doi:10.1006/jmbi.2001.4601 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **308**, 465–475

JMB

Fluorescence Characterization of the Transcription Bubble in Elongation Complexes of T7 RNA Polymerase

Cuihua Liu and Craig T. Martin*

Department of Chemistry, University of Massachusetts, Amherst, MA 01003, USA

*Corresponding author

The various kinetic and thermodynamic models for transcription elongation all require an understanding of the nature of the melted bubble which moves with the RNA polymerase active site. Is the general nature of the bubble system-dependent or are there common energetic requirements which constrain a bubble in any RNA polymerases? T7 RNA polymerase is one of the simplest RNA polymerases and is the system for which we have the highest-resolution structural information. However, there is no high-resolution information available for a stable elongation complex. In order to directly map melted regions of the DNA in a functionally paused elongation complex, we have introduced fluorescent probes site-specifically into the DNA. Like 2-aminopurine, which substitutes for adenine bases, the fluorescence intensity of the new probe, pyrrolo-dC, which substitutes for cytosine bases, is sensitive to its environment. Specifically, the fluorescence is quenched in duplex DNA relative to its fluorescence in single-stranded DNA, such that the probe provides direct information on local melting of the DNA. Placement of this new probe at specific positions in the non-template strand shows clearly that the elongation bubble extends about eight bases upstream of the pause site, while 2-aminopurine probes show that the elongation bubble extends only about one nucleotide downstream of the last base incorporated. The positioning of the active site very close to the downstream edge of the bubble is consistent with previous studies and with similar studies of the promoter-bound, pre-initiation complex. The results show clearly that the RNA:DNA hybrid can be no more than eight nucleotides in length, and characterization of different paused species suggests preliminarily that these dimensions are not sequence or position dependent. Finally, the results confirm that the ternary complex is not stable with short lengths of transcript, but persists for a substantial time when paused in the middle or at the (runoff) end of duplex DNA.

Keywords: base analog; elongation; melting; translocation; transcription bubble

Introduction

Since RNA polymerases do not carry out distributive synthesis, the stability of the enzyme-DNA-RNA complex during the elongation phase of transcription is essential to the production of full-length RNA transcripts. The stability of such complexes in both bacterial and eukaryotic RNA polymerases is proposed to follow from (heteroduplex) interactions between the transcript and the template strand of the DNA, protein-RNA interactions, and from protein ''clamp'' interactions with the downstream DNA. In the smaller bacteriophage T7-family RNA polymerases, structural and footprinting evidence suggests that the latter interaction is small, if significant at all (Ikeda & Richardson, 1986; Gunderson *et al*., 1987; Ikeda & Richardson, 1987; Chapman *et al*., 1988; Huang & Sousa, 2000). Thus in the T7 system, most of the stability must arise from interactions within the elongation bubble.

Abbreviations used: EDTA, ethylenediaminetetraacetic acid.

E-mail address of the corresponding author: CMartin@chem.umass.edu

0022-2836/01/030465–11 $35.00/0

Transcription initiation in the T7 RNA polymerase system has been well characterized biochemically (McAllister, 1993; Li *et al*., 1996; Sousa, 1996; Guajardo & Sousa, 1997; Weston *et al*., 1997; Rong *et al*., 1998), and the recent publication of crystal structures for both a binary polymerase-promoter complex and an initially transcribing (three-base) ternary complex (Cheetham *et al*., 1999; Cheetham & Steitz, 1999) provides an exciting opportunity to begin to understand the relationship between structure and function in this simple model RNA polymerase. Despite its small size, T7 RNA polymerase shows all of the fundamental features characteristic of RNA polymerases, including specific initiation, an early, unstable, abortive complex, and its conversion to a stably elongating ternary complex.

Understanding the structure of the elongation complex, including the size of the bubble and the length of the heteroduplex, is essential to the various energetic models for the elongation mechanism. There has been recent controversy over the size of the transcription bubble during elongation. While biochemical studies in the T7 system have suggested a bubble size of up to ten nucleotides (Osterman & Coleman, 1981; Tyagarajan *et al*., 1991; Sastry & Hoffman, 1995; Gopal *et al*., 1999; Huang & Sousa, 2000), modeling extended from the crystal structure has led to a proposal that the heteroduplex can be no more than about three nucleotides in length (Cheetham & Steitz, 1999). This parallels a similar controversy in *Escherichia coli* RNA polymerase (Nudler *et al*., 1997; Uptain *et al*., 1997; Nudler *et al*., 1998; Milan *et al*., 1999). While the smaller phage polymerase holds little or no structural homology with its prokaryotic and eukaryotic counterparts, if fundamental principles of nucleic acid topology and stability govern the design of an RNA polymerase, one might expect key features of the various elongation complexes to be similar (convergently evolved). In particular, von Hippel has proposed that heteroduplex energetics is a key component of the overall stability of an elongation complex (Wilson *et al*., 1999). Thus, a minimal heteroduplex length (and therefore, bubble size) is proposed to be essential for a stable elongation complex. Here, we probe the size of the elongation bubble in the T7 enzyme.

Although complexes containing RNA shorter than about nine or ten nucleotides cannot be stably isolated, a recent study has shown that complexes paused at positions 10 to 14 nucleotides from the start site are much more stable (Mentesana *et al*., 2000). A recent probing of similarly paused elongation complexes, using nucleases to estimate the footprint of the enzyme on the DNA and potassium permanganate to probe unpaired DNA bases within the bubble, provides evidence for a seven base-pair heteroduplex, with about a nine base open bubble (Huang & Sousa, 2000). The footprinting results in that study also illustrate clearly that probes which bind directly to the DNA to exert their effect, such as nucleases (and methydiumpropyl-EDTA), have the potential of perturbing the system they are attempting to observe, and such perturbations might be expected to show sequence-dependence. Recent photo crosslinking results similarly indicate an eight base-pair heteroduplex (Temiakov *et al*., 2000). Fluorescent base analogs, in contrast, report directly on their environment with minimal perturbation of the system (and functional controls can assess the impact of any perturbations arising from analog substitutions). Fluorescent probes have been used recently to map melted regions of both T7 and *E. coli* RNA polymerase, bound to their respective promoters (Újvári & Martin, 1996; Sullivan *et al*., 1997; Helmann & deHaseth, 1999; C.L. & C.T.M., unpublished results), and the conclusions from these studies in the T7 system have been confirmed by the recent high-resolution crystal structures of enzyme-promoter complexes (Cheetham *et al*., 1999; Cheetham & Steitz, 1999). More recent fluorescence measurements downstream of the start site have shown that the open complex at initiation encompasses about seven or eight base-pairs, from position −5 upstream to about position +3 downstream (C.L. & C.T.M., unpublished results).

Here, we use a new fluorescent base analog to probe stably stalled elongation complexes. The cytidine analog pyrrolo-dC (see Figure 1) is highly fluorescent, with excitation and emission maxima (350 nm and 460 nm, respectively) far from those of DNA and protein, making it ideal for probing protein-nucleic acid interactions. It pairs well with

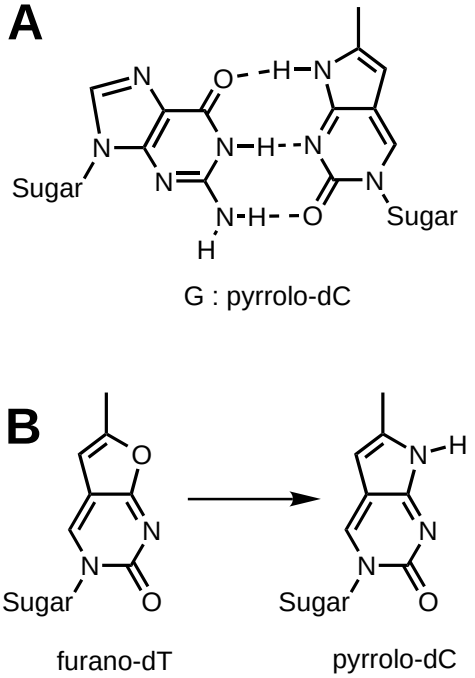

Figure 1. (a) During deprotection of a synthetic oligonucleotide, incorporated furano-dT (left) is quantitatively converted to pyrrolo-dC (right). (b) Structure of pyrrolo-dC (right) base-paired with guanine.

guanine, and like 2-aminopurine (Xu *et al*., 1994; Xu & Nordlund, 2000), shows reduced fluorescence in duplex DNA relative to its fluorescence in single-stranded DNA. This quenching of fluorescence can be used to monitor local DNA melting.

To characterize the transcription bubble in elongation complexes of T7 RNA polymerase, a series of constructs containing pyrrolo-dC or 2-aminopurine at individual positions on either the non-template or template strands has been synthesized. In addition, slight variations have been introduced in the spacing at the pause site in order to control for position or sequence-dependent effects. Parallel biochemical assays have been used to confirm that the species observed in the fluorescence measurements are in fact stably paused where expected and that the paused species can subsequently be elongated. The results indicate a melted bubble slightly larger than that at initiation (although translocated away from the start site), with a distribution of melting similarly biased upstream from the pause site. The results are consistent with a hybrid size of about eight base-pairs. These fluorescence probes now open the door to stopped-flow fluorescence studies to measure the structural dynamics of chased stalled elongation complexes.

Results

Pyrrolo-dC (Figure 1) is an ideal fluorescence probe in that excitation can be carried out far from the protein and nucleic acid absorbances. As shown in Figure 2(a), the excitation profile shows two maxima, at 260 nm and 350 nm. In the observation of melting in stalled elongation complexes presented here, excitation is exclusively at 350 nm. The emission profiles for excitation at each of these wavelengths are shown in Figure 2(b). Most importantly, as illustrated below, the fluorescence intensity is always observed to be lower in duplex DNA and higher in the single-stranded form. As for 2-aminopurine, this feature can be used to monitor polymerase-induced local melting. The extinction coefficient of the analog at 350 nm is about 5900 M^{-1} cm^{-1} (John Randolph, personal communication) and the quantum yield for the nucleoside free in solution is 0.2, which drops to about 0.03 on formation of a duplex (paired with guanine). The fluorescence typically increases about twofold in the single-stranded form.

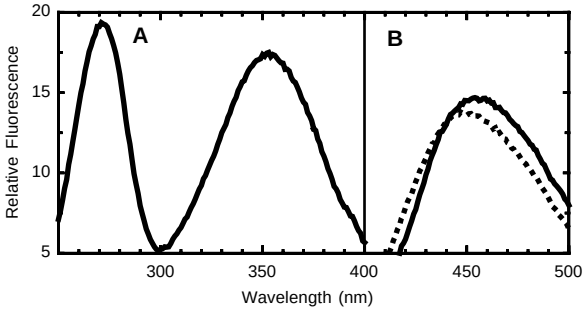

Figure 2. Representative fluorescence spectra of pyrrolo-dC in single-stranded DNA. This sample contained pyrrolo-dC placed at position +23 (relative to the start site) in the single-stranded DNA sequence shown in Figure 3. (a) Excitation spectrum observed at 460 nm. (b) Emission spectra with excitation at 350 nm (continuous line) and 260 nm (broken line). The concentration of the DNA is 1.0 μM, in standard transcription buffer.

Here, we have used a DNA sequence based on the standard ϕ10 promoter. As shown in Figure 3, this sequence allows transcription to extend to position +6 in the presence of GTP and ATP, and to position +15 in the presence of GTP, ATP, and CTP.

Mapping the elongation bubble

For each construct shown in Figure 4, fluorescence from single incorporated pyrrolo-dC or 2-aminopurine analogs was measured under a variety of conditions. In the absence of bound enzyme, the fluorescence is typically low (although still readily detected). Addition of enzyme in the absence of nucleoside triphosphates leads to fluorescence-detected changes at the start site, interpreted as melting (Jia *et al*., 1996; Sastry & Ross, 1996; Újvári & Martin, 1996). Subsequent addition of GTP and ATP allows the enzyme to proceed to position +6, but as demonstrated below, the complex stalled at this position cycles efficiently. Subsequent addition of CTP now allows the enzyme to translocate out of the classic abortive cycling region and form stably stalled complexes at position +15. Finally, addition of UTP allows the enzyme to translocate to the end of the DNA template.

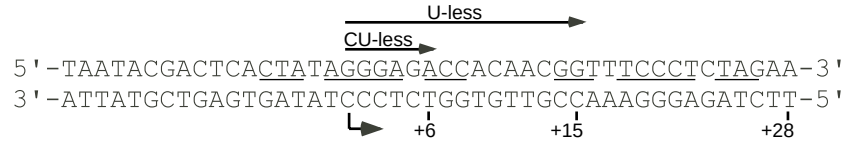

Figure 3. Base sequence used here, identical with the T7 ϕ10 promoter. The transcription start site is indicated by an arrow, and numbering is relative to that start site. The encoded transcript is ''C-less'' to position +6, and ''U-less'' to position +15. Note that the sequence from position −5 to +26 has partial palindromic character (indicated by lines between the two strands).

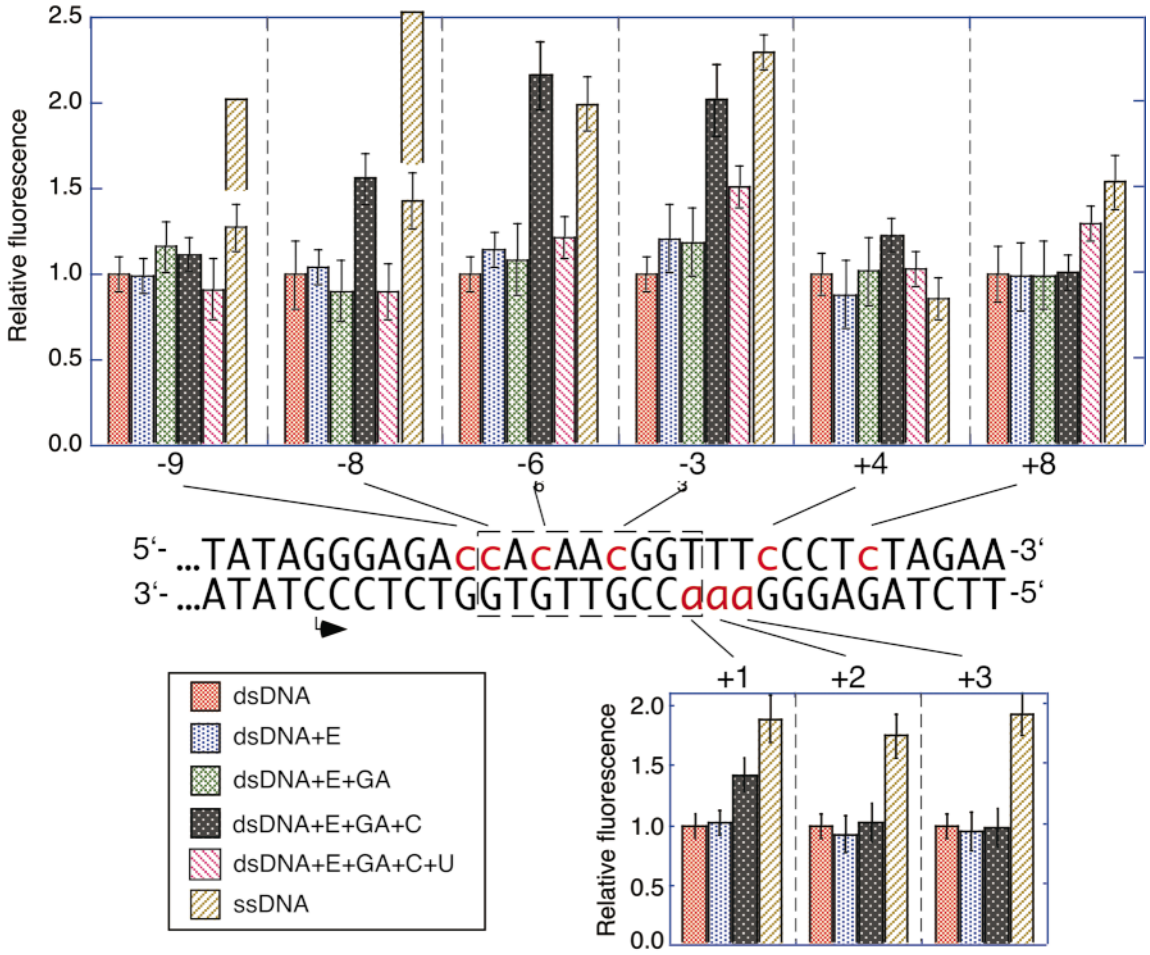

Figure 4. Fluorescence measurements from nine different constructs incorporating pyrrolo-dC (lower case *c*) or 2-aminopurine (lower case *a*) at individual positions around a pause site at position +15 relative to the transcription start site (indicated as position −1 above). Fluorescence measurements for each construct were taken (in order, from left to right within a set) as follows: (1) 1.0 μM double-stranded DNA in transcription buffer, followed by (2) addition of enzyme to 1.0 μM, then (3) addition of GTP and ATP to 1.0 mM each (allows cycling and progression to position +6), (4) addition of CTP to 1.0 mM (allows progression to position +15, with limited cycling), and (5) addition of UTP to 1.0 mM (allows runoff transcription and cycling). The last bar in each set measures fluorescence of the single (non-template) strand containing the fluorophore in the absence of enzyme and substrate. Relative fluorescence values for each group are normalized to the fluorescence in double-stranded DNA for that construct. Each bar value represents the average of three separate steady-state measurements (with standard deviation as indicated). For the first two sets, the extended bars for fluorescence from the single strands are taken from similar constructs which are predicted to adopt less secondary structure, as described in the text.

The primary focus here is the enzyme stalled at position +15. The fluorescence results for this species, with fluorophores reporting on different positions relative to the stall site, are shown in Figure 4. In this case, the numbering, by convention, is relative to the stall site at position +15. All substitutions upstream of the stall site are in the non-template strand, since base-pairing in the template strand will occur within the heteroduplex. In most cases, inclusion of pyrrolo-dC in the non-template strand has no measurable effect on transcription (data not shown). In order to probe melting at the downstream edge of the bubble, 2-aminopurine was placed in the template strand immediately past the stall site.

The increase in fluorescence at positions −8, −6, and −3 which accompanies (only) addition of GTP, ATP, and CTP shows very clearly that these bases are a part of the melted elongation bubble. In contrast, the lack of any substantial change in fluorescence from the analogs at positions −9, +2, +3, +4, and +8 provide evidence that the transcription bubble extends upstream only to position -8 and no further downstream than position +1.

In principle, fluorescence from each isolated, single-stranded oligonucleotide (the last bar in each set of data in Figure 4) should provide fluorescence comparable to that expected in a fully single-stranded melted region (within the open bubble), and for comparison, these measurements are included in the Figure. An unfortunate consequence of the use of the natural ϕ10 promoter sequence is the presence of expected secondary structure in the single-stranded control non-

template DNA (arising from intramolecular pairing with downstream DNA). Thus, as shown in Figure 3, only positions +10 and +13 (numbering here is relative to the transcription start site), which are predicted to lie in a hairpin loop, and position +23, which is predicted to be mismatched, can reasonably be expected to be unpaired in the single-stranded control. This should not, however, affect the comparisons between the double-stranded control and that same construct in binary or ternary complexes (similarly, secondary structure is not expected in the DNA of complexes paused short of the end of the DNA). As described below (see Figure 7), appropriate single-stranded controls have been constructed for the critical positions +7 and +8, and the intensities from those controls are shown extended above the actual control in Figure 4.

In the presence of all four nucleoside triphosphates, the complex shows no evidence of melting near position +15 (as expected, the complex passes rapidly through this region of the DNA such that any one intermediate species does not accumulate to a significant extent). However, the results show that the enzyme does pause at least partially at the runoff end of the DNA (in Figure 4, increased fluorescence from "+8", which now becomes −6 relative to a complex stalled at the end of the template). Fluorescence from a probe four bases farther upstream (at the position labeled "+4" in Figure 4, but now −10 relative to the complex stalled at the end of the DNA) does not increase in the presence of all four nucleoside triphosphates, confirming that the bubble in such a stalled complex does not extend ten bases upstream from the active site. These results suggest that the transcription bubble for a species paused at the end of duplex DNA is comparable in size to a bubble stalled internally.

Biochemical controls

It is important in studies such as this to remember that one is, in principle, observing fluorescence from a steady-state distribution of states. To verify that in the presence of GTP, ATP, and CTP the predominant complex giving rise to the above fluorescence is indeed the complex stalled at position +15, we carried out transcription assays under conditions which parallel the fluorescence measurements. The results presented in Figure 5(a) demonstrate steady-state transcription in the presence of different specific sets of nucleoside triphosphates. In 15 minutes of transcription, the total molar amount of RNA produced in the presence of GTP and ATP is substantially higher than the amount produced in the presence of GTP, ATP, and CTP. This demonstrates that there is substantially more turnover for a complex limited to position +6 than for a complex limited to position +15, consistent with the idea that the former is unstable, while the latter has passed out of the abortive cycling phase and is now a stable species which can be probed. Quantitative analysis of the lanes in Figure 5(a) show a total accumulation of 32 μM 6mer (and very substantial amounts of

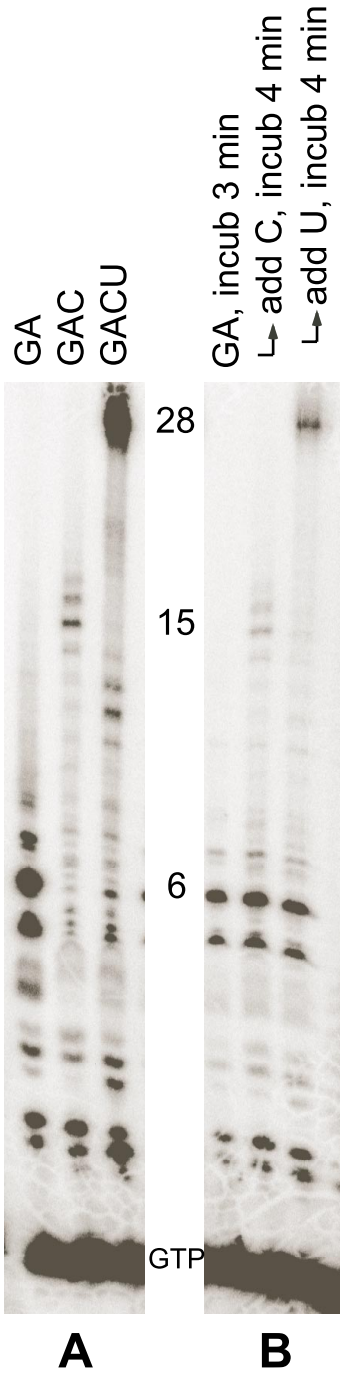

Figure 5. Transcription control for fluorescence measurements. Panel (a) continuous reaction for 15 minutes. Three independent reactions, each run for 15 minutes and quenched. Panel (b) sequential chase reactions. The sample was reacted with GTP and ATP for three minutes, then CTP was added and the reaction run for an additional four minutes. Finally, UTP was added and the reaction was run for an additional four minutes. All reactions were carried out at room temperature, with 1.0 μM each of DNA and polymerase, 1 mM each NTP.

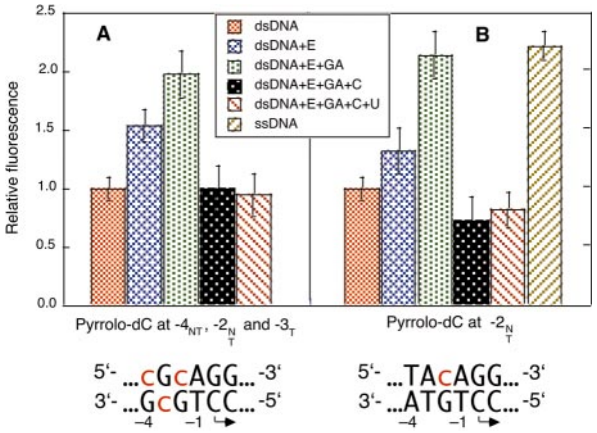

Figure 6. Determination of promoter occupancy in the presence of different subsets of nucleoside triphosphates. Assignment of the bars within a set is as in Figure 4. Panel (a) provides the added sensitivity of a construct containing three simultaneous pyrrolo-dC substitutions at positions −4, −3, and −2, while panel (b) shows the fluorescence from a less perturbed construct containing only a single pyrrolo-dC substitution at position −2. Conditions were as described in Figure 4. Due to the complex nature of construct A (three fluorophores on two different strands), a simple single-stranded control is not presented.

5mer as well) in the presence of GTP and ATP, indicating substantial cycling (production of the 6mer alone is the equivalent of 32 turnovers, or two per minute). In the presence of G, A, and C, the reaction yields only 1.3 μM 15mer (the equivalent of only a single turnover) in 15 minutes, indicating very slow complex dissociation at this length. Finally, in the presence of all four nucleoside triphosphates, the same reaction yields 15 μM full length 28mer, indicating substantially higher turnover than for the species stalled at position +15. For the latter situation, assuming that product release is rate determining (Gopal *et al*., 1999), one expects the predominant species to be the complex stalled at the last base in the template. This is consistent with the observed changes in fluorescence from base analogs expected to lie within the melted bubble of a complex stalled at the duplex terminus.

In studies such as these, it is also important to confirm that the species being probed are functional. To more closely parallel the fluorescence measurements, we carried out transcription reactions in which we halted transcription at position +6 by inclusion of only GTP and ATP, then allowed transcription to proceed to position +15 by subsequent addition of CTP. Finally, we chased the complexes at position +15 to full length by subsequent addition of UTP. The results shown in Figure 5(b) demonstrate that the complexes paused at position +15 can indeed be chased to full length. After three minutes reaction with GTP and ATP, followed by four minutes reaction with CTP, a total concentration of 0.7 μM 15mer is produced corresponding to a turnover of slightly less than one per complex (enzyme and DNA are present at 1.0 μM each). Addition of UTP four minutes later reduces the concentration of 15mer to about 0.2 μM, indicating that the stalled complexes were largely functional. This is confirmed by the data in Figure 4, which show that the fluorescence from analogs at positions −8, −6, and −3 decreases to duplex levels following the addition of UTP. The transcription bubble closes as the complex moves forward.

In summary, unstable halted complexes produce large amounts of the corresponding RNA product, confirming rapid turnover. Stably paused species generate much less of the RNA due to a lack of turnover. Complexes limited to position +6 are relatively unstable, while complexes halted at position +15 (and at the end of the duplex) are very stable, and therefore amenable to biophysical probes.

Fluorescence controls

Returning to fluorescence, the placement of probes within the initially melted region of the promoter (positions −4 to −1) has been used to characterize complexes in the absence of nucleoside triphosphates (Újvári & Martin, 1996; our own unpublished results). We now take a similar approach to demonstrate that upon stalling at position +15, the promoter DNA (from position −4 to −1) has returned to a duplex state (the enzyme has cleared the promoter). The results summarized in Figure 6 show that the non-transcribed region of the promoter remains melted in the presence of GTP and ATP. Under these conditions, the complex cycles rapidly, but never clears the promoter. However, addition of CTP to the reaction allows the complex to clear the promoter, resulting in a return to duplex fluorescence for base analogs positioned within the non-transcribed part of the melted promoter (positions −1 to −4). Subsequent addition of UTP shows that at steady-state the distribution of complexes is distinctly away from the promoter; on this construct, the elongation complex paused at the runoff end of the DNA has a substantial lifetime relative to the rates of reinitiation and promoter clearance.

Fine mapping of the upstream edge of the melted bubble

Finally, in order to account for potential sequence dependence in the fluorescence sensitivity and to determine whether the nature of the upstream edge depends on the position of the complex along the DNA (distance from the start site), we created two new promoter constructs, deleting and adding, respectively, one base-pair immediately upstream of the pause site at position +15 (yielding species which stall at positions +14 and +16 relative to the start site). This results in con-

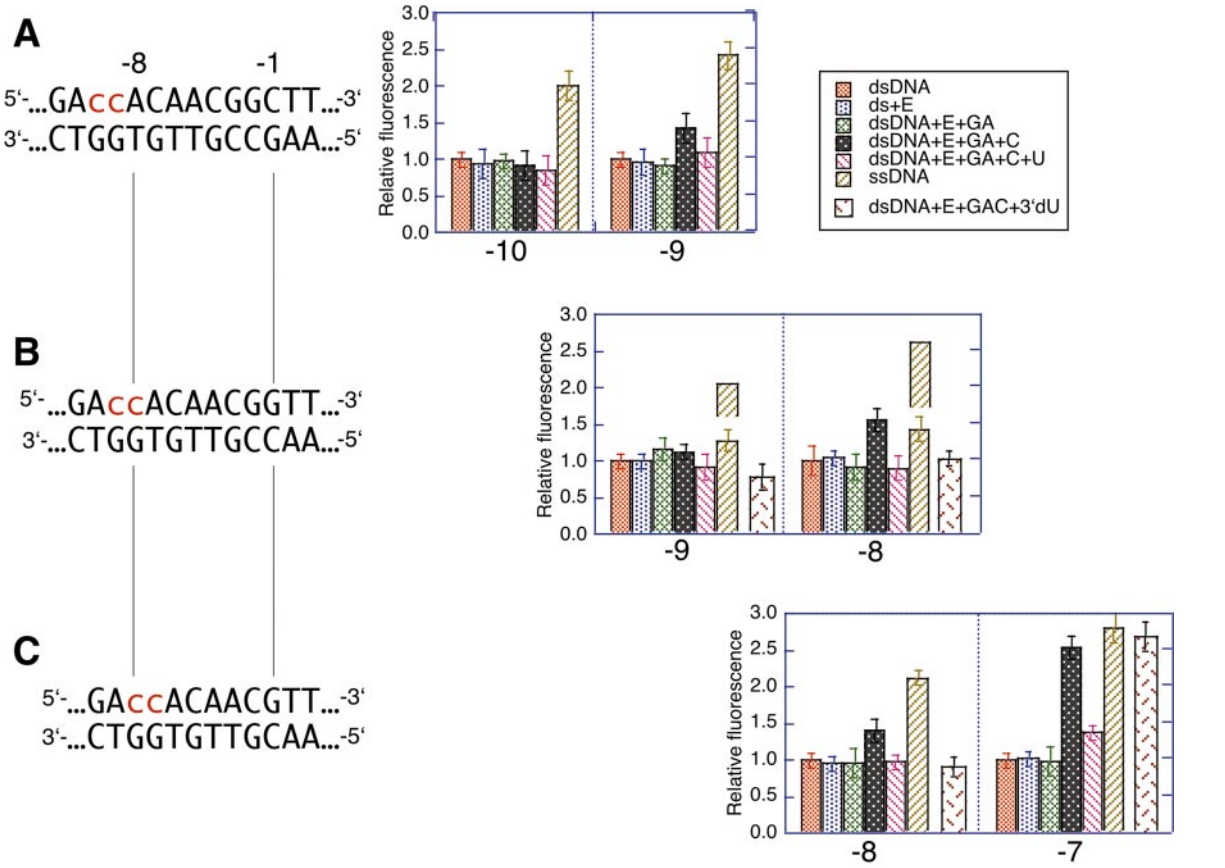

Figure 7. Confirmation of the upstream edge of the elongation bubble, in six different constructs, representing different sequence contexts and positions away from the transcription pause site. Conditions were as described for Figure 4. The first five bars in each panel represent sequential measurements on the same sample, as in Figure 4. For the sample employing 3′-deoxyUTP, enzyme and DNA were incubated, followed by the addition of GTP, ATP, and CTP to 1.0 mM each; 3′-deoxyUTP was then added to a final concentration of 1.0 mM, and the sample was incubated for three minutes.

structs which place the same sequence element, ...ACCA..., at different positions near the upstream edge of the bubble.

Fortuitously, these two constructs also disrupt the secondary structure predicted to exist at the fluorophore in the single-stranded control, as described above. Consequently, for each of the −9 and −8 constructs (panel B in Figure 7, and in Figure 4), the average fluorescence of the corresponding single-stranded controls from the new constructs is plotted as a bar above the actual (improper) ''single-stranded'' fluorescence. The sequence context (...ACCA...) is identical for the original and the two new sequence constructs.

The results summarized in Figure 7 provide multiple experimental views of the upstream edge of the bubble. The results for position −7, like those for position −6 presented in Figure 4, indicate very clearly that this position is melted in a halted complex. In contrast, the lack of any significant change in fluorescence for a probe placed at position −10 indicates that this position remains duplex in the elongation complex.

The results for position −8 show clearly that this position is melted or otherwise perturbed, but the fluorescence increase is smaller and somewhat context sensitive. The results at position −9 show only very slight evidence of a perturbation, and again the fluorescence change is context sensitive. Specifically, for both positions the change in fluorescence is greater for the probe which is at the 3′ C in the sequence 5′-ACCA-3′, suggesting that there is indeed a small environmental effect on the fluorescence intensity changes.

Finally, the addition of 3′-deoxyUTP to a complex stalled at position +15 ought to allow the incorporation of one more base into the transcript and, on this template, possible occupancy of the elongating site, moving the complex forward by one or two bases (Huang & Sousa, 2000). Consistent with this expectation, the fluorescence from probes at position −8 returns to duplex levels, as this site now becomes −9 or −10 relative to the new stall site. Fluorescence from a probe at position −7 relative to the +15 stall site does not decrease upon addition of 3′-deoxyUTP, consistent with its movement one base forward (to become position −8), but not beyond.

In summary, the consistent trend in these measurements is that the fluorescence change at

position −8 is greater than that at position −9, suggesting that this may be the base-pair step near which the DNA is resuming its duplex structure (but see Discussion below). The transcription bubble extends about eight bases upstream of the last incorporated nucleotide.

Discussion

Recent crystal structures of binary and ternary initiation complexes of T7 RNA polymerase bound to its promoter provide a clear definition of the upstream boundary of the open complex (Cheetham *et al*., 1999; Cheetham & Steitz, 1999), and the structures are consistent with a previous fluorescence mapping of that upstream boundary (Újvári & Martin, 1996). The crystal structure contains no information on the nature of the downstream edge of the transcription bubble, but subsequent fluorescence results have defined the downstream boundary (about position +3) in those same promoter-bound complexes (C.L. & C.T.M., unpublished results). Here, fluorescence is used to probe both upstream and downstream boundaries of the open complex in elongating complexes, stalled away from the promoter, a complex for which no crystallographic data exist.

Pyrrolo-dC as a probe of melting

The base analog pyrrolo-dC complements 2-aminopurine in providing a sensitive fluorescence probe of the DNA duplex in protein-DNA interactions. Unlike probes which must bind to the DNA in order to provide information (large molecules like nucleases or small molecules like methydiumpropyl EDTA), these probes do not compete with the polymerase for binding to DNA and so should not bias the equilibrium placement of the enzyme on the DNA. This is particularly important for a complex which might slide back and forth along the DNA with only moderately low energetic barriers (an elongation complex, as compared with a promoter-bound complex). Indeed, a recent study of the T7 RNA polymerase system has demonstrated clearly that external probes which bind DNA (nucleases) can effectively ''push'' the halted elongation complex, particularly when the site for the next incoming nucleoside triphosphate is unoccupied (Huang & Sousa, 2000).

An additional practical advantage specific to pyrrolo-dC as a fluorescence probe is that the optimal excitation and emission wavelengths are at substantially higher wavelengths than previous fluorescent base analogs. In this case, background subtraction of controls to correct for fluorescence from the protein and DNA are largely unnecessary, allowing for more rapid and accurate measurements of fluorescence changes from the uniquely positioned probes.

The melted bubble in an internally halted complex

Both spectroscopic and biochemical controls provide evidence that the complex limited to position +6 (''C-less'') is unstable and turns over rapidly, as observed previously (Mentesana *et al*., 2000). The turnover rate in the presence of only GTP and ATP is about two per minute at room temperature, and this number represents a lower limit for the complex dissociation, as production of shorter abortive products will decrease the observed rate of 6mer production.

The complex stalled at position +15 (''U-less'') is long-lived (biochemical assays show a turnover rate of only 0.003-0.04 min^{-1} (Gopal *et al*., 1999; Mentesana *et al*., 2000), and therefore is amenable to steady-state probes of structure. Fluorescence from probes placed at the initiation site verify that the complex has cleared the untranscribed region of the promoter (specifically, that the bubble in that region has collapsed) and that the steady-state distribution lies well away from the promoter. Indeed, recent exonuclease probes of similarly paused constructs confirm this general conclusion (Huang & Sousa, 2000). The results presented in Figure 4 indicate a melted bubble located from position −8 to about position +1, relative to the last incorporated nucleotide (−1). To refine our view of the upstream edge of the bubble, and to check for sequence-specific effects on fluorescence, a single nucleotide was either inserted or deleted at the (downstream) pause site, such that the position of the same upstream sequence was shifted forward and backward by one base relative to the pause site. This has the effect of placing the same probes in the same sequence context, but at different positions relative to the halt site. The results confirm the preliminary conclusion that the upstream edge of the bubble sits about eight base-pairs upstream. Large changes in fluorescence are observed for probes placed at positions −6 and −7, with intermediate changes for probes at position −8, and a small, barely significant change for a probe located at position −9.

These results are further supported by the addition of 3′-deoxyUTP, which allows the incorporation of one more nucleotide to the RNA, pushing the complex forward by one. Fluorescence from a probe placed at the original position −8 returns to duplex levels, as this now becomes position −9. On these constructs, which encode TTT following the stall site, 3′-deoxyUTP can also bind at the elongation site, potentially allowing a two-base translocation. Assuming that the lagging edge of the bubble follows the leading edge in this case, our results suggest preliminarily that the elongating site is not well occupied, since the original position −7 (which would become position −9) shows no sign of bubble collapse.

Two different effects should be considered in the interpretation of the fluorescence from probes placed at the expected edges of the bubble. First,

biochemical assays show some synthesis of RNA one base longer than that predicted, consistent either with misincorporation or non-templated addition of an additional base in the RNA. If the former were true, then a subpopulation of complexes might have advanced by one base. This could explain the slightly lower change in intensity for a probe placed at position −8, as the bubble might be beginning to close in those species (in other words, for that small population, the probe would now be at position −9). A second consideration concerns the underlying cause of the decreased fluorescence intensity in duplex DNA. As for 2-aminopurine, the quenching of fluorescence likely arises from stacking interactions with adjacent bases (Xu & Nordlund, 2000). Thus the decrease in fluorescence at position −9 might arise from the unstacking (melting) of the base-pair at position −8, while the base-pair at position −9 remains intact. In any of these scenarios, eight or nine nucleotides would be an upper limit for the size of the melted bubble in stalled elongation complexes. This, of course, would similarly restrict the size of the heteroduplex. These results are consistent with recent nuclease mappings of a halted complex (Huang & Sousa, 2000) and with photocrosslinking probes of the heteroduplex (Temiakov *et al*., 2000).

Finally, it has recently been proposed that stalled elongation complexes of T7 RNA polymerase, in the absence of the next incoming nucleoside triphosphate, can readily slide backwards (Huang & Sousa, 2000). While an approaching exonuclease may well be able to ''push'' the complex back, the relative sharpness and constant size of the bubble, as determined here, suggests that in the absence of external perturbations, the steady-state population of complexes back-translocated is very low. The current results could be consistent with (non-transient) sliding by at most one base.

The placement of context-sensitive fluorescent probes at specific positions within the DNA allows for the sensitive detection of structural changes in static complexes, as observed here, but in future studies will also prove useful in measuring the dynamics of bubble movement (Jia *et al*., 1996; Újvári & Martin, 1996).

Materials and Methods

Enzyme

T7 RNA polymerase was prepared from *E. coli* strain BL21 carrying the overproducing plasmid pAR1219 (kindly supplied by F. W. Studier), in which RNA polymerase is expressed under inducible control of the *lac* UV5 promoter (Davanloo *et al*., 1984). The enzyme was purified and the concentration was determined ($\varepsilon_{280} = 1.4 \times 10^5\ M^{-1}\ cm^{-1}$) as described (King *et al*., 1986). Purity of the enzyme (>95%) was verified by SDS-denaturing polyacrylamide gel electrophoresis.

Synthetic oligonucleotides

Oligonucleotides were synthesized by the phosphoramidite method on an Expedite 8909 DNA/RNA synthesizer. The standard phosphoramidites were purchased from CPG Inc. Furano-dT phosphoramidite was obtained from the Glen Research Corporation and coupled normally, with deprotection under standard conditions (30% ammonium hydroxide at 55 °C for six hours). Under these conditions, the compound is converted quantitatively to pyrrolo-dC, as shown in Figure 1 (John Randolph, personal communication). Fluorescence and transcription measurements indicate that pyrrolo-dC pairs well with guanine, but not with adenine (J. Randolph, C.L. & C.T.M., unpublished results). The results presented here are consistent with that finding.

Single strands were purified using an Amberchrom CG-16C reverse-phase resin as described (Schick & Martin, 1993). Purity of the 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 TE buffer (10 mM Tris (pH 7.8), 1 mM EDTA) at 75 °C for five minutes. The samples were then allowed to cool slowly to room temperature within two to three hours. The constructs were either used immediately or stored at −20 °C.

Steady-state fluorescence measurements

Fluorescence measurements were carried out with a PTI (Photon Technology International) L-format fluorimeter with a 75 watt arc lamp and both emission and excitation monochrometers, using a 75 µl (light path 3 × 3 mm, center 15 mm) ultramicro cell (Hellma). Fluorescence emission from pyrrolo-dC was detected at 460 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 375 nm. All fluorescence experiments were carried out in fluorescence buffer: 30 mM Hepes (pH 7.8), 15 mM magnesium acetate, 25 mM potassium glutamate, 0.25 mM EDTA and 0.05% (v/v) 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 dsDNA (1.0 µM) was first recorded alone, and then after the addition of each of the following: enzyme (to about 1.0 µM final concentration), GTP, ATP, CTP, and UTP (to 1.0 mM each), with each addition occurring at four minute intervals.

Transcription reactions

Transcription reactions were carried out at room temperature for 15 minutes in a volume of 20 µl containing 1 µM each of DNA and enzyme, 1.0 mM GTP [α-^{32}P]GTP (specific activity 800 Ci/mmole, New England Nuclear) using standard transcription buffer (30 mM Hepes (pH 7.8), 0.25 mM EDTA, 15 mM magnesium acetate, 25 mM potassium glutamate and 0.05% Tween-20). Reactions were initiated by the addition of nucleoside triphosphates and stopped by the addition of an equal volume of quenching solution (95% (v/v) formaldehyde, 50 mM EDTA, 0.01% (w/v) bromophenol, 0.01% (w/v) xylene cyanol).

For the step-wise chased transcription reactions, the enzyme and DNA (1.0 μM each) were incubated in transcription buffer (lacking nucleoside triphosphates) for four minutes at room temperature, then divided into three equivalent fractions. GTP, ATP (including [α-^{32}P]GTP) were added to the first fraction, which was then incubated for three minutes at room temperature and quenched. The second fraction was similarly incubated with GTP and ATP for three minutes, then CTP was added to a final concentration of 1.0 mM. This was then incubated for four minutes and quenched. The third fraction was incubated with GTP and ATP for three minutes, then with CTP for four minutes, and then UTP was added to 1.0 mM and reacted for four minutes before quenching. All transcription samples were heated to 95 °C for two minutes and loaded onto 20 % (w/v) polyacrylamide/6 M urea gels. Following electrophoresis, gels were dried and quantified using a Molecular Dynamics Storm 840 phosphorimager.

Acknowledgments

We thank John Randolph of Glen Research, Sterling, Virginia, for communicating mass-spectral data on the effect of deprotection and for many valuable discussions. This work was supported by grant 1R01GM55002 from the National Institutes of Health.

References

Chapman, K. A., Gunderson, S. I., Anello, M., Wells, R. D. & Burgess, R. R. (1988). Bacteriophage T7 late promoters with point mutations: quantitative footprinting and *in vivo* expression. *Nucl. Acids Res.* **16**, 4511-4524.

Cheetham, G. M. & Steitz, T. A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. *Science,* **286**, 2305-2309.

Cheetham, G. M., Jeruzalmi, D. & Steitz, T. A. (1999). Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature,* **399**, 80-83.

Davanloo, P., Rosenberg, A. H., Dunn, J. J. & Studier, F. W. (1984). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl Acad. Sci. USA,* **81**, 2035-2039.

Gopal, V., Brieba, L. G., Guajardo, R., McAllister, W. T. & Sousa, R. (1999). Characterization of structural features important for T7 RNAP elongation complex stability reveals competing complex conformations and a role for the non-template strand in RNA displacement. *J. Mol. Biol.* **290**, 411-431.

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

Gunderson, S. I., Chapman, K. A. & Burgess, R. R. (1987). Interactions of T7 RNA polymerase with T7 late promoters measured by footprinting with methidiumpropyl-EDTA-iron(II). *Biochemistry,* **26**, 1539-1546.

Helmann, J. D. & deHaseth, P. L. (1999). Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. *Biochemistry,* **38**, 5959-5967.

Huang, J. & Sousa, R. (2000). T7 RNA polymerase elongation complex structure and movement. *J. Mol. Biol.* **303**, 347-358.

Ikeda, R. A. & Richardson, C. C. (1986). Interactions of the RNA polymerase of bacteriophage T7 with its promoter during binding and initiation of transcription. *Proc. Natl Acad. Sci. USA,* **83**, 3614-3618.

Ikeda, R. A. & Richardson, C. C. (1987). Interactions of a proteolytically nicked RNA polymerase of bacteriophage T7 with its promoter. *J. Biol. Chem.* **262**, 3800-3808.

Jia, Y., Kumar, A. & Patel, S. (1996). Equilibrium and stopped-flow kinetic studies of interaction between T7 RNA polymerase and its promoters measured by protein and 2-aminopurine fluorescence changes. *J. Biol. Chem.* **271**, 30451-30458.

King, G. C., Martin, C. T., Pham, T. T. & Coleman, J. E. (1986). Transcription by T7 RNA polymerase is not zinc-dependent and is abolished on amidomethylation of cysteine-347. *Biochemistry,* **25**, 36-40.

Li, T., Ho, H. H., Maslak, M., Schick, C. & Martin, C. T. (1996). Major groove recognition elements in the middle of the T7 RNA polymerase promoter. *Biochemistry,* **35**, 3722-3727.

McAllister, W. (1993). Structure and function of the bacteriophage T7 RNA polymerase (or, the virtues of simplicity). *Cell. Mol. Biol. Res.* **39**, 385-391.

Mentesana, P. E., Chin-Bow, S. T., Sousa, R. & McAllister, W. T. (2000). Characterization of halted T7 RNA polymerase elongation complexes reveals multiple factors that contribute to stability. *J. Mol. Biol.* **302**, 1049-1062.

Milan, S., D'Ari, L. & Chamberlin, M. J. (1999). Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase: ribonuclease footprinting of the nascent RNA in complexes. *Biochemistry,* **38**, 218-225.

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

Nudler, E., Gusarov, I., Avetissova, E., Kozlov, M. & Goldfarb, A. (1998). Spatial organization of transcription elongation complex in *Escherichia coli*. *Science,* **281**, 424-428.

Osterman, H. L. & Coleman, J. E. (1981). T7 ribonucleic acid polymerase-promoter interactions. *Biochemistry,* **20**, 4884-4892.

Rong, M., He, B., McAllister, W. T. & Durbin, R. K. (1998). Promoter specificity determinants of T7 RNA polymerase. *Proc. Natl Acad. Sci. USA,* **95**, 515-519.

Sastry, S. & Hoffman, P. (1995). The influence of RNA and DNA template structures during transcript elongation by RNA polymerases. *Biochem. Biophys. Res. Commun.* **211**, 106-114.

Sastry, S. & Ross, B. (1996). A direct real-time spectroscopic investigation of the mechanism of open complex formation by T7 RNA polymerase. *Biochemistry,* **35**, 15715-15725.

Schick, C. & Martin, C. T. (1993). Identification of specific contacts in T3 RNA polymerase-promoter interactions: kinetic analysis using small synthetic promoters. *Biochemistry,* **32**, 4275-4280.

Sousa, R. (1996). Structural and mechanistic relationships between nucleic acid polymerases. *Trends Biochem. Sci.* **21**, 186-190.

Sullivan, J. J., Bjornson, K. P., Sowers, L. C. & deHaseth, P. L. (1997). Spectroscopic determination of open

complex formation at promoters for *Escherichia coli* RNA polymerase. *Biochemistry*, **36**, 8005-8012.

Temiakov, D., Mentesana, P. E., Ma, K., Mustaev, A., Borukhov, S. & McAllister, W. T. (2000). The specificity loop of T7 RNA polymerase interacts first with the promoter and then with the elongating transcript, suggesting a mechanism for promoter clearance. *Proc. Natl Acad. Sci. USA,* **97**, 14109-14114.

Tyagarajan, K., Monforte, J. A. & Hearst, J. E. (1991). RNA folding during transcription by T7 RNA polymerase analyzed using the self-cleaving transcript assay. *Biochemistry,* **30**, 10920-10924.

Újvári, A. & Martin, C. T. (1996). Thermodynamic and kinetic measurements of promoter binding by T7 RNA polymerase. *Biochemistry,* **35**, 14574-14582.

Uptain, S. M., Kane, C. M. & Chamberlin, M. J. (1997). Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* **66**, 117-172.

Weston, B. F., Kuzmine, I. & Martin, C. T. (1997). Positioning of the start site in the initiation of transcription by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **272**, 21-30.

Wilson, K. S., Conant, C. R. & von Hippel, P. H. (1999). Determinants of the stability of transcription elongation complexes: interactions of the nascent RNA with the DNA template and the RNA polymerase. *J. Mol. Biol.* **289**, 1179-1194.

Xu, D., Evans, K. & Nordlund, T. (1994). Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry,* **33**, 9592-9599.

Xu, D. G. & Nordlund, T. M. (2000). Sequence dependence of energy transfer in DNA oligonucleotides. *Biophys. J.* **78**, 1042-1058.

Edited by R. Ebright

(Received 28 November 2000; received in revised form 3 March 2001; accepted 3 March 2001)