RNA polymerase II transcription complexes may become arrested if the nascent RNA is shortened to less than 50 nucleotides

Andrea Újvári, Mahadeb Pal and Donal S. Luse†

Department of Molecular Biology
Lerner Research Institute
Cleveland Clinic Foundation
Cleveland, OH 44195

† to whom correspondence should be addressed, at:
Department of Molecular Biology, Mail Location NC20
Cleveland Clinic Foundation
9500 Euclid Ave.
Cleveland, OH 44195
Phone: 216-445-7688; FAX: 216-444-0512; email: lused@ccf.org
Running title: RNA polymerase II arrest as a function of transcript length
A significant fraction of RNA polymerase II transcription complexes become arrested when halted within a particular initially transcribed region after the synthesis of 23 to 32 nucleotide RNAs. If polymerases are halted within the same sequence at a promoter distal location, they remain elongation competent. However, when the RNAs within these promoter distal complexes are truncated to between 21 and 48 nucleotides, many of the polymerases become arrested. The degree of the arrest correlates very well with the length of the RNA in both the promoter proximal and distal complexes. This effect is also observed when comparing promoter proximal and promoter distal complexes halted over a completely different sequence. The unusual propensity of many promoter-proximal RNA polymerase II complexes to arrest may therefore be recreated in promoter distal complexes simply by shortening the nascent RNA. Thus, the transition to full elongation competence by RNA polymerase II is dependent on the synthesis of about 50 nt of RNA and this transition is reversible. We also found that arrest is facilitated in promoter distal complexes by the hybridization of oligonucleotides to the transcript between 30 and 45 bases upstream of the 3' end.
There is increasing evidence for eukaryotic gene regulation during the process of transcript elongation (reviewed in 1). It is thus important to understand in detail the elements which affect the efficiency of elongation by RNA polymerase II transcription complexes. Both biochemical (reviewed in 2) and structural studies of prokaryotic and eukaryotic RNA polymerases support a particular model of the transcript elongation complex (3-7). According to this model, ternary complexes are stabilized primarily by the presence of an 8-9 bp RNA-DNA hybrid just upstream of the catalytic center of the RNA polymerase and by a domain of the polymerase that serves as a sliding clamp, encircling the DNA just downstream of the active site and preventing dissociation of the template from the complex.

This model explains certain examples of sequence-specific loss of elongation competence, or arrest, by multisubunit RNA polymerases. For example, RNA polymerase II arrests primarily at the bold-faced Ts while traversing the T_{\text{la}} sequence (non-template strand) TTTTTTTCCCTTTTTTT from the histone H3.3 gene (8). Complexes with U-rich 3' ends contain the weakest possible RNA-DNA hybrid (U:dA). In response to this situation the polymerase can translocate upstream, carrying the transcription bubble to a location where a more stable RNA-DNA hybrid forms. Indeed, an upstream translocated footprint was reported for polymerases arrested at the histone H3.3 T_{\text{la}} site (9,10). Upstream translocation removes the 3’ end of the RNA from the active site and thus prevents transcription from continuing. Arrested complexes can be restarted by cleavage of the transcript such that the 3’ RNA end is again aligned with the enzyme catalytic center. This cleavage is catalyzed by the active site of the
polymerase and is highly stimulated by the transcription factor SII for RNA polymerase II and by GreA and GreB in the case of *E. coli* RNA polymerase (11-15).

Kerppola and Kane (8) have shown that even though arrest at the histone H3.3 T1a site occurs within the upstream run of nontemplate strand Ts, the second, downstream T-run is also required for efficient arrest. The RNA polymerase must therefore recognize downstream DNA sequences as arrest signals at the DNA level. Consistent with this, upstream translocation of RNA polymerases stalled on different DNA templates was shown to be affected by DNA sequences downstream of the point of stalling (10).

The model just discussed has provided a major advance in our understanding of transcript elongation and arrest at promoter-distal DNA sites. However, the model fails to explain the properties of newly-initiated transcription complexes. Temporary stalling of RNA polymerase II complexes in the region of +20 to +25 at one particular promoter leaves complexes strongly upstream translocated but nevertheless fully or predominantly elongation competent (10,16). Similar observations were made by Nudler et al. (17) and Komissarova and Kashlev (18) who have reported that *E. coli* RNA polymerase promoter proximal complexes also translocate upstream. These results are significant, since in vivo pausing of RNA polymerase II in the same region, after the synthesis of a 20-40 nt transcript, has been shown to be important for the regulation of expression of many genes, including hsp70, c-myc and c-fos (19-24).

Luse and Samkurashvili (25) argued that upstream translocation of promoter-proximal RNA polymerase II is not driven by weak RNA-DNA hybrids or downstream
DNA elements, since initially transcribed sequences within which upstream translocation takes place present no barrier to polymerase translocation at promoter-distal locations. Complexes with transcripts longer than about 45 nt no longer showed upstream translocation (26), so transcript length seemed to be an important determinant of lateral stability on the template. However, the earlier study could not separate effects based simply on transcript length from those which also depended upon proximity to promoter sequences.

We have now investigated directly the role of transcript length on the elongation competence of transcription complexes, independent of the presence of the promoter. Through the use of carefully controlled transcript cleavage procedures, we have prepared matched sets of transcription complexes with identical transcript length and nearly identical transcript sequence at promoter proximal and promoter distal locations. The corresponding complexes display very similar levels of transcriptional arrest. Thus, the nascent RNA itself, well upstream (i.e., 20-50 nt) of the 3’ end, facilitates continued transcript elongation by the ternary complex.

MATERIALS AND METHODS

Reagents - FPLC purified NTPs and RNaseH (Cat.#27-0894-02) were purchased from Pharmacia, $^{32}$P-labeled NTPs from NEN, Bio-Gel A1.5m was from Bio-Rad, Deep Vent DNA polymerase from NEB and ribonuclease inhibitor and restriction endonucleases were from GIBCO-BRL. Recombinant human elongation factor SII was purified as described previously (27). Chimeric 2'-O-methyl-RNA/DNA oligonucleotides,
synthesized by Oligos Etc., had the following sequences: chimera 1: 5'-
dCdGdCdGGmGmUmGmCmCm-3', chimera 2: 5'-dCdCdTdGCmCmGmCmGmGm-3',
chimera 3: 5'-dGdCdGdTcUmGmCmUmCmUm-3', chimera 4: 5'-
dCdCdCdGGmGmUmUmUmAm-3', chimera 5: 5'-dGdCdCdTCmGmUmCmGmCmCm-3'.

Plasmids- The promoters in all plasmids are based on the adenovirus major late
promoter. The construction of pML20-23 and the series of pML20-23like plasmids was
described previously (16). Plasmid pML20-23like3, which was used in this study, has
the following sequence on the nontemplate strand starting at position +1:
ACAGGAAGAGGAAGAAGCAGGCCT. The last six bases form a recognition site for the
restriction enzyme StuI; downstream of the StuI site the pML20-23like plasmids have
the same sequence as the pML20-42 construct (26). The sequence of pML20-23like3
downstream of position +2 is identical to the sequence of pML20-23 starting with
position +130 (see Fig. 1). pML20-55M was constructed by replacing the fragment
between StuI (+25) and XhoI (+39) on pML20-42 with a 66 bp synthetic fragment of the
following sequence (nontemplate strand):
5'CCTCGGCTGCGTGCGCCGTCGGGCGCGCACCTCTCTCCCCTTCTTTAAAGG
CCTCGGCGCGGC-3'

Template Preparation- Plasmid DNA was either purified by cesium chloride
centrifugation or by a Qiagen Maxi Plasmid Kit. Closed circular DNA was used as the
DNA template in experiments with the plasmid pML20-23. In the case of the pML20-
23like3 plasmid the template was produced in a PCR reaction that amplifies a 190 bp
fragment with the transcription start site 96 bp from the 5' end. DNA was purified with the Concert Rapid PCR purification System (Gibco-BRL) and then cut downstream of the promoter with the restriction enzyme HindIII to yield a 178 bp template. The DNA was then phenol/chloroform extracted and ethanol precipitated. Bead-attached templates for the experiment in Fig. 8 were made by PCR using the pML20-55M plasmid and the methods described in (16).

Assembly and Purification of Ternary Transcription Complexes- Preinitiation complexes were formed in a reaction containing 18 \( \mu \)g/ml circular plasmid or 10 \( \mu \)g/ml PCR fragment with HeLa cell nuclear extract and were purified through a Bio-Gel A1.5m column to remove contaminating NTPs as described (10). Ternary complexes on the template pML20-23 containing 151 nt or longer RNAs were synthesized in multiple steps. Complexes stalled at position +20 were generated at 30°C for 5 min with 1 mM ApC as the initiating dinucleotide, 20 \( \mu \)M each of UTP, CTP, GTP and 50 \( \mu \)M dATP as the energy source. After addition of the detergent sarkosyl to 1% complexes were gel filtered (sarkosyl rinsing- see 10). Sarkosyl rinsed U20 complexes were elongated through a U-free cassette to make C151 complexes. If C151 complexes were analyzed directly, they were generated from U20 with of 600 \( \mu \)M ATP, 100 \( \mu \)M CTP, 9 \( \mu \)M GTP and 1 \( \mu \)M \( \alpha \)-\(^32\)P-GTP for 10 min at 37°C, followed by addition of nonlabeled GTP to 600 \( \mu \)M with further incubation for 5 min at 37°C. If complexes containing RNAs longer than 151 nt were required, C151 was synthesized without radiolabel for 10 min at 37°C. The C151 complexes were gel filtered to remove NTPs. U154, G157 or A160 were then generated by the addition of 1.5 \( \mu \)M \( \alpha \)-\(^32\)P-UTP for 5 min at 37°C followed by 50 \( \mu \)M
nonlabeled UTP, or UTP and GTP, or UTP, GTP and ATP, for 5 min at 37°C (see Fig. 1 for the sequences).

Ternary complexes (G17) stalled at position +17 on the pML20-23like3 template were synthesized by incubating preinitiation complex with 1 mM ApC, 20 μM ATP, 50 μM dATP and 1 μM α-32P-GTP for 5 min at 30°C, followed by incubation for 5 min at 30°C with nonlabeled 20 μM GTP and sarkosyl rinsing. Similarly, C23 was made from preinitiation complex with 1 mM ApC, 20 μM each of ATP and GTP, 50 μM dATP and 1 μM α-32P-CTP for 5 min at 30°C followed by the addition of 20 μM nonlabeled CTP, 5 min of further incubation at 30°C, and sarkosyl rinsing. G17 complexes were elongated to C18 and C23 complexes to U26, G29 or A32 by the addition of the appropriate NTPs at 50 μM and incubation for 5 min at 25°C.

The experiment in Fig. 8 used bead-attached templates generated by PCR from the pML20-55M template. Primers for generating the PCR fragments, preparation of the bead attached templates and assembly of preinitiation complexes on these templates were as described (16). Complexes were advanced to the desired positions as given in the Fig. 8 legend.

Site specific cleavage of RNA and chase of ternary complexes—2′-O-methyl-RNA/DNA chimeras were added to ternary complexes stalled on pML20-23 at a final concentration of 0.1 to 10 μM, and were incubated in the presence of a final concentration of 0.03 U/μl RNaseH and 0.05 U/μl ribonuclease inhibitor for 10 min at 37°C. This yields the site specific cleavage of the RNA in the complex at the 5′ end of
the chimera (28,29). We believe that the faint bands which were occasionally seen above the main products (see for example lanes 2, 6, 10 of Fig. 2) are not due to variations in the location of RNAse H digestion but rather to low levels of read-through RNA in the original transcription reaction (note for example the presence of 152-mers in the C151 preparation in ref. 10). This explanation is consistent with our results with U154 complexes, which were generated from C151 with low levels of UTP and which do not appear to have any read-through products. In that case, RNAse H cleavage with the chimeric oligonucleotides gave essentially a single band (see Fig. 3). Transcript truncation for the experiment in Fig. 8 was performed using RNase T1 digestion as described in the Fig. 8 legend.

Complexes treated with RNase H were tested for transcriptional competence by the addition of 50 μM NTPs at 37°C as described in detail in the figure legends. No chimeras or RNase H were added to the complexes synthesized on pML20-23like3 but these were also incubated at 37°C for 10 min to mimic conditions of the RNAseH cleavage experiment. These complexes were then immediately chased by further incubation with 50 μM NTPs at 37°C for 5 min. When indicated a final concentration of 18 μg/ml of SII (or 29 μg/ml, for the experiment in Fig. 8) was included in the chase reactions.

It should be noted that C17 and G18 complexes showed significant transcript truncation, in the absence of added SII, during a 5 min incubation at 25°C (compare lanes 2 and 3 with lane 1 in Fig. 4). Truncation was more pronounced during mock-chase incubations at 37°C (compare lanes 2 and 3 with lanes 4 and 7).
cleavage involved RNA polymerase, since it was completely inhibited by $\alpha$-amanitin (data not shown) and was essentially confined to the C17 and G18 complexes. We did not observe breakdown with early elongation complexes walked to position +23 or later (Fig. 4), nor did we see this effect with promoter distal complexes whose transcripts had been trimmed with RNase H to 18 nt or any other length (Figs 2 and 3). We believe that the source of the cleavage is a very low level of SII from the nuclear extract that was not removed by the sarkosyl rinsing, since the addition of sarkosyl to the C17 or G18 complexes prevented RNA breakdown without inhibiting the elongation of the transcripts upon chase (data not shown). However, we cannot exclude the possibility that very early elongation complexes differ from later complexes in a way that renders their transcripts exceptionally labile to cleavage. If residual SII did remain in the promoter proximal C17 and G18 complexes (Fig. 4), this cannot explain their very low level of arrest relative to, for example, C23 promoter proximal complexes since promoter distal C18 complexes generated by RNase H cleavage, which show no RNA breakdown upon incubation, are also fully elongation competent (Fig. 2).

It is important to note that after the sarkosyl washing step, transcript initiation factors are removed (see 26,30). Thus, once transcript cleavage was performed with oligos and RNase H, there was no need to remove these components during subsequent chase reactions since the transcript segments targeted by the oligos could not be regenerated.

All transcription reactions were phenol/chloroform extracted and (except for the Fig. 7 experiment) ethanol precipitated. The reactions in Fig. 7 were concentrated by
lyophilization and loaded directly to avoid loss of the short transcript cleavage products. Samples were resolved on denaturing polyacrylamide (19:1 acrylamide to bis-acrylamide) gels with the following percentages of acrylamide: Fig. 6, 6%; Figs. 2, 3 and 4, 13%; Figs. 7 and 8, 20%. Radioactivity was visualized on a PhosphorImager and quantitation was done with the program ImageQuant (Molecular Dynamics).

RESULTS

Samkurashvili and Luse (10) reported on the exonuclease III footprints of several RNA polymerase II complexes stalled far downstream from the adenovirus major late promoter. That study employed a template, called pML20-23, which was constructed in such a way that RNA polymerase could be advanced in only two steps to position +151. The sequence TTTGGGAAACCC on the nontemplate strand immediately downstream of position +151 allowed walking of the polymerase from C151 to U154, G157, or A160 by incubating with a subset of NTPs (see Fig. 1). All of the complexes stalled from +151 to +160 were found to be fully active in chase reactions (only about 5-10 % failed to restart transcription) and their exonuclease III footprints advanced synchronously along the template with the polymerase active site.

To study the same DNA sequence in a promoter-proximal context the region downstream of position +130 on plasmid pML20-23 was moved to the beginning of the initially transcribed region of the adenovirus major late promoter, creating the pML20-23like series of plasmids (16). In this case, the TTTGGGAAACCC sequence begins only 24 bases downstream of +1 (see Fig. 1). Surprisingly, the RNA polymerase II
complexes corresponding to the +151 to +160 complexes on pML20-23, namely C23, U26, G29 and A32, were all found to be severely arrested on this template (16; also see Fig. 4). In order to determine the basis for the difference in elongation competence of these complexes, we cleaved the RNAs in the promoter-distal complexes so that they matched, in length and sequence, the transcripts in the promoter-proximal complexes.

**Transcript Truncation in Promoter Distal Complexes** - In our initial pilot experiments we used RNase A to cut the transcripts in C151 complexes made on the pML20-23 plasmid. After the RNase A digestion a prominent 21 nt long RNA remained in the ternary complexes. We found that these 21-mer complexes were predominantly arrested (data not shown). Encouraged by this observation, we employed a procedure which permits sequence-specific cleavage of transcripts in ternary transcription complexes. This involved the use of RNase H in combination with different 2'-O-methyl-RNA/DNA chimeric oligonucleotides. When such chimeras containing four DNA residues at the 5' end are annealed to RNA and the hybrids are treated with RNase H, cleavage has been show to occur exclusively at the 5' end of the chimera (28; see the Materials and Methods section).

Figure 2 shows the results of an experiment in which chimeras 1, 2 and 3 (Fig. 1) were separately annealed to body labeled RNA in the C151 complex (uncut RNA is shown in lane 1). The conditions of the RNase H digestion were optimized so that more than 90% of the input RNA was cut in 10 min at 37°C. We also confirmed that this incubation without the addition of RNase H did not deactivate the C151 complexes (see Fig. 6A, lanes 1 and 2). Digestion with RNase H yielded primarily RNAs 18, 23 and 45
nt long (lanes 2-5, 6-9 and 10-13, respectively), with a faint additional band one nt longer (see Materials and Methods).

After cleavage, the complexes were immediately chased with 50 μM of UTP or UTP and GTP at 37°C for 5 min. At the end of the first chase a final 50 μM of GTP was added to lanes 4, 8 and 12 and all tubes were incubated for 5 more minutes. The amount of arrest was calculated as the percent of the complex that could not be elongated in the chase reaction. For example, 11% of the C18 band remained in lane 3 as compared to C18 in lane 2. We consider complex C18 fully active, since about the same fraction (14 %) of its parent complex, C151, cannot be chased after a 10 min incubation at 37°C without RNase H (Fig. 6A). In contrast to the C18 complex, complexes C23 and C45 were found to be arrested (73% and 53%, respectively) in this experiment. If the elongation factor SII was included in the chase reactions nearly all (> 95%) of the complexes could be elongated (data not shown; see Fig. 4).

The same experiment was repeated with complex U154 (Fig. 3). In this case the RNA was 3' end labeled and therefore only the 3' segment is detectable after the cleavage reaction. Note the disappearance of the 5' cleavage products from the top of the gel. We observed that the U21, U26 and U48 complexes were 56, 89 and 35% arrested, respectively. These results show that the promoter distal complexes that are fully elongation competent can be converted to complexes that are easily arrested, like their promoter proximal counterparts.

During these experiments we noticed that the extent of arrest for a given complex was clearly dependent on the incubation time. For example, if C18 was chased in the
presence of both UTP and GTP (lane 5 of Fig. 2) most complexes successfully 
elongated to G24. If only UTP was included in the reaction mixture during the first 5 
minutes all complexes advanced to C21, but only a small fraction of them could be 
chased further to G24 by the addition of GTP in a second 5 min incubation (lane 4). 
Thus, the half-life of the transcriptionally active form of the C21 complex is less than 5 
min. Similar results were seen with complexes U26 and U48 in Fig. 2 (compare lanes 8 
to 9, and 12 to 13) and with complexes U21 and U26 in Fig. 3 (compare lanes 4 to 5, 
and 8 to 9).

Complexes Containing RNAs Longer Than About 50 nt Are Fully Active- About 
35% of the initial U48 complex failed to advance to G51 in a 5 min chase (lanes 10 and 
11 of Fig. 3). However, when the G51 complexes were chased in a second 5 min 
reaction, all of them extended their RNAs to A54 (lane 12). To test longer RNA 
sequences we performed an experiment in which a 2'-O-methyl-RNA/DNA 
oligonucleotide was annealed from position 123 to 133 upstream from the 3' end of the 
151 nt RNA. After the RNase H cleavage the ternary complex containing a 122 nt RNA 
remained as active as a mock-cleaved control complex (data not shown).

To sample more complexes containing different lengths of RNA after the RNase 
H cleavage reaction we repeated the above experiments on complexes G157 and 
A160. The sequences of these RNAs are shown in Fig. 1 and results are summarized 
in Fig. 5. A clear trend could be discerned from these tests. If the RNA in the ternary 
complex after cleavage was longer than about 20 nt and shorter than about 50 nt the 
complex was prone to arrest. However, complexes with shorter or longer RNAs were
as elongation competent as noncleaved controls.

Promoter-proximal complexes- As noted above, other work in our laboratory has shown that RNA polymerase II complexes stalled from 20 to 32 nt downstream of transcription start on the pML20-23like template are arrested to a significant extent (16). In order to quantitatively compare the elongation competence of promoter proximal and promoter distal complexes in the pML20-23 sequence context, it was necessary to subject the promoter proximal complexes to the same extensive 37°C incubations used in RNase H digestion of the promoter distal complexes. Figure 4 shows a representative experiment of this type with the pML20-23like3 template. The first position on this template where the polymerase complex can be stalled by NTP limitation is +17. Complex G17 was synthesized and body labeled with radioactive GTP. After sarkosyl rinsing this complex could be advanced to C18 by the addition of CTP. The G17 and C18 complexes were further incubated at 37°C for 10 min (in the absence of RNase H) to treat them identically to the RNase H cleaved samples (lanes 4-9). A subsequent chase with all four NTPs was carried out in the presence or absence of the transcription elongation factor SII. We found that both G17 and C18 were fully active under these conditions, since only 2% of the complexes could not be chased (see lanes 5 and 8). However, we obtained a very different result when we assayed complexes with longer nascent RNAs. Complex C23 was synthesized with radioactive CTP, so that all of the label was incorporated in the 3'-most 6 nt of this RNA (see Fig. 1). After sarkosyl rinsing C23 was extended to U26, G29 or A32 by the addition of the corresponding nonlabeled NTPs at 25°C for 5 min. All complexes were
incubated at 37°C for 10 min and then chased with all four NTPs at 37°C for 5 min. All of the complexes with 23 to 32 nt RNAs were found to be severely arrested in this assay (lanes 11, 14, 17 and 20). Inclusion of SII in the chase reaction resulted in the disappearance of all of the 23-32mer RNAs and the appearance of labeled cleavage products of about 10-20 nt (lanes 12, 15, 18 and 21). Almost no labeled run-off RNA was made in these SII-containing chase reactions since all of the label in the initial ternary complex RNAs was released by transcript cleavage and subsequent elongation took place with nonlabeled NTPs. We did confirm with uniformly labeled C23-A32 RNA that the transcript which remained in ternary complex was fully extended to run-off during chase in the presence of SII (data not shown).

The data on the elongation competence of halted pML20-23like3 complexes are summarized in Figure 5. There was a very close correspondence in arrest levels for the promoter proximal and promoter distal complexes with identical transcript lengths. Those complexes with RNAs shorter than about 20 nt RNA were active, while complexes with 20 to 32 nt RNAs were predominantly arrested. We did not attempt to study promoter proximal complexes containing RNAs longer than 32 nt with the techniques used here, since all possible precursor complexes, from C23 to A32, are highly arrested.

*Hybridization of oligonucleotides to upstream segments of the transcript can lead to arrest*- Transcriptional arrest by RNA polymerase II, at least in promoter distal complexes, appears to be invariably accompanied by upstream translocation of the RNA polymerase. Thus, if the upstream segment of the transcript has an anti-arrest
effect, one might expect this effect to be based on the prevention of upstream translocation, for example by direct interaction with the RNA polymerase or by the formation of secondary structure. This model predicts that interfering with the blockade of upstream translocation could lead to arrest. To test this idea we hybridized oligonucleotides to the nascent RNA over segments ranging from 19 to 59 nt upstream of the point of bond formation. We reasoned that if interactions or structures important in arrest prevention must form over this region, oligo hybridization should compete with or prevent these interactions. For this test we used the C151 complexes and the chimeric oligos previously discussed, as well as an additional oligo, chimera 5, complementary to a segment 31 to 42 nt upstream of the 3' end of the RNA. The oligos were incubated with the complexes at 37°C for 10 min as before, but in the absence of RNase H. After the incubation, reactions were chased with excess NTPs. Hybridization with chimera 2 actually reduced arrest (Fig. 6A; compare lanes 1 to 2 and 3 to 4), as expected from earlier studies (31). However, incubation with chimeras 1, 5 and 3 caused partial arrest of complex C151 (37, 42 and 35 %, respectively). This arrest could be relieved by the addition of SII (data not shown). In order to demonstrate that all of the chimeras did hybridize to the transcript, we treated aliquots of the hybridization reactions with RNase H. In all cases the expected cleavage was observed (lanes 5, 8, 11 and 14).

We have repeated this experiment several times with both complexes C151 and U154. The results are summarized in Figure 6B. The percent arrest values were normalized to the background that is measured in the absence of the chimeras. Arrest
above background was seen with oligos hybridizing from 24 to 45 nt upstream of the 3' end. The greatest effect was observed with chimera 1 and the U154 complex, where the oligo hybrid extended from 27 to 36 nt upstream. Hybridization to locations more than 46 nt upstream did not significantly diminish elongation competence. This is consistent with our observation (Fig. 5) that complete elongation competence requires a transcript at least 51 nt long.

We hypothesized that arrest from oligo hybridization results from interfering with a mechanism that would otherwise block upstream translocation by the RNA polymerase. It was therefore important to demonstrate directly that the arrest seen in Fig. 6 did result from upstream translocation. In order to do this we produced U154 complexes labeled only in the last 3 nt and incubated these complexes with chimeric oligonucleotides 1, 2, 3 or 5 (or with no oligo, as control). The complexes were then treated with SII (Fig. 7). Mock-hybridized complexes (lanes 3 and 4), which were mostly elongation competent (see Fig. 6), gave predominantly short SII cleavage products as expected (32,33). The longer (about 6-13 nt) SII cleavage products in lane 4 presumably arose from the low level of arrested complexes. Hybridization with chimera 2, which did not reduce transcriptional competence in the experiment in Fig. 6, did not substantially change the partitioning of SII cleavage fragments as compared to the mock hybridization control. However, significant increases in the proportion of large SII cleavage products and coordinate diminution of short cleavage products was seen after hybridization with chimera 5 and particularly with chimera 1; the latter oligo was the most effective in inducing arrest in the Fig. 6 assay. Thus, the induction of arrest
through oligonucleotide hybridization is accompanied by upstream translocation by the RNA polymerase.

Transcript truncation causes reduced elongation competence in another, completely different sequence context. In order to eliminate the possibility that the effects we have observed are somehow specific to the purine-rich initially transcribed sequence of pML20-23like3, we have also repeated our basic experiment using a template with a completely different initially transcribed sequence. In our earlier work we studied adenovirus major late promoter-based templates, such as pML20-42, which have pyrimidine-rich initially transcribed regions. RNA polymerases halted in promoter-proximal locations on these templates generally show much lower levels of arrest than we observed for polymerases stalled in the promoter proximal region of the pML20-23like plasmids (16). A partial exception to this rule was provided by complexes stalled at +23 on pML20-42; about 40% of these polymerases could not extend their RNA chains after a 5 min chase under the conditions of our original study (16). We therefore constructed a new template, designated pML20-55M, which is diagramed in Fig. 8A. On this template RNA synthesis will proceed to +23 when preinitiation complexes are incubated with CpA, CTP, UTP and ATP. After sarkosyl rinsing the complexes may be advanced to +54 in an ATP-less reaction, rinsed under native conditions and then walked to +77 with CTP, ATP and UTP. Note that treatment of complexes stalled at +77 with RNase T1 will result in truncation of the RNAs in the complex to 24-mers, which are identical to the RNAs present in CpA-primed, promoter proximal complexes stalled at +23 (underlined sequences in Fig. 8A). When the initial, CpA-primed
transcription reaction on pML20-55M included radiolabeled CTP, we could observe the resulting A24 complexes (lanes 1-4 of Fig. 8B). Exposure of these complexes to mock T1 digestion (incubation at 37°C for 5 min) before chase resulted in a somewhat greater level of arrest than we had seen in our earlier study (16). In the example shown in Fig. 8B, 67% of the A24 complexes became arrested (compare lanes 1 and 2). When polymerases were walked to +77 on the pML20-55M template, 32P-CTP was only included in the final step so that only the final 23 nt were labeled. The resulting A78 complexes showed a much lower level of arrest after exposure to the mock T1 digestion at 37°C, as compared with the promoter proximal A24 complexes. For the experiment shown in Fig. 8B, 22% of the A78 complexes were arrested after mock digestion (compare lanes 9 and 10; note from lane 12 that all of the complexes resumed transcription when SII was added). Treatment of the A78 complexes with RNase T1 for 5 min at 37°C cleaved essentially all of the RNA to give the collection of bands shown in lane 5 of Fig. 8B. When these complexes were chased, 67% of the A24 RNAs could not advance in the absence of SII (lane 6), while all of the RNAs were elongated with SII (lane 8). Thus, as we observed in the experiments shown in Figs. 2 and 3, promoter distal complexes whose transcripts are truncated have the elongation competence of the corresponding promoter proximal complex bearing RNA of the same length and sequence.

**DISCUSSION**

RNA polymerase II complexes paused in promoter-proximal locations have a
strong tendency to translocate upstream (16,26), an event which in many cases results in transcriptional arrest. This is an unexpected finding because the promoter proximal sequences in question do not cause arrest when traversed by the polymerase far downstream of transcription start. In this paper we show directly that this unusual property of promoter-proximal RNA polymerase II complexes may be recreated simply by shortening the nascent RNA within complexes that had transcribed to a promoter-distal location. We conclude that the transition to full elongation competence by RNA polymerase II is dependent on the synthesis of about 50 nt of RNA. Interestingly, this transition appears to be reversible.

*Current Models for Transcript Elongation Complexes Do Not Predict Our Results*- A basic model for the transcript elongation complex assigns the major role in complex stability to two features: the RNA-DNA hybrid (most frequently thought to be about 8-9 bp long) and the interaction of a segment of the polymerase (the sliding clamp) with DNA immediately downstream of the point of bond formation (reviewed in 2,34). The large body of work which supports this model involves a number of different experimental approaches with both *E. coli* RNA polymerase and RNA polymerase II (31,35-40). While some of these studies have differed on the length of the RNA-DNA hybrid needed to confer maximal stability and the relative importance of the hybrid and the sliding clamp, the two basic features of this model explain most recent observations.

A major finding of the current work is that this hybrid-plus-sliding clamp model does not predict the behavior of RNA polymerase II complexes containing RNAs of less than 50 nt. It is important to stress that the sequence context we chose for most of our
study (all experiments but those in Fig. 8) had been shown to provide no barriers to transcription in a promoter distal location. RNA polymerases stalled within this region remained transcriptionally active and did not show upstream translocated footprints (10). One would therefore predict that polymerases halted in this local sequence context have neither unfavorable interactions with downstream DNA nor particularly weak RNA-DNA hybrids. In reference to the latter point, it is worth noting that many of the complexes we studied have transcripts with GC-rich 3’ ends. For example, the C23 complex (Fig. 4) and the corresponding C151 complex trimmed back to C23 (Fig. 2) were both severely arrested, and yet both complexes should contain RNA-DNA hybrids in which 6 of the 8 base pairs, including the four base pairs at the 3' end, are CG or GC (see Fig. 1).

Our results indicate that in addition to the RNA-DNA hybrid and the sliding clamp, far-upstream nascent RNA is also an important feature in stabilizing the transcript elongation complex along the DNA template. Many earlier studies have demonstrated that the multisubunit RNA polymerases protect from 14 to 18 nt of RNA upstream of the point of bond formation from nuclease attack (36,41,42). Using *E. coli* RNA polymerase and a synthetic bubble template, Wilson et al. showed that extension of the transcript from 12 to 20 nt is important in stabilizing the elongation complex (43). The crystal structure of the yeast RNA polymerase II ternary complex (7) reveals a likely RNA exit channel which could accommodate about 10 nt upstream of the 8-9 bp RNA-DNA hybrid. However, there are no data on the path of the RNA beyond about 20 nt upstream of the active site. Thus, none of these earlier results would have predicted
that transcription complexes that do not have a very weak RNA-DNA hybrid would be strongly affected by the transcript segment from 20 to 50 nt upstream.

A recently reported observation does address a potential functional role for far-upstream RNA sequences. Kireeva et al. (44) studied yeast RNA polymerase II transcription complexes which were directly assembled from purified polymerase, synthetic template and nontemplate strands and 9 nt RNA primers. RNA polymerase recognizes and extends the preformed RNA-DNA hybrid in this system. These complexes strongly resemble promoter initiated complexes by a number of assays. Significantly, halted yeast RNA polymerase II complexes bearing 9 or 40 nt transcripts were much more likely to resume transcription than complexes containing 20, 23 or 34 nt RNAs when challenged with 1 μM of the next NTP to be added. Thus, yeast RNA polymerase II also seems to have a transient phase early in transcript elongation during which halting the RNA polymerase leaves the enzyme in danger of not being able to resume transcription.

*The Transition by RNA Polymerase II from Initiation to Elongation* - We had previously suggested (16,25) that the acquisition of lateral stability on the template, between roughly positions +20 and +50, is the last step in the commitment of RNA polymerase II to the transcript elongation phase of RNA synthesis. Among the changes which such a transition could reflect are the complete closure of the clamp structure which constrains the template within the ternary complex (7) and the formation of a transcription bubble with the dimensions characteristic of the elongation-committed polymerase (45). The net effect of these events presumably locks the polymerase in
the stable elongation configuration. While such changes may indeed be characteristic of the initiation-elongation transition, the somewhat surprising result from the present study is that these changes are reversible. The response of the transcription complex to a particular length and sequence of transcript is essentially the same, regardless of whether that complex had previously passed through the initiation-elongation transition (Fig. 5).

A major unanswered question from our work concerns the molecular basis for the strong tendency of RNA polymerase II complexes with between 20 and 50 nt of nascent RNA to translocate upstream. What destabilizing effect makes the stabilizing influence of upstream RNA necessary? We speculate that once the nascent RNA reaches a critical length (which would be about 20 nt in the sequence context of the pML20-23like3 template), there is a transient negative interaction of the RNA with the polymerase. Translocating upstream to avoid this interaction would cause arrest.

**How Does Upstream RNA Function to Prevent Arrest?** It is well established that base pairing of the transcript either with itself or with DNA oligonucleotides can prevent reverse threading of the RNA and thus block arrest (31). Interference with secondary structure in the transcript could account for the results of our oligonucleotide hybridizations in Fig. 6. It is possible that promoter distal complexes such as U154 are normally stabilized against lateral movement on the template by transcript secondary structure extending up to the point at which the transcript emerges from the polymerase. Hybridization with, for example, chimera 1 would replace the polymerase-proximal secondary structure with a chimeric oligo-transcript hybrid extending from 27 to 36 nt.
upstream of the 3' end. This would leave 8-9 nt of unpaired RNA between the polymerase and the chimera hybrid (41), thus allowing reverse threading of the transcript and upstream translocation by the polymerase. As the hybrid between the oligonucleotide and the transcript is located further and further upstream, a point will be reached at which the RNA between the polymerase and the oligo hybrid can now reform secondary structure, either by pairing with itself or with far-upstream transcript sequences. Hybridization of oligos upstream of this point would no longer provoke significant arrest.

It is important to stress that the role of the transcript in preserving elongation competence must extend beyond the simple formation of secondary structure. This point can be most easily made by considering successive complexes in early elongation which do or do not display upstream translocation when stalled. For example, the A32 complex (Fig. 4) is predominantly arrested and upstream translocated (16); however, the C35 complex on the same template is neither upstream translocated nor arrested. When RNA polymerase reaches position +35, the sequence of the RNA which has emerged from the polymerase should be 5'ACAGGAAGAGGAAGAAGC (assuming that 17 nt upstream of the 3' end remain protected by the polymerase- see 41). There is no apparent opportunity for secondary structure to form within this RNA. Similarly, while the G25 complex on the pML20-42 template is upstream translocated, the C27 complex on the same template occupies the normal template position for elongation competent complexes (26). In the case of the C27 complex, the RNA external to the polymerase should have the sequence 5'ACUCUCUUCC; again, one would not predict any
secondary structure for this RNA.

Since the nascent transcript is probably not acting to stabilize early elongation complexes by forming secondary structure, a potential alternative mechanism would involve interaction of the RNA with some protein component of the transcription complex. In this context it is interesting to recall the results of Milan et al. (42) who observed that treatment of *E. coli* RNA polymerase complexes with low levels of RNase T1 revealed partial protection over the initial 15-16 nt upstream from the 3' RNA terminus and over a second region from 30 to 45 nt upstream. This finding, in conjunction with the reported high level of structural similarity between the bacterial and eukaryotic RNA polymerases (46,47), suggests the possibility that RNA polymerase II complexes might be stabilized by the direct interaction of the transcript and the polymerase at roughly 30-45 bases upstream of the 3' end of the RNA.

*The significance of the distinctive properties of RNA polymerase II complexes with 20-50 nt nascent RNAs-* Finally, it is important to recall that in those genes in which the transition to elongation is regulated, stalled RNA polymerases are typically found 20-50 nt downstream of transcription start; for example, between +21 and +35 on the *Drosophila* hsp70 gene (reviewed in 19) and at about +45 on the human hsp70 gene (23). We propose that the cell exploits the tendency of promoter proximal RNA polymerases to arrest in order to provide an opportunity for negatively acting transcription factors to interact with the enzyme. The concerted action of negatively (e.g. DSIF, NELF) and positively (e.g. P-TEFb, FACT) acting transcription factors could then regulate the transition into productive transcript elongation (48-51).
The abbreviations used are: nt, nucleotide. DSIF, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole sensitivity-inducing factor; NELF, negative elongation factor; P-TEF, positive transcript elongation factor; FACT, facilitates chromatin transcription.

Acknowledgments- We thank Robert Landick for stimulating discussions, and David McKean for advice during the course of these experiments. This work was supported by grant GM-29487 from the National Institutes of Health.
REFERENCES


   J.Biol.Chem. 275, 6530-6536
   14704
   Science 281, 424-428
   211-217
   6940
   1194


FIGURE LEGENDS

FIG. 1  RNA sequences used in this study. The transcript of the pML20-23 template from position +97 to +163 is shown in the upper panel. The positions to which three chimeric 2'-O-methyl-RNA/DNA oligonucleotides anneal to this RNA are shown. The dotted lines indicate the expected 5' ends which should be generated when hybrids of these oligos with C151-A160 RNAs are treated with RNase H. The lengths and sequences of the RNAs which should remain in ternary complex after cleavage are also shown. The lower panel shows the sequences of 17 to 32 nt transcripts generated on the pML20-23like3 plasmid by a combination of limited nucleotide additions and gel filtration. Sequences in the upper and lower panels are aligned to show identical sequence regions. Note in particular that the sequences of C23, U26, G29 and A32 transcripts made by direct transcription (lower panel) or transcript cleavage (upper panel) are identical except for the 5' nucleotide.

FIG. 2. Truncation of the transcript in C151 complexes can lead to arrest. Complex C151 was synthesized on plasmid pML20-23 (lane 1) and the body labeled RNA was cleaved by RNase H after annealing the indicated chimeric oligonucleotide as described in Materials and Methods. Chimeras 1 and 2 were used at a concentration of 3 μM. To eliminate hybridization of chimera 3 to alternative sites, it was added only to a final concentration of 0.1 μM. Ternary complexes containing 18 (lanes 2 to 5), 23 (lanes 6 to 9) and 45 (lanes 10 to 13) nt RNAs were initially chased with 50 μM of the indicated
NTPs for 5 min at 37°C. In lanes 4, 8, and 12, an additional 5 min chase with GTP at 50 μM was performed at 37°C; in all other lanes (except lane 1) the initial chase incubation was simply continued for 5 more minutes. The percentages of the complexes which arrested were calculated by comparing the initial complex to the fraction which remained after single nucleotide chase. For example, 11% of the 18-mer in lane 2 remained in lane 3.

FIG. 3. **Truncation of the transcript in U154 complexes can lead to arrest.** Complex U154 was synthesized on plasmid pML20-23 (lane 1) with only the last 3 nt labeled. RNA was cleaved by RNase H after annealing the indicated chimeric oligonucleotides as described in Materials and Methods. Oligonucleotides were used at the same concentrations as in the experiment in Fig. 2. Transcripts remaining in complex were initially chased with 50 μM of the indicated NTPs for 5 min at 37°C. In lanes 4, 8, and 12, an additional 5 min chase with ATP at 50 μM was performed at 37°C; in all other lanes (except lane 1) the initial chase incubation was simply continued for 5 more minutes. The percentages of the complexes which arrested were calculated by comparing the initial complex to the fraction which remained after single nucleotide chase. For example, 56% of the 21-mer in lane 2 remained in lane 3.

FIG. 4. **Elongation competence of promoter proximal complexes.** Transcripts were synthesized from preinitiation complexes on the pML20-23like3 template as described in Materials and Methods. The initial transcribed sequence of this
plasmid is shown below the figure. Lanes 1-9: Sarkosyl rinsed complexes containing GTP labeled G17 RNA (lane 1) were incubated at 25°C for 5 min with (lane 3) or without (lane 2) CTP to allow advance to +18. These G17 (lanes 4-6) and C18 (lanes 7-9) complexes were incubated for 10 more minutes at 37°C to mimic the conditions of the RNase H digestion used in Figs. 2 and 3 and were then immediately chased with all four NTPs (lanes 5 and 8); SII was included in the chase in lanes 6 and 9. Complex C23 was generated from preinitiation complex with radioactive CTP, followed by sarkosyl rinsing, as described in the Materials and Methods section (lanes 10-12). These C23 complexes were elongated to U26 (lanes 13-15), G29 (lanes 16-18) or A32 (lanes 19-21) by limited (nonlabeled) nucleotide addition at 25°C for 5 min, and were further incubated at 37°C for 10 min. Chase with all four NTPs in the presence or absence of SII immediately followed at 37°C for 5 min. SII cleavage products (3' end of the cleaved RNA) can be seen in lanes 12, 15, 18 and 21. The percent of arrest for each ternary complex is indicated in the top line of the figure.

FIG. 5. **Comparison of the extent of arrest in promoter proximal and promoter distal complexes.** The fraction of complexes which failed to chase, determined as shown in Figs. 2-4, is given for promoter proximal complexes on the pML20-23like3 template (black bars) and promoter distal complexes on the pML20-23 template whose transcripts were truncated by oligonucleotide-directed RNase H cleavage (gray bars). The averages of 2 or 3 independent experiments, with standard deviations as error bars, are reported.
FIG. 6. **Induction of arrest by hybridization of oligonucleotides to the nascent RNA.** A, The indicated chimeras were annealed to body labeled C151 RNA (see Fig. 1) at 37°C for 10 min. RNase H was also added in the indicated lanes, to verify proper annealing of the oligonucleotides. In the indicated lanes, transcripts were chased by the addition of 50 µM GTP, ATP and CTP at 37°C for 10 min. The sizes of the bands are indicated to the right of the figure. B, Summary of the effects of hybridization on complexes C151 and U154. The complementary positions of the chimeras to the RNA are indicated on the figure. Each chimera annealed 3 nt more distal from the 3' end of the RNA on the U154 complexes as compared with C151. The values for percent arrest were corrected by subtraction of the percent arrest observed in the absence of the chimeras. Standard deviations were calculated from 2 to 5 independent experiments.

FIG. 7 **Hybridization of oligonucleotides well upstream of the RNA 3' end causes upstream translocation of pol II.** Complex U154 was synthesized on plasmid pML20-23 (lanes 3 to 12). The indicated chimeras were annealed to the 3' end labeled RNA at 37°C for 5 min. The complementary positions of the chimeras to the RNA are shown in Figs. 1 and 6. Incubation was continued at 37°C for 5 min in the absence or presence of SII as indicated above the lanes. The fraction of total cleavage products that were long (about 6 to 15 nt) or short (primarily dinucleotides) was determined and is indicated above the figure. The location of these cleavage products is indicated to the
left of the figure. RNA length markers of 17 (lane 14) or 23 (lane 1) nt were generated on plasmid pML20-23like3; the marker reactions in lanes 2 and 13 were treated with SII before purification of the RNAs (see ref. 17).

**FIG. 8** *Truncation of the transcript in a second sequence context also leads to greatly increased transcriptional arrest.*  

*A,* Sequence of the initially transcribed region (nontemplate strand) of the pML20-55M template. The identical promoter proximal and promoter distal sequences are underlined. The positions downstream of +1 at which the polymerase was stalled are indicated.  

*B,* Preinitiation complexes formed on bead-attached, PCR-generated pML20-55M linear templates were incubated with CpA (which pairs with the template at positions -1/+1), α-<sup>32</sup>P-CTP, UTP and ATP to obtain complexes stalled at +23. These promoter-proximal A24 complexes were rinsed with sarkosyl prior to testing for the elongation competence (lanes 1-4). A78 complexes (the precursors for promoter distal A24 complexes) were generated by incubating preinitiation complexes with the same set of NTPs used for generating the promoter proximal A24 complexes except that the CTP was not radiolabeled. After rinsing with sarkosyl, the nonlabeled A24 complexes were walked to +54 with CTP, GTP and UTP, rinsed with transcription buffer to remove free NTPs, and then advanced to +77 with α-<sup>32</sup>P-CTP, UTP and ATP. The promoter distal A24 complexes were obtained by digesting A78 complexes with RNAse T1 (10U of T1 in a 10 µl reaction for 5 min at 37°C) to liberate the upstream segment of the transcript (the 3’-most T1 cleavage site is indicated by the arrow in panel A). The T1-treated
complexes were washed with transcription buffer to remove liberated RNAs and RNAse T1. All complexes were tested under identical conditions for elongation competence (30°C for 10 min with 200 μM of all four NTPs) with or without SII (at a final concentration of 29 μg/ml) as indicated. The A78 complexes and the promoter-proximal A24 complexes were mock treated without RNAse T1 at 37°C for 5 min before chase.
Fig. 1
<table>
<thead>
<tr>
<th>Chimera</th>
<th>1st Chase</th>
<th>2nd Chase</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>- U U U G</td>
<td>- - G</td>
</tr>
<tr>
<td>73</td>
<td>- U U U G</td>
<td>- - G</td>
</tr>
<tr>
<td>53</td>
<td>- U U U G</td>
<td>- - G</td>
</tr>
</tbody>
</table>

- **Fig. 2**
% Arrested chimera
1st chase
- G G GA  - G G GA  - G G GA
2nd chase
- - A -  - - A -  - - A -

154

Fig. 3
Fig. 4
Fig. 6
<table>
<thead>
<tr>
<th>Chimera</th>
<th>SII</th>
<th>% Short SII Products</th>
<th>% Long SII Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>+</td>
<td>77</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>78</td>
<td>19</td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>49</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>69</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>73</td>
<td>23</td>
</tr>
</tbody>
</table>

Fig. 7
RNA polymerase II transcription complexes may become arrested if the nascent RNA is shortened to less than 50 nucleotides
Andrea Ujvari, Mahadeb Pal and Donal S. Luse

*J. Biol. Chem. published online June 26, 2002*

Access the most updated version of this article at doi: 10.1074/jbc.M201145200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2002/06/26/jbc.M201145200.citation.full.html#ref-list-1