BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

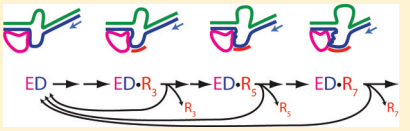
Direct Tests of the Energetic Basis of Abortive Cycling in Transcription

Ankit V. Vahia[‡] and Craig T. Martin*[,†,‡]

[†]Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003, United States

[‡]Program in Molecular & Cellular Biology, University of Massachusetts, Amherst, Massachusetts 01003, United States

ABSTRACT: Although the synthesis of RNA from a DNA template is (and must be) a generally very stable process to enable transcription of kilobase transcripts, it has long been known that during initial transcription of the first 8−10 bases of RNA complexes are relatively unstable, leading to the release of short abortive RNA transcripts. A wealth of structural data in the past decade has led to specific mechanistic models elaborating an earlier "stressed intermediate" model for initial transcription. In this study, we test fundamental predictions of each of these models in the simple model enzyme T7 RNA polymerase. Nicking or gapping the nontranscribed template DNA immediately upstream of the growing hybrid yields no systematic reduction in abortive falloff, demonstrating clearly that compaction or "scrunching" of this DNA is not a source of functional instability. Similarly, transcription on DNA in which the nontemplate strand in the initially transcribed region is either mismatched or removed altogether leads to at most modest reductions in abortive falloff, indicating that expansion or "scrunching" of the bubble is not the primary driving force for abortive cycling. Finally, energetic stress derived from the observed steric clash of the growing hybrid against the N-terminal domain contributes at most mildly to abortive cycling, as the addition of steric bulk (additional RNA bases) at the upstream end of the hybrid does not lead to predicted positional shifts in observed abortive patterns. We conclude that while structural changes (scrunching) clearly occur in initial transcription, stress from these changes is not the primary force driving abortive cycling.



Abortive cycling, the premature release of short RNA products by initially transcribing RNA polymerases, is a characteristic feature of transcription by all RNA polymerases and is thought to be a point of cellular regulation at some promoters.[1−3] Indeed, a recent study has demonstrated that abortive RNA transcripts are produced not just *in vitro* but *in vivo* as well.[4] Despite a wealth of structural studies, the energetic and mechanistic basis for the release of short abortive products has remained elusive. An early proposal (that predates structural data) suggested that an initially transcribing complex accumulates increasing stress as the RNA polymerase retains upstream promoter contacts, while moving its active site downstream along the DNA.[5] The availability of a variety of crystal structures for various polymerases and intermediates has prompted more specific proposals for the nature of this accumulating stress, although corresponding energetic characterizations have been lacking.[6−12]

Energetic Considerations. Potential energetic contributions to (in)stability in initially transcribing complexes are summarized in Figure 1. Sources of growing energetic stress might include the fact that as the polymerase initially translocates, the melted DNA bubble expands in size. In T7 RNA polymerase, the initially bound polymerase-promoter open complex contains an 8 base pair melted bubble that can accommodate up to a 3 base transcript.[6] As the RNA grows from 3 to 8 bases in length, the bubble expands by about 5 base pairs, since the downstream DNA must be melted to allow translocation, while the upstream DNA within the bubble remains open.[7,13] Clearly, melting of this additional downstream DNA must present an energetic cost. It has also been proposed more specifically that as the downstream DNA is drawn into the active site, the upstream template strand DNA must be accommodated within the active site cavity, suggesting an energetically unfavorable "scrunching" or compaction

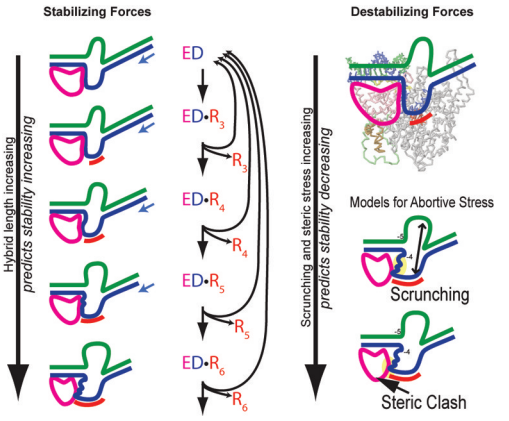


Figure 1. Initial transcription and potential energetic stresses. The cartoon upper right shows the structure of T7 RNA polymerase complex (1QLN) with a 3 base transcript. The N-terminal promoter-binding platform, which moves during initial transcription to accommodate the growing hybrid, is shown outlined in pink. Template and nontemplate DNA strands are blue and green, respectively, while nascent RNA is shown in red. Proposed structural connections to potential energetic stresses are discussed in the text.

Received: April 24, 2011
Revised: July 5, 2011
Published: July 6, 2011

of the template strand DNA.[6,9,10] Finally, as described in more detail below, growth of the hybrid "pushes" on a subdomain of the protein, forcing an internal change in the structure of the protein. All of these interactions are expected to be energetically unfavorable.

At the same time as these potentially unfavorable stresses are accumulating, initial transcription yields a steady increase in the size of the RNA–DNA hybrid. This is expected to lead to an increasing stabilization of the complex, both from the RNA-DNA interactions and potentially from protein–nucleic acid (hybrid) interactions.

Specific Predictions of Stress in T7 RNA Polymerase. Structural comparisons between T7 RNA polymerase bound to its 17 base pair consensus promoter DNA in the absence of synthesis (PDB ID: 1CEZ) and the complex with a three nucleotide nascent RNA–DNA hybrid (PDB ID: 1QLN) reveal neither movement of the N-terminal DNA binding platform nor substantial expansion of the bubble.[6,14] While upstream promoter contacts remain identical in both structures, the template strand near position −1 moves to accommodate the hybrid. The term "scrunching" was coined to describe the requirement that accumulating single-stranded template (positions −4 to −1) and nontemplate strand DNA be accommodated progressively within the enzyme active site.[6]

More recent structures of complexes halted at positions +7 and +8 verify retention of promoter contacts (as demonstrated by earlier biochemical and spectroscopic assays), accompanied by a significant rotation of the N-terminal platform to accommodate the nascent hybrid, also demonstrated earlier by spectroscopic assays.[7,8,15] Consistent with expectations, growth of the hybrid must push on the N-terminal domain, driving the latter to rotate and translate to accommodate the expanding hybrid. Thus, it has been proposed that the steric stress of this "pushing" is energetically destabilizing and may contribute to abortive release of products. These structural proposals are compelling in that as the RNA polymerase releases promoter contacts and transitions to a stable elongation complex, the now 12 (or greater) base melted bubble collapses back down to an 8 base bubble and the protein completes a structural rearrangement that now fully accommodates an 8 base pair hybrid. Coordinately, the RNA polymerase no longer releases abortive RNA products.

While the scrunching model argues that growth of the RNA–DNA hybrid contributes toward the energetic instability via proposed compaction of the single-stranded upstream template DNA, the *steric clash* model proposes that it is a more direct "push" of the growing hybrid against the N-terminal domain that leads to instability within the ternary complex.

In the current work, we have engineered mismatches, deletions, additions, and gaps into the DNA strands to test predictions of the specific stressed intermediate structure–function models. We have also introduced excess bulk at the 5′ end of the initiating RNA in the form of additional bases to test predictions of the steric clash model. The results argue that DNA scrunching and steric clash contribute at best modestly to abortive cycling and suggest other major contributors to the kinetic competition that is abortive cycling.

METHODS

Protein Expression and Purification. His-tagged T7 RNA polymerase was prepared from *E. coli* strain BL21, carrying the plasmid pBH161 (kindly supplied by William T. McAllister), and purified and characterized as described previously.[16]

DNA oligonucleotides and RNA primers were purchased commercially from Eurofins MWG Operon (DNA) and Biosynthesis Inc., respectively.

***De Novo* Transcription Assays.** Transcription reactions were carried out in a total volume of 6 μL at 37 °C for different time periods. Unless otherwise indicated, the DNA and polymerase concentrations were 1.25 and 0.25 μM, respectively (use of excess DNA prevents elongation complex instability arising from bumping[17]). The reaction buffer contained 30 mM Hepes (pH 7.8), 25 mM potassium glutamate, 15 mM magnesium acetate, 0.25 mM EDTA, and 0.05% Tween-20. The reactions were initiated by adding nucleoside triphosphates to a final concentration of 800 μM each (unless otherwise indicated) and labeled with either [α-^{32}P]GTP or [α-^{32}P]ATP.

Table 1. DNA Oligonucleotides Used To Make Modified Duplex Constructs in This Study[a]

	oligonucleotide sequence	
NT1	5′-TAATACGACTCACTATAGAGCTTGTGAGCGGATAACA-3′	−17 to +20
NT2	5′-TAATACGACTCACTATAGAG**GAA**GTGAGCGGATAACA-3′	−17 to +20
NT3	5′-TAATACGACTCACTATAGAG-3′	−17 to +3
NT4	5′-TAATACGACTCACTATA**GGA**CACCATCAACTTAACAC −3′	−17 to +20
NT5	5′-TAATACGACTCACTATAGAGATTGAATCACGACTACG-3′	−17 to +20
T1	3′-ATTATGCTGAGTGATATCTCGAAGACTCGCCTATTGA-5′	−17 to +20
T2	3′-ATTATGCTGAGTGAT-5′	−17 to −3
T3	3′-ATTATGCTGAGTG-5′	−17 to −5
T4	3′-ATCTCGAAGACTCGCCTATTGA-5′	−2 to +20
T5	3′-ATTATGCTGAGTGATATCTCTAACTTAGTGCTGATGC-5′	−17 to +20
T6	3′-ATTATGCTGAGTGA**CG**TATCTCGAAGACTCGCCTATTGA-5′	−17 to +20
T7	3′-ATTATGCTGAGTGATATCCTGTGGTAGTTGAATTGTG-5′	−17 to +20
T8	3′-TATCCTGTGGTAGTTGAATTGTG-5′	−3 to +20
T9	3′-ATTATGCTGAGTGA-5′	−17 to −4
T10	3′-ATTATGCTGAGTGAT**TACC**TGTGGTAGTTGAATTGTG-5′	−17 to +20

[a]In the first column, T*x* refers to template strands and NT*x* refers to nontemplate strands. The corresponding positions spanned (inclusive) with respect to the T7 consensus promoter are presented in the rightmost column. NT2 is mismatched with the template (T1) at positions +4 through +6; T7, T8, and T10 encode no *G*'s past position +2 (NT4 complements); T5 allows "C-less" halting at position +10 (NT3 and NT5 complement); T6 contains a 2 base insertion between positions −3 and −4; T10 prevents annealing of extra RNA bases upstream of initiating dinucleotide.

Transcription was quenched by addition of an equal volume of formamide stop solution (95% formamide, 40 mM EDTA, 0.01% bromphenol blue, and 0.01% xylene cyanol). The RNA products were resolved on 20% polyacrylamide gels containing 7 M urea and visualized with a Fuji phosphorimager.

Transcription assays using presynthesized RNA primers were carried out as previously described; however, transcription was initiated using the appropriate RNA primer and a combination of ATP, CTP, and UTP to a final concentration of 800 μM each. The RNA primers were added to a final concentration of 125 μM. In these reactions [α-^{32}P]ATP was used to radiolabel the transcribed RNA.

Percent falloff calculations were carried out by calculating the ratio of the amount of RNA released of a given length to the total amount of RNA of that length and longer.

$$\%\ \text{falloff at}\ i = \frac{[R_i]}{\sum_{j=1}^{n} R_j} \times 100\%$$

RESULTS

The stressed intermediate model for abortive cycling proposes that as RNA polymerase progresses through initial transcription, the complex becomes increasingly destabilized. Thus, the lengthening of the hybrid duplex, which should increase complex stability, is offset by increased stress in the system. Various structural models have been presented to explain potential sources of this "stressed" instability. In the following, we directly test each of these models by constructing promoters designed to either increase or decrease potential instability inferred from the structural studies. In general, there is strong sequence dependence to abortive cycling, perhaps reflecting sequence-dependent kinetics; consequently, the following experiments aim to keep the sequence as constant as possible. The key element of each test is to ask whether there is a systematic and significant change in abortive percent falloff in the direction predicted by each model.[18]

Reduce Stress in the Upstream Template DNA. In order to test the proposal that accumulating conformational stress in the template strand DNA leads to a destabilization of the initially transcribing complex, we introduced disruptions into the template strand, thereby allowing increased conformational freedom. Either "nicking" between the two bases at positions −3 and −2 (a construct formed by annealing DNA strands NT1, T2, and T4) or removing two bases entirely (the introduction of a "gap" by deleting completely bases −4 and −3) (formed from strands NT1, T3, and T4) should provide for more flexibility in the positioning of the template strand within the enzyme active site. Both should also reduce the energetic cost of the melted bubble (more so for the gapped construct). Finally, the gapped construct removes physical bulk in the active site and so should reduce steric stress from accumulation of the template strand. The simple prediction of the stressed intermediate models is that these modifications should lead to reduced abortive falloff.

The results presented in Figure 2B demonstrate that the introduction of a nick or a gap in the template strand yields only modest and nonsystematic effects on abortive cycling. Overall, transcription increases for each, possibly due to a reduction in the barrier to initial promoter melting/binding.[19] The percentages of falloff at positions 4, 6, and 7 show little

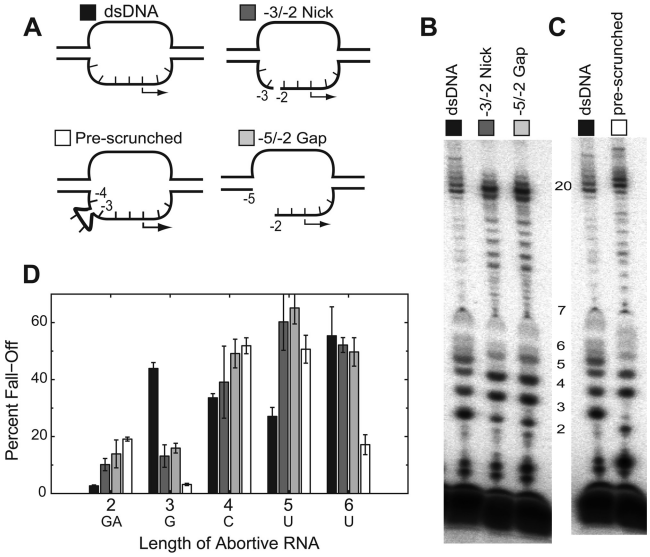


Figure 2. Constructs designed to relieve or increase scrunching stress. (A) The abortive profile from the fully double-stranded construct [annealing strands NT1 and T1] is compared with that from reduced stress constructs created with an effective nick between positions −3 and −2 [NT1, T2, and T4] or with omission of residues −4 and −3 [NT1, T3, and T4]. The increased stress construct contains two additional bases (3′-CG-5′)) inserted between positions −3 and −4 of the consensus template strand sequence [NT1 and T6]. (B, C) Reaction products (abortive and full length) assayed by denaturing polyacrylamide electrophoresis. Reactions were run for 5 min at 37 °C as described in methods. (D) Percent falloff at each translocational position, calculated from data in (B) and (C).

change on nicking or gapping, and transcripts at position 5 fall off at a higher percentage, contrary to predictions of the models. Simple analysis of products at 2 and 3 are complicated by the fact that released transcripts can be reutilized as primers in subsequent rounds of synthesis.[20] Thus, use of the simple percent falloff quantification is complicated at these positions. (In any case, while changes in 3-mer accumulation support the model, 2-mer accumulation opposes the model.) Additionally, scrunching and/or steric stress should only begin to accrue past position +3, as the initially melted bubble is sufficient for the synthesis of a 3-mer without any increase in bubble size or steric clash.

Note that for both the nicked and gapped constructs there is an increase in products 11−13 bases in length (Figure 2B). This is similar to increases observed in transcription from DNA constructs that completely lack the downstream nontemplate strand (partially single-stranded), and we have proposed that this arises from inefficient displacement of the 5′ end of the RNA in the formation of the elongation complex.[16,21]

Increased Stress in the Upstream Template DNA. If reduction in stress is predicted by the model to yield fewer abortive products, modifications that increase stress should lead to higher abortive ratios. To test this hypothesis, two additional nucleotides were added to the template strand (only), as shown in Figure 2 to form a "prescrunched" template. T7 RNA polymerase initiates primarily from the first C found in the template strand (initial G in the transcript) that is at least 4 bases downstream of the end of the bound promoter duplex at position −5.[22−24] Thus, introduction of an additional 3′-CG-5′ dinucleotide step between positions −4 and −3 places the first

dx.doi.org/10.1021/bi200620q | Biochemistry 2011, 50, 7015−7022

available template C two bases farther downstream (the introduced C is too close to the bound duplex to serve as a template). In order to initiate at the first available (consensus +1) C, the additional bases within the template strand must presumably be "crowded" into the active site upstream of the functional start site, thereby inducing a state in which the DNA is "pre-scrunched" within the enzyme. Scrunching models that invoke steric crowding of the template strand predict that initially transcribing complexes with excess DNA in this region should be unstable relative to enzyme transcribing on consensus dsDNA and should therefore increase abortive falloff at positions downstream.

The results presented in Figure 2C (and quantified in Figure 2D) show a modest increase in abortive fall off at positions +4 and +5, as predicted by the scrunching model, but a decrease in abortive falloff at position +6. Thus, the results are at best both mildly supportive of and simultaneously in contradiction to the model, suggesting that compaction of the upstream template strand is unlikely to be a major driving force for abortive cycling.

Bubble Melting Stress. Another proposed aspect of scrunching derives from the observation that the melted DNA bubble increases in size as the polymerase steps beyond position +3, prior to promoter release.[9,13] This melting must certainly require energetic input, but is this a kinetic barrier that leads to abortive products? To test this hypothesis, the melting barrier was reduced via the introduction of mismatches in the nontemplate strand at the three bases from position +4 to +6. The model predicts a decrease in abortive falloff at these positions. The results shown in Figure 3A,C reveal no significant effect at position +4, an increase in the percent falloff at position +5, and a decrease in falloff at position +6.

As an alternate approach toward answering the same question, transcription was carried out on partially single-stranded DNA (pssDNA) in which the nontemplate strand extends from −17 to only +3. It is important to note that in general, after one round of transcription on constructs single-stranded in the transcribed region, the newly synthesized RNA can anneal to the single-stranded template, restoring at least partial double-strand character in the transcribed region (in this case, an RNA−DNA hybrid). In order to limit this complication, on the pssDNA construct, the initial sequence was modified slightly to allow halting at position +10. The sequence of these constructs was C-less up to position +11, allowing halting of transcription by addition of only GTP, ATP, and UTP. Halting the complex at +10 would leave a more weakly bound (or unbound) product RNA after the first round of transcription, which should not present a large barrier to melting.

Transcription from these constructs, designed to remove the downstream barrier to melting, shows a small to moderate decrease in the abortive profile at all positions. However, abortive release is still significant at all positions (ranging from 20 to 40% at each position). Significant abortive cycling remains that cannot be explained by the energetic stress of bubble expansion.

Hybrid Steric Clash as a Source of Stress during Abortive Cycling. While bubble expansion might present a small source of energetic stress, growth of the RNA could also exert energetic stress in the form of a "pushing" of the RNA−DNA hybrid against the N-terminal platform, as described above. As the length of the hybrid increases during elongation,

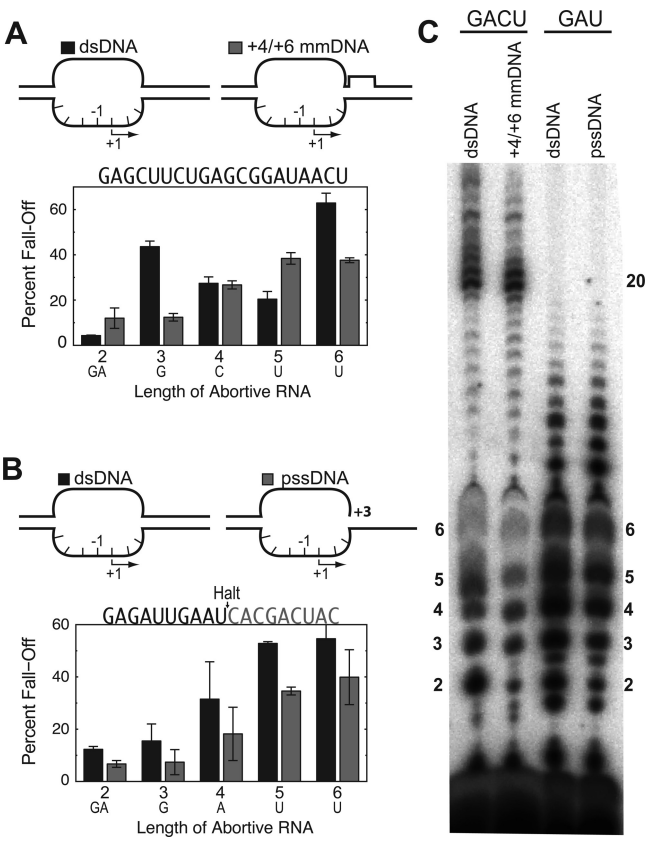


Figure 3. Constructs premelted in the initially transcribing region. To reduce the energetic cost associated with formation and extension of the DNA duplex bubble, constructs were prepared either (A) mismatching the bases at positions +4 through +6 [mm+4/+6, annealing strands NT2 and T1] or (B) fully deleting the nontemplate strand downstream of position +3, with a halt site at position +11 [pssDNA: NT3 and T5]. Abortive profile assays and percent falloff calculations as in Figure 2.

the N-terminal domain moves away from the C-terminal domain in order to accommodate the larger hybrid.[7] Thus, abortive complex instability might be a direct result of this *steric clash* of the growing RNA-DNA hybrid against the N-terminal domain of the enzyme (Figures 1 and 4B). On the basis of the structural analysis of the initiation complex, this clash is expected to become important when the hybrid increases beyond 3 bases in length.[6]

To assess whether this observed structural effect is a major energetic contributor to abortive cycling, we *increased steric clash* by adding bulk to the 5′ end of the RNA. The model predicts that if steric clash is a driving force for abortive cycling, then increasing the length of the RNA in the hybrid will lead the hybrid to push against the N-terminal domain at earlier translocational positions, resulting in increased abortive release and/or a shift toward more abortive propensity with shorter translocational positioning.

Exploiting the fact that T7 RNA polymerase effectively initiates with a dinucleotide RNA primer complementary to the first two nucleotides on the template DNA (positions +1 and +2), we initiated transcription utilizing an RNA dinucleotide with its 5′ end extended by 0, 1, 2, or 3 bases. In the experiments of Figure 4 labeled "fully double stranded DNA", these additional bases are complementary to the template strand

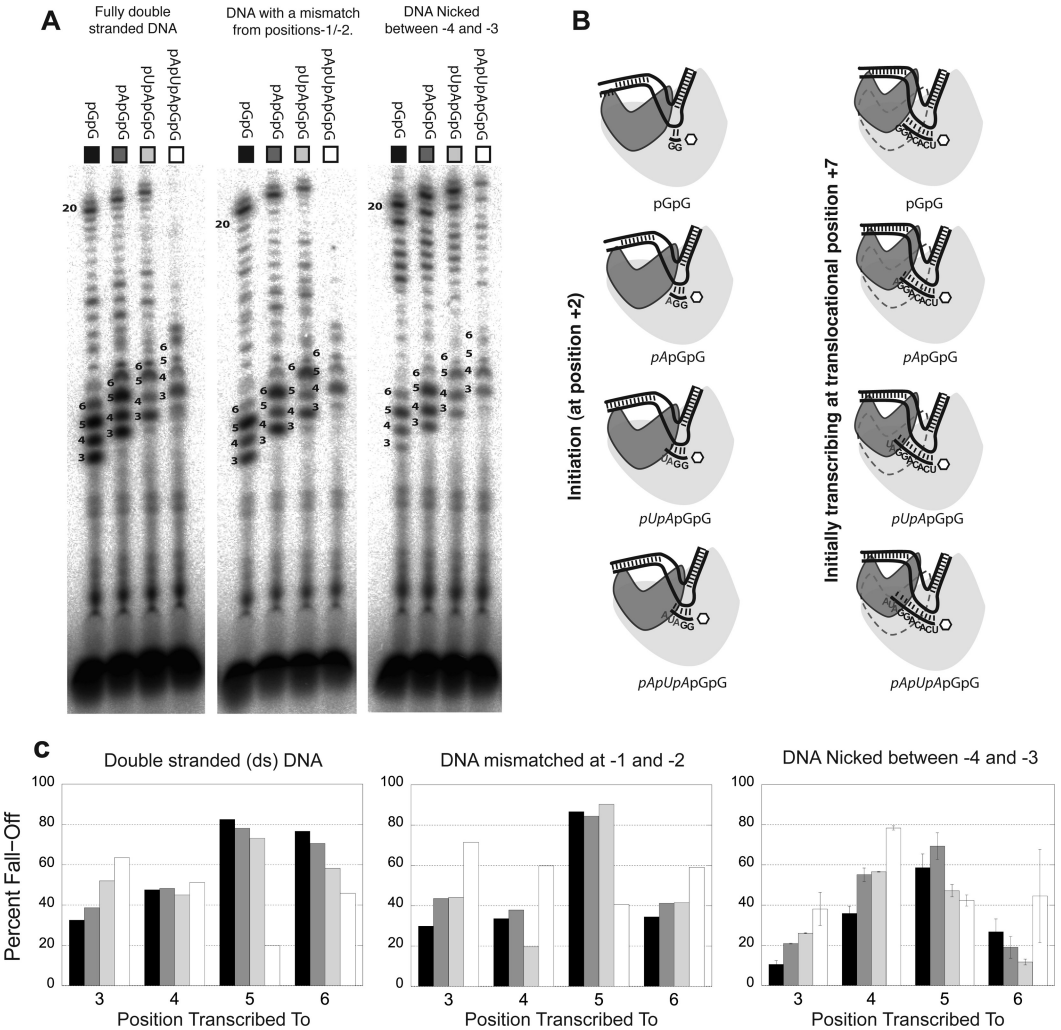


Figure 4. Transcription with additional bases at the 5′ end of an initiating RNA primer. (A) Reactions were carried out as described in Figure 2, either on fully double-stranded DNA [annealing NT4 + T7] or on DNA containing either a nick in the template strand [NT4 + T8 + T9] or with an upstream template strand sequence noncomplementary to the RNA extensions [NT4 + T10]. Translocational positions (as opposed to RNA length) are noted on the gel. (B) Cartoons based on established structures, depicting initiation complexes and complexes halted at position +7. In the latter, a dotted outline shows relative positioning of the N-terminal domain (dark gray) before rotation/translation. (C) Analysis of percent falloff on nicked constructs reveals an early increase in the abortive RNA with an increase in the length of the hybrid (at positions +3 and +4). However, at positions +5 and +6 there is no systematic increase in abortive cycling with an increase in the steric push of the DNA−RNA hybrid against the N-terminal domain.

nucleotides at positions −1, −2, and −3 of the upstream template DNA, such that the effective length of the resulting RNA−DNA hybrid can extend by as much as 0, 1, 2, or 3 bases (Figure 4B). For example, approximating normal transcription, dinucleotide primed initiation can be carried out using a 5′-pGpG-3′ RNA primer. Initiation with the 5′-pApGpG-3′ RNA primer adds an extra base (A) at the 5′ end which can base pair with the corresponding base (T) at position −1 on the template strand. To ensure that in the absence of GTP the enzyme would initiate transcription exclusively with the (GG-containing) RNA primers, in this set of experiments the template strand DNA sequence was rendered C-less except for positions +1 and +2.

The results presented in Figure 4 demonstrate initiation with all four primers, with systematic and predictable shifts in RNA lengths. For example, transcription to position +4 along the template with a primer extended by 2 bases yields an RNA of length 6. The pattern of products clearly indicates that the abortive propensities correlate with translocational position, rather than with total RNA length. For clarity, in Figure 4 we have therefore labeled the RNAs by their translocational position rather than by their length.

Quantification of percent falloff values for transcription from fully double-stranded DNA reveals no systematic change in abortive propensities. Falloff increases somewhat with primer length at translocational position +3, as predicted by the model. However, no changes are observed at position +4 and the mild trends for positions +5 and +6 are the opposite of that predicted by the model.

Given the structural requirement that the normally single-stranded template from position −4 to −1 must span the distance from the double-stranded region of the promoter to the +1 base at the active site, formation of hybrid upstream of position +1 may strain the system. In order to relax this, a template strand DNA was prepared in which the bases at positions −2 and −1 are not complementary to the RNA primer upstream extension strain In this case, a hybrid will not

dx.doi.org/10.1021/bi200620q | Biochemistry 2011, 50, 7015−7022

form, but upstream "bulk" will still be present in the form of unpaired nucleotides. Transcription with various primers on this construct again yields no consistent trends in abortive falloff.

Finally, in order to favor full upstream hybrid formation and maximize steric clash between the hybrid and the N-terminal domain, without template strand strain, a construct was prepared with a template strand nick between positions −4 and −3. The results presented in Figure 4 reveal that at positions +3 and +4 percent falloff does increase systematically (up to 2-fold) with increasing upstream bulk, suggesting a direct relation between steric clash and accumulation of abortive RNA. However, at positions +5 and +6, increasing the upstream hybrid length yields no significant increase in the percent falloff relative to the dinucleotide-initiated control. Indeed, there is some evidence for a small decrease in abortive release at each position. Thus, at positions where steric clash with the N-terminal domain might be expected to be most severe, there is little net effect of adding upstream steric bulk to the hybrid.

With the longest of these primers (pApUpApGpG), particularly on the fully double-stranded and mismatched constructs, the data in Figure 4 show overall lower accumulation of products and very large decrease in full length RNA. The former can be explained either by a reduced rate of initiation or by stably bound, but slowly extending, intermediate complexes. The complete lack of runoff complexes favors the latter. This is relieved somewhat by nicking of the template strand between positions −4 and −3. Preliminarily, these data suggest that these complexes may be deficient in transitioning to functional elongation complexes and are stuck in the initial phase.

DISCUSSION

The notion of a stressed intermediate during initial transcription has long been considered in explaining the tendency of RNA polymerases to release short, abortive products in this early phase of transcription.[5] Recent structural studies have lent credence to this model and have led to specific proposals for the source of this stress.[6,9−11,25] Although these studies provide valuable structural insights, none of them test the basic premise that accumulating stress is the energetic driving force for release of abortive transcripts. The current study is the first direct test of these structural−mechanistic models.

Bubble Expansion Is at Most a Minor Driving Force in Abortive Cycling. Footprinting studies first suggested that RNA polymerases retain upstream promoter contacts during synthesis of the first 8−10 bases of RNA.[26] Subsequent fluorescence measurements confirmed that promoter contacts are retained and the upstream edge of the bubble remains open as the downstream edge of the bubble melts to expose more template DNA—the bubble expands in size.[5,9,10,13] Most recently, the crystal structure of an initially transcribing T7 RNA polymerase poised at positions +7 and +8 has confirmed both that the promoter contacts are retained and that the upstream edge of the bubble remains melted.[7] Thus, expansion of the bubble clearly occurs during initial transcription and is a potential source of energetic stress.

While it has long been observed that T7 RNA polymerase can transcribe from DNA constructs that are single-stranded downstream of the promoter and still show abortive release, it is important to note that after the first round of successful (full or near full length) transcription the product RNA is likely to either remain bound or to rebind to the complementary (single-stranded) template DNA. Thus, multiple turnover experiments on partially single-stranded constructs do not necessarily test this hypothesis, as the bound full-length hybrid may present a similar downstream melting barrier. In order to avoid this complication, we carried out experiments shown in Figure 4 in which a limited, but key, region of the DNA is mismatched, thereby removing part of the energetic barrier to bubble expansion. The introduction of a mismatch at positions +4 through +6 yields the predicted decrease in abortive falloff at position +3 but does not reduce abortive cycling at positions +4 and +5.

As an alternate test, one can transcribe from partially single-stranded constructs, but limit downstream synthesis, so as to minimize the energetics of a persistent or reannealed downstream RNA barrier (the product RNA being too short to bind tightly). The results presented in Figure 4 show a consistent reduction in abortive fall off at each position under this scenario, but the reduction is small and abortive dissociation remains a key feature of this sequence. Together these results argue that melting of the downstream duplex is not a major energetic driving force leading to abortive cycling.

Template Strand Accumulation or "Scrunching" Is Not a Driving Force in Abortive Cycling. In the determination of the structure of an initially transcribing T7 RNA polymerase, it was noted that transcription beyond a three base transcript would lead to accumulation or "scrunching" of the template strand in the enclosed active site pocket.[6] We now know that during this initial phase the N-terminal domain rotates to accommodate movement of the template strand,[7] but conformational stress in the template strand has remained a potential source of energetic stress.

By nicking or gapping the upstream template DNA during initiation, as shown in Figure 2, we have effectively decreased or eliminated the energetic stress (scrunching) that might arise as the template strand moves to accommodate the growing RNA−DNA hybrid. Although transcription increases somewhat overall (bubble melting is presumably more facile), there is no uniform decrease in the percentage of abortive falloff at each position in these constructs. These results demonstrate clearly that scrunching of the template strand DNA is not a major contributor to abortive cycling.

As an alternative test of template strand scrunching as a determinant in abortive cycling, we designed a construct with two extra bases in the nontranscribed, upstream melted template strand DNA. In this "prescrunched" scenario, the model would predict an increase in abortive falloff at each position during initial transcription. The results presented in Figure 4 illustrate that while there is some increase in abortive falloff at positions +4 and +5, as predicted by the model, there is a decrease in abortive falloff at positions +3, +6, and +7. Thus, template strand scrunching is not a general mechanism for abortive cycling during initial transcription.

Steric Clash Contributes at Most Partially to Abortive Cycling. From the most recent structures of initially transcribing complexes poised at positions +7 and +8, it is now clear that growth of the RNA−DNA hybrid must sterically drive the observed rotation of the N-terminal domain through transcription to at least position +8.[7] The steric clash that drives this rotation is another source of potential stress in the system that could lead to instability in initially transcribing complexes and lead to the release of abortive products.

Does this steric clash lead to the proposed instability that could lead to release of abortive products? The experiments

presented in Figure 4 demonstrate that adding "bulk" at the upstream end of the hybrid does not lead to the predicted systematic increase in abortive cycling at each translocational position. This is true whether or not the added RNA bulk is complementary to the corresponding bases in the template strand. On a construct that does allow pairing, nicking of the template strand to allow the hybrid to form without the constraints of its connection to the upstream promoter bound duplex reveals the predicted pattern at positions +3 and +4 but shows the opposite pattern at positions +5 and +6.

The mutant P266L shows dramatically less abortive falloff on many abortive sequences,[27] including those used here (data not shown). The prevailing model for this this phenotype is that the mutant has a lower barrier to the structural transition. Since this model assumes steric stress as the driving force of abortive dissociation, it is likely that this mutation exerts its effect through a different mechanism.

Tang et al. have suggestive evidence that as the length of the hybrid approaches 5−8 bases in length, the rate of nucleotide addition slows progressively.[28] Since the probability of abortive falloff must be a kinetic competition between nucleotide addition and dissociation of the RNA, the increasing length of the RNA could result in slowing of both nucleotide addition and RNA release (assuming that the rate of release scales inversely with the length of the hybrid). Subtle differences in these two progressions could lead to variable abortive to productive ratios at each position. Unfortunately, this model begs the question: why would the polymerase slow progressively with increasing hybrid length? The results presented here suggest that the answer does not lie in steric clash.

Other Considerations. The most obvious energetic contributor to abortive cycling during initial transcription is an expected relatively weak binding of short RNAs in the nascent hybrid. At first glance, this predicts a decreasing probability of abortive release as the transcript grows in size and stability. However, we have recently proposed that the stability of an elongation complex derives not only from the thermodynamic stability of base pairing, but perhaps more importantly from a kinetic stability imparted by the topological "locking" of the RNA around the template DNA strand.[29] Thus, stability is not achieved until a sufficient length is achieved to provide that lock, presumably eight base pairs (plus optimally four more bases to fill the RNA exit channel). More importantly, the abortive process must be controlled by kinetics rather than by thermodynamics.[30,31] Transcript release is irreversible—products longer than about 2 bases cannot rebind and be elongated, except under specific, artificial circumstances.[20] It is also important to consider that one presumed purpose of the initially transcribing complex is to maintain the DNA bubble open while the hybrid is short and therefore (more) sensitive to collapse of the bubble. Thus, the large structural transition of the enzyme from its initial configuration to the elongation configuration should not occur too soon. Indeed, in T7 RNA polymerase that transition is not thought to occur before an RNA length of at least 9 bases, but abortive product release is generally low at RNA lengths of 7−8 bases.

AUTHOR INFORMATION

Corresponding Author

*Phone: 413-545-3299. Fax: 413-545-4490. E-mail: cmartin@chem.umass.edu.

Funding

This work was supported by National Institutes of Health Grant 1R01GM55002.

ABBREVIATIONS

T, template; NT, nontemplate; ds, double-stranded; pss, partially single-stranded; mm, mismatch.

REFERENCES

(1) Carpousis, A. J., and Gralla, J. D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry 19*, 3245−3253.

(2) Carpousis, A. J., and Gralla, J. D. (1985) Interaction of RNA polymerase with lacUV5 promoter DNA during mRNA initiation and elongation. Footprinting, methylation, and rifampicin-sensitivity changes accompanying transcription initiation. *J. Mol. Biol. 183*, 165−177.

(3) Jin, D. J., and Turnbough, C. L. Jr. (1994) An Escherichia coli RNA polymerase defective in transcription due to its overproduction of abortive initiation products. *J. Mol. Biol. 236*, 72−80.

(4) Goldman, S. R., Ebright, R. H., and Nickels, B. E. (2009) Direct detection of abortive RNA transcripts in vivo. *Science 324*, 927−928.

(5) Straney, D. C., and Crothers, D. M. (1987) A stressed intermediate in the formation of stably initiated RNA chains at the Escherichia coli lac UV5 promoter. *J. Mol. Biol. 193*, 267−278.

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

(7) Durniak, K. J., Bailey, S., and Steitz, T. A. (2008) The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. *Science 322*, 553−557.

(8) Tang, G. Q., Roy, R., Ha, T., and Patel, S. S. (2008) Transcription initiation in a single-subunit RNA polymerase proceeds through DNA scrunching and rotation of the N-terminal subdomains. *Mol. Cell 30*, 567−577.

(9) Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S., and Ebright, R. H. (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science 314*, 1144−1147.

(10) Revyakin, A., Liu, C., Ebright, R. H., and Strick, T. R. (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science 314*, 1139−1143.

(11) Cheetham, G. M., Jeruzalmi, D., and Steitz, T. A. (1998) Transcription regulation, initiation, and "DNA scrunching" by T7 RNA polymerase. *Cold Spring Harb. Symp. Quant. Biol. 63*, 263−267.

(12) Brieba, L. G., and Sousa, R. (2001) T7 promoter release mediated by DNA scrunching. *EMBO J. 20*, 6826−6835.

(13) Liu, C., and Martin, C. T. (2002) Promoter clearance by T7 RNA polymerase. Initial bubble collapse and transcript dissociation monitored by base analog fluorescence. *J. Biol. Chem. 277*, 2725−2731.

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

(15) Turingan, R. S., Theis, K., and Martin, C. T. (2007) Twisted or shifted? Fluorescence measurements of late intermediates in transcription initiation by T7 RNA polymerase. *Biochemistry 46*, 6165−6168.

(16) Gong, P., Esposito, E. A., and Martin, C. T. (2004) Initial bubble collapse plays a key role in the transition to elongation in T7 RNA polymerase. *J. Biol. Chem. 279*, 44277−44285.

(17) Zhou, Y., and Martin, C. T. (2006) Observed instability of T7 RNA polymerase elongation complexes can be dominated by collision-induced "bumping. *J. Biol. Chem. 281*, 24441−24448.

(18) Martin, C. T., Muller, D. K., and Coleman, J. E. (1988) Processivity in early stages of transcription by T7 RNA polymerase. *Biochemistry 27*, 3966−3974.

(19) Újvári, A., and Martin, C. T. (1997) Identification of a minimal binding element within the T7 RNA polymerase promoter. *J. Mol. Biol. 273*, 775−781.

(20) Kuzmine, I., and Martin, C. T. (2001) Pre-steady-state Kinetics of Initiation of Transcription by T7 RNA Polymerase: A New Kinetic Model. *J. Mol. Biol. 305*, 559−566.

(21) Martin, C. T., Esposito, E. A., Theis, K., and Gong, P. (2005) Structure and Function in Promoter Escape by T7 RNA Polymerase. *Prog. Nucleic Acid Res. Mol. Biol. 80*, 323−347.

(22) Weston, B. F., Kuzmine, I., and 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.

(23) Kuzmine, I., Gottlieb, P. A., and Martin, C. T. (2003) Binding of the priming nucleotide in the initiation of transcription by t7 RNA polymerase. *J. Biol. Chem. 278*, 2819−2823.

(24) Jiang, M., Rong, M., Martin, C., and McAllister, W. T. (2001) Interrupting the template strand of the T7 promoter facilitates translocation of the DNA during initiation, reducing transcript slippage and the release of abortive products. *J. Mol. Biol. 310*, 509−522.

(25) Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2006) Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophys. J. 90*, 1419−1431.

(26) Ikeda, R. A., and 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. U. S. A. 83*, 3614−3618.

(27) Guillerez, J., Lopez, P. J., Proux, F., Launay, H., and Dreyfus, M. (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. U. S. A. 102*, 5958−5963.

(28) Tang, G. Q., Roy, R., Bandwar, R. P., Ha, T., and Patel, S. S. (2009) Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A. 106*, 22175−22180.

(29) Liu, X., and Martin, C. T. (2009) Transcription elongation complex stability: the topological lock. *J. Biol. Chem. 284*, 36262−36270.

(30) von Hippel, P. H. (1998) An integrated model of the transcription complex in elongation, termination, and editing. *Science 281*, 660−665.

(31) von Hippel, P. H., and Yager, T. D. (1991) Transcript elongation and termination are competitive kinetic processes. *Proc. Natl. Acad. Sci. U. S. A. 88*, 2307−2311.