Cell

Leading Edge
Perspective

RNA Polymerase Backtracking in Gene Regulation and Genome Instability

Evgeny Nudler[1,*]
[1]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY 10016, USA
*Correspondence: evgeny.nudler@nyumc.org
DOI 10.1016/j.cell.2012.06.003

RNA polymerase is a ratchet machine that oscillates between productive and backtracked states at numerous DNA positions. Since its first description 15 years ago, backtracking—the reversible sliding of RNA polymerase along DNA and RNA—has been implicated in many critical processes in bacteria and eukaryotes, including the control of transcription elongation, pausing, termination, fidelity, and genome instability.

Twenty years ago, the first systematic analysis of transcription elongation complexes (ECs) revealed surprisingly irregular DNA footprints (Krummel and Chamberlin, 1992), suggesting that RNA polymerase (RNAP) shrinks and expands during elongation. This led to a provocative model for elongation called "inchworming," in which RNA synthesis is coupled with a leap-like movement of RNAP along the DNA (Chamberlin, 1994). The subsequent probing of many ECs stalled along long stretches of DNA revealed, however, that inchworming is not obligatory for elongation. Instead, the irregularities of footprints occurred only at certain DNA sites, whereas the majority of DNA positions displayed relatively monotonic movement of RNAP (Nudler et al., 1994).

Subsequently, two sets of biochemical data demonstrated that the occasional inchworming was actually reversible sliding of ECs along DNA and RNA: restricting RNA from threading back into the enzyme or stabilizing the RNA:DNA hybrid diminished footprint irregularities at inchworming sites (Komissarova and Kashlev, 1997a; Nudler et al., 1997). Moreover, analogs that destabilized the hybrid caused inchworming at sites where it previously did not exist. The conclusion was that the stability of the hybrid is the key determinant of the lateral mobility of EC (Nudler et al., 1997). This work also determined the actual length of the hybrid in EC to be 8 ± 1 base pairs and introduced the term "backtracking" to define the phenomenon of spontaneous sequence-dependent back and forth sliding of EC and to distinguish it from inchworming.

Backtracking and Gene Regulation by Pausing

During backtracking, the catalytic site becomes disengaged from the 3′ end of RNA, rendering EC inactive but stable (Nudler et al., 1997; Komissarova and Kashlev, 1997b) (Figure 1). This disengagement constitutes the mechanistic basis for a majority of regulatory pauses and arrests, although some pauses do not involve backtracking (Toulokhonov et al., 2007; Kireeva and Kashlev, 2009). Fraying of the 3′ RNA terminus, which interferes with nucleotide (nt) incorporation, was proposed to trigger both backtracked and nonbacktracked pauses (Toulokhonov et al., 2007; Sydow et al., 2009). Backtracking involves synchronized rewinding of the hybrid upstream, unwinding and rewinding of the DNA duplexes ahead and behind a transcription bubble, and threading of the single-stranded RNA through RNAP (Figure 1). Thus, the overall sequence context determines the probability of pausing at each nucleotide position (Tadigotla et al., 2006).

Promoter-proximal pauses constitute one large class of backtracking events. They usually occur within the first ~50 nt of transcribed sequences. Several features make promoter-proximal regions particularly prone to backtracking: (1) persistent contacts between the enzyme and initiation factors or promoter DNA tend to "pull" elongation complexes backward; (2) nascent RNA is not long enough to form structures that would prohibit reverse sliding; (3) there is little room for trailing ECs (and in bacteria, ribosomes) to "push" backtracked ECs forward.

Many *E. coli* and coliphage operones display early arrests that depend on the sigma initiation factor (σ^{70}) during elongation (Ring et al., 1996; Brodolin et al., 2004; Nickels et al., 2004; Hatoum and Roberts, 2008; Stepanova et al., 2009). The functional roles for most of these pauses are unknown, but σ^{70}-mediated backtracking is required for phage λ Q antitermination. Q protein is recruited to such arrested ECs and, together with the elongation factor NusA, forms an "antitermination shield" that insulates RNAP from terminators (Shankar et al., 2007).

Promoter-proximal pauses by eukaryotic RNAP II are widespread, particularly in highly active and regulated genes (Zeitlinger et al., 2007; Core et al., 2008; Rahl et al., 2010; Nechaev et al., 2010). These long-lasting pauses play a critical role in the transcriptional regulation of many genes and in RNA processing (Bentley and Groudine, 1986; Saunders et al., 2006; Chiba et al., 2010). One of the well-studied examples of promoter-proximal pausing occurs at the *Drosophila* heat-shock genes (Saunders et al., 2006). ECs located at the promoter of these genes pause as a result of backtracking (Adelman et al., 2005), poising them for rapid reactivation of transcription in response to stress. They also compete with nucleosomes at these highly regulated promoters, thereby inhibiting the formation of repressive chromatin structure and facilitating the rapid resumption of transcription (Gilchrist et al., 2010).

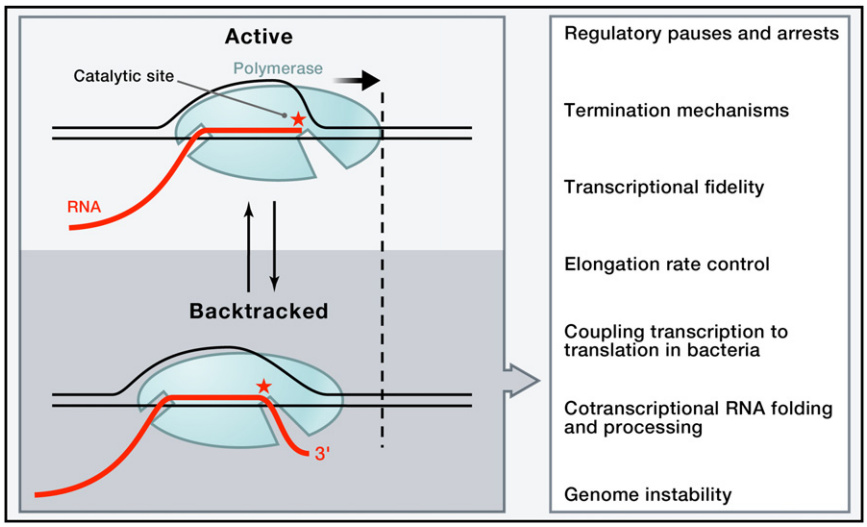

Figure 1. Multifaceted Role of RNAP Backtracking in the Cell

Schematics depict the ternary elongation complex (EC) in active and backtracked configurations. The catalytic site (star) looses the 3′-OH end of RNA (red), which is extruded through the secondary channel during backtracking.

In eukaryotic and prokaryotic cells, backtracked complexes can be rescued by transcript cleavage factors TFIIS (Reinberg and Roeder, 1987; Izban and Luse, 1992) and GreA/GreB (Borukhov et al., 1993), respectively. These factors stimulate intrinsic hydrolyzing activity of RNAP, which removes the 3′ extruded portion of the transcript to generate a new RNA 3′ end in the catalytic site, thereby reactivating the EC. The cleavage factors can relieve promoter proximal pausing (Marr and Roberts, 2000; Adelman et al., 2005; Stepanova et al., 2009).

Deep sequencing of the 3′ ends of nascent transcripts associated with yeast RNAP II revealed that backtracking-mediated pausing occurs not only near promoters, but ubiquitously throughout the transcribed sequence of any given gene (Churchman and Weissman, 2011). Among 2 × 10^5 pause sites detected in yeast genome, 75% were associated with backtracking. These pauses may not only control the rate of elongation in yeast and other organisms, but also provide a source for a wide range of microRNAs via transcript cleavage. The length of resulting RNA products would match the extension of backtracking at pause sites and vary from a few to dozens of nucleotides.

Backtracking Links the Rate of Elongation to the Efficiency of Initiation

In contrast to initiation, in which only one RNAP molecule occupies a promoter at a time, elongation often involves multiple RNAPs moving one after another along the same DNA duplex. Each molecule behaves as a ratchet machine (Bar-Nahum et al., 2005; Tagami et al., 2010). The probability of backtracking varies dramatically even at adjacent nucleotides, implying that, when leading EC backtracks, trailing EC would most likely be in the active mode, "pushing" leading EC forward. Such cooperation between ECs has been demonstrated for RNAPs in *E. coli* and yeast (Epshtein and Nudler, 2003; Epshtein et al., 2003; Jin et al., 2010; Saeki and Svejstrup, 2009).

These data suggest that the elongation phase should be considered as the effort of the entire group of RNAP molecules within the same transcription unit, which effectively links the initiation and elongation steps; the more robust the initiation, the more closely spaced the elongating RNAP molecules, and thus, the lower the probability of backtracking at any individual position. Indeed, the stronger the promoter is, the faster the elongation occurs in vitro and in vivo (Epshtein and Nudler, 2003). This cooperative mechanism explains, at least in part, why the most active genes, e.g., rRNA or heat-shock genes, have the highest elongation rates in bacteria and eukaryotes.

This cooperation effect likely extends beyond matching RNA output to promoter strength. The elongation rate modulates alternative splicing by affecting the timing at which splice sites are exposed to the splicing machinery (Kornblihtt, 2007). Moreover, paused ECs near the 3′ splice site facilitate cotranscriptional splicing (Carrillo Oesterreich et al., 2011). Thus, RNA processing can be influenced by the robustness of initiation. This kinetic mechanism could work independently or in cooperation with more specific (factor-dependent) mechanisms (Nagaike et al., 2011) in coupling transcription activation to pre-mRNA processing.

Backtracking Coordinates Transcription and Translation in Bacteria

In bacteria, transcription and translation are coupled. Translating ribosomes closely follow moving ECs. As a result, the trailing ribosome is able to "push" backtracked RNAP forward, thereby accelerating its speed (Proshkin et al., 2010; Burmann et al., 2010). This "cooperation" between ribosome and RNAP explains how the rate of transcription elongation perfectly matches the rate of translation under various growth conditions (Proshkin et al., 2010). It also explains why it depends on codon usage (i.e., the frequency of rare codons, which modulate the speed of a ribosome) (Proshkin et al., 2010). This cooperation not only conserves energy by limiting any excessive transcripts that cannot be translated in a timely manner, but it also prevents premature Rho termination by ensuring continuous coupling between transcription and translation. Thus, bacteria rely on trafficking and cooperation to finely control the expression of each individual gene in response to nutrient availability and growth phase.

Backtracking and Transcriptional Roadblocks

RNAP must traverse numerous potential roadblocks in vivo, such as nucleosomes in eukaryotes and nucleoid-associated proteins in bacteria. Although RNAP progresses relatively unimpeded in vivo, protein roadblocks readily inhibit transcription in vitro (Izban and Luse, 1991; Reines and Mote, 1993; Epshtein

et al., 2003; Walter et al., 2003; Lewis et al., 2008). Backtracking helps to resolve this apparent paradox.

Even though RNAP exerts sufficient force (Wang et al., 1998) to displace many DNA-bound proteins in its path, its propensity to backtrack complicates this displacement for any individual molecule. Indeed, upon collisions with different DNA-bound proteins, such as the lac repressor, hydrolytically defective EcoRI, or a nucleosome, RNAP backtracks in vitro and in vivo (Epshtein et al., 2003; Walter et al., 2003; Churchman and Weissman, 2011). In vitro, it can remain backtracked in front of the roadblock indefinitely (Epshtein et al., 2003). Thus, to overcome various roadblocks, RNAP must first be reactivated. Indeed, transcript cleavage factors have been shown to facilitate readthrough of the roadblocks (Reines and Mote, 1993; Epshtein et al., 2003; Walter et al., 2003). Moreover, cooperation between RNAP molecules and between RNAP and ribosomes (in bacteria) is a general and efficient mechanism of traversing the roadblocks, including nucleosomes (Epshtein et al., 2003; Proshkin et al., 2010; Jin et al., 2010).

Backtracking and Transcriptional Fidelity

Backtracking also provides a basic mechanism for transcriptional proofreading. It strongly depends on the stability of the RNA:DNA hybrid in the transcription bubble and on the nature of the 3′-terminal residue (Nudler et al., 1997; Sosunov et al., 2003): the weaker the hybrid, the higher the probability for backtracking (Nudler et al., 1997). Therefore, any mismatch in the hybrid would induce immediate backtracking, which in turn, would result in cleavage and removal of the 3′ RNA portion that contained a misincorporated nucleotide.

Indeed, GreA and TFIIS have been shown to substantially enhance transcriptional fidelity in vitro (Erie et al., 1993; Jeon and Agarwal, 1996; Thomas et al., 1998) and in vivo (Koyama et al., 2003), and the rpoB9 subunit of RNAP II, which facilitates SII-dependent transcript cleavage, contributes to fidelity in vivo (Nesser et al., 2006; Koyama et al., 2007). The mismatch during NTP insertion can also stabilize a paused state of RNAP with a frayed RNA 3′ nucleotide that inactivates RNAP and promotes backtracking and proofreading (Toulokhonov et al., 2007; Sydow et al., 2009).

Other evidence linking backtracking to fidelity comes from biochemical analysis of RNAP mutants that alter its propensity to backtrack. Backtracking-prone ("slow") RNAP usually exhibits less misincorporation, whereas backtracking-resistant ("fast") RNAPs appear to be more error prone (Bar-Nahum et al., 2005; Kireeva et al., 2008). Binding of correct NTP in the i+1 site of the catalytic center stabilizes RNAP in the posttranslocated state and suppresses backtracking (Bar-Nahum et al., 2005), most likely via substrate-induced folding of the "trigger loop" (TL) domain. By closing around the active center, TL transiently captures the correct substrate (Vassylyev et al., 2007; Kaplan et al., 2008; Kireeva et al., 2008). At the same time, it partially occludes the secondary channel (i.e., NTP delivery pore) through which RNA is extruded during backtracking (Korzheva et al., 2000). Thus, RNAP backtracking, which depends on TL conformation may assist in substrate selection; an incorrect NTP facilitates backtracking and hence its own expulsion through the secondary channel, whereas the correct NTP stabilizes the enzyme in the catalytically competent (i.e., backtracking-resistant) mode, thereby facilitating its own incorporation.

Backtracking and Transcriptional Termination

Intrinsic termination signals in bacteria consist of a GC-rich inverted repeat followed immediately by a stretch of T bases ("T stretch"). The resulting transcript forms a stable hairpin followed by several Us at the 3′ terminus. T stretches are typical backtracking signals because they create weak U:A base pairs. Irrespective of the termination hairpin sequence, the T stretch induces a brief pause precisely at the termination position (Gusarov and Nudler, 1999, 2001). This type of pausing depends on the 3′ proximal portion of the T stretch and can also be affected by bases immediately downstream of the catalytic site. The extent of this pausing determines the termination efficiency (Gusarov and Nudler, 1999, 2001).

The purpose of this pause is to provide enough time for the hairpin to fold at the right distance from the catalytic site. The termination hairpin has to fold in the closed confinement of the RNA exit channel to exert its destabilizing effect on the EC (Gusarov and Nudler, 1999; Epshtein et al., 2007) or to pull RNA from the catalytic site (known as the shearing model) (Larson et al., 2008). In either case, the pause widens the window of opportunity for the hairpin to overcome the energy barrier associated with EC destabilization. Backtracking for 1 or 2 nt in this case is sufficient to pause RNAP at the termination point without interfering with hairpin nucleation. Indeed, suppressing backtracking at the termination point either by altering its sequence (Gusarov and Nudler, 1999) or by making EC less prone to backtracking (Epshtein and Nudler, 2003; Bar-Nahum et al., 2005) inhibits termination.

T stretches are crucial elements of the termination process not only in bacteria, but also in eukaryotic cells. Eukaryotic RNAP III and archaeal RNAP pause and terminate transcription at T stretches in the absence of additional factors (Santangelo and Reeve, 2006; Werner et al., 2009). Although the actual mechanism of RNA release by either of these RNAPs has not been established, it is likely that T-stretch-associated backtracking is a component of this process.

Backtracking may also be involved in RNAP I termination, which requires a site-specific DNA-binding protein (TTF-I) that acts as a roadblock and a T-rich release element immediately upstream of the roadblocking site (Grummt, 1999; Werner et al., 2009). Termination occurs ~10–12 bp promoter proximal of the TTF-I site, suggesting extensive backtracking by halted EC.

Backtracking and Genome Instability

All dividing cells must endure frequent collisions between replication and transcription complexes, which occupy the same DNA track and function at the same time. This is particularly true of bacteria in which the rate of replisome propagation is about 20 times faster than that of RNAP. Most active and essential genes in bacteria tend to be organized codirectionally with replication, and thus, codirectional collisions should be more frequent than head-on collisions. Recent evidence demonstrates that DNA damage resulting from such collisions depends on RNAP backtracking (Dutta et al., 2011).

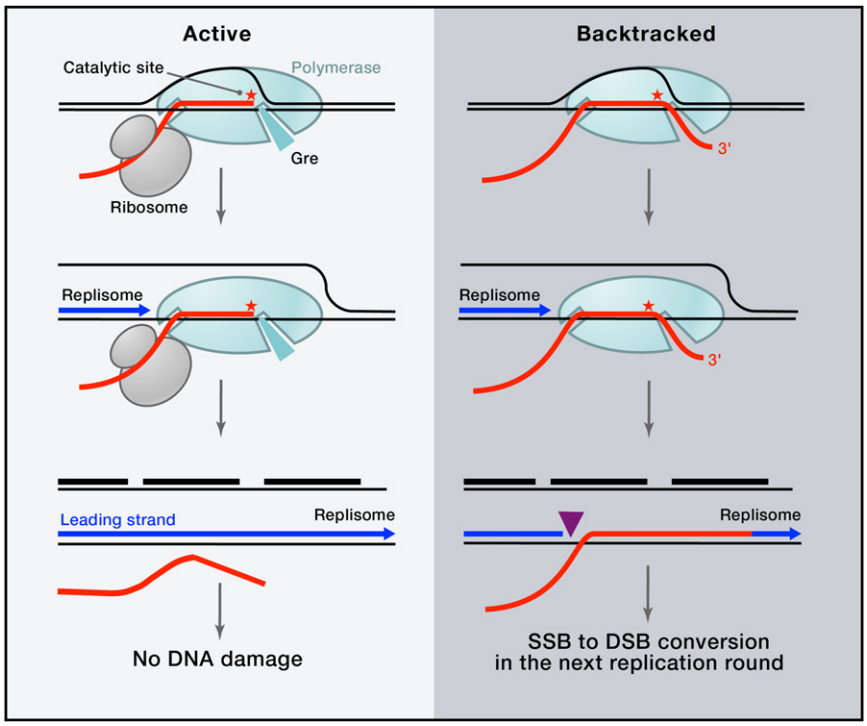

Figure 2. RNAP Backtracking and Genome Instability

Schematics show a model of double-strand break (DSB) formation as a result of codirectional collisions between the replisome and backtracked RNAP in bacteria (Dutta et al., 2011). The pink arrow indicates a single-strand break (SSB) due to replisome switching from the leading DNA strand (blue) to the RNA (red). The latter forms a stable R loop upon displacement of the backtracked elongation complex (EC). Transcript cleavage factor (Gre) in the secondary channel and trailing active ribosome prevent RNAP backtracking and R loop formation, thus preserving genome integrity.

The majority of ECs are stable protein-DNA complexes that must be dislodged by the replisome, regardless of their directionality. The structural organization of EC (Kettenberger et al., 2004; Vassylyev et al., 2007) eliminates any conceivable mechanism of replication that does not involve EC dissociation. Indeed, in vitro, the replisome kicks off bacterial RNAP by approaching it codirectionally or head on (Pomerantz and O'Donnell, 2010). In the codirectional configuration, bypassing of EC by the replisome seems to occur without a delay, whereas the head-on configuration was associated with replisome stalling.

However, codirectional collisions are less benign than traditionally thought. A genome-wide analysis of transcription-replication collisions in exponentially growing *Bacillus subtilis* revealed that codirectional collisions at ribosomal genes lead to the disruption and restart of replication (Merrikh et al., 2011). In *E. coli*, chromosomal double-strand breaks (DSBs) are the consequences of codirectional conflicts, which occur only upon collisions with backtracked ECs (Dutta et al., 2011).

Normally, such collisions are avoided by antibacktracking mechanisms that involve active ribosomes and transcript cleavage factors (Dutta et al., 2011) (Figure 2). Translating ribosomes play the primary role in preserving the integrity of protein-coding genes. If translation is compromised (e.g., by antibiotics), antibacktracking factors (GreA and GreB) as well as transcription termination factors Rho and Mfd become essential in preventing collision-related DSBs (Dutta et al., 2011; Washburn and Gottesman, 2011). For noncoding stable RNA genes, the intensive cooperation between RNAPs is likely to diminish potentially harmful backtracking.

An attractive model that explains DSBs resulting from codirectional collisions originates from in vitro observations that the synthesis of the leading DNA strand can be interrupted (Figure 2); once RNAP is displaced, DNAP (Pol III) can "jump" to the 3′ end of the nascent transcript and use it as a primer (Pomerantz and O'Donnell, 2010). This jumping by DNAP leads to a single-strand DNA break that must be repaired in vivo before the next round of replication converts it to a DSB (Figure 2). Apparently, DNA repair is not robust enough under conditions of excessive backtracking. According to the model, rapid reannealing of RNA from displaced backtracked RNAPs generates extended R loops (RNA:DNA hybrids) because backtracked RNAPs carry longer segments of RNA that are available for reannealing (the extruded 3′ portion) (Figure 2). Such extended hybrids or R loops provide accessible 3′-OH termini that could serve as primers for DNA Pol III (Figure 2). In contrast, active ECs form hybrids of only ~8 bp (Nudler et al., 1997), which are unstable and cannot survive without RNAP (Figure 3) and, therefore, do not support discontinuous replication. In support of the model, it has been shown that high levels of hybrid-specific RNase H, as well as RNAP mutations that diminish backtracking, eliminate DSBs associated with codirectional collisions in vivo (Dutta et al., 2011).

Because RNAP backtracking in bacteria provides a mechanistic link between protein synthesis and genome instability (DSBs), it has several important implications for bacterial adaptation and evolution. The ribosome is the principal sensor of metabolic fluctuations and stress. Starvation, proteotoxic challenges, and various antibiotics reduce or eliminate protein synthesis, thereby increasing the probability of RNAP backtracking and formation of DSBs. These same adverse conditions activate stress-induced mutagenesis that depends on the error-prone DSB repair process (Galhardo et al., 2007). This repair process, in turn, accelerates adaptation to environmental changes, such as acquisition of antibiotic resistance (Galhardo et al., 2007).

Thus, RNAP backtracking may contribute to stress-driven bacterial evolution. Indeed, mutation and recombination rates of Gre-deficient (i.e., backtracking-prone) cells are higher than that of wild-type cells (Dutta et al., 2011; Poteete, 2011), whereas survival of such cells depends on the SOS response and error-prone DSB repair (Dutta et al., 2011). Moreover, the likelihood

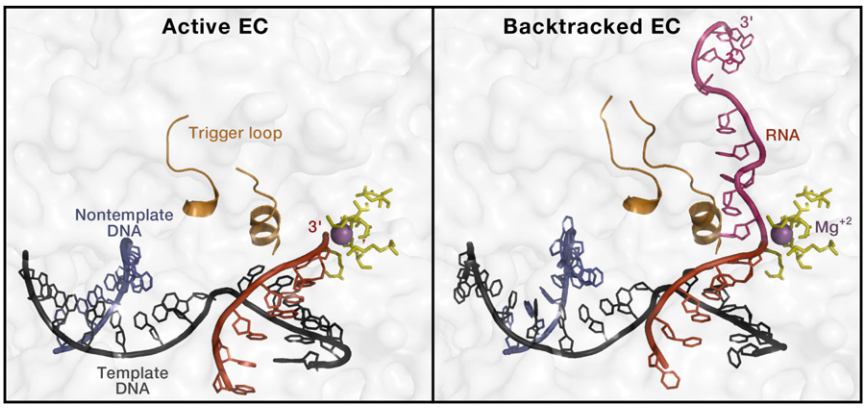

Figure 3. Structural Basis of RNAP Backtracking

Views of nucleic acids in the backtracked (PDB ID: 3PO2) and active (PDB ID: 1Y1W) RNAP II structures. Nontemplate and template DNAs are blue and black cartoons, respectively; RNA is a red cartoon with RNA extruded past the active site in hot pink. The Mg(I) ion is a magenta sphere; catalytic site residues are yellow sticks; and the trigger loop is an orange cartoon. The RNA:DNA hybrid is partially unwound and tilted in the backtracked elongation complex (EC), whereas the trigger loop is stabilized in the trapped conformation by interactions with the extruded RNA, which is also contacted by the interior of the secondary channel (Cheung and Cramer, 2011).

of RNAP backtracking increases in direct proportion to the frequency of rare codons, which modulate the rate of ribosome movement (Proshkin et al., 2010). This may explain why the mutation frequency is considerably higher for rare codons than for common codons (Alff-Steinberger, 2000) and predicts that many mutational hot spots are associated with local ribosome pausing.

Although protein synthesis is not apparently linked to the antibacktracking mechanism in eukaryotes, some of the numerous RNA-binding proteins that travel with RNAP II and participate in RNA processing and transport may function to suppress backtracking. Other antibacktracking mechanisms are clearly important. In contrast to bacterial transcript cleavage factor GreB, which is dispensable under nonstress conditions, its eukaryotic analog, TFIIS, is essential for cell viability (Sigurdsson et al., 2010). This is not surprising, considering that backtracking occurs at numerous positions within transcribed sequences of any given gene (Churchman and Weissman, 2011).

Thus, codirectional collisions with backtracked RNAP in eukaryotes are also inevitable. Considering the high evolutionary conservation between bacterial and eukaryotic replisomes and between cellular RNAPs, it seems likely that such collisions would result in DSBs via the same R-loop-dependent mechanism (Figure 2) and may constitute one source of genome instability. Some recent evidence implicating transcription-dependent R loops in genome instability support this notion (Aguilera and García-Muse, 2012; Kim and Jinks-Robertson, 2012).

Structural Basis of Backtracking

High-resolution atomic structures of backtracked ECs have become available for yeast RNAP II (Wang et al., 2009; Cheung and Cramer, 2011; see Cheung and Cramer, 2012 in this issue). These structures differ substantially from those of on-pathway ECs (Kettenberger et al., 2004; Wang et al., 2006; Brueckner and Cramer, 2008), offering some insights into the basis of the unconventional biochemical and biophysical properties of backtracked/arrested complexes (Figure 3).

As expected, the 3′ end of the nascent RNA in backtracked EC appears to be threaded into the secondary channel (i.e., the funnel) (Cheung and Cramer, 2011). Although this characteristic of backtracked EC might appear trivial in light of past biochemical data, other structural features of this complex were not anticipated.

EC backtracked by 9 nt has an unexpectedly short DNA:RNA hybrid (6 bp instead of 8–9 bp) that is tilted toward the bridge helix so that the −1 (relative to the active site) nt base of the template occupies the position normally taken by the +1 nt in active EC (Cheung and Cramer, 2011) (Figure 3). This, together with displacement of the 3′ end of the RNA from the active site (into the secondary channel), renders arrested EC incapable of NTP addition. Such a short hybrid would render nonbacktracked EC unstable (Kireeva et al., 2000). Instead, the arrested complex is exceptionally stable and apparently resistant to assisting mechanical force (Forde et al., 2002). The determinants of this exceptional stability lie not in the RNA:DNA hybrid but, rather, in the arrangement of the nucleic acid scaffold elsewhere.

One factor that is apparently responsible for stabilization of arrested EC is the binding of the RNA to the interior of the secondary channel, where it has been extruded and with which it normally does not engage (Cheung and Cramer, 2011). Specifically, in the structure reported by Cheung and Cramer, the 9-nt-long backtracked RNA makes contacts with one side of the secondary channel (dubbed "backtrack site") and a trigger loop on the other side. The latter was observed to adopt a new conformation that is different than all others reported previously (Kettenberger et al., 2004; Brueckner and Cramer, 2008; Wang et al., 2006, 2009) and was designated as "trapped" to indicate its incompatibility with on-pathway elongation (Figure 3). This new, extensive RNA-RNAP interface provides the energy necessary to stabilize this complex (undermined by the distorted and shortened RNA-DNA hybrid). At the same time, the trigger loop trapped in the nonproductive conformation by its interactions with the RNA contributes to other impediments to the spontaneous restart of elongation.

Structures of backtracked ECs also provide an explanation for why short backtracks can be as irreversible as longer ones. Tyrosine 749 in Rpb2, the second largest subunit of RNAP II, forms stacking interactions with the first backtracked RNA residue (+2) (Wang et al., 2009; Cheung and Cramer, 2011), making backtracking by 1 nt energetically neutral or even favorable (Cheung and Cramer, 2011). Further backtracking, however, would lead to disruption of RNA base stacking by this residue (hence called "gating tyrosine"), creating an activation

barrier that can be traversed only under special conditions (e.g., weak RNA:DNA hybrid, base-stacking interactions), resulting in arrest of backtracked EC stabilized by a new set of interactions with the extruded RNA (Cheung and Cramer, 2011).

Whereas "gating tyrosine" interactions with RNA appear to limit the extent of the initial backtracking (to 1 nt or more at some sites), further interactions remain largely unexplored. In their report, Cheung and Cramer (2011) provide insight into the natural upper limit for the extent of EC backtracking. Having obtained a highly defined structure of polypyrimidine RNA extruded into the secondary channel, they observed by modeling that the same backbone with purine bases produced clashes with the RNAP II. This finding not only explains the propensity of ECs to backtrack along pyrimidine-rich RNAs (Hawryluk et al., 2004), but also suggests that the first purine in the nascent RNA encountered by the EC in its retrograde motion along the nucleic acid scaffold may inhibit further backtracking due to steric clashes in the secondary channel.

Altogether, the available X-ray structures of various backtracked complexes help to explain their resistance to spontaneous reactivation. In the past, the role of the cleavage factors, such as bacterial GreA/B or eukaryotic TFIIS, was seen largely as one of remodeling or reactivating the RNAP active site through the donation of Mg^{2+}-coordinating acidic residues to stimulate cleavage of the RNA (Opalka et al., 2003; Laptenko et al., 2003; Sosunov et al., 2003). The emerging realization that backtracked EC is not equivalent to an active one simply displaced backward along the nucleic acid scaffold but is a distinct and stable conformational off-pathway state of RNAP puts the task of its reactivation beyond the mere formation of a new RNA 3′ end in the active site. It appears more likely that even the cleaved extruded portion of the RNA would exhibit a slow off rate from its extruded position (due to its extensive interactions with the secondary channel and the trigger loop trapped in an inactive conformation), delaying reactivation of the complex. TFIIS soaked into crystals formed by an arrested EC displaces the RNA from its binding site in the secondary channel and restores the trigger loop from the trapped to the "locked" conformation (Cheung and Cramer, 2011). A similar interplay between the trigger loop and GreB was proposed based on biochemical studies of backtracked *E. coli* RNAP (Roghanian et al., 2011). As a result, the remobilized RNA can be cleaved off and can dissociate, leaving EC in the state poised for NTP addition.

Conclusions

Backtracking is a fundamental property of RNAPs that allows the regulation of transcriptional elongation. It provides a means for elongation and termination factors to act on RNAP to control its local transit, overall rate, and accuracy. Backtracking-mediated pausing also plays a major role in transcription reactivation, termination, antitermination, cotranscriptional RNA folding, and processing. Finally, the association of backtracking with genome instability, at least in bacteria, provides a mechanistic link between growth conditions and cellular adaptation to stress. Backtracking is a remarkable example of how an enzyme's Brownian motion could broadly impact cellular physiology and evolution.

ACKNOWLEDGMENTS

I thank Vladimir Svetlov for his assistance in preparing the manuscript and anonymous reviewers for valuable comments. This work was supported by a grant from the NIH R01 GM58750 (E.N.).

REFERENCES

Adelman, K., Marr, M.T., Werner, J., Saunders, A., Ni, Z., Andrulis, E.D., and Lis, J.T. (2005). Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. Mol. Cell *17*, 103–112.

Aguilera, A., and García-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. Mol. Cell *46*, 115–124.

Alff-Steinberger, C. (2000). A comparative study of mutations in Escherichia coli and Salmonella typhimurium shows that codon conservation is strongly correlated with codon usage. J. Theor. Biol. *206*, 307–311.

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.

Bentley, D.L., and Groudine, M. (1986). A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. Nature *321*, 702–706.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from E. coli. Cell *72*, 459–466.

Brueckner, F., and Cramer, P. (2008). Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. Nat. Struct. Mol. Biol. *15*, 811–818.

Brodolin, K., Zenkin, N., Mustaev, A., Mamaeva, D., and Heumann, H. (2004). The sigma 70 subunit of RNA polymerase induces lacUV5 promoter-proximal pausing of transcription. Nat. Struct. Mol. Biol. *11*, 551–557.

Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rösch, P. (2010). A NusE:NusG complex links transcription and translation. Science *328*, 501–504.

Carrillo Oesterreich, F., Bieberstein, N., and Neugebauer, K.M. (2011). Pause locally, splice globally. Trends Cell Biol. *21*, 328–335.

Chamberlin, M.J. (1994). Harvey Lectures, Series *88* (New York: Wiley-Liss), pp. 1–21.

Cheung, A.C., and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature *471*, 249–253.

Cheung, A.C., and Cramer, P. (2012). A movie of RNA polymerase II transcription. Cell *149*, this issue, 1431–1437.

Chiba, K., Yamamoto, J., Yamaguchi, Y., and Handa, H. (2010). Promoter-proximal pausing and its release: molecular mechanisms and physiological functions. Exp. Cell Res. *316*, 2723–2730.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M.E., and Nudler, E. (2011). Linking RNA polymerase backtracking to genome instability in E. coli. Cell *146*, 533–543.

Epshtein, V., and Nudler, E. (2003). Cooperation between RNA polymerase molecules in transcription elongation. Science *300*, 801–805.

Epshtein, V., Toulmé, F., Rahmouni, A.R., Borukhov, S., and Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. EMBO J. *22*, 4719–4727.

Epshtein, V., Cardinale, C., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). Allosteric path to transcription termination. Mol. Cell *28*, 991–1001.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C., and von Hippel, P.H. (1993). Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. Science *262*, 867–873.

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. Proc. Natl. Acad. Sci. USA *99*, 11682–11687.

Galhardo, R.S., Hastings, P.J., and Rosenberg, S.M. (2007). Mutation as a stress response and the regulation of evolvability. Crit. Rev. Biochem. Mol. Biol. *42*, 399–435.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540–551.

Grummt, I. (1999). Regulation of mammalian ribosomal gene transcription by RNA polymerase I. Prog. Nucleic Acid Res. Mol. Biol. *62*, 109–154.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Mol Cell *3*, 495–504.

Gusarov, I., and Nudler, E. (2001). Control of intrinsic transcription termination by N and NusA: the basic mechanisms. Cell *107*, 437–449.

Hatoum, A., and Roberts, J. (2008). Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. Mol. Microbiol. *68*, 17–28.

Hawryluk, P.J., Ujvári, A., and Luse, D.S. (2004). Characterization of a novel RNA polymerase II arrest site which lacks a weak 3′ RNA-DNA hybrid. Nucleic Acids Res. *32*, 1904–1916.

Izban, M.G., and Luse, D.S. (1991). Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. Genes Dev. *5*, 683–696.

Izban, M.G., and Luse, D.S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3′----5′ direction in the presence of elongation factor SII. Genes Dev. *6*, 1342–1356.

Jeon, C., and Agarwal, K. (1996). Fidelity of RNA polymerase II transcription controlled by elongation factor TFIIS. Proc. Natl. Acad. Sci. USA *93*, 13677–13682.

Jin, J., Bai, L., Johnson, D.S., Fulbright, R.M., Kireeva, M.L., Kashlev, M., and Wang, M.D. (2010). Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. Nat. Struct. Mol. Biol. *17*, 745–752.

Kaplan, C.D., Larsson, K.M., and Kornberg, R.D. (2008). The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. Mol. Cell *30*, 547–556.

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol. Cell *16*, 955–965.

Kireeva, M.L., and Kashlev, M. (2009). Mechanism of sequence-specific pausing of bacterial RNA polymerase. Proc. Natl. Acad. Sci. USA *106*, 8900–8905.

Kireeva, M.L., Komissarova, N., Waugh, D.S., and Kashlev, M. (2000). The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J. Biol. Chem. *275*, 6530–6536.

Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purtov, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N., and Kashlev, M. (2008). Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol. Cell *30*, 557–566.

Kim, N., and Jinks-Robertson, S. (2012). Transcription as a source of genome instability. Nat. Rev. Genet. *13*, 204–214.

Komissarova, N., and Kashlev, M. (1997a). RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. J. Biol. Chem. *272*, 15329–15338.

Komissarova, N., and Kashlev, M. (1997b). Transcriptional arrest: Escherichia coli RNA polymerase translocates backward, leaving the 3′ end of the RNA intact and extruded. Proc. Natl. Acad. Sci. USA *94*, 1755–1760.

Kornblihtt, A.R. (2007). Coupling transcription and alternative splicing. Adv. Exp. Med. Biol. *623*, 175–189.

Korzheva, 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.

Koyama, H., Ito, T., Nakanishi, T., Kawamura, N., and Sekimizu, K. (2003). Transcription elongation factor S-II maintains transcriptional fidelity and confers oxidative stress resistance. Genes Cells *8*, 779–788.

Koyama, H., Ito, T., Nakanishi, T., and Sekimizu, K. (2007). Stimulation of RNA polymerase II transcript cleavage activity contributes to maintain transcriptional fidelity in yeast. Genes Cells *12*, 547–559.

Krummel, B., and Chamberlin, M.J. (1992). Structural analysis of ternary complexes of Escherichia coli RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. J. Mol. Biol. *225*, 239–250.

Laptenko, O., Lee, J., Lomakin, I., and Borukhov, S. (2003). Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. EMBO J. *22*, 6322–6334.

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.

Lewis, D.E., Komissarova, N., Le, P., Kashlev, M., and Adhya, S. (2008). DNA sequences in gal operon override transcription elongation blocks. J. Mol. Biol. *382*, 843–858.

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.

Merrikh, H., Machon, C., Grainger, W.H., Grossman, A.D., and Soultanas, P. (2011). Co-directional replication-transcription conflicts lead to replication restart. Nature *470*, 554–557.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2011). Transcriptional activators enhance polyadenylation of mRNA precursors. Mol. Cell *41*, 409–418.

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.

Nesser, N.K., Peterson, D.O., and Hawley, D.K. (2006). RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. Proc. Natl. Acad. Sci. USA *103*, 3268–3273.

Nickels, B.E., Mukhopadhyay, J., Garrity, S.J., Ebright, R.H., and Hochschild, A. (2004). The sigma 70 subunit of RNA polymerase mediates a promoter-proximal pause at the lac promoter. Nat. Struct. Mol. Biol. *11*, 544–550.

Nudler, E., Goldfarb, A., and Kashlev, M. (1994). Discontinuous mechanism of transcription elongation. Science *265*, 793–796.

Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell *89*, 33–41.

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.

Pomerantz, R.T., and O'Donnell, M. (2010). What happens when replication and transcription complexes collide? Cell Cycle *9*, 2537–2543.

Poteete, A.R. (2011). Recombination phenotypes of Escherichia coli greA mutants. BMC Mol. Biol. *12*, 12.

Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science *328*, 504–508.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432–445.

Reinberg, D., and Roeder, R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. J. Biol. Chem. *262*, 3331–3337.

Reines, D., and Mote, J., Jr. (1993). Elongation factor SII-dependent transcription by RNA polymerase II through a sequence-specific DNA-binding protein. Proc. Natl. Acad. Sci. USA *90*, 1917–1921.

Ring, B.Z., Yarnell, W.S., and Roberts, J.W. (1996). Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. Cell *86*, 485–493.

Roghanian, M., Yuzenkova, Y., and Zenkin, N. (2011). Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. Nucleic Acids Res. *39*, 4352–4359.

Saeki, H., and Svejstrup, J.Q. (2009). Stability, flexibility, and dynamic interactions of colliding RNA polymerase II elongation complexes. Mol. Cell *35*, 191–205.

Santangelo, T.J., and Reeve, J.N. (2006). Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. J. Mol. Biol. *355*, 196–210.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nat. Rev. Mol. Cell Biol. *7*, 557–567.

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.

Sigurdsson, S., Dirac-Svejstrup, A.B., and Svejstrup, J.Q. (2010). Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. Mol. Cell *38*, 202–210.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J. *22*, 2234–2244.

Stepanova, E., Wang, M., Severinov, K., and Borukhov, S. (2009). Early transcriptional arrest at Escherichia coli rplN and ompX promoters. J. Biol. Chem. *284*, 35702–35713.

Sydow, J.F., Brueckner, F., Cheung, A.C., Damsma, G.E., Dengl, S., Lehmann, E., Vassylyev, D., and Cramer, P. (2009). Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. Mol. Cell *34*, 710–721.

Tadigotla, V.R., O Maoiléidigh, D., Sengupta, A.M., Epshtein, V., Ebright, R.H., Nudler, E., and Ruckenstein, A.E. (2006). Thermodynamic and kinetic modeling of transcriptional pausing. Proc. Natl. Acad. Sci. USA *103*, 4439–4444.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., and Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. Nature *468*, 978–982.

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.

Thomas, M.J., Platas, A.A., and Hawley, D.K. (1998). Transcriptional fidelity and proofreading by RNA polymerase II. Cell *93*, 627–637.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007). Structural basis for substrate loading in bacterial RNA polymerase. Nature *448*, 163–168.

Walter, W., Kireeva, M.L., Studitsky, V.M., and Kashlev, M. (2003). Bacterial polymerase and yeast polymerase II use similar mechanisms for transcription through nucleosomes. J. Biol. Chem. *278*, 36148–36156.

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 *282*, 902–907.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell *127*, 941–954.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., and Kornberg, R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science *324*, 1203–1206.

Washburn, R.S., and Gottesman, M.E. (2011). Transcription termination maintains chromosome integrity. Proc. Natl. Acad. Sci. USA *108*, 792–797.

Werner, M., Thuriaux, P., and Soutourina, J. (2009). Structure-function analysis of RNA polymerases I and III. Curr. Opin. Struct. Biol. *19*, 740–745.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat. Genet. *39*, 1512–1516.