Cell, Vol. 89, 33–41, April 4, 1997, Copyright ©1997 by Cell Press

The RNA–DNA Hybrid Maintains the Register of Transcription by Preventing Backtracking of RNA Polymerase

Evgeny Nudler,* Arkady Mustaev,*
Evgeny Lukhtanov,† and Alex Goldfarb*
*Public Health Research Institute
New York, New York 10016
†Epoch Pharmaceuticals, Inc.
Bothell, Washington 98021

Summary

An 8–9 bp RNA–DNA hybrid in the transcription elongation complex is essential for keeping the RNA 3′ terminus engaged with the active site of E. coli RNA polymerase (RNAP). Destabilization of the hybrid leads to detachment of the transcript terminus, RNAP backtracking, and shifting of the hybrid upstream. Eventually, the exposed 3′ segment of RNA can be removed through transcript cleavage. At certain sites, cycles of unwinding–rewinding of the hybrid are coupled to reverse–forward sliding of the transcription elongation complex. This explains apparent discontinuous elongation, which was previously interpreted as contraction and expansion of an RNAP molecule (inchworming). Thus, the 3′-proximal RNA–DNA hybrid plays the dual role of keeping the active site in register with the template and sensing the helix-destabilizing mismatches in RNA, launching correction through backtracking and cleavage.

Introduction

The basic features of the ternary elongation complex (TEC) are conserved in all living organisms (Sweetser et al., 1987; Puhler et al., 1989; Darst et al., 1991; Polyakov et al., 1995), but its detailed structure is unknown. Over the past five years, models of elongation offered two alternative views on the extent and functional significance of the RNA–DNA hybrid in the transcription bubble. In the classic view, the hybrid is 12 bp long and plays a principal role in holding TEC together (Farnham and Platt, 1980; Gamper and Hearst, 1982; Hanna and Meares, 1983; Yager and von Hippel, 1986, 1991). In the revisionist view, the hybrid is 2–3 bp long and serves no other purpose than templating (Rice et al., 1991; Chamberlin, 1995). The resolution of this dilemma, which has been the subject of a lively debate (Kainz and Roberts, 1992; Altman et al., 1994; Johnson and Chamberlin, 1994; Zaychikov et al., 1995), has far-reaching implications for RNA polymerase (RNAP) structure and regulation at the level of elongation and termination (reviewed by Das, 1993; Chan and Landick, 1994; Platt, 1997; Richardson and Greenblatt, 1996). In this work, we employed RNA–DNA cross-links to measure directly the length of the hybrid and used nucleotide analogs that strengthen or weaken the hybrid to assess its functional role in TEC.

Results

The Eight to Nine Base Pair Hybrid Is Detected by RNA–DNA Cross-Linking

To determine the length of the hybrid by RNA–DNA cross-linking, we used an analog of UTP, U•, that carried the aromatic bis(2-iodoethyl)amino group attached to the fifth position of pyrimidine through a spacer arm. Normally, this group is not reactive, but upon addition of $NaBH_4$, a reactive aziridinium intermediate is formed that can cross-link to DNA or protein (Figure 1).

The extended length of the probe's arm was ~12 Å. A flexible arm of this length can in principle reach multiple sites, including the nontemplate strand of DNA. However, a U•-DNA cross-link in TEC mapped by piperidin cleavage went exclusively into a single complementary adenine in the template strand (E. N., unpublished data), which is indicative of base pairing. In the base-paired configuration, because of the need to bend across the heteroduplex major grove, the length of the arm is just enough to reach the adenine base paired to U• or an adjacent purine. In the experiments described below, the U• probe was incorporated in two different positions of the transcript. In one case, there were no purines adjacent to the adenine in the template DNA strand. In the other case, there was an adjacent guanine, but its involvement in the cross-link was excluded by piperidine mapping (E. N., unpublished data). The reactive moiety of the U• probe could also rotate away from the template DNA and cross-link to the protein.

The cross-linking of U• was analyzed in a series of consecutive TECs obtained by using immobilized RNAP (Kashlev et al., 1993; Nudler et al., 1994). This solid-state transcription system permits walking of RNAP along DNA in discrete controlled steps. RNA transcript was ^{32}P-labeled near its 5′ terminus, so that the cross-linking could be visualized by the appearance of protein–[^{32}P]RNA or DNA–[^{32}P]RNA species on the autoradiogram of PAGE slabs after the complex has been denatured.

The U• probe was incorporated either in position +21 or in position +45 (Figure 1B), and the complex was walked from +21 to +44 (Figure 1B, lanes 3–10) or from +45 to +62 (Figure 1B, lanes 12–18), respectively. In the nomenclature used, positive numbers refer to the sequence position relative to the transcription start site, whereas negative numbers indicate position in relation to the RNA 3′ terminus.

The results demonstrate that the RNA–DNA cross-link is readily formed in those TECs where U• is located from −2 to −8, relative to the RNA 3′ terminus. At the −10 position and beyond, the cross-linking to DNA is dramatically decreased. Thus, the RNA–DNA hybrid in TEC extends for the distance of 8–9 nt behind the growing tip of RNA. It should be noted that RNA–protein cross-linking displays a pattern opposite to RNA–DNA. The yield of the protein cross-link increases dramatically at −10 and beyond, suggesting that RNA–protein association becomes particularly tight upstream of the RNA–DNA hybrid region.

Unwinding of the 3′-Proximal RNA–DNA Hybrid during Transcriptional Arrest

TEC stalled at +27 (TEC27) gradually loses ability to resume elongation without releasing RNA (Markovtsov et al., 1996). This phenomenon is illustrated by the fact

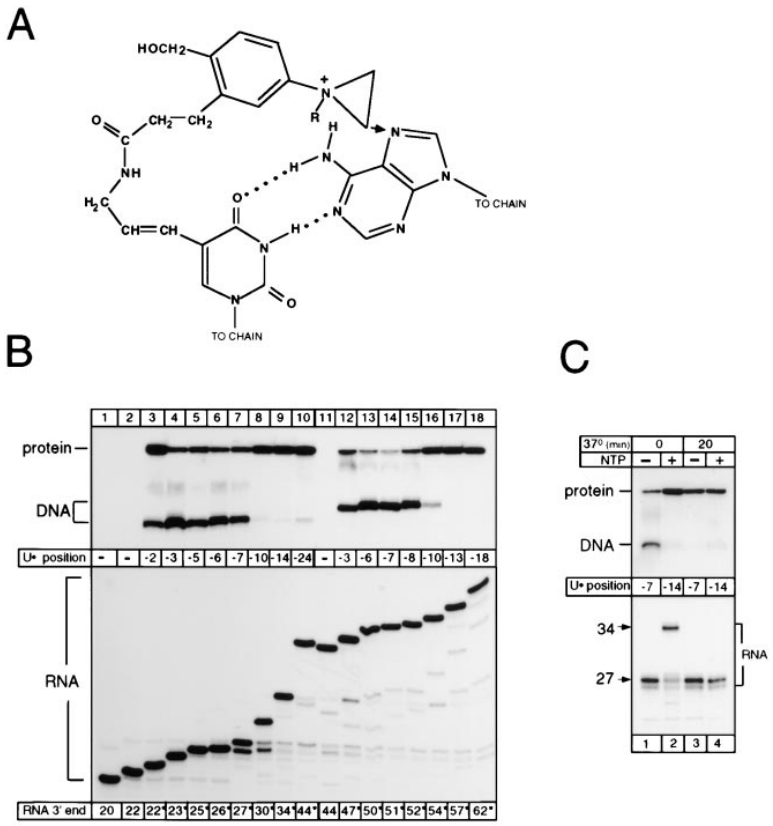

Figure 1. RNA–DNA Cross-Linking in TECs

(A) The U•-A base pair. The arrow symbolizes the attack of aziridinium intermediate at N8 of adenine, leading to a cross-link.

(B) Cross-linking in the walking TEC. The autoradiogram shows free [^{32}P]RNA transcripts (bottom panel) and cross-linked protein–[^{32}P]RNA or DNA–[^{32}P]RNA products from TEC walked to the positions indicated by the RNA 3′ end. The U• probe was incorporated in positions +20 (lanes 3–10) or +45 (lanes 12–18) as indicated by negative numbers showing distance from the 3′ end. The bottom part of the gel was underexposed to compensate for low yield of the cross-linked species. The length of transcript is designated as "RNA 3′ end"; an asterisk denotes the presence of the U• probe. The incorporation of U• resulted in slight retardation of the transcript in the gel (compare RNA 22 with 22* and 44* with 44). The doublet transcript observed at +26 and +27 (lanes 7 and 8) reflects the fact that these positions are particularly prone to spontaneous elongation arrest.

(C) Effect of arrest in TEC27 on cross-linking. Free RNA (bottom) and cross-linked species (top) were obtained before and after 20 min incubation, with or without NTP chase as indicated.

that the addition of NTP substrates to TEC27 (NTP chase) results in ~90% RNA extension to +34 at 0 min but has no effect after 20 min of TEC incubation at 37°C (Figure 1C). The arrest is accompanied by the loss of RNA–DNA cross-linking by the U• probe at +21 (−7 and −14 in TEC27 and TEC34, respectively), suggesting that the hybrid unwinds during arrest. To explore further the relationship between arrest and unwinding of the hybrid, the kinetics of formation of the arrested complex were determined in TEC27, in which the hybrid was either strengthened or weakened by incorporation of ribonucleotide analogs (Figure 2A). When RNA in the hybrid contained standard nucleotides (Figure 2A, lanes 1–5), ~50% of the complex was arrested after 1 min incubation at 37°C and ~65% after 3 min. Substitution of CMP at −3 and −4 with 5-iodoCMP (iC), an analog that stabilizes the duplex, resulted in a dramatic decrease in the rate of arrest, yielding ~15% and ~25% of arrested complex at 1 and 3 min, respectively (Figure 2A, lanes 6–10). Placement of another duplex-stabilizing analog, 5-bromo UMP (brU) at −7 had a similar albeit less profound effect, apparently because only one position in the duplex was substituted (Figure 2A, lanes 11–15). In contrast, helix-destabilizing substitutions, such as inosine (I) at −5 or 4-thiouridine (sU) at −7, dramatically increased the rate of arrest (Figure 2A, lanes 16–20 and 21–25, respectively). However, when inosine was placed at −10, i.e., upstream of the presumed RNA–DNA hybrid, the rate of arrest slightly decreased (Figure 2A, lanes 26–30). In general, incorporation of analogs upstream of −8 had the opposite effect as compared to the same substitutions in the 3′-proximal region (unpublished data). The linkage between stability of the 3′-terminal RNA–DNA hybrid and transcriptional arrest indicates that the hybrid is essential for maintaining the active state of TEC.

The RNA–DNA Hybrid Is Needed to Keep the RNA 3′ Terminus in the Active Site

We examined the arrest phenomenon in relation to the RNA–DNA hybrid at three known arrest sites that occur in the transcription template used, i.e., TEC27 (Figure 2), TEC56 (Figure 3), and TEC80 (Figure 4). Each of these sites has a particular characteristic: the arrest at +27 is irreversible, but arrest at +56 and +80 is transient (E. N., unpublished data), whereas TEC80 carries a 3′-proximal oligo(T) track that is the generic arrest signal in prokaryotes and eukaryotes (Kerppola and Kane, 1991; Gu and Reines, 1995; Nudler et al., 1995).

According to the current hypothesis, transcriptional arrest involves disengagement of the RNAP-active site from the 3′ terminus of RNA. This view is based on the argument that transcript cleavage in the arrested complexes, which in E. coli is effected by the GreB factor, is actually performed by the active site that has moved to an internal position within the transcript (Izban and Luse, 1992, 1993; Borukhov et al., 1993; Nudler et al., 1994; Reines, 1994; Rudd et al., 1994; Chamberlin, 1995; Orlova et al., 1995). To test the disengagement hypothesis directly, we took advantage of the recent observation that the Fe^{2+} ion substituting for Mg^{2+} in the RNAP active site in the binary promoter complex

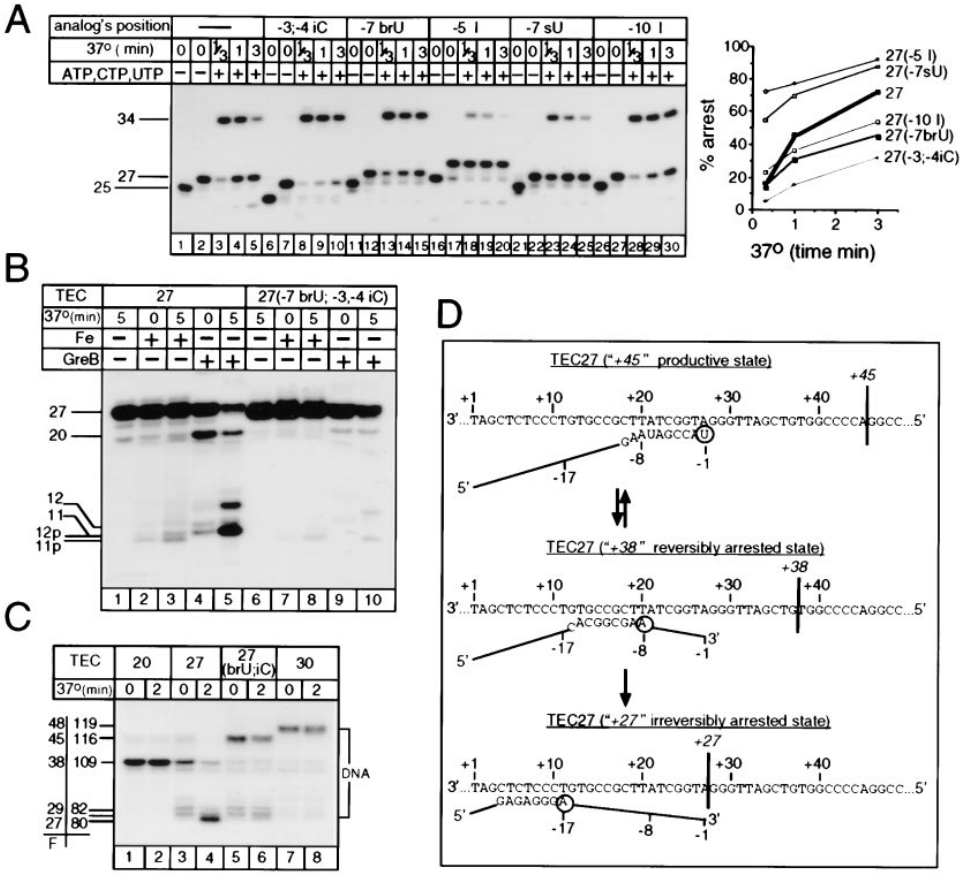

Figure 2. Elongation Arrest in TEC27 as a Function of RNA–DNA Hybrid Stability

(A) Effect of hybrid stability on arrest kinetics. TEC25 carrying [^{32}P]RNA and, where indicated, substituted ribonucleotide analogs was walked to +27 and then chased to +34 after incubation for indicated time periods. Note the effect of analogs on RNA bend migration. Failure of RNA to extend from +27 to +34 reflects accumulation of arrested complex. The graphs on the right show quantitation of the data; the bold curve represents TEC27 carrying natural nucleotides.

(B) Fe^{2+}- and GreB-induced RNA cleavage. TEC27 carrying [^{32}P]RNA was prepared with standard substrates (lanes 1–5) or NTP analogs (lanes 6–10) and exposed to Fe^{2+} and GreB before or after 5 min incubation. The faster mobility of the Fe^{2+} cleavage products is due to 3′-terminal phosphate.

(C) Front-edge mapping. The autoradiogram shows protection of ^{32}P-labeled template DNA fragment from degradation by ExoIII in TECs identified. On the table, the left column shows the position of the front edge (F) deduced from the actual size of protected DNA fragment, which is indicated on the right.

(D) Schematic representation of the three states of TEC27. The vertical bar represents the front edge; the circle represents the active site.

causes highly localized cleavage of DNA near the transcription start site. The cleavage occurs through local generation of free hydroxyl radicals (Zaychikov et al., 1996). We reasoned that in TEC, the postulated repositioning of the active site would be reflected by an upstream shift of Fe^{2+}-induced localized cleavage of RNA. We also determined how GreB- and Fe^{2+}-induced cleavage respond to changes in the hybrid stability.

In the arrested TEC27, substitution of Mg^{2+} with Fe^{2+} led to cleavage of the transcript at +11/+12, which was also the major site of GreB cleavage. Strengthening of the 3′ proximal hybrid with nucleotide analogs brU and iC suppressed cleavage both by Fe^{2+} and GreB (Figure 2B). In TEC56, GreB cleavage strictly correlated with the stability of the hybrid and with the extent of arrest (Figure 3A). In TEC80, strengthening of the oligo(U:A) hybrid with brU inhibited cleavage by both GreB and Fe^{2+}, while weakening of the hybrid with sU increased the GreB cleavage (Figure 4A). These results provide strong support for the disengagement model and demonstrate that the RNA–DNA hybrid is essential for keeping the RNA 3′ terminus in the polymerase active site.

Is the unwinding of the 3′-proximal hybrid accompanied by rewinding of the RNA–DNA hybrid in the upstream region? To answer this question, the cross-linking of the U• probe at +45 was determined in TEC56 before and after arrest (Figure 3B). In the productive TEC56, the cross-linking was small, apparently because the probe was located outside of the presumed hybrid (−12) (Figure 3B, lanes 3 and 4). In the arrested TEC56, however, cross-linking at −12 became prominent (Figure 3B, lanes 5 and 6), indicating that the hybrid in this position had reformed. These experiments suggest that elongation arrest and repositioning of the active site are associated with shifting of the RNA–DNA hybrid upstream from the 3′-proximal region.

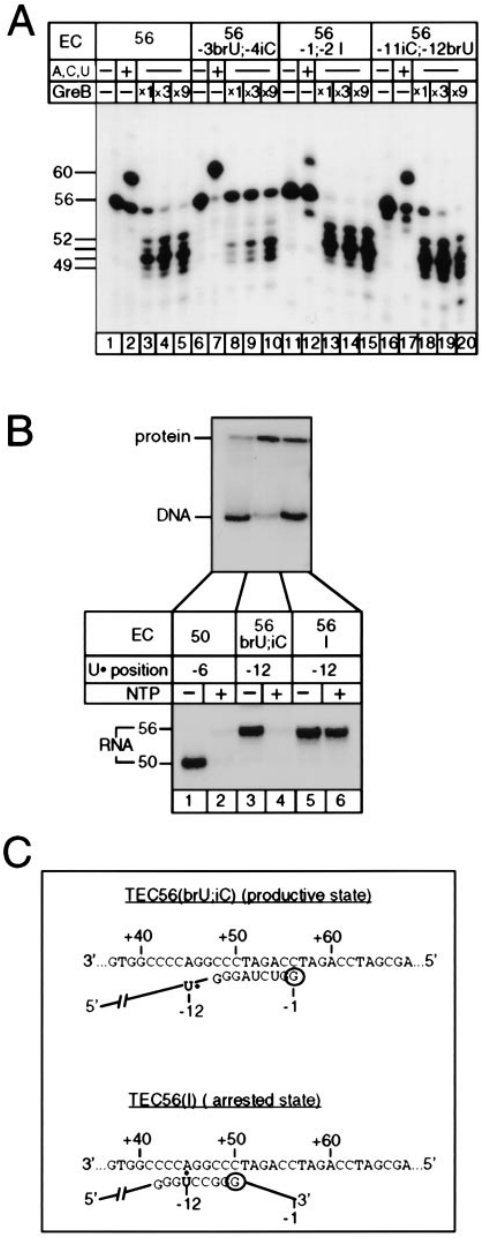

Figure 3. Elongation Arrest in TEC56

(A) Effect of ribonucleotide analog substitution on arrest rate and GreB-induced transcript cleavage. TEC56 carrying [^{32}P]RNA was prepared with standard rNTP (lane 1) or with the indicated substitutions (lanes 6, 11, and 16) and was challenged with the mixture of NTP (A, C, U-chase; lanes 2, 7, 12, and 17) for 2 min to determine the extension of RNA in productive complexes from +56 to +60. Cleavage of RNA was determined by exposing each complex to three increasing doses of GreB.

(B) Cross-linking of U• in position +45 in productive and arrested TEC56. The bottom panel shows [^{32}P]RNA in the TEC50 control and in the productive (lanes 3 and 4) and arrested (lanes 5 and 6) TEC56 before and after NTP chase. The upper panel shows respective cross-linked protein–[^{32}P]RNA or DNA–[^{32}P]RNA products. The productive TEC56 (lanes 3 and 4) carried −3BrU and −4iC, and the arrested TEC56 (lanes 5 and 6) carried −1 and −2 I substitutions, respectively.

(C) Schematic representation of the cross-linking experiment (B). Backward repositioning of the hybrid during TEC56 arrest allows the RNA–DNA cross-link.

Discontinuity of Elongation Is Suppressed by Hybrid Strengthening

In our previous work (Nudler et al., 1994, 1995), it was shown that DNA sites that induce arrest also cause characteristic irregularities of the TEC footprints. This result was interpreted as evidence of contraction followed by saltatory expansion of RNAP protein. The three characteristic features of the contracted, or strained, TEC were as follows: (i) predisposition to arrest; (ii) sensitivity to GreB-induced cleavage; and (iii) shortened distance between the 3′ end of the transcript and the front boundary of protection of DNA from degradation with exonuclease III (Nudler et al., 1994). The latter parameter was thought to reflect the distance between the active site and the front edge, or C∼F distance, in the RNAP molecule. The apparent variability of the C∼F value at certain DNA sites was the principal basis for the inchworming metaphor used to describe the movement of RNAP.

Since predisposition to arrest and sensitivity to GreB were directly linked with the stability of the 3′-proximal RNA–DNA hybrid, it seemed important to assess the relation of the hybrid to the C∼F value. This question was particularly intriguing because both the presumed contracted arrest-prone TEC27 and relaxed arrest-proof TEC52 contained an RNA–DNA hybrid of about the same length (Figure 1B), whereas the apparent C∼F value in the two complexes had been determined to be 11 and 18, respectively (Nudler et al., 1994).

In the experiment of Figure 2C, the front edge of freshly stopped TEC27 (Figure 2C, lane 3) maps at +38 (apparent C∼F = 11). Irreversible arrest of TEC27 was accompanied by backward translocation of the front edge to about +27 (apparent C∼F = 0; Figure 2C, lane 4), reflecting a long-distance reverse translocation of the whole RNAP (Komissarova and Kashlev, 1997a). Surprisingly, incorporation of hybrid-stabilizing brU and iC, the substitutions that suppress arrest (Figure 2A), led to forward translocation of the front edge to +45, yielding apparent C∼F = 18, which is the characteristic value of normal relaxed TEC. To interpret this observation, one has to assume that TEC27 exists in three alternative states, +45, +38, and +27, corresponding to the three observed positions of the front edge (Figure 2D). Of the three states, only +45 is truly elongation competent, because it contains the intact 3′-terminal hybrid and the active site attached to the 3′ terminus. The +45 and the +38 states are in equilibrium associated with unwinding–rewinding of the hybrid, whereas the conversion of +38 into +27 is irreversible. The +38 and the +27 locations of the front edge apparently correspond to relocation of the active site to +20 (−8) and + 11 (−17), respectively, as revealed by the Fe^{2+} and GreB cleavage pattern (Figure 2B). Thus, in each of the three states, the actual C∼F always equals 18, even though the apparent distance between the RNA 3′ terminus and the front edge is variable. Our earlier conclusions (Nudler et al., 1994) were in error because we equated the position of the active site with the 3′ terminus, which is not the case in complexes where the active site has relocated to an internal position in RNA.

These conclusions were confirmed by analysis of TEC80, another arrest-prone complex, containing 3′-proximal oligo(T) track (Nudler et al., 1995). In most of the

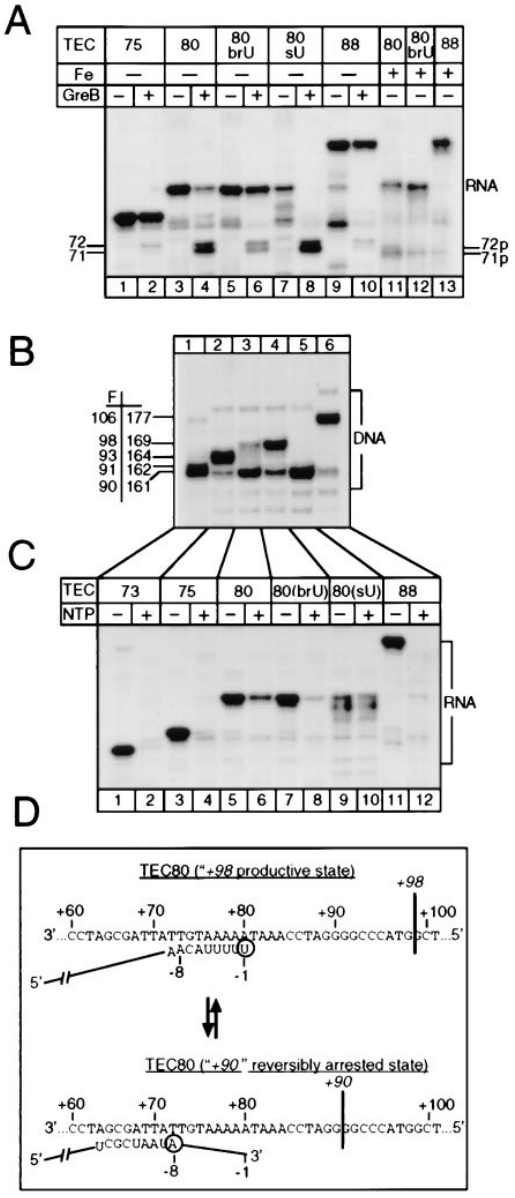

Figure 4. Elongation Arrest in TEC80

(A) Effect of brU and sU substitutions in positions +76 to +80 on Fe^{2+}- (lanes 11 and 12) and GreB-induced (lanes 3–8) cleavage of [^{32}P]RNA transcript in TEC80. Control experiments with TEC75 (lanes 1 and 2), and TEC88 (lanes 9, 10, and 13) are presented.

(B) Front-edge mapping in TEC73 (lane 1), TEC75 (lane 2), TEC80 at different analog substitutions (lanes 3, 4, and 5), and TEC88 (lane 6). The autoradiogram shows protection of terminally labeled nontemplate DNA strand from ExoIII degradation. For details, see legend to Figure 2C.

(C) Effect of brU and sU substitutions on elongation arrest. Autoradiogram shows [^{32}P]RNA in various complexes before and after NTP chase.

(D) Schematic representation of the two states of TEC80.

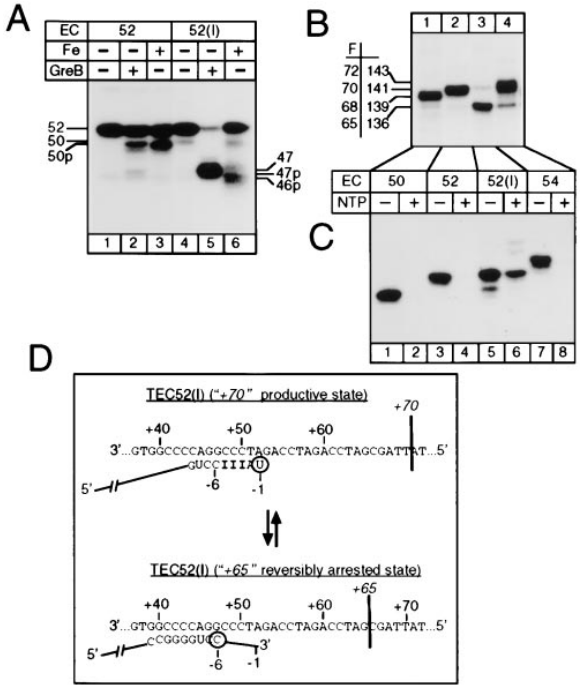

Figure 5. Induction of Discontinuous Elongation at +52 with Inosine

(A) Fe^{2+}- and GreB-induced cleavage of [^{32}P]RNA transcript in TEC52 carrying guanine (lanes 1–3) or inosine (lanes 4–6) at +48, +49, and +50.

(B and C) Front-edge mapping by ExoIII and transcriptional arrest. For details, see legends to Figures 2B and 2C.

(D) Schematic representation of the two states of inosine containing TEC52.

complex, the front edge maps at +90, with apparent C∼F = 10 (Figure 4B). Only a small fraction of the complex displayed the front edge at +98 (C∼F = 18). Stabilization of the hybrid with brU shifted the front edge in most of TEC80 to +98, while weakening of the hybrid with sU completely removed the +98 signal. Thus, the +90 and +98 states of TEC80 are interconvertible and are associated with unwinding–rewinding of the hybrid. Only the +98 state is elongation proficient, while the +90 state is reversibly arrested (Figure 4C).

The above experiments explain sequence-specific irregularities of ExoIII footprints as the consequence of the 3′-proximal hybrid melting. It can be expected, then, that destabilization of the 3′-proximal RNA–DNA hybrid at any site would lead to front-edge backtracking. To test this prediction, we replaced three successive RNA guanines at +48, +49, and +50 with inosine and determined the parameters of TEC52, which normally displays no footprint irregularities (Nudler et al., 1994). As was shown above, this complex contains a 3′-terminal RNA–DNA hybrid of at least 8 bp (Figure 1B).

As can be seen from Figure 5, incorporation of inosine into RNA leads to backtracking of the active site by ∼6 nt, as revealed by GreB- and Fe^{2+}-induced cleavage (Figure 5A). This is accompanied by retardation of the front edge for about the same distance, from +70 to +65 (Figure 5B), and elongation arrest (Figure 5C), which is reversible because most of the backtracked complex can be chased by prolonged incubation with NTP (data not shown). In other words, at a randomly chosen sequence site, artificial weakening of the hybrid caused TEC to enter into an apparently discontinuous arrest-prone phase of elongation.

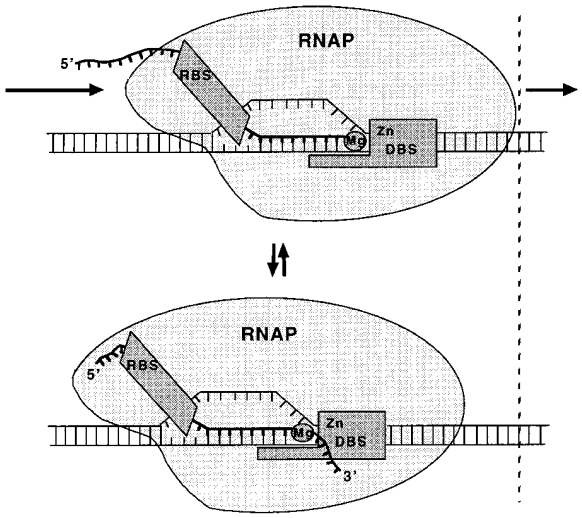

Figure 6. Model of TEC

Schematic representations of a productive and backtracked TEC are shown. Mg^{2+} denotes the active site. DBS symbolizes the front-end DNA-binding site responsible for processivity. RBS or RNA-binding site was denoted the RNA "tight binding site" (TBS) in previous models. The heteroduplex area is equivalent to RNA "loose binding site" (LBS) in previous models (Nudler et al., 1994, 1995).

Discussion

The principal results of this work are the conclusive proof of the existence of an 8–9 bp RNA–DNA hybrid in TEC and the demonstration of its functional role in keeping the complex in register with the transcribed sequence. It should be noted that the structure of the hybrid may be different from a true helix because of interaction with protein in TEC, e.g., it may be bent or partially unwound. Our results agree with the early data of Hanna and Meares (1983), who detected an RNA–DNA hybrid in ternary initiation complexes on the promoter.

Because of the central position of the hybrid in TEC, the uncertainty about its very existence was the major stumbling block to our understanding of RNAP structure and function. Resolution of the hybrid debate thus provides the basis for more definitive model building. We wish to propose a model of TEC (Figure 6) that explains elongation, pausing, and termination in structure-functional terms. In presenting the model, the role of the RNA–DNA hybrid is discussed below in the context of the following three groups of data: (i) irregularities of TEC footprints and other evidence suggesting internal plasticity of RNAP (inchworming); (ii) instances of TEC moving backward, occasionally causing arrest of elongation; and (iii) the mechanism of processivity that, as we showed previously (Nudler et al., 1996), does not involve the hybrid.

RNAP Is Not an Inchworm

The central notion of the discontinuous (a.k.a. inchworm) model of elongation was that advancing RNAP undergoes conformational transitions that are not synchronous with single-step nucleotide additions, i.e., TEC moves nonmonotonically (Chamberlin, 1995). Mechanistically, this was envisioned as contraction and expansion of advancing RNAP. Originally, it was proposed that inchworming is intrinsic to elongation, i.e., it constitutes the very mechanism of TEC advancement (Krummel and Chamberlin, 1992; Chamberlin, 1995). Subsequently, we showed that TEC advances mostly monotonically, with the exception of incidental situations when specific sites in DNA induce irregularities of ExoIII footprints (Nudler et al., 1994, 1995).

Recently, Komissarova and Kashlev (1997b) demonstrated reversible loss of catalytic activity in TEC at the inchworming sites. In such complexes, RNAP translocated backward with time after halting, as shown by fast DNA footprinting. Antisense oligonucleotides hybridized upstream of RNAP restored the initial position of the enzyme. On this basis, transient backtracking of RNAP was suggested as an alternative to inchworming. Similar suggestion has been made for yeast pol II by Reeder and Hawley (1996).

Here, we directly demonstrate that TEC backtracks as the result of weakness of the 3′-proximal RNA–DNA hybrid. The RNAP molecule that has backtracked appears to be in the same overall conformation as that in the productive elongation-competent complex, including the constant distance between the active site and the front boundary of ExoIII protection (the C∼F parameter). Moreover, in the backtracked complex, the RNA–DNA hybrid appears to have shifted to an upstream region. This conclusion is based on the observation of an RNA–DNA cross-link reemerging in the upstream region in a backtracked complex. Because of technical constraints, the reemerging cross-link could be demonstrated only for one site. The notion of hybrid shifting, however, is supported by our preliminary result that nucleotide substitutions in the upstream region affect backtracking in the direction opposite to the same substitutions in the 3′-proximal region.

A principal conclusion from these studies is that elongation-competent productive TEC is always monotonic, while apparently discontinuous advancement reflects a reversible side pathway rather than a succession of elongation intermediates (Figure 6). The irregular DNA and RNA footprints that have been observed by us and others (Krummel and Chamberlin, 1992; Nudler et al., 1994, 1995; Wang et al., 1995; Zaychikov et al., 1995; Gu et al., 1996; Samkurashvili and Luse, 1996) will have to be reinterpreted as reflections of mixed populations of TEC alternating between the productive and backtracked states. Similarly, reversible unwinding–rewinding and hybrid shifting may explain variable values of the length of the hybrid and partial protection from chemical degradation reported earlier (Kainz and Roberts, 1992; Zaychikov et al., 1995).

It should be noted that repudiation of the inchworming mechanism of elongation does not invalidate the notion of RNAP plasticity. There is compelling evidence that during the early steps of RNA chain synthesis, when the RNA–DNA hybrid is first formed, the active center domain "stretches" to accommodate the nascent RNA chain of 8–9 nt. This evidence was obtained with RNA primers secured at the 5′ face of the active center

through a cross-link (Mustaev et al., 1993) or the Rif-anchor (Mustaev et al., 1994), which excludes any possibility of RNA sliding through the protein. In view of this evidence, our present results only demonstrate that the stretched-out conformation of the active center, once established during initiation, is monotonically preserved throughout elongation and is probably maintained during backtracking. However, after transcription terminates and TEC dissociates, the active center must return to its original "compressed" state.

The RNA–DNA Hybrid and the Active Site: The Zip-Lock Analogy

In our model, during backtracking the hybrid unwinds in the 3′ proximal region and shifts upstream. We propose that channels that guide DNA and RNA chains through RNAP (Darst et al., 1991; Polyakov et al., 1995) are shaped to accommodate the RNA–DNA heteroduplex just upstream of the active site. In effect, the active site is "glued" to the heteroduplex like a zip-lock is to the locked section of the zipper. When the heteroduplex unwinds and shifts upstream, it pulls back the active site together with the whole RNAP.

What triggers and then determines the distance of backtracking? Backtracking signals are A–U rich. However, the stability of the 3′-terminal heteroduplex is not everything to it, since there are instances of A-tracks in the template that do not induce backtracking (E. N., unpublished data). Sequences behind and ahead of the heteroduplex area were reported to affect elongation arrest (Krummel and Chamberlin, 1992; Wiest et al., 1992; Nudler et al., 1994, 1995). Since backtracking seems to involve associated rewinding of the hybrid upstream, as well as unwinding–rewinding of DNA duplexes ahead and behind of a transcription bubble, the specific parameters of the event must depend on the overall sequence context. Thus, in some cases backtracking may be short and reversible (as in TEC80), or it may go too far and never come back (as in TEC27).

Backtracking may be resolved through internal RNA cleavage and resumption of elongation from the newly generated 3′ terminus. It is easy to imagine that mismatched ribonucleotides that destabilize the 3′ proximal hybrid would facilitate backtracking. Thus, in addition to keeping the active site in register with the template, our model proposes a second principal role for the hybrid, namely sensing mismatches in RNA and then launching corrective action (Erie et al., 1993; Jeon and Agarwal, 1996).

Segregation of Processivity and Fidelity Functions

In combination with our previous work (Nudler et al., 1996), these results demonstrate that separate sets of interactions are responsible for the two principal biochemical features of TEC, i.e., processivity and fidelity. The operational distinction between the two types of interactions is that the former serves to counter dissociation of TEC, whereas the latter serves to counter its wanton sliding along DNA.

Thus, in our hypothesis processivity is achieved by the front-end DNA-binding site (DBS) that locks around the double-helical DNA ahead of the bubble (Nudler et al., 1996). The locking is induced allosterically by RNA when it enters the RNA-binding site (RBS) behind the heteroduplex. Interactions in DBS and RBS prevent TEC dissociation (Nudler et al., 1996) but permit threading of DNA and RNA through RNAP, i.e., lateral translocation of TEC.

Fidelity, in our hypothesis, is achieved by heteroduplex interactions upstream of the active site, as this work shows. While these interactions are weak and insufficient for stably holding the three components of TEC together, they serve the principal role of maintaining the register of TEC. When these interactions are disrupted, TEC slides back and forth even though it does not dissociate.

Of course, the two types of interactions may cross-influence each other. Since backtracking must involve threading of RNA and DNA through RBS and DBS, respectively, it is not inconceivable that "friction" in these sites may affect backtracking. Conversely, weak protein–nucleic acid interactions in the heteroduplex area are capable of holding the ternary complex together, even when the strong processivity interactions in DBS and RBS are destroyed. Such a complex is nonprocessive and highly unstable (Nudler et al., 1996).

The notion of separation of processivity and fidelity functions has been discussed by Guajardo and Sousa (1997), who attribute stability of TEC to structural devices (thumbs, locks, etc.) and explain positioning of RNAP on the template through local free energy minima. In their model, the key role in keeping RNAP in register is played by the incoming substrate. Our views are in general agreement with theirs, except that we give the RNA–DNA hybrid the primary role in setting the free-energy equillibrium that keeps TEC in register.

Implications for Pausing and Termination

From the proposed model of TEC, simple mechanisms of pausing and termination follow. Backtracking of RNAP induced by an unstable RNA–DNA hybrid is a likely cause of a pause. Suppression of the elongation rate by inosine in eukaryotic (Matsuzaki et al., 1994) and prokaryotic (E. N., unpublished data) systems supports this notion. Since the processivity interactions in the backtracked complex are intact, such a pause would not lead to termination. Thus, there is no mechanistic difference between a paused and an arrested complex, except that the latter cannot restart because of a particularly unfavorable sequence context. The factors that accelerate elongation, such as bacteriophage antitermination factors N and Q (Yang and Roberts, 1989; Mason et al., 1992; Das, 1993), E. coli NusG (Burova et al., 1995), or eukaryotic elongin (Aso et al., 1995), may act through suppressing RNAP backtracking.

Termination must involve disruption of the processivity interactions. This is likely to occur by an allosteric mechanism triggered by detachment of RNA from RBS, e.g., through the formation of a hairpin. The resulting complex would be held together only by weak fidelity interactions in the heteroduplex area. We speculate that such a complex is a principal intermediate in the pathway leading to termination.

In conclusion, it should be emphasized that our model

combines elements from both of the initial conflicting views on TEC (see Introduction). The hybrid is clearly there, in agreement with the classical model (Yager and von Hippel, 1986, 1991), but its primary role is not holding together the components of TEC. This function is achieved through protein–nucleic acid interactions outside of the hybrid area, in accord with the revisionist view (Chamberlin, 1995).

Experimental Procedures

The T7 A1 promoter-containing template (226 bp) was obtained by polymerase chain reaction (PCR) from Template 2 (Nudler et al., 1995). Its transcribed sequence is ATCGAGAGGG ACACGGCGAA TAGCCATCCC AATCGACACC GGGGTCCGGG ATCTGGATCT GGA TCGCTAA TAACATTTTT ATTTGGATCC CCGGGTACCG AGCTCGA ATT CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAACCC TGGCG.

His-tagged RNAP was purified and immobilized on Ni-NTA-agarose as described (Nudler et al., 1996). The preparation of a start-up ternary complex with a 12 nt ^{32}P-labeled transcript and the walking reaction were performed as described (Nudler et al., 1996).

The cross-linking reagent U• is an alkylating derivative of UTP containing 3-[3′-(N,N-bis-2-iodoethyl)amino-4′-formylphenyl]propionate moiety attached to the 5 position of the pyrimidine ring through an aminoallyl spacer. The complete synthesis of U• will be published elsewhere. Upon exposure of U• to $NaBH_4$, the aromatic bis(2-iodoethyl)amino group of U• is converted to a reactive aziridinium intermediate shown in Figure 1A. The reaction proceeds through reduction of the aldehyde group of U•, which in turn activates the 2-iodoethylamino group (Gall et al., 1979).

In a typical cross-linking reaction, TEC stalled at appropriate position was washed 3 times with the buffer containing 20 mM HEPES (pH 7.25), 100 mM KCl, 2 mM $MnCl_2$. The U• probe was added to 20 μl of reaction mixture to the final concentration of 40 μM. The complex was incubated for 30 min at 20°C followed by the walking reaction in the buffer (20 mM HEPES [pH 7.25], 100 mM KCl, 10 mM $MgCl_2$). To activate the cross-link, freshly prepared solution of $NaBH_4$ (4 mg/ml, Aldrich) was added to 0.4 mg/ml for 20 min at 20°C. The reaction was stopped by adding 2 vol of the loading mix (10 mM EDTA, 12 M Urea and BPB), heated at 80°C for 1 min, and the cross-linked species were resolved on a 12% PAGE (19:1 acrylamide:bis-acrylamide, 7M urea, 0.5× TBE).

Incorporation of substrate analogs during the walking reaction (Nudler et al., 1996) was performed at 20°C for 10 min. The analog concentrations were as follows: inosine triphosphate (I; Boehringer Manheim), 20 μM; 5-bromouridine-5′-triphosphate (brU; Sigma), 10 μM; 5-iodocytidine-5′-triphosphate (iC; Sigma), 20 μM; sU (synthesized from 4-thio-UMP [Sigma] as described [Hoard and Ott, 1965]), 40 μM.

GreB cleavage and ExoIII footprinting were performed as described previously (Nudler et al., 1994), except that the time of incubation with ExoIII was 2–3 min.

In experiments with Fe^{2+}-mediated transcript cleavage, TEC was washed 6 times with the buffer containing 20 mM HEPES (pH 7.25), 50 mM KCl, followed by addition of 40 μM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ (Aldrich) and 10 mM DTT for 20 min at 20°C. The reaction was stopped and cleavage products resolved on a 12% or 6% PAGE as described above. It should be noted that Fe^{2+} cleavage of RNA is a low yield reaction that occurs concomitantly with nonspecific degradation of RNAP protein. As a consequence, the specific RNA cleavage reaction rapidly reaches a plateau and can be used only in qualitative assays.

Acknowledgments

We are grateful to E. Avetissova for His-tagged RNAP, S. Borukhov for GreB protein, M. Kozlov for advice on cross-linking techniques, and N. Komissarova and M. Kashlev, who were the first to suggest RNAP backtracking and shared their ideas. This work was supported by NIH grant GM49242.

Received December 13, 1996; revised February 18, 1997.

References

Altman, C.R., Solow-Cordero, D.E., and Chamberlin, M.J. (1994). RNA cleavage and chain elongation by Escherichia coli DNA-dependent RNA polymerase in a binary enzyme RNA complex. Proc. Natl. Acad. Sci. USA *91*, 3784–3788.

Aso, T., Lane, W.S., Conaway, J.W., and Conaway, R.C. (1995). Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II. Science *269*, 1439–1443.

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

Burova, E., Hung, S.C., Sagitov, V., Stitt, B.L., and Gottesman, M.E. (1995). Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. J. Bacteriol. *177*, 1388–1392.

Chamberlin, M.J. (1995). New models for the mechanism of transcription elongation and its regulation. Harvey Lect. *88*, 1–21.

Chan, C.L., and Landick, R. (1994). New perspectives on RNA chain elongation and termination by E. coli RNA polymerase. In Transcription: Mechanisms and Regulation, R.C. Conaway and J.W. Conaway, eds. (New York: Raven Press, Ltd.), pp. 297–321.

Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991). Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. Cell *66*, 121–128.

Das, A. (1993). Control of transcription termination by RNA-binding proteins. Annu. Rev. Biochem. *62*, 893–930.

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.

Farnham, P.J., and Platt, T. (1980). A model for transcription termination suggested by studies on the *trp* attenuator in vitro using base analogs. Cell *20*, 739–748.

Gall, A.A., Kurbatov, V.A., Mustaev, A.A., and Shishkin, G.V. (1979). Reagents for directed modification of biopolymers. Aromatic 2-chlortoaethylamines with controlled reactivity. Izv. Sib. Otd. Acad. Nauk SSSR Ser. Khim. Nauk *4*, 99–104.

Gamper, H.B., and Hearst, J.E. (1982). A topological model for transcription based on unwinding angle analysis of E. coli RNA polymerase binary, initiation and ternary complexes. Cell *29*, 81–90.

Gu, W., and Reines, D. (1995). Identification of a decay in transcription potential that results in elongation factor dependence of RNA polymerase II. J. Biol. Chem. *270*, 11238–11244.

Gu, W., Wind, M., and Reines, D. (1996). Increased accommodation of nascent RNA in a product site on RNA polymerase II during arrest. Proc. Natl. Acad. Sci. USA *93*, 6935–6940.

Guajardo, R., and Sousa, R. (1997). A model for the mechanism of polymerase translocation. J. Mol. Biol. *265*, 8–19.

Hanna, M.M., and Meares, C.F. (1983). Topography of transcription path: the leading end of nascent RNA through the Escherichia coli transcription complex. Proc. Natl. Acad. Sci. USA *80*, 4238–4242.

Hoard, D.E., and Ott, D.G. (1965). Conversion of mono- and oligodeoxyribonucleotides to 5′-triphosphates. J. Am. Chem. Soc. *87*, 1785–1789.

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

Izban, M.G., and Luse, D.S. (1993). The increment of SII-facilitated transcript cleavage varies dramatically between elongation competent and incompetent RNA polymerase II ternary complexes. J. Biol. Chem. *268*, 12874–12885.

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

Johnson, T.L., and Chamberlin, M.J. (1994). Complexes of yeast RNA polymerase II and RNA are substrates for TFIIS-induced RNA cleavage. Cell *77*, 217–224.

Kainz, M., and Roberts, J. (1992). Structure of transcription elongation complex in vivo. Science *255*, 838–841.

Kashlev, M., Martin, E., Polyakov, A., Severinov, K., Nikiforov, V., and

Goldfarb, A. (1993). Histidine-tagged RNA polymerase: dissection of the transcription cycle using immobilized enzyme. Gene *130*, 9–14.

Kerppola, T.K., and Kane, C.M. (1991). RNA polymerase: regulation of transcript elongation and termination. FASEB J. *5*, 2833–2842.

Komisarrova, N., and Kashlev, M. (1997a). Arrest of transcription: E. coli RNA polymerase translocates backward leaving the 3′ end of the RNA intact and extruded. Proc. Natl. Acad. Sci. USA *94*, in press.

Komisarrova, N., and Kashlev, M. (1997b). RNA polymerase switches between inactivated and activated states translocating back and forth along the DNA and the RNA. J. Biol. Chem. *272*, in press.

Krummel, B., and Chamberlin, M.J. (1992). Structural analysis of ternary complexes of Escherichia coli RNA polymerase. Individual complexes halted along different transcription units have distinct and unexpected biochemical properties. J. Mol. Biol. *225*, 221–234.

Markovtsov, V., Mustaev, A., and Goldfarb, A. (1996). Protein–RNA interactions in the active center of transcription elongation complex. Proc. Natl. Acad. Sci. USA *93*, 3221–3226.

Mason, S.W., Li, J., and Greenblatt, J. (1992). Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage λ. J. Biol. Chem. *267*, 19418–19426.

Matsuzaki, H., Kassavetis, G.A., and Geiduschek, E.P. (1994). Analysis of RNA chain elongation and termination by Saccharomyces cerevisiae RNA polymerase III. J. Mol. Biol. *235*, 1173–1192.

Mustaev, A., Kashlev, M., Zaychikov, E., Grachev, M., and Goldfarb, A. (1993). Active center rearrangement in RNA polymerase initiation complex. J. Biol. Chem. *268*, 19185–19187.

Mustaev, A., Zaychikov, E., Severinov, K., Kashlev, M., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1994). Topology of the RNA polymerase active center probed by chimeric rifampicin-nucleotide compounds. Proc. Natl. Acad. Sci. USA *91*, 12036–12040.

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

Nudler, E., Kashlev, M., Nikiforov, V., and Goldfarb, A. (1995). Coupling between transcription termination and RNA polymerase inchworming. Cell *81*, 351–357.

Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: RNA polymerase–DNA interactions holding together the elongation complex. Science *273*, 211–217.

Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. Proc. Natl. Acad. Sci. USA *92*, 4596–4600.

Platt, T. (1997). RNA structure in transcription elongation, termination, and antitermination. In RNA Structure and Function, R. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), in press.

Polyakov, A., Severinova, E., and Darst, S.A. (1995). Three-dimensional structure of E. coli core RNA polymerase: promoter binding and elongation conformations of the enzyme. Cell *83*, 365–373.

Puhler, G., Leffers, H., Gropp, F., Palm, P., Klenk, H.P., Lottspeich, F., Garrett, R.A., and Zillig, W. (1989). Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. Proc. Natl. Acad. Sci. USA *86*, 4569–4573.

Reeder, T.C., and Hawley, D.K. (1996). Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. Cell *87*, 767–777.

Reines, D. (1994). Transcript elongation and gene regulation in eukaryotes. In Transcription Mechanisms and Regulation, R.C. Conaway and J.W. Conaway, eds. (New York: Raven Press, Ltd.), pp. 263–278.

Rice, G.A., Kane, C.M., and Chamberlin, M.J. (1991). Footprinting analysis of mammalian RNA polymerase II along its transcript: an alternative view of transcription elongation. Proc. Natl. Acad. Sci. USA *88*, 4245–4249.

Richardson, J.P., and Greenblatt, J. (1996). Control of RNA chain elongation and termination. In Escherichia coli and Salmonella typhimurium, 2nd Ed., F.C. Neidhardt, ed. (Washington, D.C.: American Society for Microbiology), pp. 822–848.

Rudd, M.D., Izban, M.G., and Luse, D.S. (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. Proc. Natl. Acad. Sci. USA *91*, 8057–8061.

Samkurashvili, I., and Luse, D.S. (1996). Translocation and transcriptional arrest during transcript elongation by RNA polymerase II. J. Biol. Chem. *271*, 23495–23505.

Sweetser, D., Nonet, M., and Young, R.A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. USA *84*, 1192–1196.

Wang, D., Meier, T.I., Chan, C.L., Feng, G., Lee, D.N., and Landick, R. (1995). Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. Cell *81*, 341–350.

Wiest, D.K., Wang, D., and Hawley, D.K. (1992). Mechanistic studies of transcriptional arrest at the adenovirus majour late attenuation site. J. Biol. Chem. *267*, 7733–7744.

Yager, T.D., and von Hippel, P.H. (1987). Transcript elongation and termination in Escherichia coli. In Escherichia coli and Salmonella typhimurium, F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger, eds. (Washington, D.C.: American Society for Microbiology), pp. 1241–1275.

Yager, T.D., and von Hippel, P.H. (1991). A thermodynamic analysis of RNA transcript elongation and termination in E. coli. Biochemistry *30*, 1097–1118.

Yang, X., and Roberts, J.W. (1989). Gene Q antitermination protein of Escherichia coli phages 82 and λ suppress pausing by RNA polymerase at a ρ-dependent terminator and other sites. Proc. Natl. Acad. Sci. USA *86*, 5301–5305.

Zaychikov, E., Denissova, L., and Heumann, H. (1995). Translocation of the Escherichia coli transcription complex observed in the registers 11 to 20; "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble." Proc. Natl. Acad. Sci. USA *92*, 1739–1743.

Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A., and Mustaev, A. (1996). Mapping of catalytic residues in the RNA polymerase active center. Science *273*, 107–109.