RNA polymerases bind to specific sequences in DNA, melt open duplex DNA around the start site, and start transcription within the initially melted bubble. The initially transcribing complex is relatively unstable, releasing short abortive products. After synthesis of a minimal length of RNA (~10–12 bases in the T7 system), RNA polymerases complete the transition to a processive (highly stable) elongation phase and lose the initial promoter contacts. The current study strongly supports a model for T7 RNA polymerase in which initial bubble collapse from position −4 to position +3 is responsible for initiating RNA displacement in the transition process. More specifically, collapse of the bubble from position −4 to position −1 indirectly and energetically facilitates the direct strand invasion offered by collapse at positions +1 to +3. Parallel work shows that promoter release, another key event occurring during this stage of transcription, begins after translocation to position +8 and is largely complete upon translocation to about position +12. The timing of promoter release agrees with the timing of initial bubble collapse determined by our previous fluorescence studies, suggesting that these two events are closely related.

The single subunit phage T7 RNA polymerase carries out de novo transcription in a highly efficient manner, with key mechanistic features very similar to those of the structurally unrelated multi-subunit RNA polymerases (1). The polymerase binds specifically to a small duplex DNA element within the promoter (from position −17 to position −5 relative to the start site), melts open the immediate downstream region to form an initial bubble of ~7–8 base pairs, and begins transcription at the start site within the initial bubble (2–5).

Biochemical studies of T7 RNA polymerase have provided mechanistic information on the process of promoter escape. When the RNA transcript is ~2–8 nucleotides in length, the initially transcribing complex is relatively unstable and tends to release the transcript and restart transcription (6, 7). This abortive cycle is a defining feature of the initiation phase of transcription (8, 9). After the synthesis of ~12 nucleotides the RNA polymerase complex enters the elongation phase (10–15). In contrast to the initially transcribing complex, the elongation complex is highly stable, allowing transcripts to be extended over thousands of bases with little premature termination. Finally, within the elongation phase, transcription shows sequence-dependent termination, completing the synthesis of the RNA transcript (16–18).

Crystal structures of an early initially transcribing complex (with a three-base transcript) and two elongation complexes provide abundant information on the initial and final states in the transition process (19–21). Comparing all three structures, it is seen that the conformation of the enzyme, in particular in the N-terminal region (residues 1–266), changes dramatically from an early initially transcribing complex to an elongation complex. The change in structure facilitates growth of a longer RNA-DNA hybrid (~8 nucleotides), creates an RNA exit pathway, and disrupts the initial promoter binding region. This finding suggests that promoter release and initial RNA displacement are important in the transition process.

Biochemical studies also provide insight into the transition from initiation to elongation in T7 RNA polymerase (11, 13, 14, 22–25). Promoter release is thought to be required for the transition to elongation, and footprinting experiments have shown that initial promoter contacts are lost in translocation between positions +6 and +15 (22). Because there is some evidence that specific contacts with promoter bases around positions −9 and −5 are thought to be lost upon translocation to position +7 (23), the loss of the contact with promoter bases near position −17 might be the final signature of promoter release. Exonuclease footprinting experiments have shown that upon translocation to position +8 the entire DNA region from position −17 to position −5 becomes accessible for exonuclease digestion in some population (~5–20%) of the complexes. Upon translocation to positions +9 and +10, the population increases to ~40–90%, indicating that the contact around position −17 is absent or weak at this stage (26). Thus, these studies argue that promoter release occurs upon translocation beyond position +7.

Fluorescence experiments demonstrate that the collapse of the initially melted DNA bubble occurs upon translocation beyond position +8 (13). Interestingly, the size of the bubble increases from ~7–8 base pairs in the initial promoter melting to ~13 base pairs upon translocation to position +8 without any collapse at the upstream edge of the bubble. The upstream edge of the bubble begins to collapse upon translocation beyond position +8, and by the time the complex has reached position +15 (or sooner) the bubble has collapsed back to ~8 base pairs.
Initial Bubble Collapse in Promoter Escape

(13, 14, 27, 28). It seems logical that the initial bubble collapse would help to drive initial RNA displacement via direct strand invasion, and, indeed, initial RNA displacement has been shown to occur at about this point in translocation (13).

Promoter release, initial bubble collapse, and initial RNA displacement are likely associated with one another, accompanied by the conformational rearrangements of the enzyme in the final stage of the transition process. But the precise translocational timing of these three events, the relationship between them, and what movement of the enzyme drives or facilitates them remain elusive. Various results provide important insight into their relationship. Avoiding initial bubble collapse by removing or altering the relevant nontemplate DNA (by creating partially single-stranded, mismatched, or nicked constructs) impairs RNA displacement, resulting in an increase in 11–13-mer products during normal runoff transcription (7, 12). Results from similar partially single-stranded DNA constructs suggested the formation of an extended RNA-DNA hybrid in complexes stalled at position +13 (29). Thus, the 11–13-mer products seen during the above runoff experiments likely arise from improper initial RNA displacement.

The possibility that improper RNA displacement could limit forward translocation is also suggested in the mismatched bubble elongation complex models. On different nucleic acid scaffolds without complementarity in the nontemplate strand within the template-RNA hybrid region, efficient elongation of the RNA primers is possible but is limited to only 3–5 bases (30, 31), suggesting that without the assistance of DNA annealing at the upstream edge of the bubble, the polymerase has difficulty in properly initiating the process of RNA displacement. This interpretation is strengthened by a recent study in which a scaffold was generated containing a complementary nontemplate strand (32). In this case, bubble collapse and displacement of the upstream end of the RNA should be near normal, and the assembled complexes experience no limitation in elongation.

In the current study, our first goal is to address the relationship between initial bubble collapse and initial RNA displacement and to understand their importance in the transition process. Paralleling earlier studies (7, 12, 29), different DNA constructs were used in a transcription assay to investigate the importance of the nontemplate region of the DNA in initial RNA displacement. The results suggest that reannealing of the DNA from position −4 to position +3 is important for proper displacement of the 5′-end of the RNA. The second goal is to determine the timing of promoter release using a sink challenge assay (33). By allowing transcription to stall at every translocational position from position +6 to position +15, we show that promoter release begins upon translocation beyond position +8 and is largely complete upon translocation to position +12.

MATERIALS AND METHODS

Protein Expression and Purification—His-tagged wild type T7 RNA polymerase was prepared from Escherichia coli strain BL21 carrying the plasmid pBH161 (kindly supplied by William T. McAllister) and purified using a Qiagen nickel-nitrilotriacetic column (34). Protein purity was determined by SDS-PAGE analysis. The purified protein was concentrated and dialyzed against 20 mM KH$_2$PO$_4$-K$_2$HPO$_4$, 50% glycerol, 100 mM NaCl, and 1 mM Na$_2$EDTA, pH 7.8; the concentration was calculated from the measured absorbance at 280 nm using the molar extinction coefficients for 1.4 × 10$^5$ M$^{-1}$ cm$^{-1}$ (dithiothreitol) was added to a final concentration of 1 mM, and the sample was then stored at −20 °C. Activity was determined by the transcription assay described below.

Oligonucleotide Synthesis and Purification—DNA oligonucleotides were synthesized trityl-off using an Applied Biosystems Expedite 8909 DNA synthesizer, purified by denaturing polyacrylamide gel electrophoresis, excised from the gel, and eluted using an Elu-Trap$^\circledR$ device (Schleicher and Schuell Inc.). Concentrations of single-stranded DNAs were calculated using the weight sums of the three different measured molar extinction coefficients for each base at 253, 259, and 267 nm (35). Single-stranded DNAs were stored in TE buffer (10 mM Tris, pH 7.8, and 1 mM EDTA) at −20 °C.

Partially Single-stranded or Double-stranded DNA Construction—Nontemplate DNA and the relevant template DNA strands were combined in equal molar concentrations, heated to 75 °C, and then cooled slowly (about 2 h) to room temperature for annealing. Annealed DNA constructs were stored in TE buffer at −20 °C.

Transcription Assays—Transcription reactions were carried out in a total volume of 16 µl at 37 °C for 5 min before being quenched with an equal volume of formamide stop solution (95% formamide, 40 mM EDTA, and 0.02% bromphenol blue). Equal molar concentrations of DNA construct and enzyme were used at final concentrations of 0.125 µM in a reaction buffer containing 30 mM HEPES (pH 7.8), 15 mM magnesium acetate, 25 mM potassium glutamate, 0.25 mM EDTA, and 0.05% Tween 20. Reactions were initiated by the addition of NTPs to a final concentration of 400 µM each and labeled with [α-32P]GTP. RNA transcripts were resolved by 20% polyacrylamide and 7 M urea gel electrophoresis and visualized and quantified on a Storm Phosphor-Imager (Amersham Biosciences).

Sink Challenge Assays—In these studies, reactions were carried out as described above for 1 min before the addition of the sink. Buffer conditions and concentrations of DNA and enzyme before the addition of the sink were the same as described above. Reactions were initiated by the addition of GTP, ATP, and CTP to a final concentration of 400 µM each. After allowing transcription (in the absence of radiolabel) for 1 min, 2 µl of TE buffer containing 20 µM sink DNA and [α-32P]ATP was added. The reactions were quenched by the addition of an equal volume of formamide stop solution 1 min after the addition of the sink. RNA transcripts were resolved, visualized, and quantified as described above.

RESULTS

The following experiments address the interrelationship between three fundamental events in transcription, i.e. initial displacement of the 5′-end of the RNA, collapse of the upstream edge of the initial transcription bubble, and release of specific protein-promoter contacts.

Coupling between Initial Bubble Collapse and Initial RNA Displacement

The transcription bubble in T7 RNA polymerase begins as a 7–8 base pair bubble and grows to −12–13 base pairs as the initially transcribing complex synthesizes up to a 7–8-base RNA (13). The following studies explore the relationship between the initial collapse of the upstream edge of the bubble and the displacement of the 5′-end of the nascent RNA transcript. By analogy to the deficiencies in the elongation of scaffold or mismatched bubble elongation complexes, analysis of the results presented below uses the premature release of 11–13-mer RNA products as a sign of improper initial RNA displacement. More specifically, the molar ratio of 11–13-mer to 20-mer products is taken to assess the initial RNA displacement. Of the complexes that have successfully extended past position +10, this ratio represents the fraction that terminates prematurely at positions +11 to +13.

Partially Single-stranded Constructs—To address the relationship between initial bubble collapse (upstream DNA reannealing) and initial RNA displacement, a set of partially single-stranded DNA constructs was designed for comparison with the control DNA construct (fully double-stranded) in a transcription assay. All DNA constructs encode the identical runoff 20-mer RNA. As shown in Table I, partially single-stranded DNA constructs were prepared in which the nontemplate strand extends from position −22 upstream (in all constructs) to positions ranging from −5 to +3 downstream (constructs PSS[−5] to PSS[+3]). It is expected that constructs lacking nontemplate strand bases at positions where collapse is critical to invasion-mediated displacement of the 5′-end of the RNA
will be impaired in RNA displacement. As the critical nontemplate strand bases are restored, RNA displacement will be restored.

The results presented in Fig. 1 show that in a construct containing the nontemplate strand downstream only to position 5, 11–13-mer products are substantially increased relative to transcription from the fully double-stranded control, which is consistent with previous results (12). The data in lanes 1–5 of Fig. 1 show that as the nontemplate strand is extended from position 5 to 1 to 3, the levels of 11–13-mer products drop to near those of the double-stranded control (Fig. 1, lanes 7 and 8), demonstrating that the reannealing of the DNA bases from position 5 to position 3 is sufficient for proper initial RNA displacement. Extension of the nontemplate strand farther downstream does not alter the levels of 11–13-mer products. More importantly, the molar ratio of 11–13-mer to 11–20-mer products mirrors these trends, confirming that premature termination at positions 11 to 13, a property of complexes lacking the nontemplate strand, is fully rescued by extension of the nontemplate strand to only position 2.

Mismatched Bubble Constructs—To confirm the generality of the results from the partially single-stranded constructs, a set of DNA constructs was designed, each containing a window of four mismatched bases with the window centered at various positions from 4 to 20 (Table II; constructs MM[4, 1] to MM[16, +20]).

1 P. Gong and C. T. Martin, unpublished results.
Localized mismatch-containing DNA constructs

All constructs contain the same template strand as the constructs in Table I (and thus encode the same 20-base transcript). Mismatched bases are shown in lowercase letters.

<table>
<thead>
<tr>
<th>DNA</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM[−4, −1]</td>
<td>5′-GGGATTAATACGACTCACTATAGGGAGACCACCATGGTTTCC-3′</td>
</tr>
<tr>
<td></td>
<td>3′-CGCTATTATATTCGAGGTGATATCCCTCTGCTGTTGCTGCAAAAGG-5′</td>
</tr>
<tr>
<td>MM[+, +4]</td>
<td>5′-GGGATTAATACGACTCACTATAGGGAGACCACCATGGTTTCC-3′</td>
</tr>
<tr>
<td></td>
<td>3′-CGCTATTATATTCGAGGTGATATCCCTCTGCTGTTGCTGCAAAAGG-5′</td>
</tr>
<tr>
<td>MM[5, +8]</td>
<td>5′-GGGATTAATACGACTCACTATAGGGAGACCACCATGGTTTCC-3′</td>
</tr>
<tr>
<td></td>
<td>3′-CGCTATTATATTCGAGGTGATATCCCTCTGCTGTTGCTGCAAAAGG-5′</td>
</tr>
<tr>
<td>MM[9, +12]</td>
<td>5′-GGGATTAATACGACTCACTATAGGGAGACCACCATGGTTTCC-3′</td>
</tr>
<tr>
<td></td>
<td>3′-CGCTATTATATTCGAGGTGATATCCCTCTGCTGTTGCTGCAAAAGG-5′</td>
</tr>
<tr>
<td>MM[17, +20]</td>
<td>5′-GGGATTAATACGACTCACTATAGGGAGACCACCATGGTTTCC-3′</td>
</tr>
<tr>
<td></td>
<td>3′-CGCTATTATATTCGAGGTGATATCCCTCTGCTGTTGCTGCAAAAGG-5′</td>
</tr>
</tbody>
</table>

Fig. 2. Localized mismatched bases impair bubble collapse. Transcription from constructs with windows of four mismatched bases (lanes 2–6) is compared with that from a control construct (lane 1). The nomenclature is that of Table II. Transcripts are indicated as described in the legend to Fig. 1.

Constructs with the mismatched bases in the initially melted region, MM[−4, −1] and MM[+, +4], both give high ratios (−3-fold above the native level) of 11–13-mer to 11–20-mer products (Fig. 2, lanes 2 and 3), indicating an impairment in RNA displacement. The results for the MM[+, +4] construct are consistent with a model in which bubble collapse from position +1 to position +3 is essential. The observation that a construct with a mismatch bubble from position −4 to position −1 also yields high levels likely reflects that the reannealing of bases in the −4 to −1 region cooperatively facilitates reannealing of the DNA from position +1 to position +3. As the mismatched region is moved downstream as seen in constructs MM[5, +8] and MM[9, +12], the ratio of 11–13-mer to 11–20-mer products is less than that observed for mismatches at positions −4 to position +3. The energetics can be restored by replacement of the adjacent AT pairs by GC pairs. Specifically, the construct MM[+, +4] construct facilitates the strand invasion immediately downstream. To confirm this assumption, a set of DNA constructs was designed containing single or double base pair mismatches as summarized in Table III. Introducing different adjacent two-base mismatches in this critical region results in the highest (−3-fold above native level) ratios of 11–13-mer to 11–20-mer products (Fig. 3, lanes 2 and 3). Furthermore, any single base mismatch in this region also leads to a high ratio (−2-fold above native level) of 11–13-mer to 11–20-mer products (Fig. 3, lanes 4–6). These results demonstrate that even a small weakening of bubble collapse in this region is sufficient to impair proper RNA displacement.

The strong effect of the single base mismatch at position −1 supports our earlier proposal that the reannealing of DNA immediately upstream of the initial RNA displacement site (+1) facilitates the strand invasion immediately downstream. To confirm that the collapse of bases from position −4 to position −1 helps the downstream collapse, another set of DNA constructs was designed. Mismatches at either positions −4/−3 or positions −2/−1 induce a high ratio of 11–13-mer to 11–20-mer products (Fig. 3, lanes 7 and 9), implying that collapse of the bubble 1–4 bases upstream of the RNA facilitates its release. This appears to be a simple matter of energetics, however, in that the energetics can be restored by replacement of the adjacent AT pairs by GC pairs. Specifically, the construct MM[−4, −3]GC[−2, −1] shows near native ratios of 11–13-mer to 11–20-mer RNA products (Fig. 3, lane 8). Thus, collapse of the upstream edge of the initial bubble from position −4 to position −1 contributes to the overall driving force to facilitate collapse at positions downstream. Finally, replacement of the TA step at positions −2 and −1 by a GC step in an otherwise wild type construct leads to ratios of 11–13-mer to 11–20-mer...
products lower than those seen in the fully duplex wild type construct (Fig. 3, lane 10).

Translocational Timing of Promoter Release

Promoter release, initial bubble collapse, and initial RNA displacement are expected to be closely related to one another in translocational timing. The results presented here suggest that initial bubble collapse and, presumably, the consequent displacement of the 5'-end of the RNA from the initial hybrid are most critical for allowing the proper transition. But how is promoter release involved in this process? Is promoter release coupled to initial bubble collapse or initial RNA displacement or both? Determination of the translocational timing of promoter release would help to address these questions. The occurrence of promoter release was previously localized between translocational positions +7 and +15 by footprinting and UV photo cross-linking studies (22, 23). Preventing promoter release by the introduction of a covalent cross-link between the polymerase and the +17 template base of the promoter DNA results in substantial accumulation of 12–13-mer products (36), suggesting that promoter release likely happens or has to happen at some point before the complex steps to position +12 or +13.

In an attempt to precisely determine the timing of promoter release, a set of DNA constructs derived from the T7 promoter was designed and is shown in Table IV. These constructs allow transcription to stall at every position from position 6 to position 15 in the presence of GTP, ATP, and CTP. In translocation of the complex to positions 6 and 7, the promoter has not been released (22, 23). As illustrated in Fig. 4, if RNA dissociates but polymerase and DNA remain associated, the polymerase is then able to resume the starting conformation.

![Critical regions for proper RNA displacement](image)

**Fig. 3. Critical regions for proper RNA displacement.** Transcription from constructs with mismatched bases in the proposed critical DNA region (lanes 2–6) is compared with that of control construct (lane 1). Transcription from constructs with mismatched bases and/or designed GC pairs (lanes 7–10) is compared with that from the control construct (lane 1). Construct nomenclature is that of Table III (a star denotes the position of the GC dinucleotide step). Transcripts are indicated as described in the legend of Fig. 1.

**TABLE III**

<table>
<thead>
<tr>
<th>DNA</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM[-1,+1]</td>
<td>5'-GCCAGTTAGCTCATACTACAGATACCCCTGGTTGCTGCCAAAGG-3'</td>
</tr>
<tr>
<td>MM[+1,+2]</td>
<td>3'-CCCTGGTTGCTGCCAAAGG-5'</td>
</tr>
<tr>
<td>MM[-1]</td>
<td>5'-GCCAGTTAGCTCATACTACAGATACCCCTGGTTGCTGCCAAAGG-3'</td>
</tr>
<tr>
<td>MM[+1]</td>
<td>3'-CCCTGGTTGCTGCCAAAGG-5'</td>
</tr>
<tr>
<td>MM[+2]</td>
<td>5'-GCCAGTTAGCTCATACTACAGATACCCCTGGTTGCTGCCAAAGG-3'</td>
</tr>
<tr>
<td>MM[-4,-3]</td>
<td>3'-CCCTGGTTGCTGCCAAAGG-5'</td>
</tr>
<tr>
<td>MM[-4,-3]-GC[-2,-1]</td>
<td>5'-GCCAGTTAGCTCATACTACAGATACCCCTGGTTGCTGCCAAAGG-3'</td>
</tr>
<tr>
<td>MM[-2,-1]</td>
<td>3'-CCCTGGTTGCTGCCAAAGG-5'</td>
</tr>
<tr>
<td>GC[-2,-1]</td>
<td>5'-GCCAGTTAGCTCATACTACAGATACCCCTGGTTGCTGCCAAAGG-3'</td>
</tr>
</tbody>
</table>

**FIG. 4.** Transcription from constructs with mismatched bases in the region from position 4 to position 2 and/or designed GC pairs in positions 2 and 1.

**Table 3:** Mismatches very near the start site

DNA constructs with mismatched bases in the region from position -4 to position +2 and/or designed GC pairs in positions -2 and -1. Mismatched bases are shown in lowercase letters. Designed GC pairs are shown in boldface.
sequence from different position with only GTP, ATP, and CTP).

The results presented in Fig. 5 show that after initiating transcription in the absence of sink and walking to position 6, 7, or 8 to position 12 or 13. Another critical event, promoter release, has been shown by both exonuclease footprinting studies (26) and the functional studies presented here to occur at about the same stage of transcription. This correlation suggests that initial bubble collapse and promoter release are mechanistically associated. A simple model would have the 5’-end of the RNA competitively displaced from the template strand by the reannealing of the nontemplate strand DNA in that same region.

Indeed, the nontemplate DNA strand has been proposed previously to play an important role in RNA displacement (11, 29, 33). The production of 11–13-mer products from templates lacking the nontemplate strand downstream of the start site has been proposed to arise from the formation of a persistent hybrid as a result of improper RNA displacement (29). One must be careful to note, however, that in multiple turnover experiments, constructs lacking the nontemplate strand in the transcribed region (partially single-stranded constructs) convert quickly to nicked or gapped constructs containing an RNA:DNA hybrid within the transcribed regions regardless of whether the hybrid forms during the first round (persistent) or whether released RNA subsequently anneals to the partially single-stranded DNA construct. Results presented here demonstrate that both the complementarity and the integrity of the DNA region from position −5 to +3 is essential to achieve proper initial RNA displacement.

RNA products of length 11–13 bases are also observed at low levels in native transcription from fully duplex DNA, suggesting that even with intact complementary DNA strands, complexes passing through this transition have an inherent difficulty in displacing the 5’-end of the RNA. The studies presented here show that initial bubble collapse is indeed important for initial RNA displacement. More specifically, any disturbance that limits (directly or indirectly) reannealing of the DNA at positions −1 through position +2 leads to a significant increase in 11–13-mer products.

In a very recent study, a similar approach was used to investigate the role of the nontemplate strand in the transition and shows that promoter release is complete at translocational position +12 or +13.

This timing agrees well with previous fluorescent studies measuring the occurrence of bubble collapse (13). Some of these fluorescence data are overlaid onto the functional data in Fig. 5, illustrating clearly that promoter release and bubble collapse are closely related. Exonuclease footprinting studies suggest that promoter release begins upon translocation to position +8 and is complete upon translocation to position +9 or +10 (26); however, as noted in the study, the processive action of the exonuclease might drive release to occur sooner than under normal conditions.

**DISCUSSION**

Promoter Release, Upstream Bubble Collapse, and Initial RNA Displacement—Previous results and the work shown here indicate that a critical event in this process, collapse of the initial DNA bubble, starts as the complex progresses from position +8 to position +9 (13). Another critical event, promoter release, has been shown by both exonuclease footprinting studies (26) and the functional studies presented here to occur at about the same stage of transcription. This correlation suggests that initial bubble collapse and promoter release are mechanistically associated. A simple model would have the 5’-end of the RNA competitively displaced from the template strand by the reannealing of the nontemplate strand DNA in that same region.

<table>
<thead>
<tr>
<th>DNA Sequences</th>
<th>-17</th>
<th>-5</th>
<th>1</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS-6</td>
<td>5'-TAATACGACTACATAAGGAGAAGAGCGACTAC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-7</td>
<td>3'-ATATAGCGACTACATAAGGAGAAGAGCGACTAC-5'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-8</td>
<td>5'-TAATACGACTACATAAGGAGAAGAGCGACTAC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-9</td>
<td>3'-ATATAGCGACTACATAAGGAGAAGAGCGACTAC-5'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-10</td>
<td>5'-TAATACGACTACATAAGGAGAAGAGCGACTAC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-11</td>
<td>3'-ATATAGCGACTACATAAGGAGAAGAGCGACTAC-5'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-12</td>
<td>5'-TAATACGACTACATAAGGAGAAGAGCGACTAC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-13</td>
<td>3'-ATATAGCGACTACATAAGGAGAAGAGCGACTAC-5'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-14</td>
<td>5'-TAATACGACTACATAAGGAGAAGAGCGACTAC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS-15</td>
<td>3'-ATATAGCGACTACATAAGGAGAAGAGCGACTAC-5'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sink</td>
<td>5'-ACGTACTGTACAC-3'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'GGAGAATGATCTCTGTATAC-5'</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and start a second round of transcription without rebinding to the promoter. In this mode, the complex is resistant to the challenge of a competitive promoter sink (33). Upon translocation to position +15, however, all promoter contacts are lost. To start a second round of transcription, the polymerase now has to rebind to a promoter and, therefore, is susceptible to challenge by the sink.

The “sink” used in this work is a hairpin DNA construct with a consensus duplex promoter region from position −17 to position −5 but containing only the template strand from position −4 to position +3 (Table IV and Fig. 4). Constructs with similar structures have been shown to bind T7 RNA polymerase at a consensus duplex promoter region from position 1 to position 10 (37) and, thus, can efficiently inhibit transcription from normal DNA constructs.

The results presented in Fig. 5 show that after initiating transcription in the absence of sink and walking to positions +6, +7, or +8, the addition of a sink in a 20-fold excess over the normal DNA construct allows each polymerase, on average, to carry out about five cycles of transcription in 1 min. This indicates that the polymerase has not released its promoter contacts early in transcription and can efficiently inhibit transcription from normal DNA constructs.

The consensus duplex promoter region of events.

Table IV

Constructs used in the sink challenge assay

Constructs DS-6 through DS-15 contain both the template and the nontemplate strand from positions 1 to 17 (including the consensus 610 sequence from −17 to +6). The first TA pairs introduced in the encoding region are shown in boldface (which allows transcription to stall at different position with only GTP, ATP, and CTP).
from initiation to elongation. In that study, the nontemplate strand is extended to positions $5, -1, +6, +10, +15$, and $+22$. Although mostly consistent with the data presented here, that study concludes that a complementary nontemplate strand is required as far downstream as $6–10$ base pairs. In contrast, our study concludes that a complementary nontemplate strand is necessary to achieve double-stranded levels of $11–13$-mer products.

Previous studies suggest that promoter release begins as early as translocation to position $+8$ but may also occur at positions one or two bases farther on ($22, 23, 26$). Similarly, the results presented in Fig. 5 demonstrate that the sink begins to show an effect on cycling with translocation to position $+9$ but does not exert its full challenge efficiency until translocation beyond position $+12$, suggesting that promoter release does not happen at a single translocational step. The timing of promoter release shown here is similar to the occurrence of an initial bubble collapse observed in our earlier fluorescence studies ($13$). As measured by a decrease in the fluorescence signal from a fluorescent base analog at position $+9$ the fluorescence signal further decreases upon translocation to position $+10$, and only upon translocation to position $+11$ are the DNA bases at position $-2$ completely annealed.

Taken together, these findings indicate that promoter release and bubble collapse are closely coupled and that the occurrences of these two events are most likely not homogeneous (i.e. do not happen in a single translocational step). We suggest that a delay of either event in a complex can lead to some probability of improper RNA displacement with a resulting higher production of $11–13$-mer products. Recent results provide evidence to support this argument. Forced retention of promoter contacts on a DNA construct that is otherwise normal leads to a similar large increase in $12–13$-mer products ($36$). This suggests that the delay of promoter release possibly leads directly to a decrease in DNA reannealing, resulting in improper RNA displacement.

In our previous work, the translocational timing of initial RNA displacement was determined by incorporating a fluorescent base analogue into the template DNA at position $+1$ or $+2$ while having either matched or mismatched nontemplate bases opposite the fluorophore position ($13$). In complexes stalled at positions $+10$ through $+12$, results from the correctly paired constructs demonstrate that the template bases at positions $+1$ and $+2$ remain duplexed in either an RNA:DNA hybrid or a DNA:DNA duplex. Fluorescence from complexes with a single base mismatched opposite the fluorescent base analogue shows that dissociation of the $5^\prime$-end of the RNA occurs upon translocation beyond position $+10$. Complexes stalled at position $+11$ show the beginning of RNA release at position $+1$, whereas complexes stalled at position $+12$ show initial peeling off of the RNA at position $+2$. The fluorescence indicates that RNA displacement is similarly a stepwise, non-homogeneous process.

In the same study, we noted that the peeling away of the $5^\prime$-end of the RNA, if driven by reannealing of the DNA duplex, might occur later in the constructs with mismatched bases at $+1$ or $+2$ in the nontemplate strand. Our results from constructs with mismatched bases at those same positions strongly support this argument (constructs MM$[+1]$ and MM$[+2]$). With only one mismatched base in the critical region, much more of the complexes encounter difficulty in properly displacing the $5^\prime$-end of the RNA, yielding more than twice the ratio of $11–13$-mer to $11–20$-mer products compared with the fully duplex construct. We therefore conclude that in normal transcription on fully duplex DNA, displacement of the $5^\prime$-end of the RNA likely occurs earlier than what was suggested in the fluorescence study.

**Cooperative Collapse of the Bubble**—The initially melted bubble in the DNA extends from position $-4$ to about position $+3$. The results presented in the current study are consistent with a model in which the upstream edge of the bubble remains melted through translocation to position $+8$, extending the bubble downstream to about position $+9$. Subsequent translocation leads to a dramatic collapse of the upstream edge of the

---

Initial Bubble Collapse in Promoter Escape

bubble. Bases at positions −4 to +1 or +2 could collapse all at once in a cooperative fashion. Thus, changes in base pairing strength at positions −4 through −1 are expected to have an effect on the strength of the collapse at positions +1 and +2. Indeed, mismatched bases at positions −4 and −3 or at positions −2 and −1 lead to a weakening of strand displacement and a resulting increase in 11–13-mer RNA products. Similarly, strengthening of the base pairs at positions −1 and −2 is expected to facilitate collapse-mediated strand displacement and leads to a decrease in 11–13-mer products. This is seen both in an otherwise native context (GC[−2,−1]) and in the context of mismatched bases at positions −4 and −3 (MM[−4,−3]GC[−2,−1]). Interestingly, although overall transcription from the GC[−2,−1] is decreased, the ratio of abortive products to the full-length transcript is lower than for the native promoter, suggesting an approach for transcribing RNAs from constructs that otherwise produce high levels of abortive transcripts.

Why Release of 11–13-Mer RNA Transcripts?—The model discussed above proposes that the limitation of DNA bubble collapse leads to a larger portion of complexes poised at about positions +9 through +10 having improperly displaced RNA. The crystal structures of elongation complex models show the formation of an RNA exit channel in the protein in the elongation complex. We propose that complexes poised at translocational positions +9 and +10 are optimal for directing the displaced 5′-end of the RNA into the channel. Complexes that do not displace the 5′-end of the RNA at these normal translocational positions would be expected to have difficulty threading the 5′-end into the channel, as the 5′-end must be both spatially translated and rotated relative to that optimal positioning. Such complexes will then be limited sterically in further elongation as the hybrid encounters a protein block (residues 50–70) (19, 20). The results presented here suggest that complexes that fail to displace can add at most 2–4 more bases (yielding 11–13-mer RNA products). This finding is directly analogous to the observation that elongation scaffolds impaired in collapse can extend −3 to 5 bases beyond the normal 7–8 base pairs dictated by the hybrid (yielding a maximal and presumably strained 13-base pair hybrid) (30, 31). Obviously, forward translocation of the polymerase is competitively limited by premature release of the transcript. When the RNA polymerase is stalled at positions +11 to +13 on a construct similar to PSS[−5], turnover is substantially more rapid than for a similar stall on a fully duplex construct.1 This instability would presumably be similarly manifest during passage through this position during attempted runoff transcription.

In any case, complexes that achieve proper RNA displacement at positions +9 to +10 then pass through position +13 normally and enter the processive elongation phase. Once past this transition, the complexes produce runoff transcripts with high efficiency. This is true of both the majority of complexes in normal transcription and the smaller fraction of complexes in the bubble collapse-impaired constructs studied here.

In a very recent study characterizing T7 RNA polymerase elongation complexes stalled at DNA lesions, an eight-nucleotide RNA primer was successfully extended to a 133-nucleotide runoff product using a promoter-independent scaffold approach (32). However, in this case a fully complementary nontemplate DNA strand was added after the assembly of a complex including the polymerase, template DNA, and an 8-mer RNA primer. As a result, upstream bubble collapse can occur, allowing proper displacement of the 5′-end of the RNA primers and the generation of a bona fide elongation complex.

Correlation with Changes in Enzyme Structure—In T7 RNA polymerase, the transition from an initiation to an elongation complex involves a conformational change of the enzyme and a significant reorganization of the nucleic acids within the protein as promoter contacts are lost and an RNA:DNA hybrid is formed (13,14,19–21,25). The crystal structure of the early initiation transcribing complex suggests that the transition from initiation to elongation must start to happen in the earliest stages of transcription (21). In particular, a steric clash between the growing hybrid and the RNA polymerase must occur upon translocation beyond position +3. Mutagenesis studies show that the core subdomain (residues 72–151 and 206–257) and the thumb domain (residues 324–411) are probably involved in the conformational change in subsequent translocational steps (42, 43). UV photo cross-linking results suggest that contacts between the protein and the DNA at positions −5 and −9 are likely disturbed as the complex progresses to position +6 and +7 (23). These data suggest that the complex is accumulating conformational change stepwise toward the final elongation configuration.

In a proposed model for the late initiation complex stalled at position +8, the conformation of the enzyme is that of the elongation complex (the promoter binding region having both rotated and translated relative to its starting position), but the promoter has not yet been released, and the initially melted bubble region remains held open by the enzyme (19). This model would then represent the moment just before initial bubble collapse and is mostly consistent with the results from biochemical studies. A more recent model suggests an initial spatial translation of the promoter binding region of the protein without substantial rotation (44). This model also allows for translocation to position +8 prior to promoter release and bubble collapse and is also consistent with the results presented here.

Summary—The results presented here demonstrate that bubble collapse and promoter release commence as the complex progresses beyond position +8, defining the beginning of the final stage of the transition process. DNA reannealing from the upstream edge of the bubble to position +3 helps to displace the 5′-end of the nascent RNA. Those complexes that successfully complete this displacement proceed on to fully competent elongation complexes. For those complexes that do not successfully displace the 5′-end of the RNA, transcription is limited to only a few nucleotides, leading to accumulation of 11–13-mer RNAs. RNA modifications that limit initial bubble collapse lead to fewer elongation competent complexes, whereas modifications that enhance initial bubble collapse lead to an increase in elongation-competent complexes. This finding suggests that under normal transcription an intrinsic energetic barrier exists at this key transition in transcription.

REFERENCES


1. This study was supported by grants from the National Institutes of Health (grant 1RO1GM37202).
Initial Bubble Collapse in Promoter Escape


