A single-molecule view on cellular and viral RNA synthesis

Eugen Ostrofet[1], Flavia Stal Papini[1], Anssi M. Malinen[2], David Dulin[1, #]

[1] Junior Research Group 2, Interdisciplinary Center for Clinical Research, Friedrich Alexander University Erlangen-Nürnberg (FAU), Hartmannstr. 14, 91052 Erlangen, Germany.
[2] Department of Biochemistry, University of Turku, Vatselankatu 2, 20500 Turku, Finland

Correspondence may be addressed to: david.dulin@uk-erlangen.de

Abstract

RNA synthesis is at the heart of gene expression, the RNA products serving in cellular organisms as protein blueprints, structural/functional components of molecular machines or regulators of gene expression, and as the genetic material of nascent RNA viruses. Because of its importance, transcription –the process converting DNA to RNA– has been extensively studied at the single-molecule level; first in bacterial systems and quickly after in other domains of life. RNA replication in RNA viruses has recently become an important research topic in the single-molecule field, aiming not only to uncover its versatile mechanisms but also to develop new therapeutics against these devastating human pathogens. Here, we summarize the recent advances in cellular and viral RNA production using in vitro single-molecule techniques.

Key words

Single-molecule biophysics, transcription, RNA virus, RNA replication, RNA polymerase, RNA-dependent RNA polymerase, optical tweezers, magnetic tweezers, fluorescence resonance energy transfer

Introduction

Ribonucleic acid (RNA) mediates the genetic information to protein synthesis and is thus an essential component of the central dogma in molecular biology. Additionally, specific RNAs serve as structural and catalytic constituents in ribozymes, such as ribosome and spliceosome, or participate in the regulation of gene expression. Furthermore, RNA uniquely both stores and mediates the genetic information in RNA viruses (Alberts et al., 2002). RNA polymerase enzymes make RNA by joining together the ribonucleotides building blocks in a template-guided reaction. On the one hand, if the RNA polymerase uses deoxyribonucleic acid (DNA) as a template to build the complementary RNA strand, the RNA polymerase is classified as DNA-dependent RNA polymerases (DdRp) (Alberts et al., 2002). All cellular RNA polymerases (RNAPs) belong to this

class. On the other hand, if the template is RNA, e.g. RNA virus genome replication, the RNA polymerase is called RNA-dependent RNA polymerase (RdRp) (te Velthuis, 2014).

This essay focuses on the multisubunit cellular DpRps, also coined RNAPs; specifically, bacterial RNAP, archaeal RNAP and three distinct eukaryotic RNAPs (Pol I, Pol II and Pol III) (Alberts et al., 2002). The tight regulation of RNAP activity is key to gene expression, as it provides the cell the means to respond quickly to environmental change. Cellular transcription is regulated by a myriad of factors, e.g. DNA promoter sequence, transcribed sequence, the secondary structures of the RNA transcript, protein transcription factors interacting with the DNA or RNAP itself, nucleoproteins, and RNAP trafficking (Belogurov and Artsimovitch, 2015; Browning and Busby, 2016; Jonkers and Lis, 2015; Ray-Soni et al., 2016; Roberts et al., 2008; Saunders et al., 2006; Washburn and Gottesman, 2015). As a consequence, the RNAP transcriptional activity is very dynamic, with bursts of RNA synthesis being interrupted by pauses of various length and nature (Jonkers and Lis, 2015; Landick, 2006; Saunders et al., 2006). The overlapping layers of RNAP regulation make the transcription process stochastic and noisy, which eventually influences proteins production levels and the phenotype of an organism (Elowitz et al., 2002; Kaern et al., 2005; Raser and O'Shea, 2004).

The RdRp class contains two evolutionary non-related families: cellular RdRps are present in plants and produce short RNAs that are important in the developmental regulation, genome integrity maintenance and defence against pathogens (Ahlquist, 2002; Iyer et al., 2003; Zong et al., 2009), while the second family of RdRp is one of the key component of viral replication of RNA viruses (de Farias et al., 2017; Ng et al., 2008; te Velthuis, 2014). We discuss here only the latter, as no cellular RdRps single-molecule study has been published to date. Viral RdRps not only replicate the viral genome but also allow it to evolve rapidly because they have a very high error rate (Gago et al., 2009; Holmes, 2010). However, the error rate, originating from nucleotide misincorporation, is tightly controlled to simultaneously provide an evolutionary advantage over the host immune system and robust production of infectious virions (Lauring et al., 2013; Smith, 2017). RdRps replication activity is highly regulated by different factors, e.g. viral genome secondary structures, viral associated-proteins and host factors. Furthermore, because of their central role in virus proliferation and their conserved catalytic structural domains, RdRps also represent a target for broad spectrum antiviral drugs, such as nucleotide analogues (Jacome et al., 2015).

The development of single-molecule techniques has offered a unique view on the action of enzymes, including that of RNA polymerases (Dulin et al., 2013; Miller et al., 2018). The world of a biomolecule is strikingly different from the one we experience as humans: the characteristic

length of molecules is ~1-10 nm, inertial forces are negligible and cellular processes are driven by diffusion and low energy activations in the order of few k_BT (k_B being the Boltzmann constant and T the absolute temperature). For example, a polymerase that moves forward by a length of a DNA base-pair (bp) (0.34 nm) against a force of 12 pN (10^{-12} Newton) generates a mechanical work of ~1 k_BT, which equals the thermal energy provided by the environment. To give some perspective, ~25 pN hindering force is necessary to stall the bacterial RNAP (Wang et al., 1998). As it translocates with 0.34 nm DNA base pair (bp) steps, it performs up to ~2 k_BT mechanical work ($E = F \times \delta$, where E is the mechanical work, F the applied force and δ the step size of the biomolecular motor), while the Φ29 DNA packaging motor, that translocate by 2.3 bp step size, withstands a force up to ~57 pN, and performs up to ~11 k_BT mechanical work (Liu et al., 2014). The kinesin withstands forces up to 8 pN, but translocates forward with ~8 nm steps, and therefore performs up to ~16 k_BT mechanical work (Visscher et al., 1999). In the context of a long DNA sequence and ignoring the nearest neighbors influence, breaking a DNA A/T base pair costs in average ~1.5 k_BT, while breaking a G/C base pair costs ~3 k_BT (Bockelmann et al., 1998). Because the energies involved are low, positions and conformations fluctuate constantly rendering the catalytic activity of enzymes “noisy” or stochastic. Furthermore, enzymes often explore several kinetic states, with different interconversion rates and not necessarily catalytically active, making their catalytic activity like a journey through a very bumpy road. Observing these different kinetic states is difficult — if not impossible — with classic ensemble approaches, as it averages out the heterogeneity of the sample, extracting only the average behavior (Tinoco and Gonzalez, 2011). By monitoring enzyme catalysis at the single-molecule level it is possible to uncover complex and parallel kinetic pathways, where transient and rare events play important roles in the overall activity of the enzyme (English et al., 2006; Xie, 2001). In vitro single-molecule techniques can be divided in two main categories: *i)* force spectroscopy techniques, where the change in the extension of a tethered nucleic acids upon a single enzyme activity is monitored under an applied force, *ii*) fluorescence spectroscopy, where dye-labeled biomolecules are localized and followed using high-end microscopy to monitor inter- or intra-molecule interactions (Dulin et al., 2013; Miller et al., 2018). Both approaches have led to many important discoveries on the molecular mechanism of RNA synthesis in all domains of life.

In **Figure 1** and **Figure 2**, we present the single-molecule techniques used in in vitro experiments on cellular RNAPs or viral RdRps. These figures mainly serve to “visualize” the experimental assay used in the described studies, and we advise the reader to look into the numerous specialized reviews that treat the techniques from a more technical viewpoint of single-molecule biophysics (Chakraborty et al., 2017; De Vlaminck and Dekker, 2012; Dulin et al., 2013; Hohlbein et al., 2014; Joo et al., 2008; Kriegel et al., 2017; Miller et al., 2018; Moffitt et al., 2008; Neuman and Nagy, 2008; Robinson and van Oijen, 2013). In the following parts of the book

chapter, we discuss the recent literature of the in vitro single-molecule enzymology studies of RNA synthesis by cellular RNAPs and viral RdRps.

In vitro single-molecule studies of cellular RNAPs

Cellular transcription performed by large multi-subunit RNAPs can be divided into three different phases, called initiation, elongation and termination. The RNAP finds and initiates RNA synthesis at a promoter — a specifically recognized DNA sequence preceding every gene or operon (a set of adjacent co-regulated genes). The RNAP subsequently moves into the coding region of the gene and elongates the nascent RNA until it meets a termination signal at the end of the gene or operon. These three phases of transcription have been extensively studied at the single-molecule level using the RNAP from *Escherichia coli* (*E. coli*) as the model system, and more recently also using structurally more complex eukaryotic and archaeal RNAPs. We highlight here some of the key studies.

Initiation

Bacterial transcription machinery

During the initiation phase, the RNAP must (1) recognize the promoter, (2) form the transcription bubble, i.e. open the double stranded DNA, (3) initiate the synthesis of RNA, (4) stably hold the short nascent RNA in the active site, and (5) break interactions with the promoter and transcription initiation factor(s) on the way to the elongation phase (Browning and Busby, 2004, 2016; Ruff et al., 2015; Saecker et al., 2011). However, the RNAP may fail at each of these steps, making transcription initiation a highly stochastic process. Single-molecule techniques are therefore particularly suitable to uncover the determinants of the transitions from one step to the next (Kapanidis and Strick, 2009).

The bacterial transcription initiation complex — formed by the association of the core RNAP with a σ initiation factor as a holoenzyme — is the most studied transcription complex at the single-molecule level. The holoenzyme has to respond to a large variety of biochemical signals to control gene expression. The holoenzyme also needs to negotiate substantial variation in the promoter sequences that takes place ~100 bp upstream and ~20 bp downstream of the transcription start site (TSS, locates at position +1); there are ~3000 different promoters in *E. coli* each imposing a unique set of parameter (rates, stabilities) to the sub-stages of initial transcription (Browning and Busby, 2004, 2016). At the beginning of gene transcription, the holoenzyme finds the promoter and forms the RNAP-promoter closed (RP_C) complex by establishing interactions with specific elements of the promoter. The promoter search mechanism was investigated using a single-molecule fluorescence co-localization assay in combination with

total internal reflection microscopy (TIRFM) (**Figure 1AB**)(Friedman et al., 2006). The data consists of holoenzyme binding dwell times on the DNA promoter as well as the dwell times separating two binding events to the same promoter; further variables included holoenzyme concentration and the total length of DNA construct (Friedman et al., 2013). The results were consistent with a search process being dominated by three-dimensional diffusion instead of one-dimensional sliding along the DNA (Friedman et al., 2013). A similar conclusion was reached using a DNA curtains approach, where stretched bacteriophage λ-DNA containing several *E. coli* RNAP promoters was used to observe binding/unbinding RNAPs eventually converting into productive transcribing complexes (Wang et al., 2013) (**Figure 1C**). The contacts formed between the holoenzyme and promoter were recently studied using an optical tweezers assay in a dumbbell configuration (**Figure 2B**), where the promoter is encoded in the stem of an hairpin (**Figure 3A**) (Meng et al., 2017). Using this elegant experimental configuration, it is possible to determine the contact points between the promoter and the holoenzyme, even the most transient ones, by opening the hairpin with a linear increase of the force (Woodside and Block, 2014) (**Figure 3B**). By comparing the hairpin opening profiles of the free promoter and holoenzyme-bound promoter, the authors found that, besides the well-known contacts formed with the -35 and -10 elements of the promoter, the holoenzyme forms strong contacts with the spacer element — located between the -35 and -10 elements — and that the various contacts are released in a non-sequential fashion during promoter escape (Meng et al., 2017).

After recognizing the promoter, the holoenzyme opens ~12 bp stretch of the double-stranded DNA (dsDNA) to form the RNAP open promoter complex (RP_O). dsDNA melting is monitored using the topological changes taking place in a torsionally constrained DNA molecule (Revyakin et al., 2005; Rutkauskas et al., 2017). DNA has a fixed amount of twist between each nucleotide, which results in a helical pitch of ~10.5 bp/turn. When adding twist to a DNA molecule, e.g. by adding turns, and for a stretching force below ~0.5 pN, the DNA molecule eventually buckles and forms plectonems (**Figure 2C**), and therefore adds writhe, i.e. the DNA molecule crossing over itself. The sum of writhe and twist is conserved in a torsionally constrained DNA molecule (Charvin et al., 2004), and therefore the change in twist eventually leads to a change in writhe to compensate. In other words, for a negatively (or positively) supercoiled DNA molecule that has passed the buckling transition and has formed plectonems, the number of writhe decreases (or increases) upon promoter melting, i.e. decrease in twist, leading to an increase (or decrease) in the end-to-end extension of the DNA molecule. Using this property in combination with a torque spectroscopy technique, i.e. optical torque wrench (**Figure 2D**) or magnetic tweezers (**Figure 3C**), it is therefore possible to monitor promoter melting. Strick and co-workers used a magnetic tweezers assay to monitor the RP_O dynamics on a consensus and ribosomal promoters, which respectively make stable and unstable contacts with the holoenzyme (**Figure 3CD**). The authors

showed that the addition of torque in a torsionally constrained DNA molecule affects the formation and the stability of the RP_O: the addition of negative supercoil assists promoter opening by lowering the DNA melting energy penalty and promotes the formation of a stable RP_O, whereas the addition of positive supercoil increases the DNA melting energy penalty, hinders promoter opening and reduces the RP_O lifetime (Revyakin et al., 2004). A recent study using single-molecule Förster resonance energy transfer (smFRET) with TIRFM (**Figure 1DGH**) showed that in the absence of promoter supercoiling, the RP_O is very dynamic, fluctuating in millisecond timescale between the open and the closed DNA conformations; the contacts formed by the $\sigma_{3.2}$ domain of σ^{70} (the housekeeping initiation factor of *E. coli*) with the template DNA stabilizes the open form of RP_O (Duchi et al., 2018a). Noteworthy, TIRFM-smFRET and magnetic tweezers have a temporal resolutions of ~10 ms and ~1 s, respectively. Two studies indirectly assessed the submillisecond dynamics of the RP_O using confocal smFRET data (**Figure 1E**) in combination with the signal burst variance analysis (Robb et al., 2013) or the photon-by-photon hidden Markov modeling (Lerner et al., 2018). Both studies showed that the RP_O explores different transcription bubble sizes by opening more downstream DNA, which have been suggested to determine the transcription start site (TSS). The mechanistic basis for TSS selection was recently further explored using magnetic tweezers (**Figure 3CD**) and DNA-protein photocrosslinking experiments, showing that the energetics of the transcription bubble size eventually regulates the TSS selection, in practice limiting its range to the positions -1, +1 and +2 (Yu et al., 2017).

In the presence of NTPs, the RP_O quickly engages in the synthesis of the nascent RNA, forming an initially transcribing complex (ITC). Early biochemical experiments showed that the (average) position of the upstream position of the ITC does not change during the addition of the first 9–11 nucleotides to the RNA. Three different mechanisms were proposed to describe how the RNAP manages these constraints: (1) transient excursion (RNAP diffuses back and forth between subsequent abortive initiations), (2) inchworming (flexible RNAP body containing the active site stretches forward at each nucleotide incorporation) and (3) downstream DNA scrunching (the DNA bubble is extended inside and on the surface of the RNAP). Elegant confocal smFRET and magnetic tweezers studies showed that the scrunching model was the correct one (**Figure 1DE** and **Figure 3CD**, respectively), where only the downstream DNA region of the promoter is moving relative to the holoenzyme during initial transcription (Kapanidis et al., 2006; Revyakin et al., 2006).

The initial transcription leads either to successful promoter escape and synthesis of the full length RNA or the release of short (up to ~11 bases) aborted RNA products and reversion to the initial RP_O state (Browning and Busby, 2016). The overall efficiency of transcription initiation is determined at two distinct phases: either at the stage of RP_O formation if the promoter, e.g. *rrnB*,

forms unstable DNA bubble, or during initial transcription if the promoter, e.g. *lacUV5*, forms a stable RP_O. The former case was explored by Strick and co-workers and described above (Revyakin et al., 2004). The latter case was recently explored by several TIRFM-smFRET, confocal smFRET and magnetic tweezers based studies (Duchi et al., 2016; Dulin et al., 2018; Lerner et al., 2016). It has become evident that initial transcription is interrupted by two types of pauses: a short pause (half-life ~10 s) occurring after the synthesis of 6-mer RNA, originating from the clash of the nascent RNA and the σ factor blocking the RNA exit channel (**Figure 3E**) (Duchi et al., 2016), and a long pause (~100-1000 s) involving a stable backtracked complex (Lerner et al., 2016). The latest of the three studies (**Figure 3EF**) found that these two pauses are actually connected via a branched mechanism, where a fraction of the initially transcribing complexes pauses after the synthesis of an 6-mer RNA and isomerizes to a long-lived backtracked pause-state (Dulin et al., 2018). The backtracked fraction increases with the strength of the pause at 6-mer RNA, which in turn depends on the initially transcribed sequence and the NTP concentration. This study additionally showed that promoter unscrunching does not always require the release of the abortive RNA, thereby expanding the earlier models that assumed the two processes to be tightly connected (Margeat et al., 2006; Revyakin et al., 2006).

The initiation factor σ^{70}, the housekeeping σ factor in *E. coli*, was thought to impact only the initiation phase and be released upon transition to the elongation phase. However, a confocal smFRET (**Figure 1E**) study showed that σ^{70} is indeed retained during elongation, at least *in vitro* (Kapanidis et al., 2005). A more recent study from Harden and co-workers using TIRFM single-molecule co-localization (**Figure 2AB**) confirmed the retention of σ^{70} and found that σ^{70} influences the progress of the transcript elongation hundreds of bp downstream the promoter region by binding to and inducing pauses at -10 element-like DNA sequence (Harden et al., 2016).

The bacterial RNAP resembles structurally a crab claw, with the two “pincers”, formed by the β and the β’ subunits, defining the walls of the primary DNA binding cleft of the polymerase. The pincer of the β’ subunit is called the clamp, which according to the crystal structures adopts multiple conformations, including the open and closed conformations that differ at most by a 20° swinging motion of the clamp from a hinge at the base of the clamp. Chakraborty *et al.* employed unnatural amino acid mutagenesis to specifically attach fluorophores to both pincers, thus generating a FRET ruler to monitor the clamp positions (**Figure 3G**) (Chakraborty et al., 2012). Using confocal smFRET (**Figure 1DEF**), the clamp conformation in different phases of transcription was observed. The authors found that the clamp of DNA-free core enzyme, RNAP-σ^{70} holoenzyme and RNAP-σ^{54} holoenzyme have three distinct conformations at equilibrium, assigned as the open, closed and collapsed clamp, respectively. The authors further extended

the study by trapping the structural intermediates on the σ^{54}-dependent open complex formation pathway. In this experimental configuration, the clamp remains predominantly open until the promoter DNA melts and forms the transcription bubble in the open complex (RP_O); the clamp remains consistently closed in the initially transcribing complex and in elongation complex. The clamp state in the holoenzyme is modulated by antibiotics myxopyronin, corallopyronin and ripostatin, as well as by bacteriophage T7 protein Gp2, with all of them depopulating the open clamp conformation. However, the confocal smFRET experiments did not have access to the kinetics of the clamp conformations. Duchi and co-workers (Duchi et al., 2018b) thus performed further experiments using TIRFM-smFRET (**Figure 1GH**) and similarly labelled RNAP. This experimental configuration allowed to set strict selection criteria for the homogeneity of the analyzed molecules, which lead to the re-assignment of the holoenzyme clamp states as open, partly-closed and closed. A significant fraction of the holoenzymes has a dynamic clamp that interconverts between these states in a timescale of ~0.1-1 s. The binding of stringent-response alarmone ppGpp stabilizes the partly-closed clamp state of the RNAP. By combining cryo-EM based structural information with TIRFM-smFRET based data on clamp dynamics (Lin et al., 2017), it was possible to uncover the mechanism of RNAP inhibition by lipiarmycin, an antibiotic clinically used to treat *Clostridium difficile* infection. Lipiarmycin dramatically modifies the clamp by locking it in the open conformation and thereby prevents the isomerization of RP_C to RP_O.

Archaeal transcription machinery

The archaeal transcription machinery has many similarities with the eukaryotic Pol II machine as the RNAP and many associated factors are homologous (Werner and Grohmann, 2011). Because the archaeal transcription machinery is less complex and can be readily reconstituted from recombinant proteins (Schulz et al., 2015; Werner and Weinzierl, 2002), it constitutes, in addition to its inherent value, a good model system for understanding the mechanism of eukaryotic transcription. Archaeal RNAP and eukaryotic Pol II require two additional proteins for the basal level of transcription initiation: the TATA-binding protein (TBP) and the transcription factor B (known as TFB in archaea and TFIIB in eukaryotes). TBP bends the DNA promoter, associates with TFB and subsequently recruits the RNAP to form the pre-initiation complex (PIC). A TIRFM-smFRET (**Figure 1DGH**) study found that the archaeal TBP dynamically bends and unbends the promoter, whereas its eukaryotic counterpart bends the promoter into two stable populations with different bending angle; the less bended population is eventually converted into the more bended conformation upon the addition of TFIIB (Gietl et al., 2014). Nagy and coworkers used TIRFM-smFRET (**Figure 1GH**) and Nano-Positioning System analysis (Muschielok et al., 2008) (**Figure 3H**) to provide a structural model of the archaeal RP_O by determining the positions of the promoter DNA, RNAP and the transcription initiation factors TBP, TFB and the transcription factor E (TFE) (Nagy et al., 2015). In a separate study, the RNAP

binding sites of TFE and transcription elongation factor Spt4/5 were shown to overlap, which makes the two factors competing against each other to bind RNAP; this competition likely has implication in the regulation of transcription initiation and elongation. Schultz *et al.* used TIRFM-smFRET to analyze the clamp conformation in archaeal RNAP labeled with two fluorophores on the opposite sides of the DNA binding cleft (Schulz et al., 2016). The authors found that most (~80%) of the DNA-free RNAPs adopts a closed clamp conformation with a smaller amount of RNAPs having an open clamp (~20%). In contrast to the dynamic clamp study of the bacterial RNAP (Duchi et al., 2018c), the authors did not observe real-time interconversions between the two clamp states. The opening of the transcription bubble upon RP_O formation shifted the clamp equilibrium towards the open state, whereas, interestingly, the exact opposite has been observed with the bacterial transcription system (Chakraborty et al., 2012). Consistent with the mutual dependence of the transcription bubble and clamp opening in the archaeal system, they showed that the transcription initiation factor TFE, which stimulates DNA opening, increases the fraction of RNAPs with open clamp. Because the open clamp in the transcription elongation complex is also promoted by the binding of a correct (templated) nucleotide to the active site and elongation factor Spt4/5, the catalytically competent, highly processive archaeal RNAP may require a relatively open clamp conformation.

Eukaryotic transcription machineries

Eukaryotic transcription is far more complex than its bacterial counterpart, and has therefore only recently been investigated. For example, the yeast Pol II system assembles a total of 32 proteins when it forms a pre-initiation complex (PIC) on the promoter. Recently, Galburt and collegues (Tomko et al., 2017) adapted the magnetic tweezers assay pioneered by the Strick lab (**Figure 3CD**) (Revyakin et al., 2004) to measure real-time promoter melting events by the Pol II and the distributions of DNA-bubble sizes generated during different phases of initiation. They found that promoter opening is in fact a two-step process: first, the Ssl2 subunit of transcription initiation factor TFIIH pumps the downstream DNA towards the Pol II in the process generating torsional and mechanical stress that leads to the formation of an initial ~6-bp bubble; second, Pol II synthesizes the initial RNA synthesis at the transcription start site which expands the bubble to its final ~13-bp size. On the other hand, an earlier study by the Block lab using high-resolution optical tweezers (**Figure 4A**), had predicted the formation of much a bigger transcription bubble (~85-bp) by the action of Ssl2 during initial transcription (Fazal et al., 2015). Clearly, a lot of work remains to be done by single-molecule biophysicists to bring the understanding of mechanistic of the Pol II transcription initiation complex anywhere close to the bacterial one.

Transcription initiation by Pol III depends heavily on the transcription factor IIIB (TFIIIB), which is a complex formed by Brf1 (or Brf2), TBP and Bdp1. A recent study combined the use of x-ray

crystallography, TIRFM-smFRET (**Figure 1DGH**) and biochemical analysis to provide structural and functional insight into the assembly process of TFIIIB on the U6 snRNA promoter DNA (Gouge et al., 2017). smFRET in particular provided the means to monitor the TBP mediated bending of the promoter and thus the binding dynamics of TFIIIB and its subcomplexes to the promoter.

Elongation

Bacterial transcription machinery

Early single-molecule experiments following bacterial transcription elongation concluded that RNAP is a strong motor that withstands hindering force up to ~25 pN (Wang et al., 1998) and that single RNAPs progress at similar average rates (Schafer et al., 1991; Yin et al., 1995). However, improvements to the optical tweezers spatiotemporal resolution (Lang et al., 2002; Wuite et al., 2000)(**Figure 3AB**) have allowed more detailed studies and have demonstrated that the steady transcription elongation is halted by ~10-100 s pauses at specific sequence sites (Davenport et al., 2000; Forde et al., 2002). Because the probability to enter these long-lived pauses is force dependent, it has been suggested that the pauses originated from RNAP backtracking, i.e. backward sliding of the polymerase on the DNA that drives the 3'-RNA end out of register to the NTP entry channel. Having improved the optical tweezers assay sufficiently, backtracking was finally directly observed by Shaevitz *et al.* They showed that the backtracked pauses (**Figure 4C**) have long lifetimes, ranging from 20 s to above 30 min (Shaevitz et al., 2003). They also monitored that the pause duration is significantly reduced by the addition of GreA and GreB transcription factors, which bind the RNAP at the NTP entry channel (also known as secondary channel) and restore the elongation competent translocation register by stimulating the cleavage of the overhanging 3'-end of the RNA. However, all pauses did not originate from backtracked RNAP; these shorter (1-10 s), force independent pauses known as ubiquitous pauses (or elemental pauses) interrupts the progress of RNAP at any sequence position, while the probability to pause at any particular position is low (Adelman et al., 2002; Neuman et al., 2003). After the spatiotemporal resolution of optical tweezers has been pushed to its limit (Greenleaf et al., 2005; Moffitt et al., 2008), the movement of bacterial RNAP could be monitored step-by-step manner, i.e. in 1 bp resolution, at low NTPs concentration (up to ~10 µM) (Abbondanzieri et al., 2005; Righini et al., 2018). This type of data led to the formulation of a Brownian ratchet model of RNAP translocation during nucleotide addition cycles (Abbondanzieri et al., 2005). The Block lab further scrutinized the RNAP pausing and found that the ubiquitous pauses are indeed sequence dependent, similar to the long-known sequence encoded *his* and *ops* pauses (**Figure 4AB**) (Herbert et al., 2006). In combination with biochemical results, a new model has been developed (Landick, 2006), where the RNAP has a certain sequence-dependent probability to isomerize into a catalytically inactive conformation, the ubiquitous pause state, today more widely known as the

elemental pause state. The short-lived elemental pause state (typical half-life ~2 s) may further isomerize into a more stable (longer) pause by backtracking or a conformational change triggered when a RNA-hairpin forms in the RNA exit channel of RNAP (Guo et al., 2018; Kang et al., 2018). The sequence code (concensus: $G_{-10}Y_{-1}G_{+1}$; Y standing for pyrimidine) imposing the elemental/ubiquitous pause has been unraveled when it became possible to determine the exact locations of the RNAPs on the transcribed genes by massive parallel sequencing of the nascent RNAs 3'-ends, as RNAPs are enriched at pause sites. Single-molecule optical tweezers assay as well as biochemical approaches provided detailed mechanistic dissection of the consensus pause sequence (Larson et al., 2014). Interestingly, the pauses, elongation rate and processivity of the RNAP are not affected when the RNA transcript is pulled by a force up to 30 pN, i.e. twice stronger as is typically needed to melt RNA secondary structures (Dalal et al., 2006), showing that RNA structure has little influence on ubiquitous transcriptional pausing.

The pausing behavior of the bacterial RNAP is also influenced by external transcription factors. For example, NusA increases the probability and lifetime of the elemental (short-lived) and RNA hairpin stabilized (long-lived) pauses (Zhou et al., 2011), while NusG has an opposite effect (Herbert et al., 2010). The RNAP elongation rate is also affected by the amount of supercoil generated by the polymerase when it transcribes a torsionally contrained DNA molecule, e.g. the bacterial chromosome. Using optical microscopy, it was first observed that the RNAP generates torque during transcription elongation (Harada et al., 2001). Ma and collaborators used an optical torque wrench (**Figure 2D**) to control the amount of torque applied to the DNA and showed that the bacterial RNAP generates and sustains a torque up to ~11 pN.nm, i.e. enough to melt DNA; if the RNAP is stalled by excessive resisting torque, the complex eventually resumes elongation when the torque is released (Ma et al., 2013).

Bulky DNA lesions on the template strand stall RNAP. Bacteria have evolved a mechanism to utilize this RNAP property to detect the harmful DNA lesions and guide the start the repair process. When RNAP stalls on the DNA lesion, it is recognized by Mfd, an ATP-dependent DNA translocase. Mfd dissociates the RNAP from DNA and recruits the UvrABC endonuclease to cleave off and repair the damaged DNA. Recently, several single-molecule studies have dissected this transcription-coupled DNA repair pathway. Howan and coworkers used magnetic tweezers (**Figure 2C**, **Figure 3CD**) to observe the dynamic interactions of Mfd with the stalled RNAP (Howan et al., 2012). They found that Mfd binds and dissociates the RNAP in an ATP-dependent process, reaching an intermediate state where Mfd is simultaneously bound to the DNA and the RNAP. Subsequently, Mfd dissociates the RNAP from the DNA in another ATP-dependent step of a remarkable duration (~6 minutes). However, it remained unclear whether the RNAP dislodged by Mfd from the DNA lesion site remained bound to the Mfd/DNA and whether the RNAP retained the RNA transcript. To answer these questions, Graves and co-workers combined smTIRFM with their magnetic tweezers assay to simultaneously monitor both the real-

time composition (using fluorescence) and the catalytic state (using the magnetic tweezers) of the transcription-coupled repair machinery (**Figure 4E**) (Graves et al., 2015). They showed that the RNAP releases the RNA transcript when the Mfd dislodges the RNAP from the DNA lesion. Interestingly, the formed RNAP–Mfd complex is stable and translocates thousands of base-pairs on the DNA (Graves et al., 2015). The studies from the Strick lab did not investigate how Mfd translocated along the DNA molecule before finding a stall RNAP. Using an optical tweezers DNA hairpin assay (**Figure 3A**), where the hairpin is progressively opened upon Mfd progression, further revealed that the Mfd independently translocates along the DNA at ~ 7 bp/s; this rate is too slow to follow a normally transcribing RNAP, but enough to catch up with a stalled RNAP (Le et al., 2018). Collectively, the locomotive action of the Mfd assists the RNAP to either overcome translocation arrest on, e.g. a strong pause site, or to remove and terminate transcription for an RNAP stalled on an insurmountable obstacle (Le et al., 2018).

On highly expressed genes, e.g. ribosomal RNA genes, multiple RNAPs transcribe simultaneously the same gene. If the leading RNAP encounters an obstacle (e.g. a pause or DNA-bound protein) the trailing RNAPs will catch up the stalled leading RNAP, push it forward and rescue into active transcription (Epshtein and Nudler, 2003). To investigate whether the rescue of the leading RNAP by the trailing RNAPs could be linked to transcriptional bursting, Fujita and co-workers derived a smTIRFM assay that allow both monitoring the production of messenger RNAs by single-molecule fluorescence *in situ* hybridization (smFISH) and locating quantum dot labeled individual RNAPs on the template DNA (Fujita et al., 2016). Their mathematical modeling of the observed transcription dynamics supports the assumption that significant amount of transcriptional bursting simply stems from the arrest of the leading RNAP and its rescue by the trailing RNAPs.

Eukaryotic transcription machineries

In eukaryotic transcription, the elongation phase has been first studied with Pol II (Lisica and Grill, 2017). Seminal work from Galburt and coworkers used a high-resolution optical tweezers assay (**Figure 4A**) to show that Pol II molecules ceased to transcribe and were unable to recover from backtracks (**Figure 4C**) at a force of ~8 pN, only one-third of that determined for *E. coli* RNAP (Galburt et al., 2007). Most Pol II pauses were explained by backtracking. TFIIS — an eukaryotic analogue of the bacterial Gre factors — rescues Pol II from backtrack by catalyzing the cleavage of the protruding 3'-RNA end and thus allows Pol II to work against a two-fold higher hindering force. The authors suggested that there exists a full layer of transcription elongation regulation that depends on transcription factors modifying the mechanical performance of Pol II. Lisica and coworkers investigated further the mechanism of backtrack recovery by Pol I and Pol II using a similar optical tweezers assay (**Figure 4A**) (Lisica et al., 2016). Backtracking was enforced by pulling the polymerase backward with a rapidly spiked, strong hindering force. Analysis showed

that the recovery from shallow backtracks takes place via 1D diffusion of the RNAP, while recovery from deeper backtracks depends on RNA cleavage. Many transcription factors are expected to affect the elongating Pol II. For example, TFIIF —a transcription initiation factor involved in the recruitment phase of the Pol II-PIC— was shown to be also active during elongation, reducing the backtracking propensity of Pol II (Ishibashi et al., 2014). The only elongation factor found in eukaryotes, archaea and bacteria —Spt4/5 (Spt5 is homologous to bacterial NusG)— on the other hand appears to regulate Pol II transcription through the nucleosome (REF “The elongation factor Spt4/5 regulates RNA polymerase II transcription through the nucleosome” NAR 2017 by Reese lab). The GC content of the nucleic acids has also been shown to influence Pol II pausing dynamic. Indeed, it was reported that Pol II transcription elongation encounters less and shorter pauses when the DNA template is GC-rich than when it is AT-rich (Zamft et al., 2012). The authors suggest that the strong and bulky secondary structures, preferentially formed in the GC-rich RNA transcript, prevent Pol II from backtracking. Consistently, RNAse mediated degradation of the transcript abolishes the GC-content dependent pausing bias.

During eukaryotic transcription, polymerases have to pass through nucleosomes, i.e. a 146 bp stretch of DNA wrapped around a bundle of eight histone proteins. Nucleosomes form a mechanical barrier to transcription and consequently the accessibility of the DNA for transcription is also regulated by histone chemical modifications and ATP-dependent nucleosome remodeling enzymes. How an elongating Pol II bypasses a nucleosome that is placed on its path was investigated using optical tweezers (Hodges et al., 2009). The authors derived a mathematical model for the observed Pol II dynamics in the vicinity of the nucleosome and concluded that the polymerase, instead of actively separating the DNA from the histones, waits for fluctuations that locally unwrap the nucleosome and allow the Pol II to advance. In a follow-up study, the roles of various nucleosomal elements were investigated as a function of the strength and location of the barrier to transcription (Bintu et al., 2011). Specifically, the authors determined, how the trajectories of individual Pol II complexes transcribing past nucleosomes responds to the modifications in specific histone-DNA interactions or histone tails. They observed that the DNA unwrapped and re-wrapped faster around the tails-free histones, which favors Pol II movement closer to the nucleosome. In addition, they noted that point mutations compromising the DNA-histone interactions at the center of the nucleosome (dyad) decreases the local re-wrapping rate of the DNA and thus removes a barrier for Pol II to translocate forward, and that the nucleosomes amplify Pol II sequence-dependent pausing. The Block lab investigated the fundamental steps of the nucleotide addition cycle — substrate selection, catalysis and translocation — using Pol II mutants with altered trigger loop function (**Figure 4A**) (Larson et al., 2012). The trigger loops is a mobile and conserved RNA polymerase subdomain that stabilizes substrate NTPs in the active

site. The global fits to the force-velocity curves they extracted converge with a branched Brownian ratchet model for elongation, where the incoming NTP binds either the expected post-translocated state or the pre-translocated state, similarly to what was previously proposed for the bacterial RNAP (Abbondanzieri et al., 2005). The latter binding mode is expected to take place in the pre-insertion site and does not require the NTP to form interactions with the active site bound template base. Furthermore, the trigger loop was suggested to control the transitions between the pre- and post-translocated states. Another study by the Bustamante lab utilized nucleosomes as specific barriers to forward translocation – a trick that allows to determine separately both the forward and reverse translocation rates of Pol II, and further estimate all main kinetics parameters involved in the nucleotide addition and pausing phases of transcription elongation (Dangkulwanich et al., 2013). In contrast to the earlier studies that had assumed the polymerase to reach fast equilibrium between the pre- and post-translocated states prior and after each RNA elongation step (Abbondanzieri et al., 2005; Larson et al., 2012), the authors found that the forward translocation rate (88 s^{-1}) of Pol II is actually similar to the RNA extension rate (35 s^{-1}). Therefore, the translocation and RNA extension together constitute the rate-limiting steps in the nucleotide addition cycle. From these findings, a simpler, the authors proposed a linear Brownian ratchet model of transcription elongation, where the incoming NTP binds only to the post-translocated state, which is consistent with biochemical evidences.

Termination

Bacterial transcription machinery

Larson and coworkers (Larson et al., 2008) investigated the molecular mechanism of transcription termination at the single-molecule level using optical tweezers. The authors characterized three different terminators (*his*, t500, and tR2) each consisting of a GC-rich hairpin followed by a U-rich tract. Two distinct termination mechanisms for these intrinsic terminators were observed, namely termination by forward hypertranslocation or RNA:DNA hybrid shearing. When observing the forward translocation strategy, the authors found that the RNAP hypertranslocates forward by ~1.5 bp leading to shorter RNA-DNA hybrid, which destabilizes the complex enough for subsequent termination. In the RNA:DNA hybrid shearing strategy, the U-rich tract forms a weak RNA-DNA hybrid, which is easily further destabilized by the folding of an upstream RNA hairpin. By pulling the hairpin from the 5'-RNA end, two modes of action were observed: (i) at a force larger than the hairpin melting force, the termination efficiency increases because of the shearing of the RNA-DNA hybrid; (ii) at a force lower than the hairpin melting force, the termination efficiency also increases, though here because the pulling force modulates the balance of termination hairpin and other secondary structures in the RNA. The authors concluded that the most frequent cause of termination failure is the folding of the RNA into one of the competing

secondary structures. The termination by hypertranslocation likely dominates in the sites where the energy penalty of shearing the RNA-DNA hybrid is higher compared to RNAP forward hypertranslocation.

Frieda and Block developed an optical-trap based assay to monitor the co-transcriptional folding of the *pbuE* riboswitch (Frieda and Block, 2012). This riboswitch regulates the concentration of adenine in the cell by forming an adenine binding aptamer. The folding of the aptamer, which is stabilized by the binding of an adenine, prevents the formation of the competing terminator hairpin and thus allows the expression of the downstream genes. The authors determined that aptamer-dependent RNAP termination is a function of the adenine concentration and the applied force to the RNA (**Figure 4D**); they also identified the folding signature of the riboswitch. The termination vs. read-through outcome turns out to be kinetically controlled indicating that the riboswitch based regulation of gene expression is mechanistically tightly linked with the transcription elongation kinetics and the regulatory layer that controls the elongation kinetics in the cell.

In vitro single-molecule studies of viral RNA-dependent RNA polymerases

RNA viruses are particularly remarkable for their diverse genome replication strategies. The genome of RNA viruses is either positive (+), negative (-) or double-stranded (ds). The protein synthesis machinery of the host cell directly employs the positive virus genome as the template whereas the negative genome must first be copied into a complementary strand. The RNA viruses rely on the viral polymerase, formally called the RNA-dependent RNA polymerase (RdRp), to replicate and transcribe their genomes (te Velthuis, 2014). The process of genome replication and transcription is divided in two main phases – initiation and elongation. Viruses have developed many different strategies to initiate replication and transcription. For example, Φ6 bacteriophage and flaviviruses, e.g. dengue, employ *de novo* initiation (i.e. no primer) to replicate their genomes. Influenza virus initiates genome transcription by primer-extension but employs *de novo* initiation to begin the replication of its (-)RNA genome. Poliovirus primes replication of its (+)RNA genome using a VPg (viral protein genome-linked) attached at the 5’ end of the viral genome (Fields et al., 2013). The initiation phase is critical for viral survival because its specificity and efficiency must ensure sufficient synthesis of viral RNA by the RdRp to meet the demand of both viral proteins synthesis and virion assembly. After a successful initiation, the RdRp enters the elongation phase. This phase is equally important for the RNA virus because the full length genomes are necessary for correct translation and virion assembly. Viral RdRps have a relatively high misincorporation rate that serves to increase the genome diversity of the virus population, thus helping the virus to evade the host immune response (Lauring et al., 2013).

Single-molecule techniques have only recently been applied to study the genome replication/transcription of RNA viruses. Here, we describe the results from single-molecule experiments that have shedded light on the RdRp initiation mechanism on influenza A virus (IAV) and the elongation dynamics of Φ6 P2 and poliovirus RdRps.

Replication and transcription initiation in influenza A virus

IAV is a segmented (-)RNA virus, meaning that its genome is divided into eight segments of viral single-stranded RNA (vRNA). The RNA strands form a ribonucleoprotein complex with viral nucleoproteins and the partially complementary 3' and 5' RNA ends are bound by a single copy of IAV RdRp. The IAV RdRp is formed by three individual polypeptides called PB1, PB2 and PA. Each of them has a separate task in IAV genome processing: PB1 is the core polymerase, PB2 is the cap-binding domain, and PA is the metal ion-dependent endonuclease (Stubbs and Te Velthuis, 2014). This heterotrimeric complex replicates and transcribes the vRNA. The 3' and 5' ends of the vRNA are highly conserved, and hybridize to form a partially double-stranded panhandle structure that takes the shape of a corkscrew and specifically binds to the RdRp (Te Velthuis and Fodor, 2016). Though this structure has first been suggested using functional studies, a structural confirmation was lacking. In an elegant study using confocal smFRET (**Figure 1DEF**), Tomescu and co-workers have mapped the structure of the hybridized termini bound to the RdRp (Tomescu et al., 2014). The authors have studied the FRET efficiency for different FRET pair locations along the RNA, and determined the inter-dye distances from the measured FRET efficiencies. The measured distances — determined separately for the free and RdRp-bound RNA — are consistent with the RNA corkscrew structure model when RdRp is bound to the panhandle RNA structure (**Figure 5A**). Robb and co-workers have recently expanded this work (Robb et al., 2016) by characterizing the 3'-RNA end structure. They showed that the 3' end of the vRNA takes two alternative conformations upon RdRp binding, one bound on the RdRp surface in the pre-initiation state and another bound to the active site in the initiation competent state. Both conformations are present at equilibrium in the absence of NTPs, while the initiation state is favored in the presence of dinucleotide that mimicks the state of the complex after the synthesis of a 2-bp RNA (**Figure 5B**).

Once complexed with the panhandle structure of the viral genome termini, the RdRp starts either transcription or replication. In the transcription mode, the RdRp captures and cleaves a capped host 5'-mRNA to prime the mRNA synthesis from the viral genome. In the replication mode, two steps are needed to produce a new copy of the vRNA. In the first step, known as terminal initiation, a complementary RNA (cRNA) intermediate is produced from the vRNA when the RdRp initiates *de novo* by joining together the NTPs complementary to the first two residues of the 3'-

vRNA end. In the second step known as internal initiation, the initiation takes place at the cRNA positions 4 and 5 leading to the synthesis of pppAG dinucleotide, which subsequently realigns with the positions 1 and 2 and is elongated by the RdRp (Te Velthuis and Fodor, 2016). Viral RdRps that initiate replication *de novo* generally contain a priming loop domain, which stacks the 3'-RNA end of the template RNA strand and the first nucleotide of the product strand (Te Velthuis and Fodor, 2016). The IAV RdRp supports both primer-dependent and *de novo* initiation, however it remained unknown whether the priming loop is involved in the terminal and internal phases of replication initiation or in transcription initiation. te Velthuis and co-workers used *in vitro* and *in vivo* ensemble biochemical assays together with confocal smFRET to investigate the question (Te Velthuis et al., 2016). They showed that the priming loop is indeed needed to support *de novo* terminal replication initiation but not the internal replication initiation or primer-dependent transcription initiation. Interestingly, the priming loop actually represents an obstacle to transcription, as its removal being the rate-limiting step for primer-dependent transcription initiation.

The above studies used confocal smFRET to pave the way for understanding the inter- and intramolecular conformational changes occurring during IVA RdRp mediates replication and transcription initiation, thereby complementing x-ray and cryo-EM based structural studies. Future work using TIRFM-smFRET (**Figure 1DGH**) will allow the observation in real-time of the full trajectories of individual RdRp complexes engaged in the initiation of replication or transcription, which will provide the detailed dynamics of the IAV initiation mechanisms.

Φ6 P2 RdRp transcription and poliovirus RdRp replication elongation kinetics

The RdRps from RNA viruses have to replicate or transcribe the ~5-30 kb long viral RNA in order to produce new viral genomes to be incorporated into the new generation of virions or to provide templates for translation (te Velthuis, 2014). The elongation phase of replication is also very important for the viral evolution because the nucleotide misincorporations made by the RdRp are the main source of genetic diversity in the virus population (Lauring et al., 2013). Typical RdRp error incorporation rate, ~10^{-3}-10^{-4} per a nucleotide incorporation cycle, is one of the highest in all replication machineries (Holmes, 2010). However, the high error rate bears a fitness cost and must therefore be tightly balanced: an error rate too low would leave the virus unable to adapt to the host immune defense, while an error rate too high would be detrimental for the production of a sufficient amount of active virions (Smith, 2017). The kinetics of nucleotide incorporation and misincorporation have been heavily studied using fast mixing enzyme kinetics assay, such as quenched flow and stopped flow (Cameron et al., 2016). These approaches offer exquisite resolution, i.e. single nucleotide additions at millisecond timescale, however this resolution is only attainable for short templates, typically less than 10 nucleotides. Even though the

misincorporation rate of RNA viruses is high, it still remains a rare event and can only be observed in bulk assays in the absence of the correct, templated nucleotide. The viral replication also represents an important target for antiviral therapeutic strategies, currently taking the advantage from the large library of antiviral nucleotide analogues. Nucleotide analogue incorporation studies, similar to the nucleotide misincorporation studies, suffer from the limitation that these events are rare when observed in the presence of natural nucleotides. Therefore, an experimental approach compatible with the use of natural length templates, i.e. few kilobases, and discrimination power capable of distinguishing rare misincorporation or nucleotide analogue incorporation events in the background of normal replication/transcription would greatly benefit the mechanistic studies of viral RdRp mediated RNA elongation.

Single-molecule force spectroscopy approaches, especially optical tweezers and magnetic tweezers, come close to fulfill the specific technical requirements of RdRp elongation studies by offering the possibility to observe the RNA synthesis by individual RdRp's on kilobase(s) length RNA templates at ~10-100 ms temporal and near base pair spatial resolution. However, to observe events as rare as 10^{-3} per nucleotide incorporation cycle, highly multiplexed approach is needed. Unlike fluorescence spectroscopy techniques, force spectroscopy techniques have historically suffered from poor throughput. However, this limitation has been recently overcome with the development of several high-throughput techniques (Dulin et al., 2015a; Hill et al., 2017). One of the high-throughput enabling solutions involved upgrading magnetic tweezers apparatus with the latest generation of large field of view CMOS cameras with a real-time image analysis algorithm, capable of tracking hundreds of individual molecules in parallel (Berghuis et al., 2015; Cnossen et al., 2014). The real-time high-throughput magnetic tweezers were first applied to characterize the nucleotide incorporation dynamics of RdRps from bacteriophage Φ6 P2 (has dsRNA genome) and human poliovirus [(+)RNA] (Dulin et al., 2017a; Dulin et al., 2015b; Dulin et al., 2015c).

To study the viral RdRps, a double-stranded RNA tether is used to attach the magnetic bead to the surface of a microscope coverslip. When RdRp employs the dsRNA as a template for RNA synthesis, it gradually displaces the template RNA strand leaving the bead anchored to the surface via a single-stranded RNA. The progress of the RdRp action is monitored in real-time as the movement of the bead further away from the surface (**Figure 6A**). Large data sets of Φ6 P2 RdRp transcription activity were acquired at different NTP concentrations and applied force. Interestingly, Φ6 P2 RdRp shows fast bursts of nucleotide additions that are interrupted by pauses of 1-1000 s duration (**Figure 6B**). Previously developed data analysis approach characterizes separately the nucleotide addition and the pause kinetics by picking the pauses out of the traces. However, it is impossible to distinguish pauses shorter than ~1 s because of the finite spatiotemporal resolution of tweezers assay, and therefore the nucleotide addition kinetics is polluted by the missed short pauses (Dulin et al., 2015b). To overcome this issue, a dwell time

analysis combined with a Maximum Likelihood Estimation (MLE) fit has been developed to extract the elongation kinetics parameters, i.e. rates and probabilities, at once without sorting the pauses out of the traces (**Figure 6C**) (Dulin et al., 2015a). Using this dwell time analysis, the probability and average interconversion rates of the catalytic and non-catalytic states as well as the nucleotide addition rates are recovered. The dwell time distribution of Φ6 P2 RdRp synthesizing each consecutive 10 nt stretch of RNA (a limit set by the resolution of the used magnetic tweezers assay) was measured and found to be composed of four subdistributions (**Figure 6C**). The shortest dwell times (<1 s at saturating NTP concentration) is assigned to the pause-free nucleotide addition rate ((i) in **Figure 6C**). Intermediate dwell times (~1-10 s) are split into two populations exponentially distributed, representing pauses of short (Pause 1) and intermediate durations (Pause 2) ((ii) in **Figure 6C**). Finally, the distribution of the longest dwell times (>20 s) is best described by a power law ($t^{-3/2}$), suggesting a backtrack state for the polymerase (Depken et al., 2009). Furthermore, Pause 1 and Pause 2 probabilities and lifetimes are surprisingly dependent on the NTP concentration, and Pause 2 probability is affected by inosine triphosphate incorporations. These findings suggest that Pause 1 and Pause 2 are intimately linked with nucleotide misincorporation (**Figure 6D**). It was formely believed that RdRp misincorporation events were a rare incident happening along the same catalytic pathway as the correct nucleotide incorporation. The new model derived from the single-molecule data, in contrast, suggests that Φ6 P2 RdRp has two catalytic conformations: a high fidelity catalytic state (HFC) and a low fidelity catalytic state (LFC), respectively (**Figure 6D**). Majority of the RNA synthesis by RdRp takes place on the HFC pathway and leads to the rapid incorporation of the correct nucleotides to the RNA. However, the RdRp has a certain probability to isomerize into the LFC conformation, which leads to a slow nucleotide addition, i.e. Pause 1 (**Figure 6C**). The LFC has also an elevated probability (though still low in absolute terms) to elongate the RNA product strand with a wrong nucleotide. Upon misincorporation, Φ6 P2 RdRp enters an even slower catalytic state, i.e. Pause 2 or the terminal mismatched catalytic (TMC) (**Figure 6CD**), as the catalytic activity is further compromised by the mismatched 3'-RNA end.

A follow-up study focusing on poliovirus RdRp (**Figure 6E**) revealed very similar elongation kinetics compared to the Φ6 P2 RdRp; the co-existence of high and low fidelity catalytic conformations emerges thus as a general property of viral RdRps (Dulin et al., 2017a). The mechanistic model defining the intermediate pauses as misincorporation events was further corroborated by a set of experiments performed with an error-prone poliovirus RdRp mutant. Specifically, the mutator RdRp had a three-fold increased probability to enter Pause 2, an increase similar to what was determined using deep sequencing (Korboukh et al., 2014). Taking advantage of the high-throughput capability of the magnetic tweezers, the effects of five nucleotide analogues on the replication activity of poliovirus RdRp was investigated with a

physiological concentration of NTP (100 µM, saturating condition relative to the K_m). The tested compounds included the mutagenic nucleotide analogue ribavirin triphosphate (RTP), inosine triphosphate (ITP), obligatory chain terminator 3'-deoxy ATP, non-obligatory chain terminator 2'-C-met-ATP and T1106-triphosphate (T1106-TP) whose mechanism of action was unclear until then. As expected from the misincorporation–pause model, RTP and ITP specifically increase the occurrence of Pause 2. Also as expected, the chain terminators 2'-C-met-ATP and 3'-deoxy ATP decrease the processivity of the replicating RdRp, e.g. the median processivity drops from ~1200 to ~400 nt when 100 µM 3'-deoxy ATP is added to the 100 µM of natural NTPs. However, this result also demonstrates how strongly the poliovirus RdRp selects against the nucleotide analogues, i.e. ~400 correct nucleotide incorporations take place before one 3'-dATP is added to the elongated RNA. Finally, the data revealed that the addition of T1106 to the RNA chain unexpectedly triggers the RdRp to enter a unique long-lived pause, seemingly backtrack-related (**Figure 6EF**). In conclusion, high-throughput magnetic tweezers have provided new insights into the mechanisms of viral RdRps replication activity and antiviral nucleotide analogues function.

The backtracking activity of Φ6 P2 RdRp has also been characterized with high-throughput magnetic tweezers (**Figure 6G**) (Dulin et al., 2015c). The probability of Φ6 P2 RdRp to enter long-lived backtracked states decreases with the increase of the applied force. Because the force destabilizes the ds-ssRNA junction in the front of the RdRp, the dominant factor determining the RdRp backtracking appears to be the base pairing energy at the dsRNA fork. Surprisingly, it was also found that the extensively backtracked 3'-RNA end of the newly synthesized product strand may be used as a template for *de novo* initiation by another RdRp. Eventually, the second RdRp pushes the first RdRp backward, all the way to the upstream end of the product strand, which produces the "reversal" traces shown in **Figure 6GH**. One possible biological function of the reversal mechanism could be assisting viral RNA recombination — another important evolutionary pathway. The ~2-fold higher rate of the reversal transcription compared to the forward transcription (Dulin et al., 2015c) may also provide the virus with a more efficient viral RNA production pathway in the host cell.

Perspective

Single-molecule techniques have offered a complete new angle on the understanding of the molecular mechanisms of cellular and viral RNA polymerases. The unique power of single-molecule approaches largely arises from its ability to resolve individual steps in complex reaction pathways, competing reaction pathways and multiple co-existing conformations. The force spectroscopy methods additionally allow nanomanipulation of the biological molecules (pushing, pulling and twisting) thus creating a versatile experimental tool that can be used to steer the energy landscape of biomolecular reactions. Technical improvements in observation

parallelization (Dulin et al., 2015a; Hill et al., 2017) and spatiotemporal resolution (Laszlo et al., 2017), or the combination of fluorescence and force spectroscopy (Comstock et al., 2011; Heller et al., 2013; Madariaga-Marcos et al., 2018; van Loenhout et al., 2012) will allow to monitor the activities and structural dynamics of individual RNA polymerase molecules in ever more accurate and complex settings.

Funding

DD was supported by the Interdisciplinary Center for Clinical Research (IZKF) at the University Hospital of the University of Erlangen-Nuremberg. A.M.M. was supported by the Academy of Finland (grant number: 307775).

Figure Captions

Figure 1: In vitro single-molecule fluorescence spectroscopy techniques in transcription studies. (A) Single-molecule fluorescence co-localization microscopy monitors in real-time binding-unbinding events of interacting molecules (Churchman and Spudich, 2012; Friedman et al., 2006; Selvin et al., 2007). Total internal reflection fluorescence microscopy (TIRFM) is used to image surface-attached biomolecules (Axelrod, 2001). In TIRFM, the excitation laser is reflected at the glass-water interface, generating an evanescent wave (green shade) above the coverslip (CS) that excites dyes within ~100 nm from the coverslip top surface. Non-specific adsorption of the labeled biomolecules is blocked with specific surface chemistry, e.g. polyethylene glycol coating (Chandradoss et al., 2014). **(B)** Observed variation in fluorescence at a specific location where two biomolecules successively bind (1) and dissociate (2) from the nucleic acid molecule as depicted in panel (A). **(C)** By stretching the nucleic acids, as in flow stretching (Geertsema et al., 2015) or in DNA curtains (Collins et al., 2014), one is able to localize the DNA binding molecules with ~100 nm accuracy. **(D)** In single-molecule Forster Resonant Energy Transfer (smFRET), a fraction of the excitation energy of the donor dye (green) is transferred non-radiatively to the acceptor dye (red) with an efficiency E_{FRET} that decreases when the distance between the two dyes increases (Ha et al., 1996). Using this molecular ruler, distances ranging from 2 to 10 nm can be determined (Hohlbein et al., 2014). **(E)** In confocal microscopy, dye-labeled biomolecules at low concentration diffuse freely in the solution and a short burst of photons is detected when a labeled molecule crosses the confocal detection volume. To increase sensitivity, photons emitted outside the confocal volume are spatially filtered out in the imaging optical path. smFRET combined with alternative-laser excitation (ALEX) in confocal microscopy is a powerful tool to access conformational equilibrium and complex formation by biomolecules in solution (Hohlbein et al., 2014; Kapanidis et al., 2004). **(F)** Confocal smFRET provides the prevalence and the number of different conformations, which are revealed by the amplitude and the width of the normal distributions describing the E_{FRET} histogram, respectively; however, solution smFRET cannot obtain the rates of conformational change. **(G)** To observe the conformational changes happening in an individual biomolecule, which has to be immobilized to the surface of a coverslip. **(H)** E_{FRET} (grey) time trace, typically obtained using TIRFM-based smFRET, shows the biomolecule to interconvert between two distinct conformations. The kinetic constants defining the stabilities of the two states are recovered using a Hidden Markov Model (blue) (van de Meent et al., 2014). Single-molecule FRET can also determine accurate distances between the donor and acceptor dyes; the obtained distances help to model the 3D structure of the biomolecule (Beckers et al., 2015; Kalinin et al., 2012).

Figure 2: Single-molecule force spectroscopy techniques applied to transcription studies. **(A)** In an optical tweezers assay, one possible configuration consists to attach a nucleic acid molecule (NA) from one end to the glass coverslip of the flow chamber and from the other end to a polystyrene bead trapped in a focused laser beam. The bead position can be controlled in three dimensions by optically moving the trap. Displacement of the bead from the equilibrium position, i.e. the center of the trap, increases linearly the force F experienced by the NA tether. F ranges from ~0.1 pN to hundreds of pN (Neuman and Block, 2004). **(B)** Most modern optical tweezers utilize two optical traps to pull the NA from both ends. The configuration produces a signal with smaller drift artefact and thus a resolution high enough to distinguish translocation steps at ~0.34 nm/s velocity (Greenleaf et al., 2005). **(C)** In a magnetic tweezers assay, double or single stranded NA is attached from one end to the coverslip and from the other end to a magnetic bead. The force F (from ~10 fN to ~1 nN) is generated by pulling the bead with a pair of magnets located above the flow chamber. A reference bead on the coverslip surface is used to correct for the mechanical vibrations caused drift in the position of the sample bead (Vilfan et al., 2009). Camera-based detection allows the simultaneous tracking of hundreds of beads at near base pair resolution. The magnetization (m_0) of the bead permits its rotation, by rotating the magnetic field originating from the magnets. Rotation of the bead adds supercoiling to the torsionally constrained double stranded NA, which eventually leads to the formation of plectonems. Torque-dependent behavior of protein-NA interactions can thus be studied (De Vlaminck and Dekker, 2012). **(D)** In an optical torque wrench, a birefringent particle (here a cylinder) is trapped by a

polarized laser beam. By rotating the polarization of the laser, the birefringent cylinder rotates and induces supercoiling in the NA (La Porta and Wang, 2004; Santybayeva and Pedaci, 2017). The "sticky" ends of the NA, necessary in optical tweezers and magnetic tweezers assays, are typically generated with biotin and digoxigenin labeled nucleotides that bind very stably to streptavidin/neutravidin and anti-digoxigenin antibody, respectively, coating the bead or coverslip surface (Janissen et al., 2014).

Figure 3: In vitro single-molecule studies of transcription initiation by multisubunit RNA polymerases. (A) The free energy landscape of the RNAP–DNA promoter interaction is manipulated and characterized by progressively unwinding the promoter sequence present in the DNA hairpin by moving the optical traps further apart. **(B)** Force extensions curves from the experiments presented in (A) without (green) and with RNAP (blue) bound to the promoter in the hairpin stem. Figure adapted from Ref. (Meng et al., 2017). **(C)** In magnetic tweezers experiment, a torsionally-constrained DNA molecule is supercoiled by rotating a pair of permanent magnets above the flow chamber, which eventually induces plectonems in the DNA molecule. The conservation of the linking number — the sum of twist (the number of turns the two DNA strands make around each other) and writhe (the number of turns the double stranded DNA makes over itself) — implies that removing one twist in the DNA double helix must be compensated in change in the number of writhe. The unwinding of ~n twists (one twist is 10.6 bp of DNA) by the RNAP when it opens the transcription bubble during RNAP open-promoter complex (RP_O) formation results in the addition of n writhe in a positively supercoiled DNA. This leads to a decrease in the end-to-end distance of the DNA molecule by $n.\Delta z$, where Δz is the change in the distance of the DNA molecule upon the addition of one writhe at a given force. DNA scrunching during initial RNA synthesis (initial transcription) further unwinds the transcription bubble therefore increasing the number of writhe added to the DNA. **(D)** Conformational dynamics of RP_O (left) and abortive synthesis of 8-mer RNAs by the initially transcribing complex (ITC, right) was monitored by magnetic tweezers. Adapted from Ref. (Revyakin et al., 2006). **(E)** TIRFM-based smFRET assay can record the dynamics of initial transcription. RP_O complex is formed with a promoter containing a donor dye (green sphere) and an acceptor dye (red sphere) upstream and downstream of the transcription bubble, respectively. The RP_O is immobilized to the coverslip surface with an antibody. RNA synthesis is coupled to the promoter scrunching and movement of the downstream DNA towards the RNAP, leading to the change in the dye pair distance and E_{FRET}. **(F)** E_{FRET} was continuously measured with the assay described in (F) to monitor the magnitude of promoter scrunching and thus the progress of initial transcription. (F, G) are adapted from Ref. (Dulin et al., 2017b). **(G)** Model of alternative clamp positions in the bacterial RNAP was developed based on confocal microscopy smFRET experiments. Red, yellow and green indicate the positions explored by the β' clamp domain and the donor dye (spheres), respectively. Black sphere indicates the immobile position of the acceptor dye on the opposite side of the DNA binding cleft. Adapted from Ref. (Chakraborty et al., 2012). **(H)** Nanopositioning system was used to map the structure of archaeal pre-initiation complex by determining multiple distances between the transcription initiation factors (TBP, TFB and TFE) and the RNAP as well as between the non-template DNA strand and RNAP. The inter-dye distances were calculated from the measured E_{FRET} values. The donor and acceptor dye locations used in the study are indicated with green and red stars, respectively. Adapted from Ref. (Nagy et al., 2015).

Figure 4: Optical and magnetic tweezers assays to study transcription elongation and termination by multisubunit RNA polymerases. (A) Optical tweezers based assay allows to subject the RNAP either to assisting or hindering force depending on which direction the transcription is designed to progress on the template DNA. **(B)** Transcription activity traces for individual RNAPs obtained with the high-resolution optical tweezers assay depicted in (A). Adapted from Ref. (Herbert et al., 2006). **(C)** A close-up of a transcription activity trace shows the RNAP to backtrack on the template DNA in the optical tweezers assay. Adapted from Ref. (Shaevitz et al., 2003). **(D)** Optical tweezers transcription assay where the assisting force is applied to the nascent RNA. **(E)** Magnetic tweezers was combined with TIRFM to study transcription-coupled repair in Ref. (Fan et al., 2016; Graves et al., 2015). The bead position is

affected by the size of the transcription bubble thus transmitting information on the RNAP occupancy on the DNA and the stage of transcription. The binding of dye-labeled Mfd to the DNA or DNA-bound RNAP is inferred from the sudden appearance of a fluorescent spot on the coverslip surface. Because the strength of evanescent field decays exponentially with the distance from the surface, the position of Mfd on the DNA can be extracted from the intensity of the fluorescent spot. **(F)** Schematic of the optical tweezers assay monitoring Pol II transcription past a nucleosome.

Figure 5: Solution smFRET studies of *de novo* replication initiation by influenza virus RNA-dependent RNA polymerase. (A) Different configurations observed for the nucleic acid scaffold mimicking the 3' and 5' ends of the influenza RNA genome. The green and red spheres indicate the donor and acceptor dye positions, respectively. **(B)** Model for *de novo* replication initiation by influenza RdRp. Adapted from Ref. (Robb et al., 2016).

Figure 6: In vitro single-molecule studies of RdRp transcription elongation. (A) Magnetic tweezers assay can be used to study the dynamics of RdRp transcription elongation. The magnetic bead is tethered to the coverslip surface by a double stranded RNA that experiences a constant force. A short non-hybridized segment of the RNA template presents a free 3' end for RdRps to perform *de novo* initiation. To study primer dependent initiating RdRps, such as poliovirus RdRp, the 3' end of the template RNA is modified to contain a short priming hairpin. Following successful initiation, the RdRp elongates the RNA product strand, unwinding the template strand and converting the tether to ssRNA. In the process, the end-to-end distance of the tether changes, thus reporting on the RdRp activity. **(B)** 52 traces of transcribing Φ6 P2 RdRps were acquired in a single experiment using high-throughput magnetic tweezers (Cnossen et al., 2014). Adapted from Ref. (Dulin et al., 2015c). **(C)** Probability density distribution of the dwell times corresponding to the synthesis of ten consecutive nucleotides stretches of RNA. Four distinct dwell time distributions are fitted; these correspond to the pause-free nucleotide incorporation (nucleotide addition, green), short pauses (Pause 1, dark blue), intermediate pauses (Pause 2, light blue), and long pauses caused by polymerase backtracking (backtrack, red). Example trace snapshots above illustrate each dwell time type. Adapted from Ref. (Dulin et al., 2015b). **(D)** Nucleotide error incorporation model explains the dwell time distribution of the Φ6 P2 RdRp. The model details are explained in the main text. HFS, the high fidelity catalytic (HFC) state; LFC, low fidelity catalytic state; TMC, terminal mismatched catalytic state. **(E)** Poliovirus RdRp replication traces in the presence of 100 µM of NTPs and **(F)** 100 µM of NTPs with 10 µM of antiviral nucleotide analogue T1106-triphosphate. (E) and (F) are adapted from Ref. (Dulin et al., 2017a). **(G)** A fraction of Φ6 P2 RdRp transcription traces displayed "reversal" activity (arrows). **(H)** The reversal activity originates from a backtracked RdRp that presents a protruding 3' end of the product RNA strand, which is used by another RdRp as a template of transcription. The second RdRp pushes back the first RdRp resulting in the rehybridization of the original template and non-template strands. The shortening of the end-to-end distance of the tether is thus detected as a "reversal" trace. (G) and (H) are adapted from Ref. (Dulin et al., 2015c).

References

Abbondanzieri, E.A., Greenleaf, W.J., Shaevitz, J.W., Landick, R., and Block, S.M. (2005). Direct observation of base-pair stepping by RNA polymerase. Nature *438*, 460-465.

Adelman, K., La Porta, A., Santangelo, T.J., Lis, J.T., Roberts, J.W., and Wang, M.D. (2002). Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. Proceedings of the National Academy of Sciences of the United States of America *99*, 13538-13543.

Ahlquist, P. (2002). RNA-dependent RNA polymerases, viruses, and RNA silencing. Science (New York, N.Y *296*, 1270-1273.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Molecular biology of the cell, 4th edn (New York: Garland Science).

Axelrod, D. (2001). Total internal reflection fluorescence microscopy in cell biology. Traffic *2*, 764-774.

Beckers, M., Drechsler, F., Eilert, T., Nagy, J., and Michaelis, J. (2015). Quantitative structural information from single-molecule FRET. Faraday discussions.

Belogurov, G.A., and Artsimovitch, I. (2015). Regulation of Transcript Elongation. Annu Rev Microbiol *69*, 49-69.

Berghuis, B.A., Dulin, D., Xu, Z.Q., van Laar, T., Cross, B., Janissen, R., Jergic, S., Dixon, N.E., Depken, M., and Dekker, N.H. (2015). Strand separation establishes a sustained lock at the Tus-Ter replication fork barrier. Nature chemical biology *11*, 579-585.

Bintu, L., Kopaczynska, M., Hodges, C., Lubkowska, L., Kashlev, M., and Bustamante, C. (2011). The elongation rate of RNA polymerase determines the fate of transcribed nucleosomes. Nature structural & molecular biology *18*, 1394-1399.

Bockelmann, U., Essevaz-Roulet, B., and Heslot, F. (1998). DNA strand separation studied by single molecule force measurements. Phys Rev E *58*, 2386-2394.

Browning, D.F., and Busby, S.J. (2004). The regulation of bacterial transcription initiation. Nat Rev Microbiol *2*, 57-65.

Browning, D.F., and Busby, S.J. (2016). Local and global regulation of transcription initiation in bacteria. Nat Rev Microbiol *14*, 638-650.

Cameron, C.E., Moustafa, I.M., and Arnold, J.J. (2016). Fidelity of Nucleotide Incorporation by the RNA-Dependent RNA Polymerase from Poliovirus. Enzymes *39*, 293-323.

Chakraborty, A., Meng, C.A., and Block, S.M. (2017). Observing Single RNA Polymerase Molecules Down to Base-Pair Resolution. Methods Mol Biol *1486*, 391-409.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R*., et al.* (2012). Opening and closing of the bacterial RNA polymerase clamp. Science (New York, N.Y *337*, 591-595.

Chandradoss, S.D., Haagsma, A.C., Lee, Y.K., Hwang, J.H., Nam, J.M., and Joo, C. (2014). Surface passivation for single-molecule protein studies. Journal of visualized experiments : JoVE.

Charvin, G., Allemand, J.F., Strick, T.R., Bensimon, D., and Croquette, V. (2004). Twisting DNA: single molecule studies. Contemp Phys *45*, 383-403.

Churchman, L.S., and Spudich, J.A. (2012). Colocalization of fluorescent probes: accurate and precise registration with nanometer resolution. Cold Spring Harbor protocols *2012*, 141-149.

Cnossen, J.P., Dulin, D., and Dekker, N.H. (2014). An optimized software framework for real-time, high-throughput tracking of spherical beads. The Review of scientific instruments *85*, 103712.

Collins, B.E., Ye, L.F., Duzdevich, D., and Greene, E.C. (2014). DNA curtains: novel tools for imaging protein-nucleic acid interactions at the single-molecule level. Methods in cell biology *123*, 217-234.

Comstock, M.J., Ha, T., and Chemla, Y.R. (2011). Ultrahigh-resolution optical trap with single-fluorophore sensitivity. Nature methods *8*, 335-340.

Dalal, R.V., Larson, M.H., Neuman, K.C., Gelles, J., Landick, R., and Block, S.M. (2006). Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. Molecular cell *23*, 231-239.

Dangkulwanich, M., Ishibashi, T., Liu, S., Kireeva, M.L., Lubkowska, L., Kashlev, M., and Bustamante, C.J. (2013). Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. eLife *2*, e00971.
Davenport, R.J., Wuite, G.J., Landick, R., and Bustamante, C. (2000). Single-molecule study of transcriptional pausing and arrest by E. coli RNA polymerase. Science (New York, N.Y *287*, 2497-2500.
de Farias, S.T., Dos Santos Junior, A.P., Rego, T.G., and Jose, M.V. (2017). Origin and Evolution of RNA-Dependent RNA Polymerase. Front Genet *8*, 125.
De Vlaminck, I., and Dekker, C. (2012). Recent advances in magnetic tweezers. Annu Rev Biophys *41*, 453-472.
Depken, M., Galburt, E.A., and Grill, S.W. (2009). The origin of short transcriptional pauses. Biophysical journal *96*, 2189-2193.
Duchi, D., Bauer, D.L.V., Fernandez, L., Evans, G., Robb, N., Hwang, L.C., Gryte, K., Tomescu, A., Zawadzki, P., Morichaud, Z.*, et al.* (2016). RNA Polymerase Pausing during Initial Transcription. Molecular Cell *63*, 939-950.
Duchi, D., Gryte, K., Robb, N.C., Morichaud, Z., Sheppard, C., Brodolin, K., Wigneshweraraj, S., and Kapanidis, A.N. (2018a). Conformational heterogeneity and bubble dynamics in single bacterial transcription initiation complexes. Nucleic acids research *46*, 677-688.
Duchi, D., Mazumder, A., Malinen, A.M., Ebright, R.H., and Kapanidis, A.N. (2018b). The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. Nucleic acids research.
Duchi, D., Mazumdera, A., Malinen, A.M., Ebright, R.H., and Kapanidis, A.N. (2018c). The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. BioRxiv.
Dulin, D., Arnold, J.J., van Laar, T., Oh, H.S., Lee, C., Perkins, A.L., Harki, D.A., Depken, M., Cameron, C.E., and Dekker, N.H. (2017a). Signatures of Nucleotide Analog Incorporation by an RNA-Dependent RNA Polymerase Revealed Using High-Throughput Magnetic Tweezers. Cell Rep *21*, 1063-1076.
Dulin, D., Bauer, D.L.V., Malinen, A.M., Bakermans, J.J.W., Kaller, M., Morichaud, Z., Petushkov, I., Depken, M., Brodolin, K., Kulbachinskiy, A.*, et al.* (2017b). Pausing controls branching between productive and non-productive pathways during initial transcription. BioRxiv.
Dulin, D., Bauer, D.L.V., Malinen, A.M., Bakermans, J.J.W., Kaller, M., Morichaud, Z., Petushkov, I., Depken, M., Brodolin, K., Kulbachinskiy, A.*, et al.* (2018). Pausing controls branching between productive and non-productive pathways during initial transcription in bacteria. Nature communications *9*, 1478.
Dulin, D., Berghuis, B.A., Depken, M., and Dekker, N.H. (2015a). Untangling reaction pathways through modern approaches to high-throughput single-molecule force-spectroscopy experiments. Curr Opin Struct Biol *34*, 116-122.
Dulin, D., Lipfert, J., Moolman, M.C., and Dekker, N.H. (2013). Studying genomic processes at the single-molecule level: introducing the tools and applications. Nature reviews. Genetics *14*, 9-22.
Dulin, D., Vilfan, I.D., Berghuis, B.A., Hage, S., Bamford, D.H., Poranen, M.M., Depken, M., and Dekker, N.H. (2015b). Elongation-Competent Pauses Govern the Fidelity of a Viral RNA-Dependent RNA Polymerase. Cell Rep *10*, 983-992.
Dulin, D., Vilfan, I.D., Berghuis, B.A., Poranen, M.M., Depken, M., and Dekker, N.H. (2015c). Backtracking behavior in viral RNA-dependent RNA polymerase provides the basis for a second initiation site. Nucleic acids research.
Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science (New York, N.Y *297*, 1183-1186.
English, B.P., Min, W., van Oijen, A.M., Lee, K.T., Luo, G., Sun, H., Cherayil, B.J., Kou, S.C., and Xie, X.S. (2006). Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. Nature chemical biology *2*, 87-94.
Epshtein, V., and Nudler, E. (2003). Cooperation between RNA polymerase molecules in transcription elongation. Science (New York, N.Y *300*, 801-805.
Fan, J., Leroux-Coyau, M., Savery, N.J., and Strick, T.R. (2016). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. Nature *536*, 234-237.

Fazal, F.M., Meng, C.A., Murakami, K., Kornberg, R.D., and Block, S.M. (2015). Real-time observation of the initiation of RNA polymerase II transcription. Nature *525*, 274-277.
Fields, B.N., Knipe, D.M., and Howley, P.M. (2013). Fields virology, 6th edn (Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins).
Forde, N.R., Izhaky, D., Woodcock, G.R., Wuite, G.J., and Bustamante, C. (2002). Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by Escherichia coli RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America *99*, 11682-11687.
Frieda, K.L., and Block, S.M. (2012). Direct observation of cotranscriptional folding in an adenine riboswitch. Science (New York, N.Y *338*, 397-400.
Friedman, L.J., Chung, J., and Gelles, J. (2006). Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. Biophysical journal *91*, 1023-1031.
Friedman, L.J., Mumm, J.P., and Gelles, J. (2013). RNA polymerase approaches its promoter without long-range sliding along DNA. Proceedings of the National Academy of Sciences of the United States of America *110*, 9740-9745.
Fujita, K., Iwaki, M., and Yanagida, T. (2016). Transcriptional bursting is intrinsically caused by interplay between RNA polymerases on DNA. Nature communications *7*, 13788.
Gago, S., Elena, S.F., Flores, R., and Sanjuan, R. (2009). Extremely high mutation rate of a hammerhead viroid. Science (New York, N.Y *323*, 1308.
Galburt, E.A., Grill, S.W., Wiedmann, A., Lubkowska, L., Choy, J., Nogales, E., Kashlev, M., and Bustamante, C. (2007). Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. Nature *446*, 820-823.
Geertsema, H.J., Duderstadt, K.E., and van Oijen, A.M. (2015). Single-molecule observation of prokaryotic DNA replication. Methods Mol Biol *1300*, 219-238.
Gietl, A., Holzmeister, P., Blombach, F., Schulz, S., von Voithenberg, L.V., Lamb, D.C., Werner, F., Tinnefeld, P., and Grohmann, D. (2014). Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. Nucleic acids research *42*, 6219-6231.
Gouge, J., Guthertz, N., Kramm, K., Dergai, O., Abascal-Palacios, G., Satia, K., Cousin, P., Hernandez, N., Grohmann, D., and Vannini, A. (2017). Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. Nature communications *8*, 130.
Graves, E.T., Duboc, C., Fan, J., Stransky, F., Leroux-Coyau, M., and Strick, T.R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. Nature structural & molecular biology *22*, 452-457.
Greenleaf, W.J., Woodside, M.T., Abbondanzieri, E.A., and Block, S.M. (2005). Passive all-optical force clamp for high-resolution laser trapping. Physical review letters *95*, 208102.
Guo, X., Myasnikov, A.G., Chen, J., Crucifix, C., Papai, G., Takacs, M., Schultz, P., and Weixlbaumer, A. (2018). Structural Basis for NusA Stabilized Transcriptional Pausing. Mol Cell *69*, 816-827 e814.
Ha, T., Enderle, T., Ogletree, D.F., Chemla, D.S., Selvin, P.R., and Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. Proceedings of the National Academy of Sciences of the United States of America *93*, 6264-6268.
Harada, Y., Ohara, O., Takatsuki, A., Itoh, H., Shimamoto, N., and Kinosita, K., Jr. (2001). Direct observation of DNA rotation during transcription by Escherichia coli RNA polymerase. Nature *409*, 113-115.
Harden, T.T., Wells, C.D., Friedman, L.J., Landick, R., Hochschild, A., Kondev, J., and Gelles, J. (2016). Bacterial RNA polymerase can retain sigma70 throughout transcription. Proceedings of the National Academy of Sciences of the United States of America *113*, 602-607.
Heller, I., Sitters, G., Broekmans, O.D., Farge, G., Menges, C., Wende, W., Hell, S.W., Peterman, E.J., and Wuite, G.J. (2013). STED nanoscopy combined with optical tweezers reveals protein dynamics on densely covered DNA. Nature methods *10*, 910-916.
Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., and Block, S.M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell *125*, 1083-1094.

Herbert, K.M., Zhou, J., Mooney, R.A., Porta, A.L., Landick, R., and Block, S.M. (2010). E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. Journal of molecular biology *399*, 17-30.
Hill, F.R., Monachino, E., and van Oijen, A.M. (2017). The more the merrier: high-throughput single-molecule techniques. Biochem Soc Trans *45*, 759-769.
Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. Science (New York, N.Y *325*, 626-628.
Hohlbein, J., Craggs, T.D., and Cordes, T. (2014). Alternating-laser excitation: single-molecule FRET and beyond. Chem Soc Rev *43*, 1156-1171.
Holmes, E.C. (2010). Evolution in health and medicine Sackler colloquium: The comparative genomics of viral emergence. Proceedings of the National Academy of Sciences of the United States of America *107 Suppl 1*, 1742-1746.
Howan, K., Smith, A.J., Westblade, L.F., Joly, N., Grange, W., Zorman, S., Darst, S.A., Savery, N.J., and Strick, T.R. (2012). Initiation of transcription-coupled repair characterized at single-molecule resolution. Nature *490*, 431-434.
Ishibashi, T., Dangkulwanich, M., Coello, Y., Lionberger, T.A., Lubkowska, L., Ponticelli, A.S., Kashlev, M., and Bustamante, C. (2014). Transcription factors IIS and IIF enhance transcription efficiency by differentially modifying RNA polymerase pausing dynamics. Proceedings of the National Academy of Sciences of the United States of America *111*, 3419-3424.
Iyer, L.M., Koonin, E.V., and Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. BMC Struct Biol *3*, 1.
Jacome, R., Becerra, A., Ponce de Leon, S., and Lazcano, A. (2015). Structural Analysis of Monomeric RNA-Dependent Polymerases: Evolutionary and Therapeutic Implications. PloS one *10*, e0139001.
Janissen, R., Berghuis, B.A., Dulin, D., Wink, M., van Laar, T., and Dekker, N.H. (2014). Invincible DNA tethers: covalent DNA anchoring for enhanced temporal and force stability in magnetic tweezers experiments. Nucleic acids research *42*, e137.
Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nature reviews. Molecular cell biology *16*, 167-177.
Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C., and Ha, T. (2008). Advances in single-molecule fluorescence methods for molecular biology. Annual review of biochemistry *77*, 51-76.
Kaern, M., Elston, T.C., Blake, W.J., and Collins, J.J. (2005). Stochasticity in gene expression: from theories to phenotypes. Nature reviews. Genetics *6*, 451-464.
Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and Seidel, C.A. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nature methods *9*, 1218-1225.
Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. Mol Cell *69*, 802-815 e801.
Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. Proceedings of the National Academy of Sciences of the United States of America *101*, 8936-8941.
Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., and Ebright, R.H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science (New York, N.Y *314*, 1144-1147.
Kapanidis, A.N., Margeat, E., Laurence, T.A., Doose, S., Ho, S.O., Mukhopadhyay, J., Kortkhonjia, E., Mekler, V., Ebright, R.H., and Weiss, S. (2005). Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis. Mol Cell *20*, 347-356.
Kapanidis, A.N., and Strick, T. (2009). Biology, one molecule at a time. Trends in biochemical sciences *34*, 234-243.

Korboukh, V.K., Lee, C.A., Acevedo, A., Vignuzzi, M., Xiao, Y., Arnold, J.J., Hemperly, S., Graci, J.D., August, A., Andino, R.*, et al.* (2014). RNA virus population diversity, an optimum for maximal fitness and virulence. The Journal of biological chemistry *289*, 29531-29544.
Kriegel, F., Ermann, N., and Lipfert, J. (2017). Probing the mechanical properties, conformational changes, and interactions of nucleic acids with magnetic tweezers. Journal of structural biology *197*, 26-36.
La Porta, A., and Wang, M.D. (2004). Optical torque wrench: angular trapping, rotation, and torque detection of quartz microparticles. Physical review letters *92*, 190801.
Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.
Lang, M.J., Asbury, C.L., Shaevitz, J.W., and Block, S.M. (2002). An automated two-dimensional optical force clamp for single molecule studies. Biophysical journal *83*, 491-501.
Larson, M.H., Greenleaf, W.J., Landick, R., and Block, S.M. (2008). Applied force reveals mechanistic and energetic details of transcription termination. Cell *132*, 971-982.
Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Nayak, D., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R., and Weissman, J.S. (2014). A pause sequence enriched at translation start sites drives transcription dynamics in vivo. Science (New York, N.Y *344*, 1042-1047.
Larson, M.H., Zhou, J., Kaplan, C.D., Palangat, M., Kornberg, R.D., Landick, R., and Block, S.M. (2012). Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. Proceedings of the National Academy of Sciences of the United States of America *109*, 6555-6560.
Laszlo, A.H., Derrrington, I.M., and Gundlach, J.H. (2017). Subangstrom Measurements of Enzyme Function Using a Biological Nanopore, SPRNT. Methods in enzymology *582*, 387-414.
Lauring, A.S., Frydman, J., and Andino, R. (2013). The role of mutational robustness in RNA virus evolution. Nat Rev Microbiol *11*, 327-336.
Le, T.T., Yang, Y., Tan, C., Suhanovsky, M.M., Fulbright, R.M., Jr., Inman, J.T., Li, M., Lee, J., Perelman, S., Roberts, J.W.*, et al.* (2018). Mfd Dynamically Regulates Transcription via a Release and Catch-Up Mechanism. Cell *172*, 344-357 e315.
Lerner, E., Chung, S., Allen, B.L., Wang, S., Lee, J., Lu, S.W., Grimaud, L.W., Ingargiola, A., Michalet, X., Alhadid, Y.*, et al.* (2016). Backtracked and paused transcription initiation intermediate of Escherichia coli RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America *113*, E6562-E6571.
Lerner, E., Ingargiola, A., and Weiss, S. (2018). Characterizing highly dynamic conformational states: The transcription bubble in RNAP-promoter open complex as an example. The Journal of chemical physics *148*, 10.
Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y.W., Ebright, R.Y., Sineva, E., Gigliotti, M.*, et al.* (2017). Structural basis of transcription inhibition by fidaxomicin (lipiarmycin A3). BioRxiv.
Lisica, A., Engel, C., Jahnel, M., Roldan, E., Galburt, E.A., Cramer, P., and Grill, S.W. (2016). Mechanisms of backtrack recovery by RNA polymerases I and II. Proceedings of the National Academy of Sciences of the United States of America *113*, 2946-2951.
Lisica, A., and Grill, S.W. (2017). Optical tweezers studies of transcription by eukaryotic RNA polymerases. Biomol Concepts *8*, 1-11.
Liu, S., Chistol, G., Hetherington, C.L., Tafoya, S., Aathavan, K., Schnitzbauer, J., Grimes, S., Jardine, P.J., and Bustamante, C. (2014). A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills. Cell *157*, 702-713.
Ma, J., Bai, L., and Wang, M.D. (2013). Transcription under torsion. Science (New York, N.Y *340*, 1580-1583.
Madariaga-Marcos, J., Hormeno, S., Pastrana, C.L., Fisher, G.L.M., Dillingham, M.S., and Moreno-Herrero, F. (2018). Force determination in lateral magnetic tweezers combined with TIRF microscopy. Nanoscale *10*, 4579-4590.
Margeat, E., Kapanidis, A.N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R.H., and Weiss, S. (2006). Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. Biophysical journal *90*, 1419-1431.

Meng, C.A., Fazal, F.M., and Block, S.M. (2017). Real-time observation of polymerase-promoter contact remodeling during transcription initiation. Nature communications *8*, 1178.
Miller, H., Zhou, Z., Shepherd, J., Wollman, A.J.M., and Leake, M.C. (2018). Single-molecule techniques in biophysics: a review of the progress in methods and applications. Rep Prog Phys *81*, 024601.
Moffitt, J.R., Chemla, Y.R., Smith, S.B., and Bustamante, C. (2008). Recent advances in optical tweezers. Annual review of biochemistry *77*, 205-228.
Muschielok, A., Andrecka, J., Jawhari, A., Bruckner, F., Cramer, P., and Michaelis, J. (2008). A nano-positioning system for macromolecular structural analysis. Nature methods *5*, 965-971.
Nagy, J., Grohmann, D., Cheung, A.C., Schulz, S., Smollett, K., Werner, F., and Michaelis, J. (2015). Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. Nature communications *6*, 6161.
Neuman, K.C., Abbondanzieri, E.A., Landick, R., Gelles, J., and Block, S.M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. Cell *115*, 437-447.
Neuman, K.C., and Block, S.M. (2004). Optical trapping. Review of Scientific Instruments *75*, 2787-2809.
Neuman, K.C., and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nature methods *5*, 491-505.
Ng, K.K., Arnold, J.J., and Cameron, C.E. (2008). Structure-function relationships among RNA-dependent RNA polymerases. Current topics in microbiology and immunology *320*, 137-156.
Raser, J.M., and O'Shea, E.K. (2004). Control of stochasticity in eukaryotic gene expression. Science (New York, N.Y *304*, 1811-1814.
Ray-Soni, A., Bellecourt, M.J., and Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. Annual review of biochemistry *85*, 319-347.
Revyakin, A., Ebright, R.H., and Strick, T.R. (2004). Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. Proceedings of the National Academy of Sciences of the United States of America *101*, 4776-4780.
Revyakin, A., Ebright, R.H., and Strick, T.R. (2005). Single-molecule DNA nanomanipulation: improved resolution through use of shorter DNA fragments. Nature methods *2*, 127-138.
Revyakin, A., Liu, C., Ebright, R.H., and Strick, T.R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. Science (New York, N.Y *314*, 1139-1143.
Righini, M., Lee, A., Canari-Chumpitaz, C., Lionberger, T., Gabizon, R., Coello, Y., Tinoco, I., Jr., and Bustamante, C. (2018). Full molecular trajectories of RNA polymerase at single base-pair resolution. Proceedings of the National Academy of Sciences of the United States of America.
Robb, N.C., Cordes, T., Hwang, L.C., Gryte, K., Duchi, D., Craggs, T.D., Santoso, Y., Weiss, S., Ebright, R.H., and Kapanidis, A.N. (2013). The transcription bubble of the RNA polymerase-promoter open complex exhibits conformational heterogeneity and millisecond-scale dynamics: implications for transcription start-site selection. Journal of molecular biology *425*, 875-885.
Robb, N.C., Te Velthuis, A.J., Wieneke, R., Tampe, R., Cordes, T., Fodor, E., and Kapanidis, A.N. (2016). Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA. Nucleic acids research.
Roberts, J.W., Shankar, S., and Filter, J.J. (2008). RNA Polymerase Elongation Factors. Annu Rev Microbiol *62*, 211-233.
Robinson, A., and van Oijen, A.M. (2013). Bacterial replication, transcription and translation: mechanistic insights from single-molecule biochemical studies. Nat Rev Microbiol *11*, 303-315.
Ruff, E.F., Record, M.T., Jr., and Artsimovitch, I. (2015). Initial events in bacterial transcription initiation. Biomolecules *5*, 1035-1062.
Rutkauskas, M., Krivoy, A., Szczelkun, M.D., Rouillon, C., and Seidel, R. (2017). Single-Molecule Insight Into Target Recognition by CRISPR-Cas Complexes. Methods in enzymology *582*, 239-273.
Saecker, R.M., Record, M.T., Jr., and Dehaseth, P.L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. Journal of molecular biology *412*, 754-771.
Santybayeva, Z., and Pedaci, F. (2017). Optical Torque Wrench Design and Calibration. Methods Mol Biol *1486*, 157-181.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nature reviews. Molecular cell biology *7*, 557-567.
Schafer, D.A., Gelles, J., Sheetz, M.P., and Landick, R. (1991). Transcription by single molecules of RNA polymerase observed by light microscopy. Nature *352*, 444-448.
Schulz, S., Gietl, A., Smollett, K., Tinnefeld, P., Werner, F., and Grohmann, D. (2016). TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. Proceedings of the National Academy of Sciences of the United States of America *113*, E1816-1825.
Schulz, S., Kramm, K., Werner, F., and Grohmann, D. (2015). Fluorescently labeled recombinant RNAP system to probe archaeal transcription initiation. Methods *86*, 10-18.
Selvin, P.R., Lougheed, T., Tonks Hoffman, M., Park, H., Balci, H., Blehm, B.H., and Toprak, E. (2007). Fluorescence Imaging with One-Nanometer Accuracy (FIONA). CSH protocols *2007*, pdb top27.
Shaevitz, J.W., Abbondanzieri, E.A., Landick, R., and Block, S.M. (2003). Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. Nature *426*, 684-687.
Smith, E.C. (2017). The not-so-infinite malleability of RNA viruses: Viral and cellular determinants of RNA virus mutation rates. PLoS pathogens *13*, e1006254.
Stubbs, T.M., and Te Velthuis, A.J. (2014). The RNA-dependent RNA polymerase of the influenza A virus. Future virology *9*, 863-876.
te Velthuis, A.J. (2014). Common and unique features of viral RNA-dependent polymerases. Cellular and molecular life sciences : CMLS *71*, 4403-4420.
Te Velthuis, A.J., and Fodor, E. (2016). Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. Nat Rev Microbiol *14*, 479-493.
Te Velthuis, A.J., Robb, N.C., Kapanidis, A.N., and Fodor, E. (2016). The role of the priming loop in influenza A virus RNA synthesis. Nature microbiology *1*, 16029.
Tinoco, I., Jr., and Gonzalez, R.L., Jr. (2011). Biological mechanisms, one molecule at a time. Genes & development *25*, 1205-1231.
Tomescu, A.I., Robb, N.C., Hengrung, N., Fodor, E., and Kapanidis, A.N. (2014). Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. Proceedings of the National Academy of Sciences of the United States of America *111*, E3335-3342.
Tomko, E.J., Fishburn, J., Hahn, S., and Galburt, E.A. (2017). TFIIH generates a six-base-pair open complex during RNAP II transcription initiation and start-site scanning. Nature structural & molecular biology *24*, 1139-1145.
van de Meent, J.W., Bronson, J.E., Wiggins, C.H., and Gonzalez, R.L., Jr. (2014). Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments. Biophysical journal *106*, 1327-1337.
van Loenhout, M.T., de Grunt, M.V., and Dekker, C. (2012). Dynamics of DNA supercoils. Science (New York, N.Y *338*, 94-97.
Vilfan, I.D., Lipfert, J., Koster, D.A., Lemay, S.G., and Dekker, N.H. (2009). Magnetic Tweezers for Single-Molecule Experiments. Handbook of Single-Molecule Biophysics, 371-395.
Visscher, K., Schnitzer, M.J., and Block, S.M. (1999). Single kinesin molecules studied with a molecular force clamp. Nature *400*, 184-189.
Wang, F., Redding, S., Finkelstein, I.J., Gorman, J., Reichman, D.R., and Greene, E.C. (2013). The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion. Nature structural & molecular biology *20*, 174-181.
Wang, M.D., Schnitzer, M.J., Yin, H., Landick, R., Gelles, J., and Block, S.M. (1998). Force and velocity measured for single molecules of RNA polymerase. Science (New York, N.Y *282*, 902-907.
Washburn, R.S., and Gottesman, M.E. (2015). Regulation of transcription elongation and termination. Biomolecules *5*, 1063-1078.
Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol *9*, 85-98.
Werner, F., and Weinzierl, R.O. (2002). A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription. Mol Cell *10*, 635-646.
Woodside, M.T., and Block, S.M. (2014). Reconstructing folding energy landscapes by single-molecule force spectroscopy. Annu Rev Biophys *43*, 19-39.

Wuite, G.J., Davenport, R.J., Rappaport, A., and Bustamante, C. (2000). An integrated laser trap/flow control video microscope for the study of single biomolecules. Biophysical journal *79*, 1155-1167.
Xie, S.N. (2001). Single-molecule approach to enzymology. Single Mol *2*, 229-236.
Yin, H., Wang, M.D., Svoboda, K., Landick, R., Block, S.M., and Gelles, J. (1995). Transcription against an applied force. Science (New York, N.Y *270*, 1653-1657.
Yu, L., Winkelman, J.T., Pukhrambam, C., Strick, T.R., Nickels, B.E., and Ebright, R.H. (2017). The mechanism of variability in transcription start site selection. eLife *6*.
Zamft, B., Bintu, L., Ishibashi, T., and Bustamante, C. (2012). Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. Proceedings of the National Academy of Sciences of the United States of America *109*, 8948-8953.
Zhou, J., Ha, K.S., La Porta, A., Landick, R., and Block, S.M. (2011). Applied force provides insight into transcriptional pausing and its modulation by transcription factor NusA. Molecular cell *44*, 635-646.
Zong, J., Yao, X., Yin, J., Zhang, D., and Ma, H. (2009). Evolution of the RNA-dependent RNA polymerase (RdRP) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene *447*, 29-39.

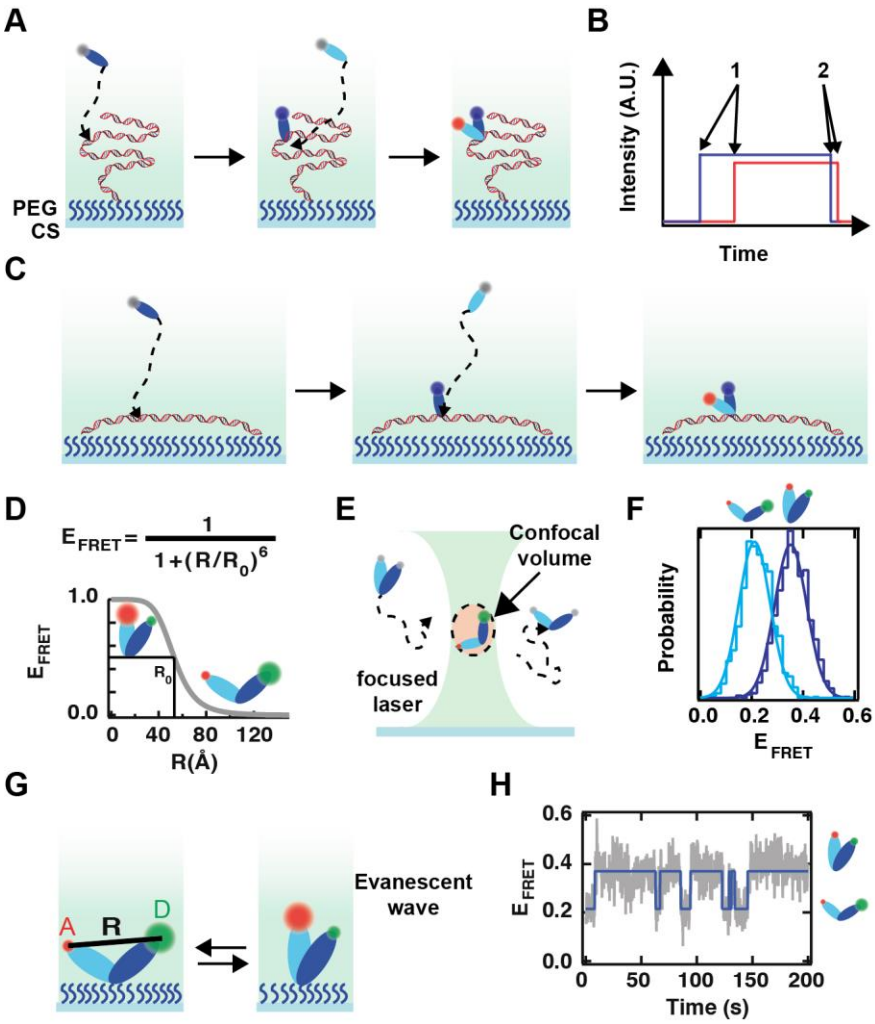

Figure 1: In vitro single-molecule fluorescence spectroscopy techniques in transcription studies. (A) Single-molecule fluorescence co-localization microscopy monitors in real-time binding-unbinding events of interacting molecules (Churchman and Spudich, 2012; Friedman et al., 2006; Selvin et al., 2007). Total internal reflection fluorescence microscopy (TIRFM) is used to image surface-attached biomolecules (Axelrod, 2001). In TIRFM, the excitation laser is reflected at the glass-water interface, generating an evanescent wave (green shade) above the coverslip (CS) that excites dyes within ~100 nm from the coverslip top surface. Non-specific adsorption of the labeled biomolecules is blocked with specific surface chemistry, e.g. polyethylene glycol coating (Chandradoss et al., 2014). **(B)** Observed variation in fluorescence at a specific location where two biomolecules successively bind (1) and dissociate (2) from the nucleic acid molecule as depicted in panel (A). **(C)** By stretching the nucleic acids, as in flow stretching (Geertsema et al., 2015) or in DNA curtains (Collins et al., 2014), one is able to localize the DNA binding molecules with ~100 nm accuracy. **(D)** In single-molecule Forster Resonant Energy Transfer (smFRET), a fraction of the excitation energy of the donor dye (green) is transferred non-radiatively to the acceptor dye (red) with an efficiency E_{FRET} that decreases when the distance between the two dyes increases (Ha et al., 1996). Using this molecular ruler, distances ranging from 2 to 10 nm can be determined (Hohlbein et al., 2014). **(E)** In confocal microscopy, dye-labeled biomolecules at low concentration diffuse freely in the solution and a short burst of photons is detected when a labeled molecule crosses the confocal detection volume. To increase sensitivity, photons emitted outside the confocal volume are spatially filtered out in the imaging optical path. smFRET combined with alternative-laser excitation (ALEX) in confocal microscopy is a powerful tool to access conformational equilibrium and complex formation by biomolecules in solution (Hohlbein et al., 2014; Kapanidis et al., 2004). **(F)** Confocal smFRET provides the

prevalence and the number of different conformations, which are revealed by the amplitude and the width of the normal distributions describing the E_{FRET} histogram, respectively; however, solution smFRET cannot obtain the rates of conformational change. **(G)** To observe the conformational changes happening in an individual biomolecule, which has to be immobilized to the surface of a coverslip. **(H)** E_{FRET} (grey) time trace, typically obtained using TIRFM-based smFRET, shows the biomolecule to interconvert between two distinct conformations. The kinetic constants defining the stabilities of the two states are recovered using a Hidden Markov Model (blue) (van de Meent et al., 2014). Single-molecule FRET can also determine accurate distances between the donor and acceptor dyes; the obtained distances help to model the 3D structure of the biomolecule (Beckers et al., 2015; Kalinin et al., 2012).

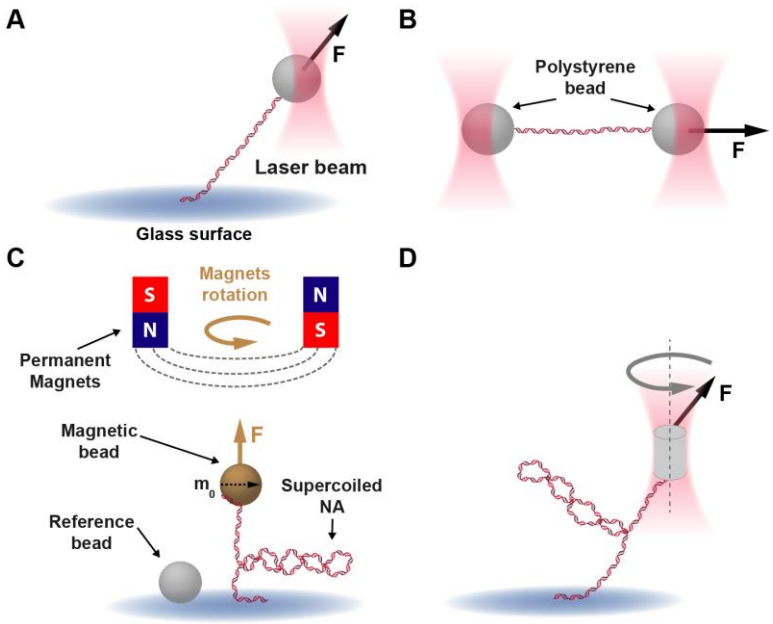

Figure 2: Single-molecule force spectroscopy techniques applied to transcription studies. **(A)** In an optical tweezers assay, one possible configuration consists to attach a nucleic acid molecule (NA) from one end to the glass coverslip of the flow chamber and from the other end to a polystyrene bead trapped in a focused laser beam. The bead position can be controlled in three dimensions by optically moving the trap. Displacement of the bead from the equilibrium position, i.e. the center of the trap, increases linearly the force F experienced by the NA tether. F ranges from ~0.1 pN to hundreds of pN (Neuman and Block, 2004). **(B)** Most modern optical tweezers utilize two optical traps to pull the NA from both ends. The configuration produces a signal with smaller drift artefact and thus a resolution high enough to distinguish translocation steps at ~0.34 nm/s velocity (Greenleaf et al., 2005). **(C)** In a magnetic tweezers assay, double or single stranded NA is attached from one end to the coverslip and from the other end to a magnetic bead. The force F (from ~10 fN to ~1 nN) is generated by pulling the bead with a pair of magnets located above the flow chamber. A reference bead on the coverslip surface is used to correct for the mechanical vibrations caused drift in the position of the sample bead (Vilfan et al., 2009). Camera-based detection allows the simultaneous tracking of hundreds of beads at near base pair resolution. The magnetization (m_0) of the bead permits its rotation, by rotating the magnetic field originating from the magnets. Rotation of the bead adds supercoiling to the torsionally constrained double stranded NA, which eventually leads to the formation of plectonems. Torque-dependent behavior of protein-NA interactions can thus be studied (De Vlaminck and Dekker, 2012). **(D)** In an optical torque wrench, a birefringent particle (here a cylinder) is trapped by a polarized laser beam. By rotating the polarization of the laser, the birefringent cylinder rotates and induces supercoiling in the NA (La Porta and Wang, 2004; Santybayeva and Pedaci, 2017). The “sticky” ends of the NA, necessary in optical tweezers and magnetic tweezers assays, are typically generated with biotin and digoxigenin labeled nucleotides that bind very stably to streptavidin/neutravidin and anti-digoxigenin antibody, respectively, coating the bead or coverslip surface (Janissen et al., 2014).

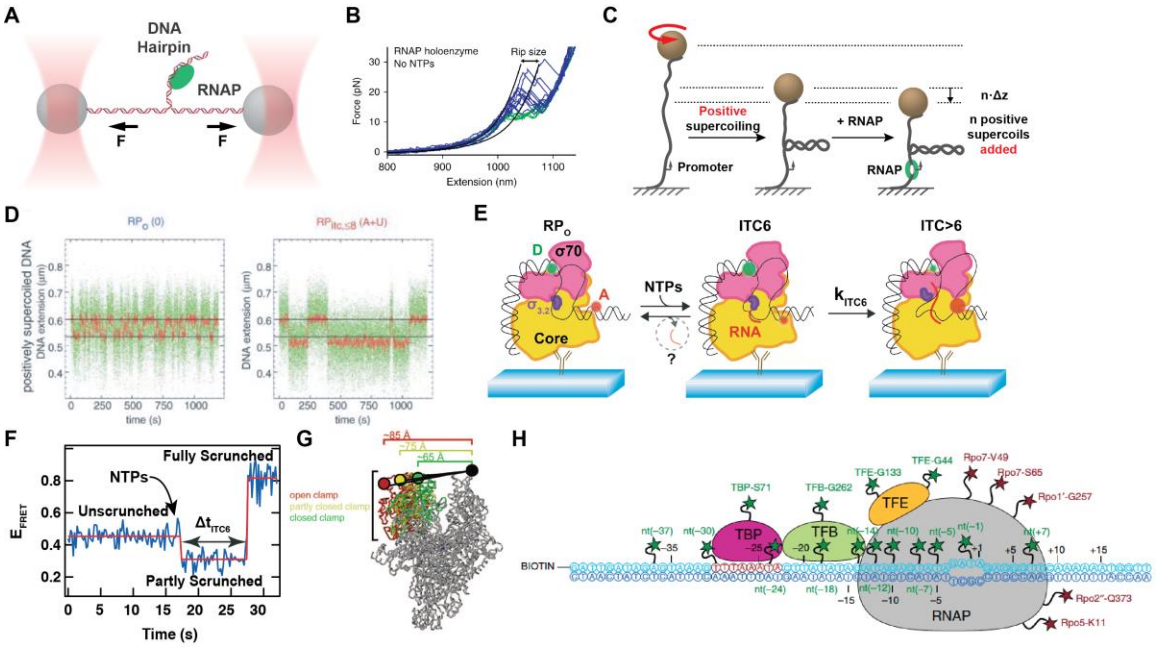

Figure 3: In vitro single-molecule studies of transcription initiation by multisubunit RNA polymerases. (A) The free energy landscape of the RNAP–DNA promoter interaction is manipulated and characterized by progressively unwinding the promoter sequence present in the DNA hairpin by moving the optical traps further apart. **(B)** Force extensions curves from the experiments presented in (A) without (green) and with RNAP (blue) bound to the promoter in the hairpin stem. Figure adapted from Ref. (Meng et al., 2017). **(C)** In magnetic tweezers experiment, a torsionally-constrained DNA molecule is supercoiled by rotating a pair of permanent magnets above the flow chamber, which eventually induces plectonems in the DNA molecule. The conservation of the linking number — the sum of twist (the number of turns the two DNA strands make around each other) and writhe (the number of turns the double stranded DNA makes over itself) — implies that removing one twist in the DNA double helix must be compensated in change in the number of writhe. The unwinding of ~n twists (one twist is 10.6 bp of DNA) by the RNAP when it opens the transcription bubble during RNAP open-promoter complex (RP_O) formation results in the addition of n writhe in a positively supercoiled DNA. This leads to a decrease in the end-to-end distance of the DNA molecule by n.Δz, where Δz is the change in the distance of the DNA molecule upon the addition of one writhe at a given force. DNA scrunching during initial RNA synthesis (initial transcription) further unwinds the transcription bubble therefore increasing the number of writhe added to the DNA. **(D)** Conformational dynamics of RP_O (left) and abortive synthesis of 8-mer RNAs by the initially transcribing complex (ITC, right) was monitored by magnetic tweezers. Adapted from Ref. (Revyakin et al., 2006). **(E)** TIRFM-based smFRET assay can record the dynamics of initial transcription. RP_O complex is formed with a promoter containing a donor dye (green sphere) and an acceptor dye (red sphere) upstream and downstream of the transcription bubble, respectively. The RP_O is immobilized to the coverslip surface with an antibody. RNA synthesis is coupled to the promoter scrunching and movement of the downstream DNA towards the RNAP, leading to the change in the dye pair distance and E_{FRET}. **(F)** E_{FRET} was continuously measured with the assay described in (F) to monitor the magnitude of promoter scrunching and thus the progress of initial transcription. (F, G) are adapted from Ref. (Dulin et al., 2017b). **(G)** Model of alternative clamp positions in the bacterial RNAP was developed based on confocal microscopy smFRET experiments. Red, yellow and green indicate the positions explored by the β' clamp domain and the donor dye (spheres), respectively. Black sphere indicates the immobile position of the acceptor dye on the opposite side of the DNA binding cleft. Adapted from Ref. (Chakraborty et al., 2012). **(H)** Nanopositioning system was used to map the structure of archaeal pre-initiation complex by determining multiple distances between

the transcription initiation factors (TBP, TFB and TFE) and the RNAP as well as between the non-template DNA strand and RNAP. The inter-dye distances were calculated from the measured E_{FRET} values. The donor and acceptor dye locations used in the study are indicated with green and red stars, respectively. Adapted from Ref. (Nagy et al., 2015).

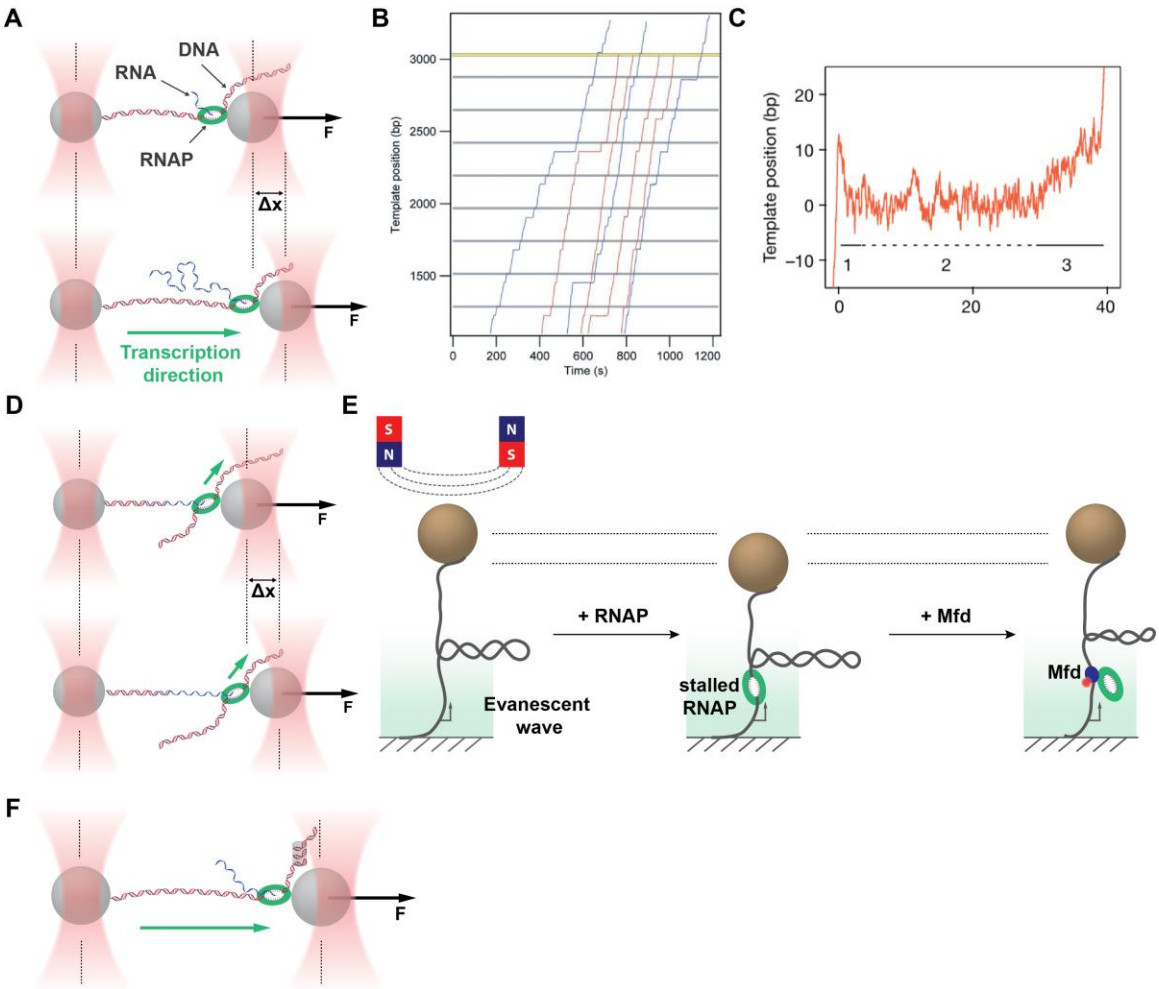

Figure 4: Optical and magnetic tweezers assays to study transcription elongation and termination by multisubunit RNA polymerases. (A) Optical tweezers based assay allows to subject the RNAP either to assisting or hindering force depending on which direction the transcription is designed to progress on the template DNA. **(B)** Transcription activity traces for individual RNAPs obtained with the high-resolution optical tweezers assay depicted in (A). Adapted from Ref. (Herbert et al., 2006). **(C)** A close-up of a transcription activity trace shows the RNAP to backtrack on the template DNA in the optical tweezers assay. Adapted from Ref. (Shaevitz et al., 2003). **(D)** Optical tweezers transcription assay where the assisting force is applied to the nascent RNA. **(E)** Magnetic tweezers was combined with TIRFM to study transcription-coupled repair in Ref. (Fan et al., 2016; Graves et al., 2015). The bead position is affected by the size of the transcription bubble thus transmitting information on the RNAP occupancy on the DNA and the stage of transcription. The binding of dye-labeled Mfd to the DNA or DNA-bound RNAP is inferred from the sudden appearance of a fluorescent spot on the coverslip surface. Because the strength of evanescent field decays exponentially with the distance from the surface, the position of Mfd on the DNA can be extracted from the intensity of the fluorescent spot. **(F)** Schematic of the optical tweezers assay monitoring Pol II transcription past a nucleosome.

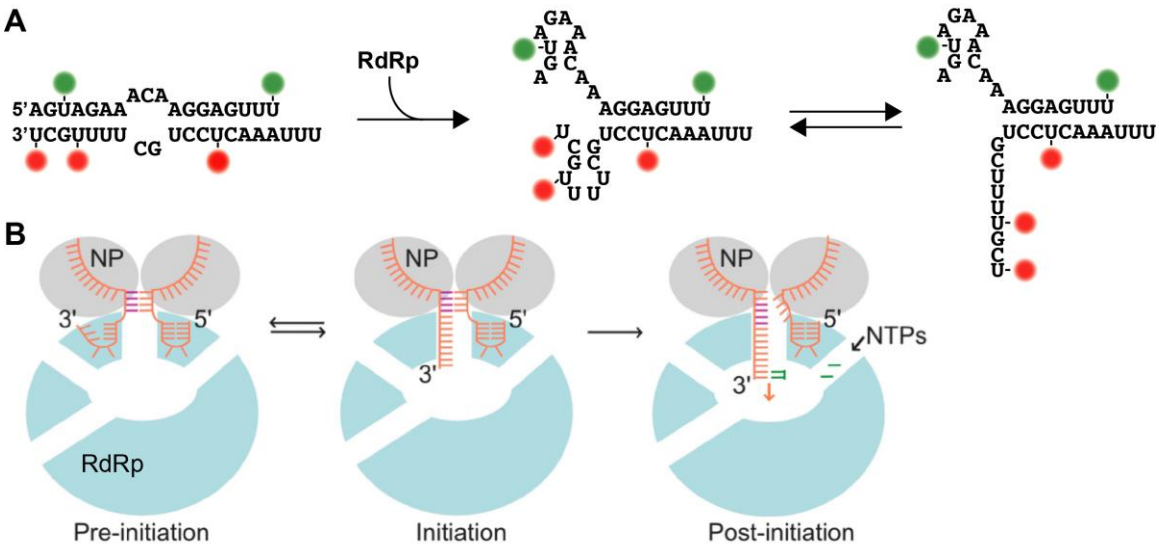

Figure 5: Solution smFRET studies of *de novo* replication initiation by influenza virus RNA-dependent RNA polymerase. (A) Different configurations observed for the nucleic acid scaffold mimicking the 3' and 5' ends of the influenza RNA genome. The green and red spheres indicate the donor and acceptor dye positions, respectively. **(B)** Model for *de novo* replication initiation by influenza RdRp. Adapted from Ref. (Robb et al., 2016).

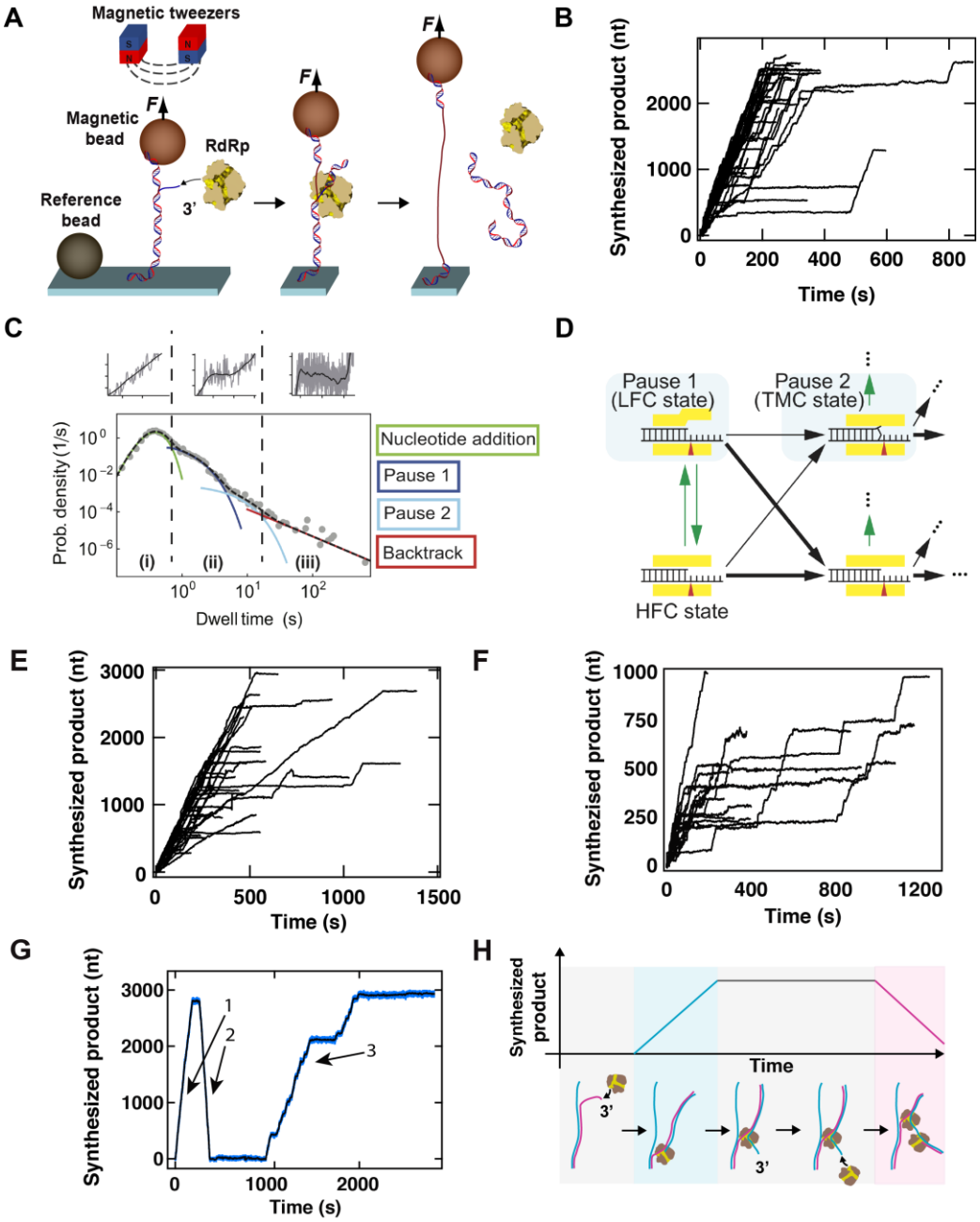

Figure 6: In vitro single-molecule studies of RdRp transcription elongation. (A) Magnetic tweezers assay can be used to study the dynamics of RdRp transcription elongation. The magnetic bead is tethered to the coverslip surface by a double stranded RNA that experiences a constant force. A short non-hybridized segment of the RNA template presents a free 3' end for RdRps to perform *de novo* initiation. To study primer dependent initiating RdRps, such as poliovirus RdRp, the 3' end of the template RNA is modified to contain a short priming hairpin. Following successful initiation, the RdRp elongates the RNA product strand, unwinding the template strand and converting the tether to ssRNA. In the process, the end-to-end distance of the tether changes, thus reporting on the RdRp activity. **(B)** 52 traces of transcribing Φ6 P2 RdRps were acquired in a single experiment using high-throughput magnetic tweezers (Cnossen et al., 2014). Adapted from Ref. (Dulin et al., 2015b). **(C)** Probability density distribution of the dwell times corresponding to the synthesis of ten consecutive nucleotides stretches of RNA. Four distinct dwell time distributions are fitted; these correspond to the pause-free nucleotide incorporation (nucleotide addition, green), short pauses (Pause 1, dark blue), intermediate

pauses (Pause 2, light blue), and long pauses caused by polymerase backtracking (backtrack, red). Example trace snapshots above illustrate each dwell time type. Adapted from Ref. (Dulin et al., 2015a). **(D)** Nucleotide error incorporation model explains the dwell time distribution of the Φ6 P2 RdRp. The model details are explained in the main text. HFS, the high fidelity catalytic (HFC) state; LFC, low fidelity catalytic state; TMC, terminal mismatched catalytic state. **(E)** Poliovirus RdRp replication traces in the presence of 100 µM of NTPs and **(F)** 100 µM of NTPs with 10 µM of antiviral nucleotide analogue T1106-triphosphate. (E) and (F) are adapted from Ref. (Dulin et al., 2017a). **(G)** A fraction of Φ6 P2 RdRp transcription traces displayed “reversal“ activity (arrows). **(H)** The reversal activity originates from a backtracked RdRp that presents a protruding 3’ end of the product RNA strand, which is used by another RdRp as a template of transcription. The second RdRp pushes back the first RdRp resulting in the rehybridization of the original template and non-template strands. The shortening of the end-to-end distance of the tether is thus detected as a “reversal” trace. (G) and (H) are adapted from Ref. (Dulin et al., 2015b).

Axelrod, D. (2001). Total internal reflection fluorescence microscopy in cell biology. Traffic *2*, 764-774.
Beckers, M., Drechsler, F., Eilert, T., Nagy, J., and Michaelis, J. (2015). Quantitative structural information from single-molecule FRET. Faraday discussions.
Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R*., et al.* (2012). Opening and closing of the bacterial RNA polymerase clamp. Science (New York, NY *337*, 591-595.
Chandradoss, S.D., Haagsma, A.C., Lee, Y.K., Hwang, J.H., Nam, J.M., and Joo, C. (2014). Surface passivation for single-molecule protein studies. Journal of visualized experiments : JoVE.
Churchman, L.S., and Spudich, J.A. (2012). Colocalization of fluorescent probes: accurate and precise registration with nanometer resolution. Cold Spring Harbor protocols *2012*, 141-149.
Cnossen, J.P., Dulin, D., and Dekker, N.H. (2014). An optimized software framework for real-time, high-throughput tracking of spherical beads. The Review of scientific instruments *85*, 103712.
Collins, B.E., Ye, L.F., Duzdevich, D., and Greene, E.C. (2014). DNA curtains: novel tools for imaging protein-nucleic acid interactions at the single-molecule level. Methods in cell biology *123*, 217-234.
De Vlaminck, I., and Dekker, C. (2012). Recent advances in magnetic tweezers. Annu Rev Biophys *41*, 453-472.
Dulin, D., Arnold, J.J., van Laar, T., Oh, H.S., Lee, C., Perkins, A.L., Harki, D.A., Depken, M., Cameron, C.E., and Dekker, N.H. (2017a). Signatures of Nucleotide Analog Incorporation by an RNA-Dependent RNA Polymerase Revealed Using High-Throughput Magnetic Tweezers. Cell Rep *21*, 1063-1076.
Dulin, D., Bauer, D.L.V., Malinen, A.M., Bakermans, J.J.W., Kaller, M., Morichaud, Z., Petushkov, I., Depken, M., Brodolin, K., Kulbachinskiy, A*., et al.* (2017b). Pausing controls branching between productive and non-productive pathways during initial transcription. BioRxiv.
Dulin, D., Vilfan, I.D., Berghuis, B.A., Hage, S., Bamford, D.H., Poranen, M.M., Depken, M., and Dekker, N.H. (2015a). Elongation-Competent Pauses Govern the Fidelity of a Viral RNA-Dependent RNA Polymerase. Cell Rep *10*, 983-992.
Dulin, D., Vilfan, I.D., Berghuis, B.A., Poranen, M.M., Depken, M., and Dekker, N.H. (2015b). Backtracking behavior in viral RNA-dependent RNA polymerase provides the basis for a second initiation site. Nucleic acids research.
Fan, J., Leroux-Coyau, M., Savery, N.J., and Strick, T.R. (2016). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. Nature *536*, 234-237.
Friedman, L.J., Chung, J., and Gelles, J. (2006). Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. Biophysical journal *91*, 1023-1031.
Geertsema, H.J., Duderstadt, K.E., and van Oijen, A.M. (2015). Single-molecule observation of prokaryotic DNA replication. Methods Mol Biol *1300*, 219-238.
Graves, E.T., Duboc, C., Fan, J., Stransky, F., Leroux-Coyau, M., and Strick, T.R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. Nature structural & molecular biology *22*, 452-457.
Greenleaf, W.J., Woodside, M.T., Abbondanzieri, E.A., and Block, S.M. (2005). Passive all-optical force clamp for high-resolution laser trapping. Physical review letters *95*, 208102.
Ha, T., Enderle, T., Ogletree, D.F., Chemla, D.S., Selvin, P.R., and Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. Proceedings of the National Academy of Sciences of the United States of America *93*, 6264-6268.
Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., and Block, S.M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell *125*, 1083-1094.
Hohlbein, J., Craggs, T.D., and Cordes, T. (2014). Alternating-laser excitation: single-molecule FRET and beyond. Chem Soc Rev *43*, 1156-1171.
Janissen, R., Berghuis, B.A., Dulin, D., Wink, M., van Laar, T., and Dekker, N.H. (2014). Invincible DNA tethers: covalent DNA anchoring for enhanced temporal and force stability in magnetic tweezers experiments. Nucleic acids research *42*, e137.

Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and Seidel, C.A. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nature methods *9*, 1218-1225.
Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. Proceedings of the National Academy of Sciences of the United States of America *101*, 8936-8941.
La Porta, A., and Wang, M.D. (2004). Optical torque wrench: angular trapping, rotation, and torque detection of quartz microparticles. Physical review letters *92*, 190801.
Meng, C.A., Fazal, F.M., and Block, S.M. (2017). Real-time observation of polymerase-promoter contact remodeling during transcription initiation. Nature communications *8*, 1178.
Nagy, J., Grohmann, D., Cheung, A.C., Schulz, S., Smollett, K., Werner, F., and Michaelis, J. (2015). Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. Nature communications *6*, 6161.
Neuman, K.C., and Block, S.M. (2004). Optical trapping. Review of Scientific Instruments *75*, 2787-2809.
Revyakin, A., Liu, C., Ebright, R.H., and Strick, T.R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. Science (New York, NY *314*, 1139-1143.
Robb, N.C., Te Velthuis, A.J., Wieneke, R., Tampe, R., Cordes, T., Fodor, E., and Kapanidis, A.N. (2016). Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA. Nucleic acids research.
Santybayeva, Z., and Pedaci, F. (2017). Optical Torque Wrench Design and Calibration. Methods Mol Biol *1486*, 157-181.
Selvin, P.R., Lougheed, T., Tonks Hoffman, M., Park, H., Balci, H., Blehm, B.H., and Toprak, E. (2007). Fluorescence Imaging with One-Nanometer Accuracy (FIONA). CSH protocols *2007*, pdb top27.
Shaevitz, J.W., Abbondanzieri, E.A., Landick, R., and Block, S.M. (2003). Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. Nature *426*, 684-687.
van de Meent, J.W., Bronson, J.E., Wiggins, C.H., and Gonzalez, R.L., Jr. (2014). Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments. Biophysical journal *106*, 1327-1337.
Vilfan, I.D., Lipfert, J., Koster, D.A., Lemay, S.G., and Dekker, N.H. (2009). Magnetic Tweezers for Single-Molecule Experiments. Handbook of Single-Molecule Biophysics, 371-395.