin pause and backtracking pause. See the text for additional details.

force is exerted on the RNA transcript to eliminate any potential secondary structure, implying that RNA hairpins do not contribute to elemental pauses.[83] On the other hand, elemental pauses by yeast Pol II have been inferred to be caused by backtracking.[48,84,85] It is not clear whether this difference results from the different RNAP species or from how the data are interpreted.[86] Elemental pauses are also believed to be the precursor of stabilized pauses, which can result from RNAP backtracking, hairpin formation in the nascent RNA transcript or from interactions with transcription regulators. High-resolution optical trap experiments revealed that during a long stabilized pause, RNAP backtracks with a moderate hindering force at a frequency close to the rNTP misincorporation rate during transcription.[46] This result suggests that some stabilized pauses may reflect RNA proofreading, where the polymerase backtracks as a result of misincorporation and waits for transcription factors (GreA and GreB in prokaryotes, or TFIIS in eukaryotes) to cleave the 3′ end of RNA before resuming elongation.[21]

Termination

RNAP is highly processive, but transcription must eventually terminate. Termination prevents read-through into downstream genes and can also occur in regulatory regions where it is used to control expression of a particular gene. Bacteria use two strategies, either intrinsic termination or rho-dependent termination (Fig. 6). For intrinsic termination, transcription stops when the newly synthesized RNA product forms a GC-rich hairpin followed by a series of seven to nine uridine molecules, which weaken the RNA–DNA hybrid within the EC. In rho-dependent termination, the helicase rho translocates in the 5′→3′ direction along the nascent transcript. When rho catches up to the polymerase, it destabilizes the entire EC and releases RNAP for the next transcription cycle.[87]

Three models have been proposed as general mechanisms encompassing both intrinsic and rho-dependent termination (Fig. 6). The first model proposes that the formation of an RNA hairpin or the action of rho pulls the nascent transcript out of RNAP by shearing the RNA–DNA hybrid without forward translocation of RNAP.[88] The second model hypothesizes that formation of a hairpin or translocation of rho along the RNA transcript pushes RNAP forward. This forward push disengages the polymerase active site from the 3′ end of the transcript and shortens the RNA–DNA hybrid, which subsequently destabilizes the EC.[89] The third model proposes that the hairpin or rho induces a conformational change in RNAP, which allosterically disrupts the entire EC.[90] These three models are not mutually exclusive.

Single-molecule techniques have recently been used to try to differentiate between these models for termination. Larson *et al.* used optical traps to explore the mechanism of intrinsic termination by pulling either the DNA template or RNA transcript, while observing how termination efficiency changed with force.[49] They found that the kinetics of transcript release were altered for one intrinsic terminator (referred to as t500), and termination efficiency was altered for one mutant of this terminator when force was applied to the DNA, which was most consistent with the forward translocation model. However, termination efficiency was insensitive to force applied to DNA for two other intrinsic terminators (*his* and tR2). Instead, termination at *his* and tR2 varied with force applied to the transcript, consistent with the RNA–DNA shearing model. Furthermore, application of force sufficient to rupture two to three base pairs at the

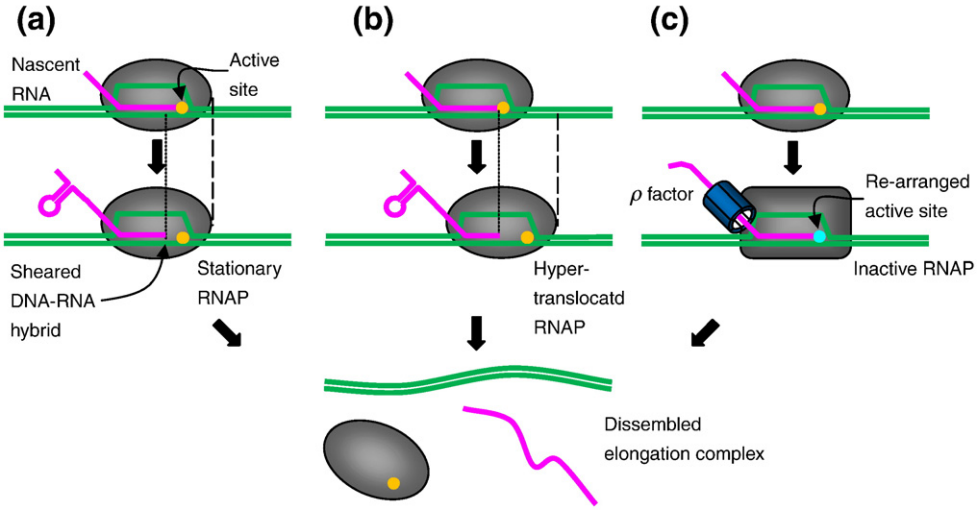

Fig. 6. Model for transcription termination. (a) In the "RNA shearing" model, the hairpin formation pulls the nascent RNA and shears the DNA–RNA hybrid. (b) In the "forward translocation" model, the hairpin formation (or translocation of rho factor) pushes RNAP forward without rNTP incorporation and disengages the polymerase active site. (c) In the "allosteric regulation" model, rho factor (or RNA hairpin) induces a conformational change of RNAP that inactivates the enzyme.

base of the RNA hairpin reduced termination efficiency, implying that these base pairs were necessary to disrupt the RNA–DNA hybrid. Based on these data, the authors derived a quantitative model that predicted termination efficiency as a function of DNA sequence for different terminators. In contrast to this conclusion, a recent bulk biochemical study proposed an allosteric mechanism for rho-dependent termination.[91] Epshtein *et al.* showed that rho associates with RNAP via protein–protein interactions throughout the transcription cycle without requiring the nascent transcript for initial binding.[91] Moreover, they also suggested that rho-dependent termination involves an allosterically induced rearrangement of the RNAP catalytic center through the evolutionarily conserved mobile trigger loop domain. They further demonstrated that the 3′ RNA end remained in the catalytic center of the termination complex, which is contrary to both the shearing and forward translocation models, and found that the efficiency of EC inactivation was insensitive to sequences downstream of the active site, arguing against a forward translocation model that requires the unwinding of downstream sequence. In addition, the efficiency of EC inactivation was higher for a GC-rich sequence than an AT-rich sequence, arguing against a shearing model. Finally, they showed that an EcoRI^{E111Q} roadblock situated immediately downstream of the termination site did not affect the rate of EC dissociation by rho, which again argues against the forward translocation model. The similarities and differences between rho-dependent and intrinsic termination pathways remain important subjects of investigation.

Emerging Trends in Single-Molecule Studies of Transcription

As single-molecule studies of transcription advance into their third decade, new trends are beginning to emerge that will have broad implications for our understanding of RNAP and transcriptional regulation. Up until now, most single-molecule studies have been confined to relatively simple systems involving one molecule of *E. coli* RNAP transcribing a naked DNA template. In contrast, transcription in cells is subject to extensive regulation and new single-molecule experiments are just beginning to probe the effects of transcription factors on RNAP. Moreover, transcription in eukaryotes occurs within the context of chromatin and involves a higher level of regulation and a multitude of transcription factors, and emerging single-molecule studies are now beginning to address these scenarios. In addition, new tools are being established allowing single-molecule studies to be performed *in vivo*, offering the future promise of direct insights into transcription within living organisms. Below we highlight recent examples that illustrate these emerging trends and try to offer a glimpse into future developments.

Eukaryotic RNAP

One major goal in coming years will be to apply single-molecule technologies to eukaryotic polymerases. Eukaryotic RNA polymerase II (Pol II) is a 550-kDa complex composed of 12 subunits and is reliant upon a much wider range of regulatory

factors compared to its prokaryotic counterparts. Pol II also faces the challenge of transcribing with the context of chromatin and must overcome the hindering effects of nucleosomes. Bulk biochemical studies of Pol II are inherently complicated because of the large number of components required to effectively recapitulate these reactions *in vitro*. However, the bulk biochemistry has reached a high level of sophistication, and single-molecule studies can now begin tackling these processes.

One limitation affecting the study of eukaryotic Pol II, is the poor efficiency with which active Pol II ECs can be prepared *in vitro*. This problem was overcome through the work of Kashlev, who developed a very clever approach for selectively eliminating inactive ECs by pausing active Pol II over a restriction site.[92] ECs that failed to transcribe could be eliminated through a simple restriction digest. Using this approach to enrich for active ECs, Galburt *et al.* characterized individual Pol II complexes with optical traps.[48] This study demonstrated that Pol II translocated at a velocity (12.2 ± 4.5 bp/s) comparable to that of its prokaryotic counterpart, but stalled at a much lower force (7.5 *versus* 35 pN). This stall force may not be the maximum force that Pol II can generate, but may reflect a limit of some other process that disrupts the forward translocation. The authors suggested that Pol II backtracked at 7.5 pN and that this backtracking led to misalignment of the Pol II active site with the 3′ end of the nascent transcript, which in turn arrested the forward translocation. Pol II also displayed frequent pauses, the distribution of which were consistent with the hypothesis that backtracking causes pauses and that during the pauses Pol II diffuses by a random walk until the active center realigned with the 3′ end of the transcript. This backtracking model is different from what has been proposed for prokaryotic RNAP pauses, which are not thought to involve backtracking.[80]

Crystal structures of Pol II ECs revealed the course of template and nontemplate DNA strands and the nascent RNA.[93–96] However, the upstream DNA, some of the nontemplate strand within the transcription bubble, and the exiting RNA were not observed. To help resolve these structural details, Andrecka *et al.* used spFRET to probe the EC.[97] These authors determined unknown positions within the EC by measuring the distances to these positions from at least three (or more) known positions through a process referred to as triangulation. By repeating this for different sites along the DNA and RNA, they mapped the path of the exiting RNA, the upstream DNA duplex, and the nontemplate DNA within the Pol II EC.[97] This work provided direct evidence that RNA leaves the polymerase active center via the previously proposed exit tunnel, and an independent single-molecule work using a similar method also supported this finding.[98] These results, together with recent crystal structures of Pol II and transcription factor TFIIB (a factor present in the Pol II preinitiation complex), provided insights into the transition from transcription initiation to elongation.[99,100] Crystal structures show that TFIIB binds on the top of Pol II and reaches into the active center, but comparison of the crystal structures with the single-molecule data reveals potential clashes between TFIIB and the EC. Part of TFIIB occupies the RNA exit channel and would clash with transcripts longer than 8 nt, and another segment of TFIIB would clash with upstream DNA and nontemplate DNA.[97] These steric clashes may cause TFIIB to be released during the transition from initiation to elongation, and this scheme may be a strategy coupling transcript length to initiation factor release.

Transcription factors and their effects on RNAP

Until recently, most single-molecule studies have been limited to simple systems lacking regulatory transcription factors, but some single-molecule studies have now begun to delve into investigations of different transcription factors, as well as the interplay between RNAP and various transcription factors. For example, the target search mechanisms of the transcrption factors *lac* repressor and p53 have been visualized with single-molecule imaging,[101,102] and the effects of GreA, GreB, and TFIIS on transcription have also been studied at the single-molecule level.[46,48]

Much like RNAP, transcription factors must also scan the genome for specific binding sites. *lac* repressor is a prokaryotic transcription that has long served as a paradigm for the study of transcriptional regulation and protein–DNA interactions.[103] Riggs *et al.* demonstrated that *lac* repressor could locate its target site (the *lac* operator) up to 100-fold faster than predicted by a simple 3D diffusion search process,[104] and this seminal study helped lead to great interest in the general problem of target search mechanisms.[59] This problem was first studied at the single-molecule level by Wang *et al.*, who used total internal reflection fluorescence microscopy (TIRFM) to visualize GFP-tagged *lac* repressor as it slid along DNA molecules.[102] Not only did this work confirm that *lac* repressor could slide along DNA, but it also yielded quantitative values for the 1D diffusion coefficients (2.3×10^{-4} to 1.3×10^{-1} $\mu m^2\ s^{-1}$) and the distances covered by the protein in each binding event (120 to 2920 nm).[102] More recently, Tafvizi *et al.* used a similar TIRFM approach to look at the DNA binding behavior of the eukaryotic transcription factor p53.[101] These authors concluded that p53 moved along DNA via a two-state search mechanism wherein the C-terminal domains mediate fast sliding along the DNA, while the core domain sampled the DNA for binding sites

through frequent binding and dissociation events. This proposed two-state mechanism is facilitated by the multidomain nature of p53, which allows the protein to readily switch between different binding modes.[101] More important, these types of single-molecule search studies have not been reported with DNA molecules harboring specific targets for the DNA-binding proteins under investigation, so it will be important in future work to visualize what happens with proteins such as *lac* repressor or p53 when they actually encounter their target sites.

GreA, GreB, and TFIIS are transcription factors that enable promoter escape, enhance transcription fidelity, and suppress pausing and arrest by stimulating the intrinsic endonucleolytic activity of RNAP.[105] This endonucleolytic activity cleaves 2–18 nt from the RNA 3′ end in backtracked ECs, allowing transcription to resume from the newly generated 3′ end. Using a two-bead optical trap, Shaevitz *et al.* observed backtracking and recovery by single molecules of RNAP.[46] They found that backtracking was associated with long pauses and that the addition of GreB decreased the pause duration, whereas GreA did not appreciably change the duration. GreB removed most backtracking pauses, yielding a backtracking distance close to zero (0.5 ± 0.8 bp), whereas the average distance backtracked in the presence of GreA increased to 6.6 ± 0.7 bp. These findings are consistent with the known functions of the two factors; GreB accelerates the removal of larger RNA fragments, whereas GreA stimulates the cleavage of dinucleotides.[106]

Galburt *et. al.* carried out similar experiments with Pol II and TFIIS, but found that the effect of TFIIS on Pol II is not the same as that of Gre factors on *E. coli* RNAP.[48] First, Pol II translocated against a hindering force of 16.9 ± 3.4 pN in the presence of TFIIS (*versus* 7.5 ± 2.0 pN in its absence). Second, the probability of exiting from a backtracked pause increased at high force in the presence of TFIIS (>7 pN), which may explain the increased resistance to stall force. These two studies mark the first single-molecule forays into the effects of transcription factors on RNAP, and one can envision future work directed at dissecting the effects of other factors that effect transcriptional regulation.

NusG is another transcription factor in *E. coli* that increases the rate of transcription and influences termination. To understand the transcription enhancement mechanism of NusG, Herbert *et. al.* studied transcription by a single RNAP molecule under various rNTP concentrations, applied loads, and temperatures, using a two-bead optical trap.[107] They found that NusG increased the pause-free velocity by 17% and reduced the density of backtracking pauses by ~3-fold, but did not reduce the lifetime of short pauses. These observations are consistent with a model where NusG promotes translocation of RNAP along DNA. Interestingly, a recent study showed that the NusG N-terminal domain interacts with RNAP, whereas its C-terminal domain forms a complex with either rho or with the transcription factor NusE, which is a protein component of the small ribosomal subunit.[108] Thus NusG may act as a physical link between transcription and translation, and the role of NusG in coupling transcription and translation will be an interesting subject of further studies.

Effects of nucleosomes on transcription

Eukaryotic genomes are compacted into chromatin, and nucleosomes have to be disrupted or modified to allow gene expression. How polymerases function on chromatin is a long-standing question that has recently been tackled at the single-molecule level by Hodges *et al.*, who used optical traps to probe Pol II as it transcribed through a nucleosome.[50] They found that while transcribing through a nucleosome, the probability for Pol II to pause tripled, the median pause duration doubled, and the transcriptional velocity dropped by 40%. They proposed a model in which pauses were associated with backtracking, with elongation resuming only after the active site of Pol II can realign with the 3′ end of the transcript. They further proposed that nucleosomes fluctuate rapidly between two states—one in which the DNA is partially unwrapped in front of the polymerase and another in which it is completely wrapped around the nucleosome core—and suggested that Pol II moved forward only when the nucleosome was partially unwrapped. This model could predict transcriptional and pausing kinetics measured by their experiments and was consistent with previous ensemble work.[109] In addition, Hodges *et al.* suggested that nucleosomes transferred to DNA behind the progressing polymerase through a transient DNA loop, a model that had been previously proposed by Studitsky *et al.* on the basis of bulk biochemical measurements.[110,111] It will be important to determine whether TFIIS can promote Pol II transcription through the nucleosome, since it rescues backtracked complexes and helps Pol II translocate against higher force.

Transcript elongation *in vivo* can be relatively fast, even in the presence of nucleosomes, compared with *in vitro* transcription with naked DNA.[112,113] One mechanism that could account for this difference is the cooperative action of several RNAP molecules. Previous ensemble experiments demonstrated that multiple RNAPs originating from the same promoter alleviated transcriptional pauses and arrests, an effect that was attributed to the assistance of the trailing RNAP in pushing the leading RNAP forward.[114] To further address this issue Jin *et al.* unzipped DNA molecules with an optical trap to locate RNAP after transcription, which in turn enabled them to investigate the synergistic action

of multiple *E. coli* RNAPs transcribing through a eukaryotic nucleosome.[45] When one RNAP was on the template DNA, it paused upon encountering a nucleosome, but when there were two RNAPs on the same DNA, the trailing RNAP pushed the leading RNAP forward, increasing transcription through the nucleosome. This result was consistent with gel-based ensemble experiments[114] and indicated that synergistic action of multiple RNAPs can help overcome nucleosomes, which may help explain how polymerases can transcribe through these barriers *in vivo*. Although *E. coli* RNAP would never normally encounter nucleosomes, the use of *E. coli* RNAP as a proxy for eukaryotic polymerases is justified by bulk biochemical experiments, which have shown that *E. coli* RNAP and eukaryotic pol II transcribe through nucleosomes using similar mechanistic principles.[115]

Future single-molecule studies will need to address how RNAP navigates DNA templates bound not by single isolated nucleosomes, but rather are occupied by long arrays of nucleosomes or even higher-order chromatin structures. In addition, as a longer-term goal, future studies will need to address how transcription is coupled to other cellular processes such as chromatin remodeling and nucleosome assembly, DNA replication as well as translation in prokaryotes, and RNA processing in eukaryotes (Fig. 7).[116–118]

In vivo single-molecule studies

One major challenge in the single-molecule field is to visualize individual molecules in living cells. Realization of this goal will help unravel the biophysical behavior of the transcription machinery in real-life situations where the full complexity of the entire cellular system contributes to gene expression. As highlighted below, the pioneering efforts of a few research groups have now begun to tackle this goal by visualizing fluorescently tagged proteins and/or RNAs *in vivo*.

The Xie laboratory was the first to directly observe single fluorescently tagged transcription factors inside living cells.[28] They achieved this remarkable feat by expressing *lac* repressor (LacI) fused to YFP. LacI moved so fast that initial images of the protein were blurred, and to overcome this blurring, they used stroboscopic excitation with 5-ms laser illumination cycles. This strategy enabled them to pinpoint the location of LacI at any given time, and by adding or removing IPTG, they were able to follow LacI as it searched for its operator sequence. They first measured its diffusion throughout the cells and reported an effective *in vivo* diffusion coefficient (D_{eff}) of 0.4 $\mu m^2\ s^{-1}$. Next they measured the 3D ($D_{3D} \sim 3\ \mu m^2\ s^{-1}$) and 1D diffusion coefficients ($D_{1D} \sim 0.046\ \mu m^2\ s^{-1}$) with *in vitro* single-molecule techniques, and by comparing these values to D_{eff} they concluded that LacI spends 90% of the time diffusing in 1D along nonspecific DNA while scanning the genome for targets. This groundbreaking study presents the framework necessary to begin probing other aspects of transcription *in vivo* with single-molecule sensitivity.

It is also now possible to visualize single RNA molecules *in vivo*, and the Singer laboratory has pioneered single mRNA detection by FISH to study gene expression.[29,119–121] Using this technique, Zenklusen *et al.* investigated transcriptional activity

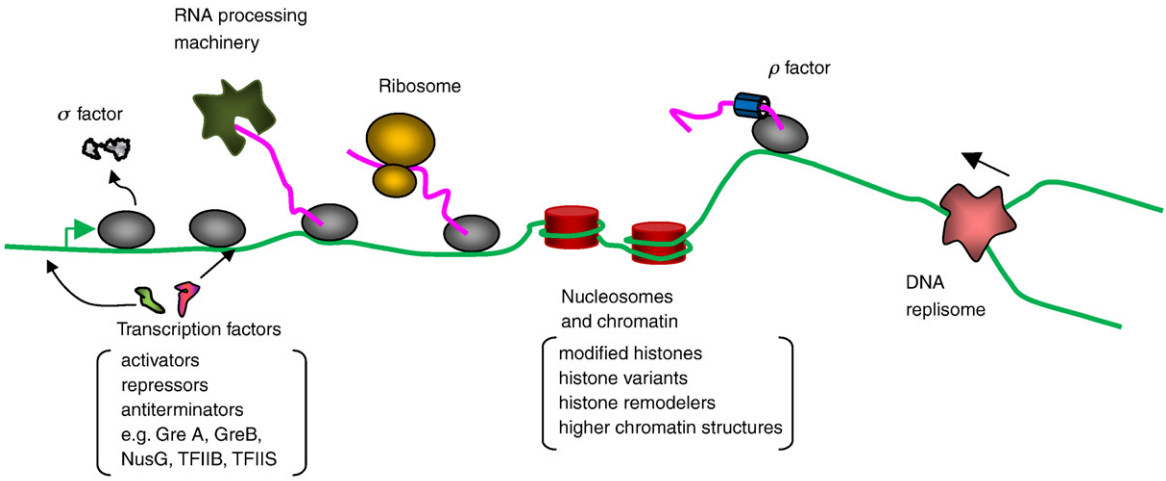

Fig. 7. Single-molecule transcription with multiple reaction components. Future *in vitro* single-molecule studies will advance our knowledge about transcription by inquiring into not only how RNAPs interact with other RNAP molecules and transcription factors, but also how transcription interacts with other cellular processes, such as DNA replication, translation (prokaryotes), chromatin structure, and RNA processing (eukaryotes).

within single *Saccharomyces cerevisiae* cells and found that expression levels of particular genes (e.g., *PDR5*, a gene regulated by the transcription coactivator complex SAGA) can vary substantially.[29] Surprisingly, 50% of cells contained either no PDR5 mRNA or only a single transcript, whereas other cells showed up to 11 transcripts, which was explained as bursts of transcriptional activity. But the variability for most constitutively expressed genes was much smaller, and the mRNA levels of housekeeping genes fell within a narrow range, which could be explained by uncoordinated transcription initiation. These data argue for the existence of multiple expression modes in cells. Variation of mRNA levels in mammalian cells has also been studied using FISH, and Raj *et al.* showed that at any given time, ~24% of CHO cells displayed large clusters of recently transcribed mRNA molecules (>200 per cell) densely packed around the gene from which they were transcribed.[122] But the majority of cells displayed <50 mRNA molecules, which were dispersed throughout the cell. The large variation in the number of transcripts per cell across the entire population further demonstrated that gene expression can occur through bursts of transcription. If transcription did not occur in bursts, then one would expect less variation in the number of transcripts per cell. They next tested the source of these bursts by comparing two reporter genes (M1 and M2) either positioned adjacent to each other or positioned at different chromosomal loci. When the two genes were next to one another they exhibited coordinated bursts of transcription, but when they were at different loci the transcription bursts occurred independently. They concluded that bursts reflected random switching of the M1 and M2 genes between inactive and active states and that the bursts were not induced by extrinsic factors, such as M1- or M2-specific transcriptional activators. They also observed the same behavior for the gene encoding RNAP. It is surprising that cells produce and tolerate such big variations in the mRNA level of such an essential housekeeping gene, but in the case of M1, the researchers showed that protein stability buffers the consequences of stochastic mRNA production. Whether this variation in mRNA production is subject to other evolutionary forces remains to be answered.

Transcription has also been studied in prokaryotes with single-molecule sensitivity. Golding *et al.* engineered an artificial gene (RFP; red fluorescent protein) harboring a tandem array of 96 sites for the site-specific RNA-binding protein MS2.[30] They then coexpressed this gene with a GFP-MS2 fusion protein in *E. coli*, which allowed them to count the number of mRNA molecules in real time in living cells, and in doing so they were able to directly demonstrate that transcription occurred in short bursts. Most recently, Thompson *et. al.* adapted this MS2-GFP tagging strategy to help develop new methods for tracking the movement of individual mRNA molecules in real time within living *S. cerevisiae*.[123] They were able to follow the 3D movement of single mRNAs at 15-ms temporal resolution with 25-nm precision in the x–y plane and 50-nm resolution in the z direction by fitting the images to a double-helical point spread function (DH-PSF).[123] This new approach demonstrates the ability to track single mRNAs (and potentially other molecules) inside of living cells with incredibly high spatial and temporal resolution and should offer future insights into transcription *in vivo*.

Finally, in another recent effort, Taniguchi *et al.* developed methods for quantitative gene expression analysis in single *E. coli* cells and reported the absolute numbers of mRNA for 137 highly expressed genes and the numbers of proteins for 1018 genes.[37] They used a library of genes tagged with fluorescent protein and an automated single-molecule imaging system to count proteins, and they employed FISH with a universal probe to observe the corresponding mRNA. They found 0.1 to 10,000 protein molecules and 0.05 to 5 mRNAs per cell for each gene. It was also found that essential genes and genes transcribed in the same direction as the replication fork had more mRNA, and the underlying cause for the latter observation is worthy of further research. Surprisingly, the average number of mRNAs and proteins per cell were well correlated, but the exact numbers of these two in individual cells were not. This discrepancy can be explained by the fact that mRNA is typically degraded within minutes, whereas proteins can survive many hours. Therefore the number of mRNA molecules present within any given cell reflects recent transcription activity, whereas proteins can accumulate over longer periods of time. This work represents the marriage of single-molecule detection with systems biology and for the first time allows variance in gene expression to be investigated at the level of the transcriptome and proteome.

Outlook

Single-molecule methods are allowing transcription to be studied in unprecedented detail, and as the field moves forward, researchers will be able to probe more and more complex aspects of these processes. As highlighted throughout this review, advances in single-molecule research are often driven by the development of new technologies. At the forefront of the field are new types of optical microscopy that may greatly facilitate *in vivo* imaging of single molecules. Of particular interest are new classes of super-resolution microscopy, such as PALM (photoactivation localization microscopy), STORM (stochastic optical reconstruction micro-

scopy), and STED (stimulated emission depletion) microscopy. These microscopy techniques use various optical tricks to overcome classical optical resolution limits (~200–300 nm in *x*–*y* and ~500–700 nm in *z*) and offer the ability to locate single molecules inside of cells with ~20-nm precision (reviewed in Refs. 124 and 125). Currently, these powerful new technologies are largely limited to fixed cells due to relatively slow image acquisition rates. However, with continued advances, including improved spatial resolution and the development of more rapid imaging capabilities, each of these new methods may help pave the way toward future nanometer-scale studies of transcription *in vivo* while advancing our understanding of gene expression and regulation.

Acknowledgements

We thank members of the Greene laboratory for insightful discussions and careful reading of the manuscript. This work was supported by grants from the NIH (GM074739 and GM082848) and by a National Science Foundation PECASE (Presidential Early Career Awards for Scientists and Engineers) award. E.C.G is an Early Career Scientist with the Howard Hughes Medical Institute.

References

1. Jacquier, A. (2009). The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat. Rev. Genet.* **10**, 833–844.
2. Repoila, F. & Darfeuille, F. (2009). Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol. Cell*, **101**, 117–131.
3. Carthew, R. W. & Sontheimer, E. J. (2009). Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642–655.
4. Kim, H. D., Shay, T., O'Shea, E. K. & Regev, A. (2009). Transcriptional regulatory circuits: predicting numbers from alphabets. *Science*, **325**, 429–432.
5. Balleza, E., Lopez-Bojorquez, L. N., Martinez-Antonio, A., Resendis-Antonio, O., Lozada-Chavez, I., Balderas-Martinez, Y. I. *et al.* (2009). Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol. Rev.* **33**, 133–151.
6. Panning, B. & Taatjes, D. J. (2008). Transcriptional regulation: it takes a village. *Mol. Cell*, **31**, 622–629.
7. Jiang, C. & Pugh, B. F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172.
8. Williams, S. K. & Tyler, J. K. (2007). Transcriptional regulation by chromatin disassembly and reassembly. *Curr. Opin. Genet. Dev.* **17**, 88–93.
9. Umlauf, D., Fraser, P. & Nagano, T. (2008). The role of long non-coding RNAs in chromatin structure and gene regulation: variations on a theme. *Biol. Chem.* **389**, 323–331.
10. McAllister, W. T. (1993). Structure and function of the bacteriophage T7 RNA polymerase (or, the virtues of simplicity). *Cell. Mol. Biol. Res.* **39**, 385–391.
11. Grummt, I. (1999). Regulation of mammalian ribosomal gene transcription by RNA polymerase I. *Prog. Nucleic Acid Res. Mol. Biol.* **62**, 109–154.
12. Willis, I. M. (1993). RNA polymerase III. Genes, factors and transcriptional specificity. *Eur. J. Biochem.* **212**, 1–11.
13. Sawadogo, M. & Sentenac, A. (1990). RNA polymerase B (II) and general transcription factors. *Annu. Rev. Biochem.* **59**, 711–754.
14. Sikorski, T. W. & Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. *Curr. Opin. Cell Biol.* **21**, 344–351.
15. Martinez, E. (2002). Multi-protein complexes in eukaryotic gene transcription. *Plant. Mol. Biol.* **50**, 925–947.
16. Bell, S. D. & Jackson, S. P. (2001). Mechanism and regulation of transcription in archaea. *Curr. Opin. Microbiol.* **4**, 208–213.
17. Steitz, T. A. (2004). The structural basis of the transition from initiation to elongation phases of transcription, as well as translocation and strand separation, by T7 RNA polymerase. *Curr. Opin. Struct. Biol.* **14**, 4–9.
18. Hsu, L. M. (2002). Promoter clearance and escape in prokaryotes. *Biochim. Biophys. Acta*, **1577**, 191–207.
19. Boeger, H., Bushnell, D. A., Davis, R., Griesenbeck, J., Lorch, Y., Strattan, J. S. *et al.* (2005). Structural basis of eukaryotic gene transcription. *FEBS Lett.* **579**, 899–903.
20. Margaritis, T. & Holstege, F. C. (2008). Poised RNA polymerase II gives pause for thought. *Cell*, **133**, 581–584.
21. Sydow, J. F. & Cramer, P. (2009). RNA polymerase fidelity and transcriptional proofreading. *Curr. Opin. Struct. Biol.* **19**, 732–739.
22. Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**, 1062–1066.
23. Nudler, E. & Gottesman, M. E. (2002). Transcription termination and anti-termination in *E. coli*. *Genes Cells*, **7**, 755–768.
24. Ciampi, M. S. (2006). Rho-dependent terminators and transcription termination. *Microbiology*, **152**, 2515–2528.
25. Buratowski, S. (2005). Connections between mRNA 3′ end processing and transcription termination. *Curr. Opin. Cell Biol.* **17**, 257–261.
26. Andrecka, J., Lewis, R., Bruckner, F., Lehmann, E., Cramer, P. & Michaelis, J. (2008). Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl Acad. Sci. USA*, **105**, 135–140.
27. Kim, J. H. & Larson, R. G. (2007). Single-molecule analysis of 1D diffusion and transcription elongation of T7 RNA polymerase along individual stretched DNA molecules. *Nucleic Acids Res.* **35**, 3848–3858.
28. Elf, J., Li, G. W. & Xie, X. S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, **316**, 1191–1194.

29. Zenklusen, D., Larson, D. R. & Singer, R. H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* **15**, 1263–1271.
30. Golding, I., Paulsson, J., Zawilski, S. M. & Cox, E. C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell*, **123**, 1025–1036.
31. Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J. *et al.* (2005). Retention of transcription initiation factor σ^{70} in transcription elongation: single-molecule analysis. *Mol. Cell*, **20**, 347–356.
32. Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S. & Ebright, R. H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, **314**, 1144–1147.
33. Tang, G. Q., Roy, R., Bandwar, R. P., Ha, T. & Patel, S. S. (2009). Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase. *Proc. Natl Acad. Sci. USA*, **106**, 22175–22180.
34. Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E. & Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl Acad. Sci. USA*, **101**, 8936–8941.
35. Kabata, H., Kurosawa, O., Arai, I., Washizu, M., Margarson, S. A., Glass, R. E. & Shimamoto, N. (1993). Visualization of single molecules of RNA polymerase sliding along DNA. *Science*, **262**, 1561–1563.
36. Finkelstein, I., Visnapuu, M. L. & Greene, E. (2010). Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature*, **468**, 983–987.
37. Taniguchi, Y., Choi, P. J., Li, G. W., Chen, H., Babu, M., Hearn, J. *et al.* (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, **329**, 533–538.
38. Maamar, H., Raj, A. & Dubnau, D. (2007). Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science*, **317**, 526–529.
39. Schafer, D. A., Gelles, J., Sheetz, M. P. & Landick, R. (1991). Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature*, **352**, 444–448.
40. Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. (2005). Direct observation of base-pair stepping by RNA polymerase. *Nature*, **438**, 460–465.
41. Yin, H., Landick, R. & Gelles, J. (1994). Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophys. J.* **67**, 2468–2478.
42. Sakata-Sogawa, K. & Shimamoto, N. (2004). RNA polymerase can track a DNA groove during promoter search. *Proc. Natl Acad. Sci. USA*, **101**, 14731–14735.
43. Harada, Y., Ohara, O., Takatsuki, A., Itoh, H., Shimamoto, N. & Kinosita, K., Jr. (2001). Direct observation of DNA rotation during transcription by *Escherichia coli* RNA polymerase. *Nature*, **409**, 113–115.
44. Thomen, P., Lopez, P. J., Bockelmann, U., Guillerez, J., Dreyfus, M. & Heslot, F. (2008). T7 RNA polymerase studied by force measurements varying cofactor concentration. *Biophys. J.* **95**, 2423–2433.
45. Jin, J., Bai, L., Johnson, D. S., Fulbright, R. M., Kireeva, M. L., Kashlev, M. & Wang, M. D. (2010). Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. *Nat. Struct. Mol. Biol.* **17**, 745–752.
46. Shaevitz, J. W., Abbondanzieri, E. A., Landick, R. & Block, S. M. (2003). Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature*, **426**, 684–687.
47. Herbert, K. M., La Porta, A., Wong, B. J., Mooney, R. A., Neuman, K. C., Landick, R. & Block, S. M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell*, **125**, 1083–1094.
48. Galburt, E. A., Grill, S. W., Wiedmann, A., Lubkowska, L., Choy, J., Nogales, E. *et al.* (2007). Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. *Nature*, **446**, 820–823.
49. Larson, M. H., Greenleaf, W. J., Landick, R. & Block, S. M. (2008). Applied force reveals mechanistic and energetic details of transcription termination. *Cell*, **132**, 971–982.
50. Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M. & Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science*, **325**, 626–628.
51. Skinner, G. M., Baumann, C. G., Quinn, D. M., Molloy, J. E. & Hoggett, J. G. (2004). Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase. A single-molecule view of the transcription cycle. *J. Biol. Chem.* **279**, 3239–3244.
52. Deufel, C., Forth, S., Simmons, C. R., Dejgosha, S. & Wang, M. D. (2007). Nanofabricated quartz cylinders for angular trapping: DNA supercoiling torque detection. *Nat. Methods*, **4**, 223–225.
53. Revyakin, A., Ebright, R. H. & Strick, T. R. (2004). Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. *Proc. Natl Acad. Sci. USA*, **101**, 4776–4780.
54. Mangiarotti, L., Cellai, S., Ross, W., Bustamante, C. & Rivetti, C. (2009). Sequence-dependent upstream DNA–RNA polymerase interactions in the open complex with λPR and λPRM promoters and implications for the mechanism of promoter interference. *J. Mol. Biol.* **385**, 748–760.
55. Guthold, M., Zhu, X., Rivetti, C., Yang, G., Thomson, N. H., Kasas, S. *et al.* (1999). Direct observation of one-dimensional diffusion and transcription by *Escherichia coli* RNA polymerase. *Biophys. J.* **77**, 2284–2294.
56. Ando, T., Uchihashi, T., Kodera, N., Yamamoto, D., Miyagi, A., Taniguchi, M. & Yamashita, H. (2008). High-speed AFM and nano-visualization of biomolecular processes. *Pfluegers Arch.* **456**, 211–225.
57. Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B. *et al.* (2009). Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One*, **4**, e7526.
58. Cho, B., Zengler, K., Qiu, Y., Park, Y., Knight, E., Barrett, C. *et al.* (2009). The transcription unit

architecture of the *Escherichia coli* genome. *Nat. Biotechnol.* **27**, 1043–1049.
59. Berg, O. G., Winter, R. B. & von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*, **20**, 6929–6948.
60. Gorman, J. & Greene, E. (2008). Visualizing one-dimensional diffusion of proteins along DNA. *Nat. Struct. Mol. Biol.* **15**, 768–774.
61. Halford, S. E. & Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* **32**, 3040–3052.
62. Ricchetti, M., Metzger, W. & Heumann, H. (1988). One-dimensional diffusion of *Escherichia coli* DNA-dependent RNA polymerase: a mechanism to facilitate promoter location. *Proc. Natl Acad. Sci. USA*, **85**, 4610–4614.
63. Roe, J. H., Burgess, R. R. & Record, M. T., Jr. (1984). Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λPR promoter. *J. Mol. Biol.* **176**, 495–522.
64. Harada, Y., Funatsu, T., Murakami, K., Nonoyama, Y., Ishihama, A. & Yanagida, T. (1999). Single-molecule imaging of RNA polymerase-DNA interactions in real time. *Biophys. J.* **76**, 709–715.
65. Singer, P. & Wu, C. W. (1987). Promoter search by *Escherichia coli* RNA polymerase on a circular DNA template. *J. Biol. Chem.* **262**, 14178–14189.
66. deHaseth, P. L., Zupancic, M. L. & Record, M. T., Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.* **180**, 3019–3025.
67. Rivetti, C., Guthold, M. & Bustamante, C. (1999). Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex. *EMBO J.* **18**, 4464–4475.
68. Davis, C. A., Bingman, C. A., Landick, R., Record, M. T., Jr. & Saecker, R. M. (2007). Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **104**, 7833–7838.
69. Cellai, S., Mangiarotti, L., Vannini, N., Naryshkin, N., Kortkhonjia, E., Ebright, R. H. & Rivetti, C. (2007). Upstream promoter sequences and αCTD mediate stable DNA wrapping within the RNA polymerase–promoter open complex. *EMBO Rep.* **8**, 271–278.
70. Carpousis, A. J. & Gralla, J. D. (1985). Interaction of RNA polymerase with lacUV5 promoter DNA during mRNA initiation and elongation. Footprinting, methylation, and rifampicin-sensitivity changes accompanying transcription initiation. *J. Mol. Biol.* **183**, 165–177.
71. Straney, D. C. & Crothers, D. M. (1987). A stressed intermediate in the formation of stably initiated RNA chains at the *Escherichia coli lac* UV5 promoter. *J. Mol. Biol.* **193**, 267–278.
72. Krummel, B. & Chamberlin, M. J. (1989). RNA chain initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry*, **28**, 7829–7842.
73. Pal, M., Ponticelli, A. S. & Luse, D. S. (2005). The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell*, **19**, 101–110.
74. Revyakin, A., Liu, C., Ebright, R. H. & Strick, T. R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, **314**, 1139–1143.
75. Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H. & Weiss, S. (2006). Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophys. J.* **90**, 1419–1431.
76. Mooney, R. A., Darst, S. A. & Landick, R. (2005). Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell*, **20**, 335–345.
77. Shimamoto, N., Kamigochi, T. & Utiyama, H. (1986). Release of the σ subunit of *Escherichia coli* DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. *J. Biol. Chem.* **261**, 11859–11865.
78. Mooney, R. A. & Landick, R. (2003). Tethering σ^{70} to RNA polymerase reveals high in vivo activity of σ factors and σ^{70}-dependent pausing at promoter-distal locations. *Genes Dev.* **17**, 2839–2851.
79. Wade, J. T. & Struhl, K. (2004). Association of RNA polymerase with transcribed regions in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **101**, 17777–17782.
80. Neuman, K. C., Abbondanzieri, E. A., Landick, R., Gelles, J. & Block, S. M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell*, **115**, 437–447.
81. Greenleaf, W. J. & Block, S. M. (2006). Single-molecule, motion-based DNA sequencing using RNA polymerase. *Science*, **313**, 801.
82. Rajala, T., Hakkinen, A., Healy, S., Yli-Harja, O. & Ribeiro, A. S. (2010). Effects of transcriptional pausing on gene expression dynamics. *PLoS Comput. Biol.* **6**, e1000704.
83. Dalal, R. V., Larson, M. H., Neuman, K. C., Gelles, J., Landick, R. & Block, S. M. (2006). Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. *Mol. Cell*, **23**, 231–239.
84. Depken, M., Galburt, E. A. & Grill, S. W. (2009). The origin of short transcriptional pauses. *Biophys. J.* **96**, 2189–2193.
85. Galburt, E. A., Grill, S. W. & Bustamante, C. (2009). Single molecule transcription elongation. *Methods*, **48**, 323–332.
86. Herbert, K. M., Greenleaf, W. J. & Block, S. M. (2008). Single-molecule studies of RNA polymerase: motoring along. *Annu. Rev. Biochem.* **77**, 149–176.
87. Banerjee, S., Chalissery, J., Bandey, I. & Sen, R. (2006). Rho-dependent transcription termination: more questions than answers. *J. Microbiol.* **44**, 11–22.
88. Toulokhonov, I. & Landick, R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol. Cell*, **12**, 1125–1136.
89. Santangelo, T. J. & Roberts, J. W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol. Cell*, **14**, 117–126.
90. Toulokhonov, I., Artsimovitch, I. & Landick, R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science*, **292**, 730–733.

91. Epshtein, V., Dutta, D., Wade, J. & Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. *Nature*, **463**, 245–249.
92. Kireeva, M. L., Lubkowska, L., Komissarova, N. & Kashlev, M. (2003). Assays and affinity purification of biotinylated and nonbiotinylated forms of double-tagged core RNA polymerase II from Saccharomyces cerevisiae. *Methods Enzymol*, **270**, 138.
93. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science*, **292**, 1876–1882.
94. Kettenberger, H., Armache, K. J. & Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell*, **16**, 955–965.
95. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. (2004). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science*, **303**, 1014–1016.
96. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. (2004). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell*, **119**, 481–489.
97. Andrecka, J., Treutlein, B., Arcusa, M. A., Muschielok, A., Lewis, R., Cheung, A. C. *et al.* (2009). Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. *Nucleic Acids Res.* **37**, 5803–5809.
98. Chen, C. Y., Chang, C. C., Yen, C. F., Chiu, M. T. & Chang, W. H. (2009). Mapping RNA exit channel on transcribing RNA polymerase II by FRET analysis. *Proc. Natl Acad. Sci. USA*, **106**, 127–132.
99. Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M. & Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature*, **462**, 323–330.
100. Liu, X., Bushnell, D. A., Wang, D., Calero, G. & Kornberg, R. D. (2010). Structure of an RNA polymerase II–TFIIB complex and the transcription initiation mechanism. *Science*, **327**, 206–209.
101. Tafvizi, A., Huang, F., Fersht, A. R., Mirny, L. A. & van Oijen, A. M. (2011). A single-molecule characterization of p53 search on DNA. *Proc. Natl Acad. Sci. USA*, **108**, 563–568.
102. Wang, Y., Austin, R. & Cox, E. (2006). Single molecule measurements of repressor protein 1D diffusion on DNA. *Phys. Rev. Lett.* **97**, 048302.
103. Lewis, M. (2005). The *lac* repressor. *C. R. Biol.* **328**, 521–548.
104. Riggs, A., Bourgeois, S. & Cohn, M. (1970). The *lac* repressor–operator interaction. 3. Kinetic studies. *J. Mol. Biol.* **53**, 401–417.
105. Fish, R. N. & Kane, C. M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta*, **1577**, 287–307.
106. Borukhov, S., Sagitov, V. & Goldfarb, A. (1993). Transcript cleavage factors from *E. coli*. *Cell*, **72**, 459–466.
107. Herbert, K. M., Zhou, J., Mooney, R. A., Porta, A. L., Landick, R. & Block, S. M. (2010). *E. coli* NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *J. Mol. Biol.* **399**, 17–30.
108. Burmann, B. M., Schweimer, K., Luo, X., Wahl, M. C., Stitt, B. L., Gottesman, M. E. & Rosch, P. (2010). A NusE:NusG complex links transcription and translation. *Science*, **328**, 501–504.
109. Kireeva, M. L., Hancock, B., Cremona, G. H., Walter, W., Studitsky, V. M. & Kashlev, M. (2005). Nature of the nucleosomal barrier to RNA polymerase II. *Mol. Cell*, **18**, 97–108.
110. Studitsky, V., Clark, D. & Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. *Cell*, **76**, 371–382.
111. Studitsky, V., Clark, D. & Felsenfeld, G. (1995). Overcoming a nucleosomal barrier to transcription. *Cell*, **83**, 19–27.
112. O'Brien, T. & Lis, J. T. (1993). Rapid changes in *Drosophila* transcription after an instantaneous heat shock. *Mol. Cell Biol.* **13**, 3456–3463.
113. Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S. M., Phair, R. D. & Singer, R. H. (2007). In vivo dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* **14**, 796–806.
114. Epshtein, V. & Nudler, E. (2003). Cooperation between RNA polymerase molecules in transcription elongation. *Science*, **300**, 801–805.
115. Walter, W., Kireeva, M., Studitsky, V. & Kashlev, M. (2003). Bacterial polymerase and yeast polymerase II use similar mechanisms for transcription through nucleosomes. *J. Biol. Chem.* **278**, 36148–36156.
116. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science*, **328**, 504–508.
117. Reed, R. (2003). Coupling transcription, splicing and mRNA export. *Curr. Opin. Cell Biol.* **15**, 326–331.
118. Pandit, S., Wang, D. & Fu, X. D. (2008). Functional integration of transcriptional and RNA processing machineries. *Curr. Opin. Cell Biol.* **20**, 260–265.
119. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. (1998). Visualization of single RNA transcripts in situ. *Science*, **280**, 585–590.
120. Levsky, J. M., Shenoy, S. M., Pezo, R. C. & Singer, R. H. (2002). Single-cell gene expression profiling. *Science*, **297**, 836–840.
121. Chubb, J. R., Trcek, T., Shenoy, S. M. & Singer, R. H. (2006). Transcriptional pulsing of a developmental gene. *Curr. Biol.* **16**, 1018–1025.
122. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309.
123. Thompson, M., Casolari, J., Badieirostami, M., Brown, P. & Moerner, W. (2010). Three-dimensional tracking of single mRNA particles in *Saccharomyces cerevisiae* using a double-helix point spread function. *Proc. Natl Acad. Sci. USA*, **107**, 17864–17871.
124. Hell, S. W. (2009). Microscopy and its focal switch. *Nat. Methods*, **6**, 24–32.
125. Huang, B., Bates, M. & Zhuang, X. (2009). Super-resolution fluorescence microscopy. *Annu. Rev. Biochem.* **78**, 993–1016.