Promoter Clearance by T7 RNA Polymerase

INITIAL BUBBLE COLLAPSE AND TRANSCRIPT DISSOCIATION MONITORED BY BASE ANALOG FLUORESCENCE*

Received for publication, September 13, 2001, and in revised form, October 31, 2001
Published, JBC Papers in Press, November 1, 2001, DOI 10.1074/jbc.M108856200

Cuihua Liu and Craig T. Martin‡

From the Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003

Footprinting, fluorescence, and x-ray structural information from the initial, promoter-bound complex of T7 RNA polymerase describes the very beginning of the initiation of transcription, whereas recent fluorescence and biochemical studies paint a preliminary picture of an elongation complex. The current work focuses on the transition from an initially transcribing, promoter-bound complex to an elongation complex clear of the promoter. Fluorescence quenching is used to follow the melted state of the DNA bubble, and a novel approach using a locally mismatched fluorescent base analog reports on the local structure of the heteroduplex. Fluorescent base analogs placed at positions −2 and −1 of the promoter indicate that this initially melted, non-transcribed region remains melted as the polymerase translocates through to position +8. In progressing to position +9, this region of the DNA bubble begins to collapse. Probes placed at positions +1 and +2 of the template strand indicate that the 5′ end of the RNA remains in an heteroduplex as the complex translocates to position +10. Subsequent translocation leads to sequential dissociation of the first 2 bases of the RNA. These results show that the initially transcribing complex bubble can reach a size of up to 13 base pairs and a maximal heteroduplex length of 10 base pairs. They further indicate that initial bubble collapse precedes dissociation of the 5′ end of the RNA.

The recent past has seen a number of structures of RNA polymerases, ranging from RNA polymerase II (1, 2) to a smaller bacterial RNA polymerase (3) to the simplest well studied system: the single-subunit T7 RNA polymerase (4–7). Structures are available for polymerases without DNA, with DNA bound, and with either a very short initial transcript or with longer transcripts, corresponding to elongation. Lacking in any system, however, is x-ray structural information on the critical transition from an unstable initially transcribing complex (ITC)1 to a stable elongation complex. This transition occurs at about 10 base pairs (bp) in all RNA polymerases, suggesting that the transition is a fundamental feature of transcription, independent of the specific system (8). What is the nature of the transcription bubble at this critical, early transition point?

For the single-subunit RNA polymerase from T7, a crystal structure is available with a 3-base RNA transcript at the active site (7). Modeling from that structure, Cheetham and Steitz (7) predicted that as the polymerase transcribes forward, the enzyme could accommodate no more than a 3-base heteroduplex. Recent results have shown, however, that in a stably elongating complex stalled clear of the promoter, near position +15, the transcription bubble extends about 8–9 bp upstream of the stall site, consistent with a heteroduplex size of −8 bp (9–11). Cross-linking studies report that in a complex stalled at position +23, RNA at position −9 relative to the stall site cross-links to RNA polymerase but not to DNA, whereas RNA at positions closer to the last incorporated base cross-links to DNA, again consistent with an 8-bp heteroduplex (10).

Footprinting studies have indicated that complexes stalled at position +6 retain promoter occupancy, as in the crystal structure, but complexes stalled at position +15 have cleared the promoter (9, 12). It is reasonable to expect a substantial difference in these structures; perhaps the initially transcribing complex has different structural constraints.

We have previously used fluorescent base analogs to map DNA melting in both the initial promoter-bound complex and in a stably elongating complex (beyond the transition) in the model enzyme T7 RNA polymerase (11, 13). The details of the model for the initial melted complex derived from the early fluorescence study have been confirmed by recent crystal structures (6), and the results from the latter are consistent with biochemical probes of a similarly stalled elongation complex (9). In the current work, we turn this approach toward mapping the collapse of the initially melted, untranscribed region of the DNA and the initial peeling away of the RNA from the heteroduplex as the ITC progresses through and beyond position +10.

The new finding that pyrrolo-dC in a mismatch context has typical fluorescence values higher than that of the same probe in single-stranded DNA (which, in turn, has higher fluorescence than in double-stranded DNA), coupled with the fact that fluorescence is quenched in a heteroduplex, now allows us to probe not only the melted state of the DNA but also for the presence of a heteroduplex. The results present a detailed structural picture of promoter clearance in the T7 RNA polymerase model system.

MATERIALS AND METHODS

Enzyme—T7 RNA polymerase was prepared from Escherichia coli strain BL21, carrying the overproducing plasmid pR1219 (kindly supplied by F. W. Studier), in which RNA polymerase is expressed under inducible control of the lac UV5 promoter (14). The enzyme was purified, and the concentration was determined (εmax = 1.4 × 105 M/cm) as described previously (15). Purity of the enzyme (>95%) was verified by SDS-denaturing polyacrylamide gel electrophoresis.

Synthetic Oligonucleotides—Oligonucleotides were synthesized by
the phosphoramide method on an Expedite 8909 DNA/RNA synthesizer. The fluorescent cytidine analog pyrrolo-dC was incorporated into oligonucleotides using furano-DT phosphoramidite (Phen Research Corp.) as described previously (11). Single strands were purified using an AmBisorb CG-18C reverse phase resin as described (16). Purification of single-stranded oligonucleotides was confirmed by denaturing (urea) gel electrophoresis of 5’-end labeled single strands. Double-stranded templates were prepared by heating a 1:1 mixture of complementary single strands in Tris-EDTA buffer (10 mM Tris, pH 7.8, 1 mM EDTA) at 75°C for 5 min. The samples were then allowed to cool slowly to room temperature. The constructs were either used immediately or stored at −20°C.

**Steady-state Fluorescence Measurements**—Fluorescence measurements were carried out with a Photon Technology International L-format fluorimeter with a 75-watt arc lamp and both emission and excitation monochromators using a 75-μl (light path, 3 × 3 mm; center, 15 mm) ultramicrowell cell (Hellma). Fluorescence emission from pyrrolo-dC was detected at 480 nm, with excitation at 350 nm; slits on both channels were set to 5 nm. Fluorescence from 2-aminopurine was obtained similarly but with excitation at 315 nm and emission detected at 370 nm. For measurements of fluorescence from 2-aminopurine, fluorescence is reported after subtraction of the enzyme background. All fluorescence experiments were carried out in fluorescence buffer: 30 mM HEPEs, pH 7.8, 15 mM magnesium acetate, 25 mM potassium glutamate, and 300 μM EDTA, and 0.05% Tween 20 (Calbiochem; 10% protein grade) in a sample compartment thermostated at 25°C. The fluorescence changes recorded in all experiments represent an average of three measurements. In the chase experiments, the fluorescence from double-stranded DNA (1.0 μM) was first recorded alone and then after the addition of each of the following: enzyme (to ~1.0 μM final concentration), GTP and ATP, 3’-dCTP (or CTP), and UTP (to 1000 μM each), with all incubations occurring at 3 min intervals. For samples with two fluorophores, the same protocol was followed, except that during the measurement, the monochromators were switched, under computer control, alternately between the excitation and emission pairs above.

**Transcription Assays**—Transcription reactions were carried out in the above fluorescence buffer at 25°C in a volume of 20 μl containing nucleoside triphosphates as described below and (α-32P)GTP (specific activity, 800 Ci/mmol; PerkinElmer Life Sciences). Reactions were stopped by the addition of an equal volume of quenching solution (95% formaldehyde, 50 mM EDTA, 0.01% bromphenol blue, 0.01% xylene cyanol). The enzyme and DNA (1.0 μM each) were incubated in transcription buffer (lacking nucleoside triphosphates) for 4 min at room temperature and then divided into two equivalent fractions. GTP and ATP (including α-32P(GTP) were added to the first fraction, which was then incubated for 3 min at room temperature and quenched. The second fraction was similarly incubated with GTP and ATP for 3 min, and then 3’-dCTP was added to a final concentration of 1000 μM. This was then incubated for 3 min and quenched. All transcription samples were heated to 95°C for 2 min and loaded onto 20% polyacrylamide-6 M urea gels. After electrophoresis, gels were dried and quantified using a Molecular Dynamics Storm 840 PhosphorImager.

**RESULTS**

In the initial formation of the open complex, base pairs at positions −4 through +3 are melted open (6, 13). As the polymerase translocates through to position +3, the upstream edge (position −4) of this initial bubble remains open (7). However, translocation to position +15 results in the closure of the initial bubble and the loss of promoter contacts (11, 12). In a complex stalled at position +15, the heteroduplex extends 18 bp upstream of the last base incorporated (9–11). Between these two states, the RNA polymerase must clear the promoter and become a stable elongation complex. At what point between positions +3 and +15 does the initial bubble collapse? At what point does the RNA begin to peel away from the template?

Fluorescence from base analogs placed site-specifically within the DNA reports on the melted state near the analog. To monitor fluorescence changes at specific positions as the RNA polymerase translocates along the template DNA, DNA constructs (summarized in Table I) were prepared, which allow walking of the RNA polymerase to positions +7–12 in the presence of only GTP and ATP as substrates. Subsequent addition of 3’-dCTP allows further translocation by only 1 base pair, to positions +8–13, ensuring that misincorporation does not lead to heterogeneity in the stalled complexes. In some experiments below, subsequent addition of UTP allows binding

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>DNA sequence downstream of position −5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2AP−dCTGA (7)</td>
<td>. . .TATAGGAAGAGGCTTACCAATCATGATC</td>
</tr>
<tr>
<td>2AP−dCTGA (8)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>2AP−dCTGA (9)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>2AP−dCTGA (10)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>2AP−dCTGA (11)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>m-pyC−dCTGA (10)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>m-pyC−dCTGA (11)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>m-pyC−dCTGA (12)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>m-pyC−dCTGA (10)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>m-pyC−dCTGA (11)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
<tr>
<td>m-pyC−dCTGA (12)</td>
<td>. . .TATAGGAAGAGCCTTACCAATCATGATC</td>
</tr>
</tbody>
</table>
FIG. 1. Observation of the collapse of the initial bubble by following the fluorescence quenching of 2-aminopurine placed near the start site, at positions −2 of the template strand. Controls representing fully duplex (double-stranded DNA (dsDNA)) and fully single-stranded DNA (ssDNA) are shown. In all cases, relative fluorescence is normalized to that of the double-stranded DNA control. Below the controls, the sequences of the DNA templates are such that transcription in the presence of GTP and ATP will walk the complex out to the position shown in the top number of each pair along the y axis. Subsequent addition of 3′-dCTP (3′dC) allows the complex to walk 1 base pair further, as indicated by the bottom number of each pair. High fluorescence, as for complexes stalled at positions +7 and +8, indicates a melted bubble. Low fluorescence, as at positions +10 and beyond, indicates collapse to a duplex. Similar results (not shown) were observed for constructs with 2-aminopurine (2AP) at position −1 of the non-template strand. Construct nomenclature is that of Table I. Enz, enzyme.

of the next substrate nucleotide (and translocation by 1 more base), but without covalent bond formation.

Collapse of the Initially Melted Region—The DNA sequence TATA at positions −4 through −1 of the promoter is melted in the initial open complex (6, 7, 13). Placement of 2-aminopurine at position −2 of the template strand provides a probe of the melted state of this region, as shown in earlier studies (13, 17, 18). A set of data showing closing of the bubble is presented in Fig. 1. Different DNA constructs were prepared, which in the presence of GTP and ATP allow the steady-state accumulation of complexes stalled at various positions (indicated by the upper member of each pair of numbers along the y axis). Subsequent addition of 3′dCTP allows translocation by one nucleotide (as indicated by the lower number of each pair). The results clearly that in complexes translocated to position +7 or +8, the initially melted region (positions −2 and −1) remains melted, as evidenced by the high fluorescence relative to the double-stranded control. In complexes stalled at position +9, on each of two different constructs, the fluorescence is intermediate between that of double- and single-stranded DNA. Finally, on translocation to position +10, the fluorescence is close to that of the fully double-stranded control. The latter result shows that by position +10, the TATA region, as represented by probes at positions −2 and −1, has collapsed. In contrast, on translocation up to position +8, the initially melted region remains melted, consistent with the retention of promoter contacts.

Looking more carefully at the transition itself, these data are consistent with one of three, more detailed, models. In one model, on translocation to position +9, the complexes are homogeneous, but the stacking interactions at position −2 have been altered to yield an intermediate level of fluorescence. In a clearly distinct model, on translocation to position +9, the initially melted region collapses fully in −40–50% of the complexes, with near complete collapse occurring on translocation to position +10. A third model suggests that on stalling at position +9, there is a distribution of translocational complexes. In this most likely scenario, the complexes are distributed between those that have incorporated the base at +9 but have not moved forward and those that have moved forward to expose the elongating base in the template strand. In this scenario, collapse of the bubble would occur on translocation to position +10.

Measurement of Turnover—A potential issue in these types of studies involves the fact that walked complexes can undergo turnover (returning to the initial promoter positioning) during the measurement, such that steady-state measurements under any given set of substrates will yield a distribution of complexes, including complexes poised for initiation (19). It is important to note that the time scale of the fluorescence measurement (nanoseconds) is much faster than potential translocational events, such that this approach provides an accurate measurement of any such distribution. Direct fluorescence measurements can easily be carried out within 3–4 min of addition of the substrate NTPs, limiting overall turnover (11). To assess the extent of turnover within this time frame, steady-state measurements of RNA synthesis were made under conditions identical to the fluorescence studies. The results, summarized in Table II, show that turnover (maximally 1 min−1 for complexes stalled at position +7) is slower than the rate of initiation, providing evidence that the expected stalled complex should be the predominant species in the fluorescence measurements. These results also demonstrate that, as expected, the turnover rate steadily decreases with increasing position of the stall. Given that the sequence context at the stall site, as well as the initial few bases of the transcript, remain constant across these constructs, these results also demonstrate that the inherent stability of the ITC increases steadily as the enzyme translocates from position +7 to +10.

Mapping the Downstream Edge of the Bubble—Our previous study of an elongation complex stalled at position +15 showed that the downstream edge of the melted bubble lies very close to the last incorporated base (11). We now ask whether this is true in complexes that have just cleared the promoter (as evidenced by collapse of the upstream TATA bubble). The data in Fig. 1 follow fluorescence from probes near the start site, in complexes stalled 7–11 nucleotides downstream. In these same stalled complexes, we can use fluorescence from probes placed just downstream of the stall site. The results presented in Fig. 2A show that in complexes stalled at position +10, fluorescence from a probe placed at position +11 on the non-template strand increases slightly, indicating only minor melting at this position just 1 base downstream of the stall site. Walking these complexes forward to position +15, to place the probe fully within the expected bubble (position −5 relative to the stall site), yields an increase in fluorescence to single-stranded levels, consistent with full opening.

TABLE II
Characterization of turnover on templates encoding different length RNAs under conditions that mimic the fluorescence measurements
As described under “Materials and Methods,” 1.0 μM polymerase and 1.0 μM double-stranded promoter were allowed to react for 3 min at 25°C in the presence of 1000 μM GTP, 1000 μM ATP, and [γ32P]GTP, conditions that mimic those in the fluorescence measurement. Reactions were quenched and electrophoresed, and bands were quantified. The concentrations of RNA (R) of stall length are presented in column GA. In an identical experiment, at 3 min, 3′-dCTP was added to a final concentration of 1000 μM, and the reaction was incubated for another 3 min. The concentrations of RNAs of length n and n + 1 are presented in the columns GA + 3′-dC. On all of these templates, abortive products beyond 3 mer were minimal. Construct nomenclature is that of Table I.

<table>
<thead>
<tr>
<th>DNA construct</th>
<th>nGA</th>
<th>GA + 3′-dC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[R]</td>
<td>[R]</td>
</tr>
<tr>
<td>2AP−2GA(7)</td>
<td>+7</td>
<td>2.0</td>
</tr>
<tr>
<td>2AP−2GA(8)</td>
<td>+8</td>
<td>1.7</td>
</tr>
<tr>
<td>2AP−2GA(9)</td>
<td>+9</td>
<td>1.0</td>
</tr>
<tr>
<td>2AP−2GA(10)</td>
<td>+10</td>
<td>0.8</td>
</tr>
</tbody>
</table>
Promoter Clearance by T7 RNA Polymerase

**Fig. 2.** Mapping the downstream edge of the bubble in complexes stalled at position +10, following the fluorescence quenching of pyrrolo-dC placed at position +11 of the nontemplate strand (A) and at position +13 of the template strand (B). Complexes were stalled as described in the legend to Fig. 1; construct nomenclature is that of Table I. Note that in these experiments, (ribo)CCTP was used, such that subsequent addition of UTP allowed full-length runoff synthesis. dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; Enz, enzyme.

In a complementary experiment presented in Fig. 2B, placing the label in the template strand at position +13 and walking to position +10 yields no increase in fluorescence, consistent with the retention of duplex DNA downstream. Interestingly, walking this stalled complex to position +15 also results in no increase in fluorescence. In this case, the DNA must be melted, but the probe in the template strand is now in an RNA-DNA heteroduplex.

**Simultaneous Placement of Fluorescent Probes**—One cannot, from the above, rule out the possibility that some stalled complexes, at position +8, for example, might transiently backtrack to the promoter, thus yielding a distribution of products at steady state. In other words, the actual footprint of an individual enzyme on the DNA might be smaller than implied by the data in Fig. 1 but reflect a steady-state scenario in which some complexes are situated at the promoter with their -2 base pairs open and their +8 base pairs closed, whereas others are situated at the +8 stall site, with their -2 bases closed and their +8 bases open. One might also worry that the introduction of base analogs in two separate constructs results in different behavior for the two constructs.

To address these concerns quantitatively, we prepared a series of DNA constructs in which 2-aminopurine was placed at position -2 of the template strand, whereas pyrrolo-dC was simultaneously placed immediately downstream of the stall site on the nontemplate strand. Excitation at 315 nm with observation at 370 nm yielded a signal from 2-aminopurine, whereas excitation at 350 nm with observation at 460 nm yielded a signal from 2-aminopurine placed at position -2 and +8, inconsistent with transient sliding of a (smaller) fixed length bubble.

**Fluorescence changes in complexes containing two fluorophores:** 2-aminopurine at the start site and pyrrolo-dC 1 base downstream of the initial stall site. Three DNA constructs were used, allowing initial stalling at positions +7–9. Each of these can be walked by subsequent addition of 3’dCTP, followed by UTP. A, fluorescence of 2-aminopurine in the initially melted region. B, fluorescence from pyrrolo-dC 1 base downstream of the initial (GA) stall site, as described in Table I. dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; Enz, enzyme.

This result can be used to our advantage in assessing the state of the bubble and of the heteroduplex. Correct pairing of the probe pyrrolo-dC in the template strand opposite guanine in the original DNA duplex, should give low fluorescence when the probe is in either a DNA:DNA duplex or in an RNA:DNA heteroduplex. Fluorescence will be high only when the template strand is in a single stranded environment. Construction of an identical DNA construct, with the exception that pyrrolo-dC is mispaired opposite adenine, provides complementary data. In this case, fluorescence will be low in the (correctly paired) RNA:DNA heteroduplex, but formation of the (mispaired) DNA:DNA duplex should show fluorescence intensities higher than that of the same probe in single stranded DNA. Presumably, the mismatch imposes constraints on stacking not present in the single strand.

Fluorescence is likely the true measure of a fully melted (at the stall site) complex, such that close to 100% of the complexes are stalled as anticipated. Thus, under conditions designed to stall transcription at position +8, the majority of the complexes have melted DNA at positions -2 and +8, inconsistent with transient sliding of a (smaller) fixed length bubble.

**A Valuable Tool: A Mismatched Probe Has Unusually High Fluorescence**—The changes in fluorescence intensity observed in 2-aminopurine are thought to arise from changes in the stacking interactions surrounding the fluorophore (20–22). In duplex DNA, the bases are well stacked, providing for efficient quenching of fluorescence. In contrast, in single-stranded DNA, stacking is still present but is presumably less well ordered. These results have been used to probe melted regions within DNA, both for 2-aminopurine and, more recently, for pyrrolo-dC (11, 13, 17, 23). The results shown in Fig. 4 demonstrate that the fluorescence of pyrrolo-dC in an engineered mismatched setting (opposite adenine) is higher than that of the same probe in single stranded DNA. Presumably, the mismatch imposes constraints on stacking not present in the single strand.

Peeling Away of the 5’ End of the RNA—To determine at what point the 5’ end of the RNA initially separates from the heteroduplex and the DNA reanneals, we placed fluorescent base analogs at positions +1 and +2, either correctly paired opposite guanine or mispaired with adenine in the nontemplate strand. The results presented in Fig. 5, A and C, suggest that with one possible exception, probes placed at position +1 or +2 of the template strand remain in a duplex environment as the polymerase translocates from positions +10–12. The fluorescence is close to that of the control DNA duplex and signifi-
DISCUSSION

All RNA polymerases show a transition from an abortive cycling phase to a more stable elongation phase at ~10 base pairs, a region midway between the initial promoter-bound complex and an elongation complex stalled clear of the promoter (24). Understanding this transition is critical to understanding both the fundamental machinery of transcription and the various regulatory processes that exploit this phenomenon. That the transition occurs at about the same position in both the multisubunit eukaryotic and prokaryotic polymerases and in the much smaller single-subunit RNA polymerase (27). In particular, 2-aminopurine and pyrrolo-dC placed individually at specific positions along the DNA, we now present a thorough study following the transition of the initially elongating complex. In separate experiments, probes were monitored at various positions from -2 to +13. The data are summarized in Fig. 6.

As summarized pictorially in Fig. 6, the initially melted bubble (as reported by probes at positions -2 and -1 on the template and nontemplate strands, respectively) remains single-stranded through translocation to position +8. As the complex translocates through position +9 to position +10, the initially melted upstream edge of the bubble reanneals. This indicates that a maximal bubble size of 10–12 nucleotides is reached on translocation to position +8. The collapse of the bubble at this point suggests a dramatic reorganization of the DNA on the protein. Retention of the upstream specificity contacts (Fig. 6, yellow) as the TATA region (~4 to ~1) resumes a duplex topology would further strain the template strand connectivity. Thus we predict that promoter contacts are lost on translocation beyond position +8.

Mismatched Fluorescent Base Analogs: A New Tool—The quenching of the fluorescence of base analogs in DNA is thought to arise primarily from neighbor stacking interactions (22). Such interactions are substantial in single-stranded DNA but increase significantly in the more ordered environment of a DNA duplex. The results presented here show that quenching is similarly higher in an RNA-DNA heteroduplex relative to the control single-stranded DNA. This is useful for probing template strand involvement in a hybrid but presents a problem in that one cannot readily distinguish an RNA-DNA duplex from a DNA-DNA duplex. The new observation that quenching is generally less in an otherwise duplex mismatch than it is in single-stranded DNA presents a tool to allow one to distinguish between an RNA-DNA duplex and a reannealed homoduplex.

In the case of the placement of the fluorescent base analog in the template strand opposite a correctly matched base in the nontemplate strand, low fluorescence indicates the presence of either a homo- or a heteroduplex. For a parallel experiment in which the base analog in the template strand is mismatched to an incorrect base in the nontemplate strand, low fluorescence can arise only from a (correctly paired) heteroduplex.

The result that quenching in a mismatched but otherwise duplex environment is less than that in single-stranded DNA is consistent with thoughts about the structures of mismatched DNA. Combined with parallel measurements in a correctly paired construct, this approach provides a powerful tool to probe for the presence of a local heteroduplex in a transcription bubble.

Characterization of Complexes Transiting from the Abortive to the Processive Phase—Using both 2-aminopurine and pyrrolo-dC placed individually at specific positions along the DNA, we now present a thorough study following the transition of the initially elongating complex. In separate experiments, probes were monitored at various positions from -2 to +13. The data are summarized in Fig. 6.

As summarized pictorially in Fig. 6, the initially melted bubble (as reported by probes at positions -2 and -1 on the template and nontemplate strands, respectively) remains single-stranded through translocation to position +8. As the complex translocates through position +9 to position +10, the initially melted upstream edge of the bubble reanneals. This indicates that a maximal bubble size of 10–12 nucleotides is reached on translocation to position +8. The collapse of the bubble at this point suggests a dramatic reorganization of the DNA on the protein. Retention of the upstream specificity contacts (Fig. 6, yellow) as the TATA region (~4 to ~1) resumes a duplex topology would further strain the template strand connectivity. Thus we predict that promoter contacts are lost on translocation beyond position +8.

The Initial Heteroduplex Reaches a Maximal Length of ~10 Base Pairs—Through translocation to position +10, beyond the point at which the initial bubble has collapsed, probes placed in the template strand at positions +1 and +2 remain in a heteroduplex, indicating that RNA is bound to DNA all the way back to its 5’ end (the RNA has not yet begun to dissociate). As the complex progresses to position +11, the RNA at position −1 (the 5’ end) begins to peel away. On translocation to position +12, the RNA base at position +2 relative to the start site dissociates (position +2 is no longer a heteroduplex), and the template strand base at position +1 returns to a DNA-DNA duplex.

These results suggest a smooth and sequential peeling off of the RNA and indicate that the DNA bubble collapses very close behind the exiting RNA, consistent with maintaining optimal melting energetics. At this point, the heteroduplex is expected to be ~10 base pairs in length (which, with data to date,
appears to be an upper limit for the size of the heteroduplex). Interestingly, similar probes of complexes stalled near position +15 suggest a maximal heteroduplex length of 8 (9, 11), suggesting a minor transition in the nature of the bubble between translocational positions +11 and +15.

The introduction of a mismatched base pair at position +1 or +2 raises the possibility that the peeling away of the 5' end of the RNA, if driven by reannealing of the DNA duplex, might occur later (at a slightly longer length of heteroduplex) in these constructs than in normal transcription. Preliminary results indicate a similar peeling away on constructs completely lacking the nontemplate strand, suggesting that heteroduplex length is limited by other factors.

Conclusions—These results for the single-subunit T7 RNA polymerase are consistent with those from some recent studies.

of the more complex multisubunit enzymes. The elongation complexes of both E. coli RNA polymerase and yeast Pol II are thought to contain a 8–9-bp hybrid with the bubble melted only slightly ahead of the stalled complex (2, 28–30), very similar to the elongation complex in T7 RNA polymerase revealed by recent fluorescence and biochemical studies (9, 11).

The current study of the ITC in T7 RNA polymerase suggests that during this initial phase, the heteroduplex can reach a length slightly larger, up to 10 nucleotides, and that the bubble (extending upstream of the start site) can reach a maximum length of 10–12 nucleotides. These results are very similar to those from a permanganate footprinting study of eukaryotic Pol II stalled on the adenovirus major late promoter (31). In the much larger eukaryotic system, the downstream edge of the bubble similarly extends uniformly, whereas the initial upstream bubble (positions −9 to −2) remains open. The latter closes abruptly on translocation to position +11, very similar to the transition observed here near position +10. Given the very different structures (and sizes) of the RNA polymerases and that fact that Pol II requires ATP hydrolysis for initial melting, whereas T7 RNA polymerase does not, this suggests that the transition near position +10 is a fundamental feature of the torsional events occurring in the nucleic acids.

These results for the ITC are consistent with expectations of the stressed intermediate model of promoter clearance (32). The initial promoter interactions are retained through translocation to position +8, as was recently observed in the crystal structure of T7 RNA polymerase stalled at position +3 (7). The results are also consistent with a DNA scrunching model, in which the accumulation of DNA and a heteroduplex upstream of the active site ultimately leads to collapse of the bubble and loss of promoter contacts; however, they are not consistent with the proposal that the heteroduplex is at most 3 base pairs in length (7). The 5′ end of the RNA does not begin to peel away until the enzyme has translocated to position +11, indicating that heteroduplex separation does not occur until after the bubble has collapsed.

These fluorescence snapshots of complexes stalled along the pathway from open complex through promoter clearance provide detailed and site-specific information on the nature of the bubble and the heteroduplex. The probes are relatively nonperturbing and very local in their reporting and therefore provide a powerful approach to the study of movement in transcription. The results presented here with a very simple single-subunit RNA polymerase provide a picture similar in detail to that of the behavior of much larger, multisubunit RNA polymerases, suggesting a common fundamental mechanism.

REFERENCES