

SURVEY AND SUMMARY

Basic mechanisms and kinetics of pause-interspersed transcript elongation

Jin Qian, David Dunlap* and Laura Finzi¹*

Physics, Emory University, Atlanta, GA 30307, USA

Received September 14, 2020; Revised November 16, 2020; Editorial Decision November 18, 2020; Accepted November 19, 2020

ABSTRACT

RNA polymerase pausing during elongation is an important mechanism in the regulation of gene expression. Pausing along DNA templates is thought to be induced by distinct signals encoded in the nucleic acid sequence and halt elongation complexes to allow time for necessary co-transcriptional events. Pausing signals have been classified as those producing short-lived elemental, long-lived backtracked, or hairpin-stabilized pauses. In recent years, structural microbiology and single-molecule studies have significantly advanced our understanding of the paused states, but the dynamics of these states are still uncertain, although several models have been proposed to explain the experimentally observed pausing behaviors. This review summarizes present knowledge about the paused states, discusses key discrepancies among the kinetic models and their basic assumptions, and highlights the importance and challenges in constructing theoretical models that may further our biochemical understanding of transcriptional pausing.

INTRODUCTION

RNA polymerases (RNAPs) constitute a class of molecular motors that consume chemical energy to incorporate nucleotide triphosphates to synthesize RNA. After initial stages of promoter recognition, double strand opening and polymerization of an approximately 8–12 nucleotide long transcript (1,2), RNA elongation is carried out by a transcription elongation complex (TEC) that features a DNA bubble separating upstream and downstream DNA duplexes, an 8–9 nucleotide DNA/RNA hybrid, and an emerging nascent RNA chain. Recent advances in x-ray

crystallography and cryo-electron microscopy have revealed further structural features of TEC, such as the trigger loop (TL) and the bridge helix (BH), which are thought to be involved in translocation and proofreading of the 3' end of the RNA (3–5), the lid and the flap domains that contact the RNA and upstream DNA (Figure 1A and B) (6,7). This structural information, along with experimental data from single-molecule assays, has contributed to our understanding of the mechanism of transcript elongation (8–12). Although some details are still under debate, a Brownian ratchet mechanism, in which TEC forward motion is stochastically generated by thermal fluctuation and then stabilized by molecular pawls (Figure 1C), is favored over other models for the multi-subunit bacterial RNA polymerase. This model gained support mostly because it features a 1-nucleotide translocation step (compared to the power-stroke model in which the step-size is less than 1 nucleotide), and it predicts the experimentally observed force-velocity relationship for transcription under different NTP conditions (13–15).

TECs are processive machines capable of producing long nascent transcripts, yet, transcriptional elongation must be highly regulated in order to respond to abnormal events (e.g. nucleotide misincorporation, transcriptional roadblocks), coordinate co-transcriptional events (e.g. coupling with translation and splicing), and produce biologically meaningful transcripts (e.g. terminating at correct position) (16–19). Regulation of transcription is achieved in large part by pauses that interrupt forward translocation.

During elongation, at every nucleotide coding position, RNA polymerase may follow various kinetically competing reaction pathways. In the event of misincorporation, forward translocation is compromised and various correction pathways, involving backsliding and removal of the 'wrong' nucleotide, become kinetically preferred. Alternatively, and very slowly, synthesis past the misincorporated base resumes, resulting in a 'mutated' RNA. Finally, tran-

*To whom correspondence should be addressed. Tel: +1 404 727 4930, Email: lfinzi@emory.edu
Correspondence may also be addressed to David Dunlap. Email: ddunlap@emory.edu

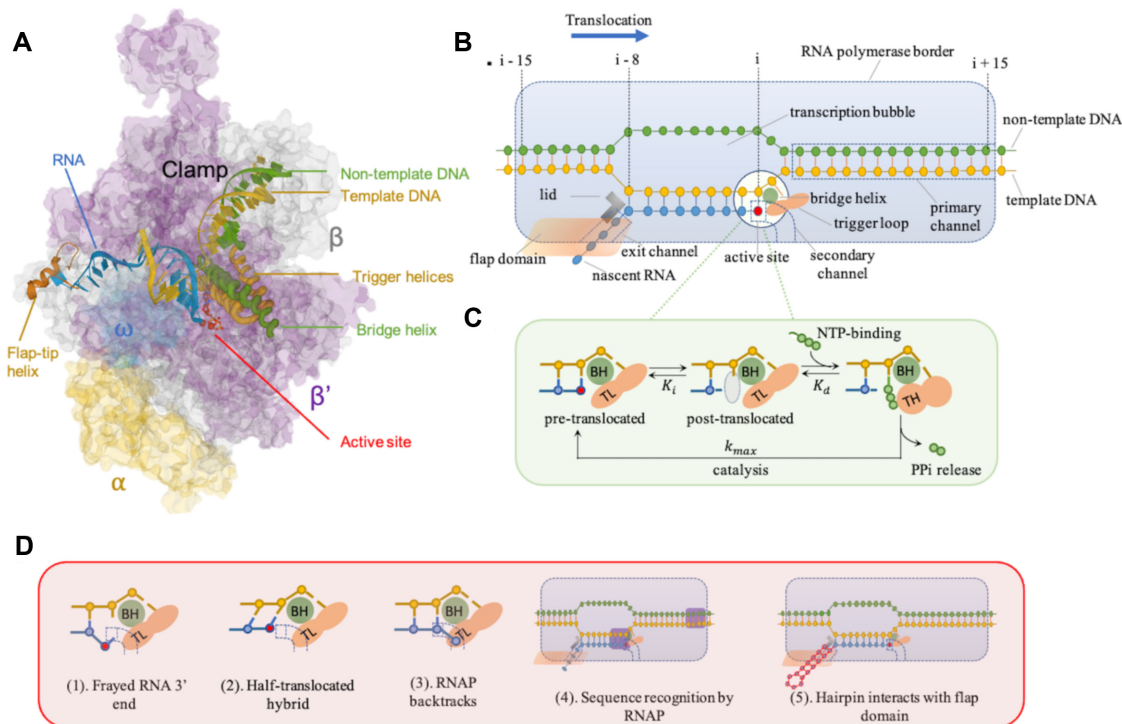


Figure 1. (A) Crystal structure of the *Thermus thermophilus* RNA polymerase elongation complex with the bound NTP substrate (PDB: 205J) (3). (B) Cartoon of an elongation complex. An elongation complex covers 30–35 base pairs (bp) of DNA, including ~14 bp of the DNA downstream of the active site, 10–11-bp of the transcription bubble, and nascent RNA which binds the template strand to form an 8–9-bp RNA–DNA hybrid. The RNAP primary channel accommodates the downstream dsDNA; the secondary channel serves as the site for NTP entry and RNA extrusion when RNAP backtracking occurs; the nascent RNA emerges through the exit channel near the flap after the RNA/DNA hybrid strand separates at the lid. The structure is stabilized by the interaction of specific RNAP domains, such as the bridge helix (BH) and trigger loop (TL) that comprise the active center, the lid and flap that interact with nascent RNA. (C) A nucleotide addition cycle involves a Brownian ratchet at the active center. RNAP shifts between the pre- and post-translocated registers until an incoming NTP (green) occupies the active site (gray). There, NTP reacts to form a phosphodiester linkage to 3'-OH group of the growing RNA chain and release inorganic pyrophosphate (PPi). During the process, the TL folds into trigger helices (TH) and positions an NTP for catalysis. (D) Off-pathway states proposed to explain elongational pauses include: (1) fraying of the DNA–RNA hybrid at the active site detected by crosslinking, although structural data are not consistent with this; (2) incomplete template DNA strand translocation, with a pre-translocated DNA strand and a post-translocated RNA strand, that precludes NTP addition; (3) RNAP backtracking upstream and extruding the 3' end of nascent RNA into/through the catalytic site; (4) RNAP recognizing pause signals encoded in DNA/RNA hybrid and/or downstream DNA sequences (purple); (5) a hairpin structure forming in nascent RNA that interacts with the RNAP exit channel and flap domain to cause a global conformational change that disrupts elongation. States 1, 2 and 4 are proposed as elemental pauses. State 3 represents the backtracked paused complex, and state 5 represents the hairpin-stabilized paused complex.

script termination and release from the transcription complex may also occur. Various pausing mechanisms direct these choices.

Although the functional roles of transcriptional pausing are still under investigation, aberrant pause and release in the human Pol II system has been shown to relate with various human pathologies and reviewed elsewhere (20–23). In prokaryotic systems, transcriptional pauses are well-studied events with lifetimes spread over a broad range, from brief ‘elemental’ pauses (a.k.a. ubiquitous pauses) that last from milliseconds to a few seconds (24–26) to longer-lived pauses that can last for minutes (26–28). Artsimovitch and Landick first proposed that the entry into long paused states begins with the formation of an intermediate, elemental paused state, that can be stabilized mechanistically according to different types of signals and converted into either a ‘backtracked’ or ‘hairpin-stabilized’ paused state (26). These pauses originate from mechanistically different pausing signals and likely have distinct roles in transcription. For example, the backtracked pause is thought to

control gene expression at promoter-proximal sites and it is the state leading to correction of misincorporated bases (20,27,29); the hairpin-stabilized pause is thought to guide the folding of leader RNA structures (30). In the recent two decades, single-molecule experimentation and cryo-EM structures of paused TECs have refined our understanding the co-factors that influence the entry into and the escape from paused states (4,11,31,32).

Transcriptional pausing is largely probabilistic, and a vast majority of pausing sites are not 100% efficient (33–35). Thus, the roles of different paused states in transcriptional regulation and the mechanisms that lead to integration of various transcriptional pauses and generation of pause-interspersed transcription are difficult to assess. Nonetheless, many studies have revealed the dynamics of transcript elongation. Here, we review the mechanisms of transcriptional pausing with a focus on the kinetics of the paused states, especially with regard to entry into, and exit from, these states. We hope to provide a mechanistic perspective of how highly regulated transcription develops from a ki-

netic competition between forward translocation and various paused states of TECs.

The elementally paused elongation complex (ePEC)

Although the short-lived elemental pauses have been observed in various contexts ranging from ensemble measurements to single-molecule assays, the origin of the ePEC is a subject of debate. In the Brownian-ratchet model, the ePEC stems from the thermodynamics of the RNA–DNA scaffold that permit toggling between the pre-translocated and post-translocated states until NTP binding favors the post-translocated state (Figure 1C) (36,37). This intuitively acceptable explanation was first proposed by Yager & von Hippel and Guajardo & Sousa (38,39). Bai *et al.* used this concept in a quantitative sequence-dependent transcription model which predicted pre-translocated pauses with lifetimes resembling experimentally observed lifetimes of ePECs (40). Indeed, the consensus elements of brief pauses have been identified as a GC-rich segment at the upstream edge of the RNA–DNA hybrid and a pyrimidine at the pause site followed immediately by a G (G-10Y-1G+1), such that G/C at –10 and +1 will disfavor forward translocation relative to the less stable A/T as the Brownian-ratchet model predicts (12,33,34). Although other evidence showed that the elemental pausing also depends on the DNA–RNA hybrid and downstream DNA duplex sequences (33,35,41), the contributions from sequences in the upstream and downstream fork-junctions of the transcriptional bubble demonstrate that the Brownian-ratchet model broadly reproduces what is known about the ePEC.

The Brownian-ratchet model also predicts that the ePEC is an on-pathway intermediate. Experimentally, using core RNAPs that freely transcribe along an optimal pause sequence with a 17-nucleotide downstream segment, Bochkareva *et al.* observed pause events in the pre-translocated state that emphasized the on-pathway mechanism (42). However, this result is disputed by Saba *et al.*, who re-examined kinetic modeling of the pause events along a similar scaffold with an extended downstream segment. They concluded that the obligatory pausing events observed by Bochkareva *et al.* were due to the truncated downstream DNA (41). Overall, the on-pathway mechanism of the ePEC, although intuitively attractive, lacks experimental support.

In fact, experimental evidence seems to place the ePEC off-pathway. Single-molecule experiments have shown that external load on either nascent RNA or motor enzymes has little effect on duration and probability of an ePEC, placing it off the translocation pathway where it is less exposed to mechanical loads (24,43). In experiments with high temporal resolution as short as 1 s, elemental pauses occurred with frequencies well below 100%, implying that the elementally paused state is off the translocation pathway (12,31). Recent cryo-EM structures of the ePEC revealed a post-translocated RNA and pre-translocated DNA forming a tilted RNA/DNA hybrid, along with a rearrangement of the RNAP trigger loop (Figure 1D(2)). This half-translocated state may be an off-pathway state attributable to an elementally paused state and may exhibit longer dwell

times than on-pathway intermediates (41,44). Besides the half-translocated register, other conformational arrangements that could inhibit RNAP forward translocation were proposed to explain the formation of the ePEC, such as a frayed 3' end of the nascent RNA at the active site (Figure 1D(1)), or sequence-dependent nucleic acid and RNA polymerase interactions (Figure 1D(4)) (26,42,45). Backtracked RNAP has been proposed to explain short transcriptional pauses, although this proposition has been disfavored by experimental results (24,46,47).

However, studies on the elongating structure suggest that the translocation step of bacterial RNAP may not act as a simple ratchet driven by NTP-binding affinity. Yin and Steitz, based on the crystal structures of RNAP binding NTP substrates, described a mechanism of translocation in which a helix subdomain of RNAP experiences substantial conformational changes at every cycle of single nucleotide addition. They postulated that RNAPs remain at the pre-translocation position, due to the coordination of the R627 and D537 residues with pyrophosphate, until the rotation of the O helix promoted by the release of pyrophosphate translocates RNAP to the post-translocation position (48). This finding suggests that the release of PP_i determines the translocation register rather than the presence of NTPs and contradicts the mechanistic assumptions of most models on transcriptional pauses. In agreement with this finding, a dual-ratchet model was proposed, in which the helix acts as a first reciprocating pawl pushing RNAP forward relative to the nucleic acid scaffold, while the incoming NTP substrate acts as a second stationary pawl preventing RNAP from slipping backward (49).

Backtracked, paused elongation complex (bPEC)

Backtracking, a reverse motion of RNAP, can induce pausing by extending the 3' end of a nascent transcript into/through the catalytic site (Figures 1D(3), 2A) (26,28). A subsequent endonucleolytic cleavage event, either intrinsic or induced by accessory factors (e.g. GreA and GreB for *Escherichia coli* and SII/TFIIS for Pol II), can rescue a bPEC (50–52). In *E. coli*, backtracking is known to occur at operon polarity suppressor (*ops*) sites and is responsible for promoter-proximal pausing (26,29). For Pol II in eukaryotes, high G/C content followed by A/T-rich sequence near a DNA promoter is thought to produce unstable RNA–DNA hybridization and polymerase backtracking (53). Using optical tweezers assays, Shaevitz *et al.* reported that backtracking events by RNAP associated with pauses lasting 20 s to >30 min were observed not only at consensus sequences but at locations throughout DNA templates (27).

The formation of bPECs is force sensitive, and the pathway leading to bPECs is relatively well understood. Assisting or opposing loads respectively inhibit or facilitate the occurrence of a bPEC at a pause site. This indicates that bPECs are energetically stable states produced by reverse translocation of RNAP (54). Saba *et al.* proposed that an ePEC could equilibrate rapidly among the pre-translocated, half-translocated and one-base-pair backtracked states, given that the energy associated with a single base pair is a minimal barrier (41). Other work using high-resolution optical trapping assays also demonstrated

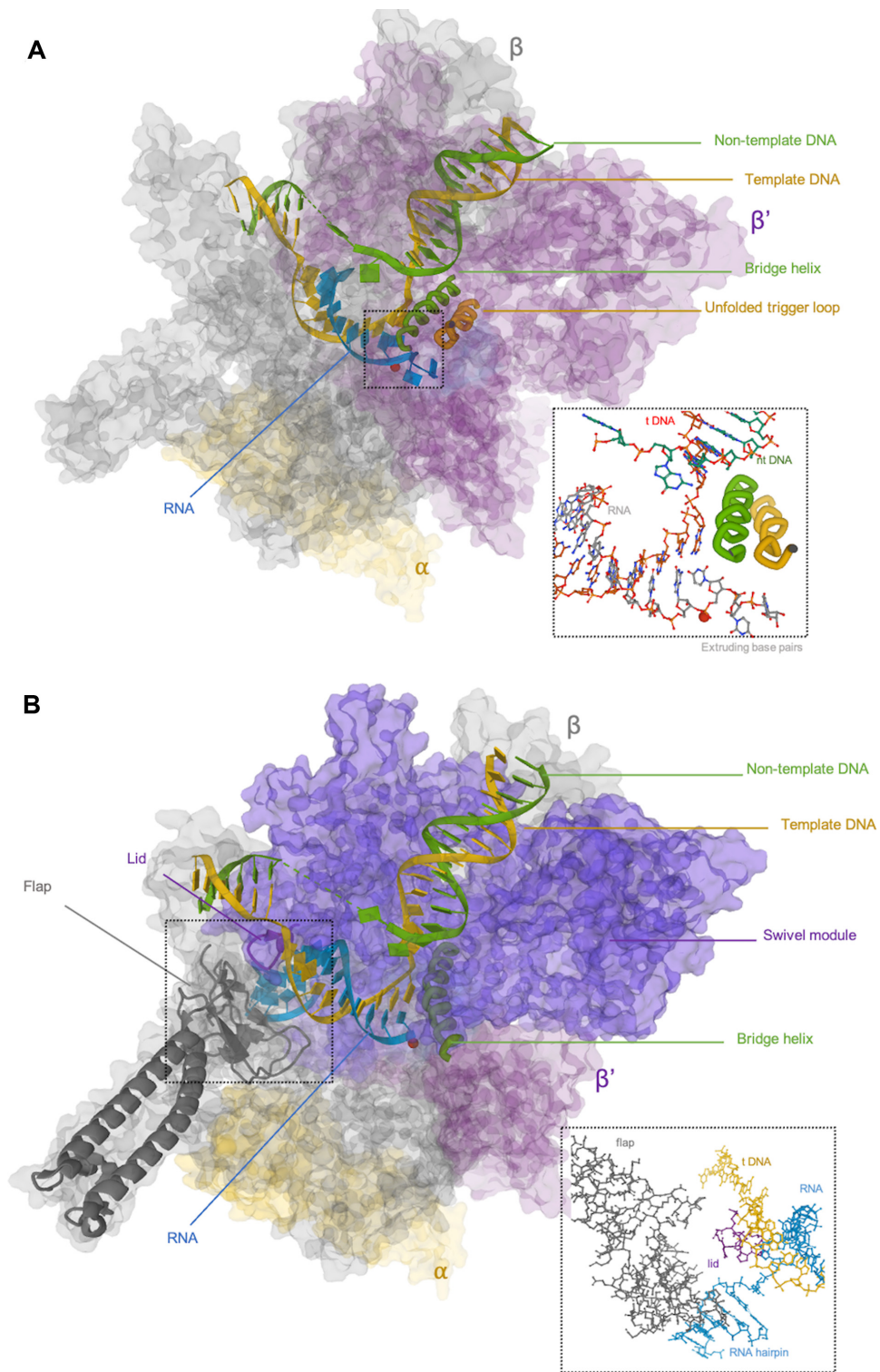


Figure 2. (A) Crystal structure of a backtracked *E. coli* RNA polymerase elongation complex (PDB: 6RI9) (86). Notice the RNA (cyan) bases extruded beyond the catalytic site in the enlargement. (B) Crystal structure of a hairpin-stabilized *E. coli* RNA polymerase elongation complex (PDB: 6ASX) (44). Notice the flap domain (gray) interacts with the hairpin (cyan) and stabilizes the swivel module (purple).

the ease with which a one-base-pair bPEC forms (37,55). More extended backtracking is less frequently observed, but is characterized by longer dwell times, so the formation of deeper bPECs likely involves greater activation barriers. Also, there is evidence that a conformational change might be associated with backtracking and drive the bPEC into a state resistant to rescue by external loads and RNA cleavage events (45,56,57).

Recovery from backtracking is thought to occur by either 1D diffusion or cleavage of RNA blocking the catalytic site. An optical trapping assay on RNAP II has shown that the distribution of backtracked pauses of less than 10 s follows a $t^{-3/2}$ power law, implying backtracked RNAP II diffuses following a 1D unbiased random walk with one nucleotide steps in the absence of RNA cleavage events (54). Lisica *et al.* reported that the choice of the recovery mechanism is determined by a kinetic competition between the random walker and the RNA cleavage event (58). A bPEC in a shallow energetic trap, such as one-base-pair, tends to recover by 1D diffusion, while a bPEC in a deep energetic trap is more likely to recover through RNA cleavage. In recent work using high-throughput magnetic tweezers, Janissen and colleagues supported this by showing that recovery of PECs backtracked by >4-bp is predominantly achieved by intrinsic cleavage, in the absence of cleavage factors that significantly facilitate recovery (57).

Hairpin-stabilized paused elongation complex (hsPEC)

An important class of long-lived pausing signals is encoded in RNA secondary structures (Figures 1D(5), 2B). A nascent RNA structure, perhaps a pseudo knot or a hairpin, can interact with the flap domain near the exit channel and inhibit nucleotide addition in the active site 65 Å away (6,59–61). According to an allosteric model supported by cryo-EM reconstruction, upon hairpin-flap interaction, RNAP adopts a global conformational rearrangement that stabilizes a swivel module and prevents trigger loop folding, disrupting the active nucleotide addition cycle (7,44). Transcriptional factor NusA is known to enhance hairpin-mediated pausing by providing a positively charged cavity in the RNAP exit channel for the formation of RNA secondary structure and the stabilization of RNA–RNAP interaction (7). Note that an RNA hairpin also helps disrupt the RNA–DNA hybrid during intrinsic transcriptional termination. Hairpins may induce similar conformational rearrangements of polymerases in both hairpin-stabilized pausing and intrinsic termination (26,62–64).

Although the mechanisms underlying the hsPEC are not fully understood, many features of the hsPEC have been revealed experimentally. Recently, experimental data have separately shown that the RNAP flap domain, trigger loop, and RNA hairpin are indispensable elements for the formation of a hsPEC, in agreement with the mechanism suggested by the cryoEM hsPEC structure (55,65–67). A four-nucleotide separation between the hairpin and the RNA–DNA hybrid is optimal for the formation of an hsPEC. Decreasing this spacer to 2 nt substantially reduces pausing (6). In addition, the stability and size of the hairpin affect the formation of the hsPEC. Chauvier *et al.* found that a stable stem favors formation while the size of the end-loop is less

important (67). However, overall dimensions are important, and Touloukhonov *et al.* found that hsPEC formation drops substantially for RNA hairpins with longer stems. Surprisingly, an artificial hairpin formed by hybridization of an oligonucleotide to the nascent transcript cannot mimic the effect of the nascent RNA hairpin (68).

Kinetics of pause-interrupted transcription

The Brownian-ratchet kinetic model of the active elongation pathway consists of a translocation step, an NTP binding step, and a rate-limiting NTP nucleation step (Figures 1C and 3A) (14,40,69). The overall kinetics is described by the Michaelis-Menten equation.

$$k_{\text{forward}} = \frac{k_{\text{max}} [\text{NTP}]}{K_d (1 + K_i) + [\text{NTP}]} \quad (1)$$

where K_i is the equilibrium rate between the pre- and post-translocated state of TEC at position i , K_d is NTP-specific parameters for NTP dissociation and k_{max} is the reaction rate of an irreversible step including NTP catalysis, PPi release and phosphodiester bond formation (Figure 3A). To measure the values of the NTP-specific parameters in Equation (1), Bai *et al.* recorded the transcriptional rates of *E. coli* RNAP along DNA segments lacking one of the four nucleotides and reported the fitted parameters in which K_d ranged from 7 μM for CTP to 62 μM for GTP and k_{max} ranged from 18 s⁻¹ for UTP to 50 s⁻¹ for ATP (14). Force (F) perturbations would be expected to modulate the kinetics of the equilibrium of translocation step by $K_i = \exp[(\Delta G(i, \text{post}) - \Delta G(i, \text{pre}) - F\delta)/k_B T]$, where δ is a 1 nucleotide step by the ratchet-like TEC. However, this simplest ratchet model does not address the roles of the folding and unfolding of the BH and the TL, thus may not fully represent the mechanistic origin of transcriptional kinetics. A more universal and detailed model that considered the folded/unfolded states of bridge helix and trigger loop further classified six interconnected states within a nucleotide addition cycle (70).

The ePEC, bPEC and hsPEC were depicted as thermal states branching from the active elongation complex. As illustrated in Figure 3, forward elongation dynamics result from the probabilities of transition across activation barriers connecting the paused and active states (71). This mechanism suggests that an elongation complex favors pausing at low free energy positions. Indeed, Ruckenstein and colleagues simulated the occurrence of pausing by evaluating the local minimum in free energy landscape over a finite region in the vicinity of translocation position (72). For the formation of the ePEC, an on-pathway Brownian-ratchet interpretation leads to pre-translocated pauses at sites where $\Delta G(i, \text{post})$ is much greater than $\Delta G(i, \text{pre})$ (Figure 3A) (14,40). In contrast, if the ePEC is considered an off-pathway intermediate, entry into the ePEC can be achieved by either forward translocation from the $(i - 1, \text{post})$ state, or by inactivation from the (i, pre) state. Indeed, the latter (Figure 3C) is supported by previous cross-linking experiments (6,31,34,73) and the fact that stabilization of post-translocated TECs would reduce the efficiency of elemental pause sites. Ruckenstein and colleagues reported

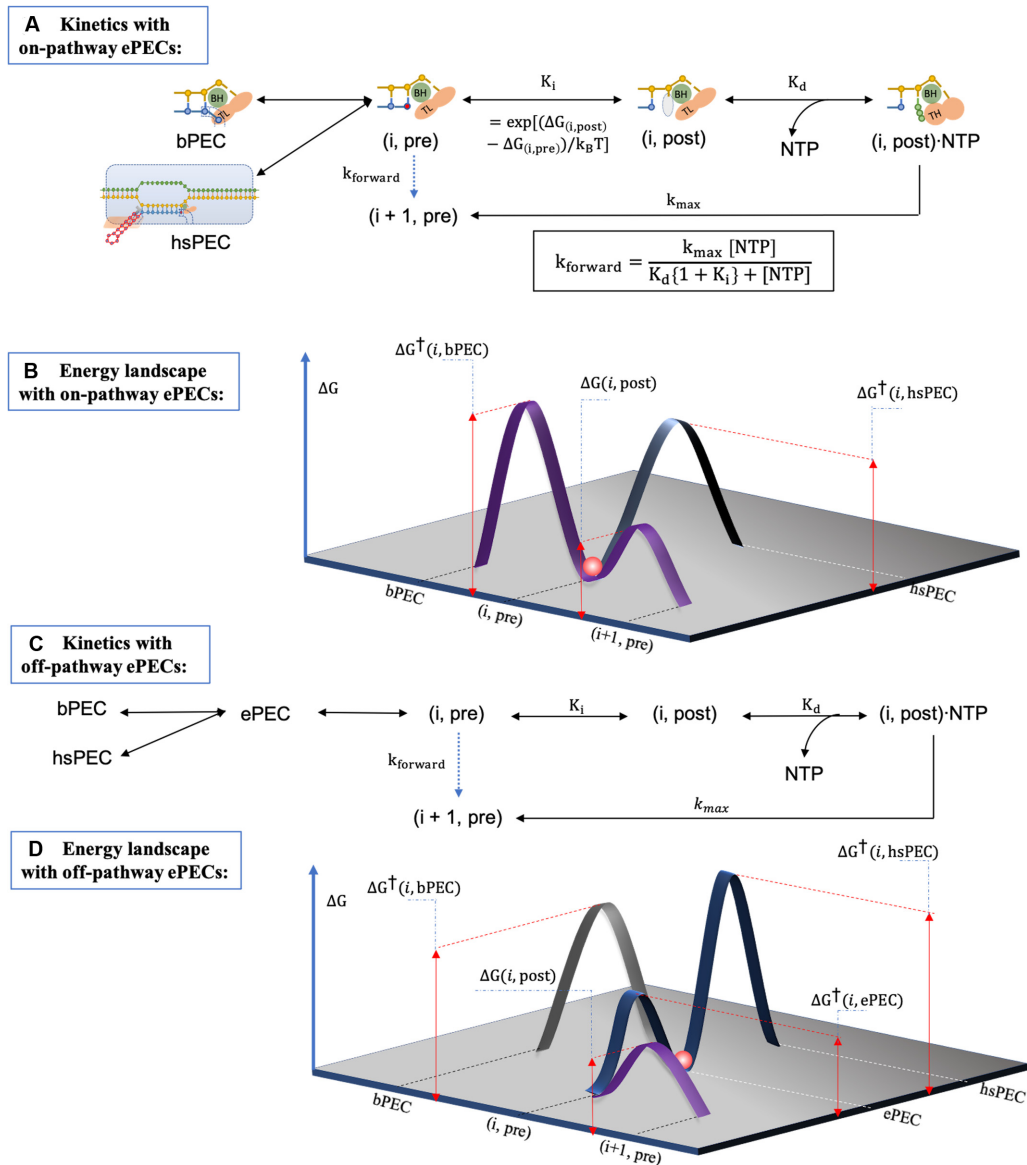


Figure 3. (A) A kinetic scheme of pause-interrupted elongation, with an on-pathway elementally paused complex (ePEC). The equilibrium constant K_i is determined by the free energies of the pre- and post-translocated configurations. K_d is the NTP-dependent dissociation constant, and k_{max} is the NTP-specific reaction rate of an irreversible step including NTP catalysis, PPi release and possibly other steps to catalyze phosphodiester bond formation between the 3' end of the nascent transcript and an incoming NTP. The overall forward rate from the (i, pre) state to the $(i+1, pre)$ state, indicated by the dashed arrow as a virtual connection, is given by the Michaelis-Menten equation with parameters K_i , K_d and k_{max} . (B) A qualitative energy landscape for the kinetic scheme in A. The elementally paused complex (red sphere) and hairpin-stabilize complex are translationally similar to the (i, pre) state. The backtracked, paused complex is translationally distinct from both the ePEC and the (i, pre) state. (C) A kinetic diagram with an off-pathway ePEC state that is conformationally different from the pre-translocated complex state and can further convert to the bPEC or hsPEC states. (D) A qualitative energy landscape associated with the kinetic scheme described in C.

that in models considering solely the pre-translocated pausing (with an energy landscape shown in Figure 3B), transcription tends to arrest quickly and the frequency of pausing is below that observed experimentally, while addition of an intermediate state preceding the entry to the backtracked pause (i.e. an off-pathway, elementally paused state) can resolve these problems (69). Also, they reported that the pre-translocated pause still comprises the majority of short pauses even if an intermediate state is included in the model. These findings indicate that pre-translocated and intermediate paused states may indeed coexist (and form an energy

landscape as shown in Figure 3D) but remain indistinguishable in experimental data.

The entry to the backtracked pause is usually modeled as an off-pathway event due to the upstream movement of RNAP across an energetic barrier. Although the energetic barriers of deeper backtracked complexes are likely larger than those of one-base-pair backtracked complexes, they are taken as constant in most cases (40,46). Backtracking recovery is usually simulated as a force-biased continuous or discrete random walk, where the backtrack dwell time is described as a first-passage time to reach an active reg-

istry (40,74). RNA cleavage functions as a stochastic resetting process that competes with the random walker (75). Moreover, the secondary structure of nascent transcripts is often considered a barrier that precludes deep backtracking (72,76). For *E. coli* genome pause sites, Bai *et al.* fitted a backtracking barrier of $46.2k_B T$ and incorporated only a 1D diffusive recovery pathway, while Douglas *et al.* estimated the barrier as a normal distribution centered at $5k_B T$ and incorporated a cleavage pathway for backtracks of <10 nucleotides (40,77). Despite significantly different energy barriers and recovery pathways, both models accurately predicted pause positions and dwell time distributions, indicating the difficulty in determining a biologically accurate model for the manifold transcription process.

As a counterpart to bPECs, hypertranslocated PECs, in which a TEC may translocate ahead without concomitant 3' RNA extension, were considered in several models. Hypertranslocated pauses have been reported experimentally and were thought to be the rate-limiting step for hsPEC until a more plausible mechanism was revealed from cryo-EM structures (26,78). The approach to model hypertranslocated pauses is similar to that of the backtracked pauses, with an energetic barrier to entry and escape through 1D thermally-driven diffusion (40,77).

The hsPEC is a class of transcriptional pauses for which the kinetics are least understood. No reported models depict the kinetics of the entry to and the escape from the hsPEC, due in large part to the difficulty of quantitatively detecting interactions between RNA and RNAP residues. Indeed, Dalal *et al.*, who used optical tweezers to pull directly on the nascent mRNA and inhibit folding during transcription, found that the kinetics of pausing were not affected by the perturbation (43). Future single-molecule experiments with lesser mechanical perturbations to the polymerase than used so far may be necessary to reveal hsPECs which nonetheless may comprise a minor fraction of paused states.

DISCUSSION

Currently, our understanding of the kinetics of transcript elongation trails our understanding of the mechanics. Although cryo-EM structures of paused TECs gave insight into the mechanisms of distinct pausing signals, these static snapshots cannot reveal the kinetics of pause states. For kinetics, we rely on single molecule experiments that reveal the dynamics of transcript elongation in conjunction with modeling that can decompose the overall dynamics into contributions from individual components. A Brownian ratchet model accurately reflects the rate and the force-dependence of the processive elongation observed in experiments. However, the kinetics of paused states are less clear: (i) in different proposed models, the ePEC are represented as either on-pathway or off-pathway states, and the models do not seem to completely fit the mechanistic origin of the ePEC; (ii) kinetics of the bPEC are well characterized, but fit-determined parameters vary according to experimental constructs; (iii) and finally, experiments to analyze the kinetics of the hsPEC are difficult. Overall, the difficulty in characterizing the kinetics of the paused states arises from distinguishing the observed pauses experimentally,

especially for the elemental and hairpin-stabilized PECs which cannot be distinguished translocationally in single-molecule transcription assays. Even for bPECs, which are characterized by reverse motion of TECs, efficient identification requires base-pair resolution and high signal-to-noise ratio experiments that are difficult to achieve in most laboratories.

The Brownian-ratchet model predicts the configuration of a transcription bubble to be a subtle but critical element that could significantly affect the kinetics. Nucleotides 1–2 positions proximal to the catalytic site and the unpaired nucleotides at the edges of the bubble are important, because they chiefly determine the relative stability of the pre- and post-translocated states, and hence the possibility of entering a paused state. Many models employ a fixed-length transcription bubble and a fixed-length DNA/RNA hybrid (40,77,79). These assumptions conflict with changes in the size of the transcription bubble detected experimentally and may introduce thermally unfavorable bubble configurations that might change spontaneously. A statistical mechanics approach was implemented in some model constructs to account for the variation in bubble/hybrid size (72).

The Brownian-ratchet model of RNAP transcription was proposed 30 years ago and has been subsequently refined. However, a significant defect of the model is a probable over-simplification of the real transcription mechanism, such as neglect of a potential allosteric nucleotide binding site that the elongation complex may contain, as proposed by Foster *et al.* (80). In addition, the effects of transcriptional modulators are overlooked in most models, and none address heterogeneity of pausing among species. For example, in *E. coli*, NusG is an anti-pausing factor that could stimulate forward translocation and prevent RNAP backtracking, while in *Bacillus subtilis*, NusG induces pausing by shifting RNAP to the post-translocation register (26,81–84). Another example is that *E. coli* RNAP recognizes a well-characterized hairpin-stabilized *his* pause site, while *B. subtilis* RNAP and mammalian Pol II do not respond to this hairpin-mediated signal (26,85).

Despite the limitations, models of transcript elongation have produced new insights about the kinetics of mechanistically identified paused states. For example, in modelling the distribution of pause times of bacterial TEC, Janissen *et al.* identified three interconnected pause states, two of which appear to be backtracked PECs. They found that the recovery from one occurs 20 times slower than that from the other and cannot be accelerated by cleavage factor GreB (57). Therefore, they postulated that a bPEC could undergo conformational changes to enter a longer-lived RNA-cleavage-resistant state. Douglas *et al.*, by comparing models with and without the intermediate state, backtracking, hypertranslocation and RNA folding, found that these factors are not necessary for predicting the locations and frequencies of pauses (77). Thus, they concluded that occurrence of pauses is chiefly facilitated by the relative stability between the pre- and post-translocated states, while the off-pathway events only serve to extend the pauses.

In the future, we hope to see a unified kinetic model of transcript elongation that concurs with our biochemical understanding of the effects of other factors on paused complexes and predicts the experimentally characterized pause

sites and pause frequencies. An immediate difficulty is the construction of distinctive models for the hsPEC and other mechanistically similar states and experimental methods with which to detect them. Moreover, such models must be consistent with the effects of transcriptional modulators that accentuate particular mechanisms of elongation and pausing from prokaryotes to eukaryotes.

FUNDING

National Institutes of Health (NIH) [R01 GM084070]. Funding for open access charge: NIGMS [R01GM084070]. *Conflict of interest statement.* None declared.

REFERENCES

- Huang, J. and Sousa, R. (2000) T7 RNA polymerase elongation complex structure and movement. *J. Mol. Biol.*, **303**, 347–358.
- Ikeda, R.A. and Richardson, C.C. (1986) Interactions of the RNA polymerase of bacteriophage T7 with its promoter during binding and initiation of transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 3614.
- Vassilyev, D.G., Vassilyeva, M.N., Perederina, A., Tahirov, T.H. and Artsimovitch, I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, **448**, 157–162.
- Toulokhonov, I., Zhang, J., Palangat, M. and Landick, R. (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell*, **27**, 406–419.
- Mazumder, A., Lin, M., Kapanidis, A.N. and Ebright, R.H. (2020) Closing and opening of the RNA polymerase trigger loop. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 15642.
- Toulokhonov, I., Artsimovitch, I. and Landick, R. (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science*, **292**, 730–733.
- 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.
- Larson, M.H., Landick, R. and Block, S.M. (2011) Single-molecule studies of RNA polymerase: one singular sensation, every little step it takes. *Mol. Cell*, **41**, 249–262.
- Dangkulwanich, M., Ishibashi, T., Bintu, L. and Bustamante, C. (2014) Molecular mechanisms of transcription through single-molecule experiments. *Chem. Rev.*, **114**, 3203–3223.
- Kireeva, M.L. and Kashlev, M. (2009) Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 8900.
- Weixlbaumer, A., Leon, K., Landick, R. and Darst, S.E. (2013) Structural basis of transcriptional pausing in bacteria. *Cell*, **152**, 431–441.
- Herbert, K.M., 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.
- 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.
- Bai, L., Fulbright, R.M. and Wang, M.D. (2007) Mechanochemical kinetics of transcription elongation. *Phys. Rev. Lett.*, **98**, 068103–068103.
- Thomen, P., Lopez, P.J. and Heslot, F. (2005) Unravelling the mechanism of RNA-Polymerase forward motion by using mechanical force. *Phys. Rev. Lett.*, **94**, 128102.
- Naftelberg, S., Schor, I.E., Ast, G. and Kornblihtt, A.R. (2015) Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annu. Rev. Biochem.*, **84**, 165–198.
- Saldi, T., Cortazar, M.A., Sheridan, R.M. and Bentley, D.L. (2016) Coupling of RNA polymerase II transcription elongation with Pre-mRNA splicing. *J. Mol. Biol.*, **428**, 2623–2635.
- 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*, **344**, 1042–1047.
- Gong, F. and Yanofsky, C. (2003) A transcriptional pause synchronizes translation with transcription in the tryptophanase operon leader region. *J. Bacteriol.*, **185**, 6472–6476.
- Mayer, A., Landry, H.M. and Churchman, L.S. (2017) Pause & go: from the discovery of RNA polymerase pausing to its functional implications. *Curr. Opin. Cell Biol.*, **46**, 72–80.
- Liu, X., Kraus, W.L. and Bai, X. (2015) Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. *Trends Biochem. Sci.*, **40**, 516–525.
- Gaertner, B. and Zeitlinger, J. (2014) RNA polymerase II pausing during development. *Development*, **141**, 1179.
- Levine, M. (2011) Paused RNA polymerase II as a developmental checkpoint. *Cell*, **145**, 502–511.
- 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.
- Landick, R. (2009) Transcriptional pausing without backtracking. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 8797.
- Artsimovitch, I. and Landick, R. (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 7090–7095.
- 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.
- Komissarova, N. and Kashlev, M. (1997) Transcriptional arrest: Escherichia coli RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 1755–1760.
- Perdue, S.A. and Roberts, J.W. (2010) A backtrack-inducing sequence is an essential component of Escherichia coli σ 70-dependent promoter-proximal pausing. *Mol. Microbiol.*, **78**, 636–650.
- Shankar, S., Hatoum, A. and Roberts, J.W. (2007) A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. *Mol. Cell*, **27**, 914–927.
- 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. *Mol. Cell*, **44**, 635–646.
- Kang, J.Y., Mishanina, T.V., Landick, R. and Darst, S.E. (2019) Mechanisms of transcriptional pausing in bacteria. *J. Mol. Biol.*, **431**, 4007–4029.
- 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*, **344**, 1042–1047.
- Vvedenskaya, I.O., Vahedian-Movahed, H., Bird, J.G., Knoblauch, J.G., Goldman, S.R., Zhang, Y., Ebright, R.H. and Nickels, B.E. (2019) Interactions between RNA polymerase and the core recognition element counteract pausing. *Science*, **344**, 1285–1289.
- Imashimizu, M., Takahashi, H., Oshima, T., McIntosh, C., Bubunenkov, M., Court, D.L. and Kashlev, M. (2015) Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases *in vivo*. *Genome Biol.*, **16**, 98.
- Kireeva, M., Trang, C., Matevosyan, G., Turek-Herman, J., Chasov, V., Lubkowska, L. and Kashlev, M. (2018) RNA-DNA and DNA-DNA base-pairing at the upstream edge of the transcription bubble regulate translocation of RNA polymerase and transcription rate. *Nucleic Acids Res.*, **46**, 5764–5775.
- 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.
- Yager, T.D. and Von Hippel, P.H. (1991) A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. *Biochemistry*, **30**, 1097–1118.
- Guajardo, R. and Sousa, R. (1997) A model for the mechanism of polymerase translocation. *J. Mol. Biol.*, **265**, 8–19.
- Bai, L., Shundrovsky, A. and Wang, M.D. (2004) Sequence-dependent kinetic model for transcription elongation by RNA polymerase. *J. Mol. Biol.*, **344**, 335–349.
- Saba, J., Chua, X.Y., Mishanina, T.V., Nayak, D., Windgassen, T.A., Mooney, R.A. and Landick, R. (2019) The elemental mechanism of transcriptional pausing. *eLife*, **8**, e40981.

42. Bochkareva, A., Yuzenkova, Y., Tadigotla, V.R. and Zenkin, N. (2012) Factor-independent transcription pausing caused by recognition of the RNA–DNA hybrid sequence. *EMBO J.*, **31**, 630–639.
43. 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. *Mol. Cell*, **23**, 231–239.
44. 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.
45. Mukherjee, S., Brieba, L.G. and Sousa, R. (2003) Discontinuous movement and conformational change during pausing and termination by T7 RNA polymerase. *EMBO J.*, **22**, 6483–6493.
46. Depken, M., Galburt, E.A. and Grill, S.W. (2009) The origin of short transcriptional pauses. *Biophys. J.*, **96**, 2189–2193.
47. Mejia, Y.X., Mao, H., Forde, N.R. and Bustamante, C. (2008) Thermal probing of E. coli RNA polymerase off-pathway mechanisms. *J. Mol. Biol.*, **382**, 628–637.
48. Yin, Y.W. and Steitz, T.A. (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell*, **116**, 393–404.
49. Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A. and Nudler, E. (2005) A ratchet mechanism of transcription elongation and its control. *Cell*, **120**, 183–193.
50. Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W. and Darst, S.A. (2003) Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell*, **114**, 335–345.
51. Cheung, A.C.M. and Cramer, P. (2011) Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature*, **471**, 249–253.
52. Marr, M.T. and Roberts, J.W. (2000) Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol. Cell*, **6**, 1275–1285.
53. Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y. and Adelman, K. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of pol II in *Drosophila*. *Science*, **327**, 335–338.
54. 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.
55. Gabizon, R., Lee, A., Vahedian-Movahed, H., Ebricht, R.H. and Bustamante, C.J. (2018) Pause sequences facilitate entry into long-lived paused states by reducing RNA polymerase transcription rates. *Nat. Commun.*, **9**, 2930.
56. Forde, N.R., Izhaky, D., Woodcock, G.R., Wuite, G.J.L. and Bustamante, C. (2002) Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci.*, **99**, 11682–11687.
57. Janissen, R., Eslami-Mossallam, B., Artsimovitch, I., Depken, M. and Dekker, N. (2019) A unifying mechanistic model of bacterial transcription with three interconnected pause states and non-diffusive backtrack recovery. bioRxiv doi: <https://doi.org/10.1101/762468>, 16 September 2019, preprint: not peer reviewed.
58. Lisica, A., Engel, C., Jahn, M., Roldán, É., Galburt, E.A., Cramer, P. and Grill, S.W. (2016) Mechanisms of backtrack recovery by RNA polymerases I and II. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 2946–2951.
59. Kyzer, S., Ha, K.S., Landick, R. and Palangat, M. (2007) Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J. Biol. Chem.*, **282**, 19020–19028.
60. Chan, C.L. and Landick, R. (1993) Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. *J. Mol. Biol.*, **233**, 25–42.
61. Widom, J.R., Nedialkov, Y.A., Rai, V., Hayes, R.L., Brooks, C.L. III, Artsimovitch, I. and Walter, N.G. (2018) Ligand modulates cross-coupling between riboswitch folding and transcriptional pausing. *Mol. Cell*, **72**, 541–552.
62. Roberts, J.W. (2019) Mechanisms of bacterial transcription termination. *J. Mol. Biol.*, **431**, 4030–4039.
63. Yakhnin, A.V. and Babitzke, P. (2010) Mechanism of NusG-stimulated pausing, hairpin-dependent pause site selection and intrinsic termination at overlapping pause and termination sites in the *Bacillus subtilis* trp leader. *Mol. Microbiol.*, **76**, 690–705.
64. Ray-Soni, A., Bellecourt, M.J. and Landick, R. (2016) Mechanisms of bacterial transcription termination: all good things must end. *Annu. Rev. Biochem.*, **85**, 319–347.
65. Touloukianov, I. and Landick, R. (2003) The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol. Cell*, **12**, 1125–1136.
66. Windgassen, T.A., Mooney, R.A., Nayak, D., Palangat, M., Zhang, J. and Landick, R. (2014) Trigger-helix folding pathway and S13 mediate catalysis and hairpin-stabilized pausing by *Escherichia coli* RNA polymerase. *Nucleic Acids Res.*, **42**, 12707–12721.
67. Chauvier, A., Nadon, J.-F., Grondin, J.P., Lamontagne, A.-M. and Lafontaine, D.A. (2019) Role of a hairpin-stabilized pause in the *Escherichia coli* thiC riboswitch function. *RNA Biol.*, **16**, 1066–1073.
68. Artsimovitch, I. and Landick, R. (1998) Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev.*, **12**, 3110–3122.
69. Maoiléidigh, D., Tadigotla, V.R., Nudler, E. and Ruckenstein, A.E. (2011) A unified model of transcription elongation: what have we learned from Single-Molecule experiments? *Biophys. J.*, **100**, 1157–1166.
70. Zuo, Y. and Steitz, T.A. (2017) A structure-based kinetic model of transcription. *Transcription*, **8**, 1–8.
71. Greive, S.J. and von Hippel, P.H. (2005) Thinking quantitatively about transcriptional regulation. *Nat. Rev. Mol. Cell Biol.*, **6**, 221–232.
72. Tadigotla, V.R., Maoiléidigh, D.Ó., Sengupta, A.M., Epshtein, V., Ebricht, R.H., Nudler, E. and Ruckenstein, A.E. (2006) Thermodynamic and kinetic modeling of transcriptional pausing. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 4439–4444.
73. Korzhova, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A. and Darst, S.A. (2000) A structural model of transcription elongation. *Science*, **289**, 619–625.
74. Voliotis, M., Cohen, N., Molina-Paris, C. and Liverpool, T.B. (2008) Fluctuations, pauses, and backtracking in DNA transcription. *Biophys. J.*, **94**, 334–348.
75. Roldán, É., Lisica, A., Sánchez-Taltavull, D. and Grill, S.W. (2016) Stochastic resetting in backtrack recovery by RNA polymerases. *Phys. Rev. E*, **93**, 062411.
76. Zamft, B., Bintu, L., Ishibashi, T. and Bustamante, C. (2012) Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. *Proc. Natl. Acad. Sci.*, **109**, 8948–8953.
77. Douglas, J., Kingston, R. and Drummond, A.J. (2020) Bayesian inference and comparison of stochastic transcription elongation models. *PLoS Comput. Biol.*, **16**, e1006717.
78. Nedialkov, Y.A., Nudler, E. and Burton, Z.F. (2012) RNA polymerase stalls in a post-translocated register and can hyper-translocate. *Transcription*, **3**, 260–269.
79. Wilson, K.S., Conant, C.R. and von Hippel, P.H. (1999) Determinants of the stability of transcription elongation complexes: interactions of the nascent RNA with the DNA template and the RNA polymerase I. Edited by K. Yamamoto. *J. Mol. Biol.*, **289**, 1179–1194.
80. Foster, J.E., Holmes, S.F. and Erie, D.A. (2001) Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation. *Cell*, **106**, 243–252.
81. Yakhnin, A.V., FitzGerald, P.C., McIntosh, C., Yakhnin, H., Kireeva, M., Turek-Herman, J., Mandell, Z.F., Kashlev, M. and Babitzke, P. (2020) NusG controls transcription pausing and RNA polymerase translocation throughout the *Bacillus subtilis* genome. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 21628.
82. Sevostyanova, A., Belogurov, Georgiy A., Mooney, Rachel A., Landick, R. and Artsimovitch, I. (2011) The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol. Cell*, **43**, 253–262.
83. Yakhnin, A.V., Kashlev, M. and Babitzke, P. (2020) NusG-dependent RNA polymerase pausing is a frequent function of this universally conserved transcription elongation factor. *Crit. Rev. Biochem. Mol. Biol.*, **55**, 716–728.
84. 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. *J. Mol. Biol.*, **399**, 17–30.
85. Keene, R.G., Mueller, A., Landick, R. and London, L. (1999) Transcriptional pause, arrest and termination sites for RNA polymerase II in mammalian N- and c-myc genes. *Nucleic Acids Res.*, **27**, 3173–3182.
86. Abdelkareem, M.M., Saint-André, C., Takacs, M., Papai, G., Crucifix, C., Guo, X., Ortiz, J. and Weixlbaumer, A. (2019) Structural basis of transcription: RNA polymerase backtracking and its reactivation. *Mol. Cell*, **75**, 298–309.