Published in final edited form as:

Biochemistry. 2019 May 07; 58(18): 2339–2352. doi:10.1021/acs.biochem.9b00049.

RNA Polymerase: Step by Step Kinetics and Mechanism of Transcription Initiation

Kate L. Henderson¹, Claire E. Evensen¹, Cristen M. Molzahn¹, Lindsey C. Felth¹, Sarah Dyke¹, Guanyu Liao¹, Irina A. Shkel^{1,2}, and M. Thomas Record Jr.^{1,2}

¹Department of Biochemistry, University of Wisconsin – Madison, Madison, Wisconsin, 53706, USA

²Department of Chemistry, University of Wisconsin – Madison, Madison, Wisconsin, 53706, USA

Abstract

To determine the step-by-step kinetics and mechanism of transcription initiation and escape by *E. coli* RNA polymerase from the λ PR promoter, we quantify the accumulation and decay of transient short RNA intermediates on the pathway to promoter escape and full-length (FL) RNA synthesis over a wide range of NTP concentrations by rapid-quench mixing and phosphorimager analysis of gel separations. Experiments are performed at 19 °C, where almost all short RNAs detected are intermediates in FL-RNA synthesis by productive complexes or end-products in non-productive (stalled) initiation complexes, and not from abortive initiation. Analysis of productive-initiation kinetic data yields composite second-order rate constants for all steps of NTP binding and hybrid extension up to the escape point (11-mer). The largest of these rate constants is for incorporation of UTP into the dinucleotide pppApU in a step which does not involve DNA opening or translocation. Subsequent steps, each of which begins with reversible translocation and DNA opening, are slower with rate constants that vary more than ten-fold, interpreted as effects of translocation stress on the translocation equilibrium constant. Rate constants for synthesis of 4- and 5-mer, 7-mer to 9-mer and 11-mer are particularly small, indicating that RNAP-promoter interactions are disrupted in these steps. These reductions in rate constants are consistent with the previously-determined 9 kcal cost of escape from λ PR. Structural modeling and previous results indicate that the three groups of small rate constants correspond to sequential disruption of in-cleft, –10 and –35 interactions. Parallels to escape by T7 RNAP are discussed.

Corresponding Authors Tel: +1 608 262 5332; Fax: +1 608 262 3453; mtrecord@wisc.edu, khenderson4@wisc.edu.

Present Addresses

Cristen M. Molzahn, Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada
Lindsey C. Felth, Department of Pharmacology, University of California – Davis, Davis, California, 95616, USA

Author Contributions

K.L.H., C.M.M., L.C.F., and M.T.R. designed research; K.L.H., C.M.M., L.C.F., C.E.E., S.D., and G.L. performed research; K.L.H., C.M.M., C.E.E., I.A.S., and M.T.R. analyzed data; and K.L.H. and M.T.R. wrote the paper.

ASSOCIATED CONTENT

Additional Methods, Figures (S1–S7), and Tables (S1–S8) are available free of charge in the Supporting Information.

Accession Numbers

P0A7Z4, subunit α ; P0A8V2, subunit β ; P0A8T7, subunit β' ; P0A800, subunit ω ; P00579 subunit σ .

Introduction

Substantial progress has been made recently to understand the mechanism of transcription initiation by RNA polymerases (RNAP), including DNA opening, stabilization of the initial open complex, start-site selection, the relationship between abortive and productive initiation, and pause points in initiation^{1–18}. For *E. coli* RNAP^{2, 5, 7, 8, 19}, eukaryotic pol II RNAP²⁰, and phage T7 RNAP⁹, initiation complexes formed by NTP binding and synthesis of a short RNA-DNA hybrid before promoter escape are found to be functionally heterogeneous. Roughly equal populations are observed of productive complexes that escape from the promoter to make a full-length RNA and of non-productive complexes that stall after synthesis of a short RNA-DNA hybrid. Subsequently, stalled non-productive complexes slowly release the small RNA from the hybrid and reinitiate synthesis of another small RNA in a cycle of abortive synthesis. Slow conversion of a small fraction (< 20%) of non-productive complexes to productive complexes is observed in some studies^{7–9}, on a longer time scale than that of many initiation assays.

Translocation of the RNA-DNA hybrid into the cleft during initiation generates stress in the initiation complex, increasing its free energy and destabilizing RNAP-promoter contacts. Two such stresses are steric stress from contacts of the growing hybrid with $\sigma 3.2$ in the cleft^{21, 22, 8} and scrunching stress from deforming the single strands of promoter DNA in the cleft^{3, 23, 24}. More translocation stress should be needed for RNAP to escape from a more stable OC; hence escape is predicted to be slower and to require synthesis of a longer RNA-DNA hybrid. Results of recent studies of initiation and escape from WT and variant λ PR and T7A1 promoters are consistent with these predictions^{2, 5}.

For productive complexes, the RNA-DNA hybrid length at which RNAP escapes from the promoter increases as OC lifetime and stability increase⁵. For two OCs with the λ PR discriminator (stability ~ -16 kcal/mol), escape occurs upon extending the 10-mer hybrid. For two less-stable (~ -13 kcal/mol) OCs with the T7A1 discriminator, escape occurs upon extending the 7-mer hybrid. From these results, we deduced that the build-up of scrunching and/or steric stress from translocation of the hybrid into the cleft increases the free energy of the initiating complex by ~ 1 kcal/mol per nucleotide added to the hybrid. A similar result was obtained from single-molecule measurements²⁵. Stress build up from 9 translocation steps therefore provides ~ 9 kcal/mol of driving force for escape of RNAP from the λ PR promoter. Interactions of σ^{70} with the discriminator²⁶, -10 and -35 and upstream regions of promoter DNA^{21, 22, 27}, and other OC interactions not present in the elongation complex, must be disrupted prior to escape, release of σ^{70} , and the transition to elongation.

A second prediction is that build-up of translocation stress in a step of initiation should be expressed as a reduction in rate constant for NTP incorporation and hybrid extension in that step. Translocation stress provides the driving force for disruption of the specific RNAP-promoter contacts necessary for escape^{21–23, 28}, but its consequences for the step-by-step kinetics and energetics of RNA-DNA hybrid extension by productive complexes during the initiation phase in single-round transcription have not been investigated. One focus of the research reported here is to determine which steps of initiation are most affected by

translocation stress and when these stresses are relieved by disruption of specific interactions of RNAP with promoter DNA.

Analysis of the kinetics and mechanism of elongation steps provide a useful comparison, since the basic mechanism is the same but there is no translocation stress. Studies of the kinetics and mechanism of elongation by bacterial²⁹ and eukaryotic^{20, 30} RNAP show elongation consists of reversible translocation, reversible NTP binding, and catalysis, made irreversible by pyrophosphate (PPi) release at low PPi concentration^{31,32}. Translocation in elongation is reversible and intrinsically unfavorable, with the pre-translocated state favored in the absence of the correct NTP^{20, 29, 33}. In elongation, NTP binding shifts the equilibrium to favor the post-translocated state. Translocation is proposed to be even more unfavorable in initiation because of steric and scrunching translocation stress^{3, 8, 21–24}.

In this study, we determine the kinetics of initiation at the λ PR promoter using rapid quench mixing over a range of NTP concentrations. By reducing the temperature, we stabilize short RNA-DNA hybrids at stalled non-productive complexes, preventing additional rounds of short RNA synthesis by these complexes and drastically reducing the total amount of short RNA from non-productive complexes. By subtracting the amount of short RNA produced by stalled, non-productive complexes from the total RNA of that length synthesized, we determine the transient build-up and decay of each RNA length on the pathway to FL RNA. Particularly significant transient populations of 3-, 6- and 10-mer are observed at all different NTP conditions examined, indicating that the rate constants for forming these intermediates are large and the rate constants for extending these intermediates are small. We interpret the variation of large and small rate constants in terms of the unfavorable costs of translocation-induced disruption of specific in-cleft and upstream RNAP-promoter contacts, driven by NTP binding, that lead to promoter escape.

MATERIALS AND METHODS

Reagents, Buffers, Gels

Reagents used in buffers and stock solutions were purchased in the highest grade available and used as received. All solutions were prepared using 18 M Ω deionized water. NTPs and dNTPs (Thermo Fisher Scientific, Waltham, MA) used in transcription assays and PCR reactions are 99% pure and used as received. Enzymes for PCR reactions were purchased from New England Biolabs (NEB, Ipswich, MA) and used according to the manufacturer's protocols.

Storage buffer (SB) for core RNAP, σ^{70} , and RNAP holoenzyme is 50% v/v glycerol, 0.01 M Tris, 0.1 M NaCl, 0.1 mM EDTA, 0.1 mM DTT. Transcription buffer (TB) is 40 mM Tris (adjusted to pH 8.0 at 19 °C), 5 mM MgCl₂, 60 mM KCl, 1 mM DTT, and 0.05 mg/mL BSA.

For rapid quench flow (RQF) transcription assays with α -³²P-GTP, initiation solution (IS) is 400 μ M ATP, 400 μ M UTP, 0.10 mg/mL heparin in TB, plus unlabelled GTP and α -³²P-GTP. Mixing 1:1 with preformed OC reduces these concentrations to half their initial value.

For a final GTP concentration of 10 μ M, IS includes 20 μ M GTP and 35 nM α -³²P-GTP. Compositions of all IS are tabulated in Supplemental Table S1.

Quench solution (QS) for initiation assays is 8 M urea, 15 mM EDTA in TB. Quench solution with added dyes (QSD) used for polyacrylamide gel electrophoresis (PAGE) has 0.05% w/v xylene cyanol and 0.05% bromophenol blue in QS. TBE buffer for gel electrophoresis is 90 mM Tris-borate (pH 8.3), 2 mM Na₂EDTA. All transcription gels are 20% acrylamide-bisacrylamide (19:1) and were made using the UreaGel System (National Diagnostics, Atlanta, GA).

Enzyme and DNA Preparation

RNA polymerase core enzyme ($\alpha_2\beta\beta'\omega$) is overexpressed and purified from pVS10 plasmid in XJbDE3 cells and stored in SB at -80°C ³⁴. Specificity subunit σ^{70} is overexpressed and purified using Ni-affinity chromatography from pIA586 plasmid in BL21(DE3) cells and stored in SB at -80°C ³⁵. Holoenzyme is assembled by incubating σ^{70} and RNAP core enzyme in a 2:1 ratio in SB at 37°C for 1 hr and stored at -20°C . Filter binding activity assays³⁶ on the holoenzyme preparation used in all experiments reported here show that $50 \pm 10\%$ of RNAP molecules form a stable open complex with the λ PR promoter at 37°C . All RNAP concentrations reported here refer to this active fraction.

λ PR promoter DNA was prepared as described previously⁵. The 125 bp linear promoter fragment has the natural λ PR promoter sequence from -60 to $+1$ with a modified initial transcribed region (ITR). Short primers (non-promoter DNA) were used to extend the sequence to -82 upstream and $+42$ downstream. Table S2 lists sequences of the primers and templates used in this research. Single base exchanges in the λ PR ITR (specified for the nontemplate strand) were made at positions $+5$ (C to G) and $+17$ (G to C) to stop or pause transcription after synthesis of a 16-mer RNA by withholding CTP.

Escape of RNAP from λ PR promoter is found to occur in extension of the RNA from 10-mer to 11-mer^{5, 37}. Hence 10-mer and shorter RNA lengths are designated short RNAs and 11-mer and longer RNAs are designated full-length (FL) RNA. Post-escape transcription is single round; dissociation of a paused RNAP is slow and heparin was added to sequester any dissociated RNAP. Significant misincorporation occurs at the 16-mer stop, as reported previously in other contexts^{8, 31, 38}, resulting in readthrough to a second C stop at $+32$.

Single Round Transcription Assays

λ PR promoter DNA (fragment length -82 to $+42$; 100 nM final concentration) is incubated with a twofold excess of active RNAP holoenzyme in TB for 1 h at 19°C to form open complexes (OC). Initiation kinetic experiments are performed at 19°C with a KinTek Rapid Quench Flow (RQF). At zero time, λ PR OCs and the appropriate RQF IS are rapidly mixed 1:1 in the RQF, reducing concentrations of λ PR OC, NTPs and heparin to half their initial values. Experiments are quenched at the desired time points (0.5 – 150 s) by rapid mixing with ~ 10 -fold excess volume of QS and collected in 1.5 mL tubes. For separation of RNAs by length, 8.8 μ L of the quenched reaction mixture is combined with 1.2 μ L of QSD and a 5 μ L aliquot is loaded onto a 20% acrylamide gel.

Gels from transcription assays are run in TBE and transferred to a phosphorimaging cassette for 18 hours of exposure. Gels are imaged using a Typhoon 9000 phosphorimager and analyzed using ImageQuant software. Statistical corrections for ^{32}P incorporation are made as reported previously to obtain RNA concentrations⁵. Quantitative results reported here for experiments with IS1 and IS5 are the average of 4 RQF experiments at each NTP concentration, and experiments with IS3 are the average of 3 RQF experiments. The uncertainties reported for these sets of experiments are one standard deviation from the mean. Quantitative results reported for experiments with IS2 and IS4 are the average of 2 RQF experiments at each NTP concentration, and the uncertainty is estimated as half the difference between the two results.

Analysis of Transient Species on Pathway to FL RNA

Amounts of FL RNA (>10-mer) and of the eight short RNA species (3-mer to 10-mer) formed prior to promoter escape are calculated as a function of time from phosphorimager line scans of transcription gels with α - ^{32}P -labeled GTP or UTP in IS at low (10 μM or less) concentration of the NTP used for labeling and 75 or 200 μM ATP and the other NTP. The 2-mer (pppApU) is not labeled in experiments using α - ^{32}P -GTP and is not separated sufficiently from the large peak of unincorporated labeled monomer to resolve in experiments using α - ^{32}P -UTP. Results from 2–5 independent initiation kinetics experiments at each NTP condition are averaged. Amounts of each short RNA synthesized by nonproductive (stalled) complexes as a function of time during the fast phase of FL RNA synthesis are determined as described in Results and SI. These amounts are subtracted from the total amount of that RNA to obtain the amount of each short RNA intermediate synthesized by productive complexes on the pathway to FL RNA synthesis as a function of time.

Time courses of appearance and disappearance of these transient populations of true intermediates in FL RNA synthesis and of formation of FL RNA at the above NTP conditions are globally fit to an eleven-step minimal mechanism (Scheme 1) in which all steps after ordered binding of the first and second initiating NTP are irreversible because catalysis is effectively irreversible for the conditions studied. In this fitting, the dissociation equilibrium constant $K_{d,1}^{\text{ATP}}$ for the initiating ATP is fixed at $5 \times 10^{-4} \text{ M}$ ³⁹, while the individual on and off rate constants are floated, subject to the constraint of this $K_{d,1}^{\text{ATP}}$. The off-rate constant is found to be very large ($> 1000 \text{ s}^{-1}$) indicating that this step is in rapid equilibrium on the time scale of UTP binding. At this promoter, neither DNA opening nor translocation is necessary for UTP binding and synthesis of pppApU⁴⁰.

In all steps after pppApU synthesis, reversible translocation and NTP binding precede catalysis. In elongation, translocation is unfavorable and is driven by NTP binding, resulting in apparent Michaelis constants (K_m values) for these reversible steps which are much larger than for NTP binding alone²⁰. As developed in Results and Discussion below, translocation in initiation is made more unfavorable because of translocation stress, resulting in apparent K_m values for steps of initiation (here designated $K_{m,i}^{\text{obs}}$) involving translocation which significantly exceed the NTP concentrations examined ($K_{m,i}^{\text{obs}} \gg [\text{NTP}_i]$). This situation allows

data at different [NTP] to be fit using only a second-order composite rate constant (analog of k_{cat}/K_m) for each step involving unfavorable translocation (see Results and Discussion). Since pppApU synthesis does not involve translocation⁴⁰, this process has a much smaller K_m and requires separate parameters for reversible UTP binding (quantified by K_m) and catalysis of 2-mer synthesis (quantified by k_{cat}) to fit kinetic data for 3-mer and larger RNAs at the very different UTP concentrations examined.

RESULTS

Gel Visualization of the Progression of Short RNA Intermediates in Single-Round Full-Length RNA Synthesis

Representative gel electrophoretic separations of individual short and FL RNA species present in the time range from 0.5 s to 90 or 150 s during transcription initiation and escape of RNAP from the λ PR promoter at 19 °C are shown in Figure 1 A–C for three different sets of NTP concentrations. The transcript sequence is pppA⁺¹pUpGpUpApG⁺⁶pUpApApG pG⁺¹¹pApGpGpUpU⁺¹⁶pC... which is modified from that of λ PR so that the first C in the transcript is at position +17, stopping elongation at a 16-mer RNA when CTP is withheld. Transcription is initiated by rapid mixing of preformed OC with an initiation solution (IS) containing ATP, GTP and UTP. Low concentrations (typically 10 μ M) of either unlabeled GTP (with 35 nM α -³²P-GTP; Fig 1 A, B) or of unlabeled UTP (with 35 nM α -³²P-UTP; Fig. 1 C) are used to obtain efficient incorporation of α -³²P-label in the transcript. Even though significant read-through of the +17 stop is observed at longer times in these experiments, a second stop-point after synthesis of a 31-mer prevents run-off of RNAP. Heparin, added with the NTP mixture, binds tightly to free RNAP and prevents re-initiation by any RNAP that does dissociate, as well as any possible interference from inactive RNAP which is incapable of forming open complexes at the promoter. Samples are quenched at the indicated times. The different RNA species present at each time in initiation are separated on a polyacrylamide gel (Fig. 1) and quantified by phosphorimager as detailed in Methods.

Gels in Figure 1 panels A and B show α -³²P-GTP-detected time courses of short and FL RNA synthesis in initiation at relatively high concentrations (75 μ M, panel A; 200 μ M, panel B) of the two initiating NTP (ATP, UTP) and low (10 μ M) GTP. Time courses of short and FL RNA synthesis at lower GTP concentrations (3 μ M, 1 μ M) at 200 μ M ATP and UTP are shown in Fig. S1. All of these low-GTP gels show a clear progression of transient buildup and decay of short RNAs on the time course of FL RNA synthesis, beginning with 3-mer (which is a maximum at or before ~ 0.5 s at 10 μ M GTP, 75 or 200 μ M ATP and UTP) and continuing from 4-mer to 10-mer with increasing time (1 s to 10 s). This progression reveals the extension of the RNA-DNA hybrid to make FL RNA. The initial dinucleotide pppApU is not labeled and so cannot be detected. Escape of RNAP from the highly stable λ P_R OC and the transition to elongation occurs at the 10- to 11-mer step^{5,37}, and all longer RNA products are designated full-length (FL) RNA. FL RNA is predominantly 16-mer, with some 13-mer at shorter times and a slow read-through to 31-mer. At these NTP concentrations, the first appearance of FL RNA is at 2–3 s and FL synthesis is complete by 30 s.

The gel in Figure 1 panel C shows an α -³²P-UTP-detected time course of short and FL RNA synthesis in initiation at a much lower concentration of UTP (10 μ M) and relatively high

concentrations (200 μ M) of ATP and GTP. Several aspects of the time course at low UTP (Fig. 1 C) differ significantly from those at low GTP (Fig 1 A, B). At the low UTP condition, only three of the eight short RNAs (3-, 6-, 10-mer) exhibit the progression of significant transient buildup and decay observed for all eight short RNAs (3- to 10mer) at the low-GTP conditions of Fig. 1 A, B. Also, the kinetics of accumulation and decay of these species and the kinetics of FL RNA synthesis are slower than at the low-GTP condition of Fig. 1 A, B. Though labeled, the initial dinucleotide pppApU is not resolved from the labeled α -³²P-UTP. Synthesis of FL RNA in Panel C occurs after a 5 s lag and is complete at about 45 s.

Changes in Short RNA Populations with Time in Single-Round FL RNA Synthesis

The gels of Fig. 1 and SI Fig S1 reveal that, at all NTP concentrations investigated, amounts of 3-, 6- and 10-mer RNAs increase to a maximum and then decrease during the time course of full-length RNA synthesis. Phosphorimager analyses of two to four gels from independent experiments at each set of NTP concentrations were averaged to quantify the concentrations (normalized per OC) of FL RNA and of each short RNA as a function of time after mixing NTP with OC. For the three sets of NTP concentrations investigated in Fig. 1, amounts of 3-, 6- and 10-mer RNAs (normalized per open complex) are plotted in Fig. 2 as a function of time over the interval 0 – 90 s. Amounts of all other short RNA species (4- to 9-mer) observed in these experiments are plotted as a function of time in SI Figures S2–S4.

At all sets of NTP concentrations studied in Fig 1, single-round synthesis of FL RNA (quantified in 1A-C of Fig. 2) exhibits a lag of 2–4 s after mixing before increasing to a plateau value of about 0.4–0.5 FL RNA/OC at 30 s, similar to the plateau value for FL RNA/OC in single round assays at 37 °C at this promoter⁵. This subpopulation of OC that initiates and synthesizes FL RNA is called productive. Other OC, called nonproductive, were found to stall after making a first short RNA in this time range⁵. At 37 °C, short RNAs are slowly released from non-productive OC, allowing re-initiation and the start of a cycle of abortive initiation which does not result in FL RNA synthesis⁵.

By analogy with previous observations at 37 °C, changes in concentration of each short RNA on the time scale of FL RNA synthesis in experiments at 19 °C like those of Fig. 1 are expected to result from the net effect of synthesis and extension of that RNA length by both productive and nonproductive complexes. Our primary interest here is in the kinetics of synthesis and extension of short RNAs by productive complexes, since those short RNAs are the intermediates in FL RNA synthesis. To obtain this information it is necessary to separate contributions from productive and nonproductive complexes to the amount of each short RNA per OC as a function of time in FL RNA synthesis. The remaining panels of Fig. 2, and SI Figs. S2–S4 indicate how this is accomplished.

For all short RNAs plotted in Fig. 2, maxima in the amounts of RNA/OC are observed at times in the range 0.5 – 5 s, before much FL RNA synthesis is observed. In experiments at 10 μ M GTP, 75 μ M or 200 μ M ATP and UTP, such maxima are observed for all short RNA lengths from 3- to 10-mer (Fig 2A, B and SI Figs S2–S3). In experiments at 10 μ M UTP, 200 μ M ATP and GTP, analogous maxima are observed for some but not all RNA lengths (particularly 3-, 6-, 10-mer; Fig. 2C and SI Fig. S4). Three generalizations can be drawn

from these results. i) For all three sets of NTP concentrations investigated, the time at which the maximum is attained increases with increasing RNA length, consistent with sequential formation and subsequent extension of each intermediate RNA length up to the escape point in FL RNA synthesis. ii) Notably, instead of decreasing to zero, amounts of short RNAs per OC decrease to significant, non-zero values which are maintained to the time where FL RNA synthesis is complete (20–30 s at 19 °C). iii) At longer times than that required for completion of FL RNA synthesis, amounts of the shorter RNAs (especially 3- and 4-mer; Fig. 2 and SI Figs S2–S4) increase slowly and linearly, as expected for abortive synthesis of these very short RNAs. In this time range, no significant change is observed in the amounts of the longer RNAs (5- to 10-mer).

The long-time synthesis of 3- and 4-mer (iii above) is consistent with previous observations for all short RNAs (3- to 10-mer) in manual mixing (10 s deadtime) assays of transcription initiation at 37 °C⁵. A single round of FL RNA synthesis by about half the population of OC (the productive subpopulation) was completed within the deadtime of these 37 °C assays. The other half of the OC population, designated nonproductive, stalled after rapid (< 10 s) synthesis of a first short RNA, slowly released this short RNA and reinitiated RNA synthesis (abortive cycling) but was unable to escape and synthesize FL RNA on the 480 s time scale of these experiments. Steady-state rates of abortive synthesis of 3- and 4-mer, obtained from the slopes, are listed in Table S4. For the set of NTP concentrations investigated at both 37 °C and 19 °C (10 μM UTP, 200 μM ATP and GTP), the abortive rate for 3-mer is ~30% as large at 19 °C as at 37 °C, while the abortive rate for 4-mer is only ~10% as large at 19 °C as at 37 °C. Results in Fig. 2 and SI Figs S2–S4 show there is no detectable abortive synthesis of longer RNAs (5- to 10-mer) at 19 °C. We conclude that nonproductive complexes that stall after synthesis of 5- to 10-mer RNA at 19 °C do not release their RNA and reinitiate, and hence show extreme product inhibition. The above observations on λPR nonproductive complexes are completely consistent with earlier observations of abortive rates in initiation at the lacUV5 promoter⁴¹ and can be interpreted in terms of the expected large differences in thermal stability and lifetime of the RNA-DNA hybrid in a stalled nonproductive initiation complex as its length increases. Both stability and lifetime of nucleic acid duplexes increase with increasing length and with decreasing temperature⁴².

While the times of the maxima in Fig. 2 and SI Figs S2–S4 increase with increasing RNA length, the amounts of RNA/OC at the maxima do not exhibit any simple progression. For example, at the low-GTP condition of Fig 2B, the maximum in the amount of 10-mer RNA (panel B4; 0.19 10-mer RNA/OC, occurring at 5 s) is almost twice as large as those observed for 3-mer and 6-mer (0.11 RNA/OC, occurring at 0.5 s (panel B2) and 1.5 s (panel B3)). At the low-UTP condition of Fig 2C, the maximum in the amount of 10-mer RNA (panel C4; 0.02 10-mer RNA/OC, occurring at 5 s) is much smaller than the maxima in 3-mer (panel C2; 0.45 RNA/OC, occurring at 2.5 s) and 6-mer RNAs (panel C5; 0.12 RNA/OC, occurring at 5 s). The explanation is that the height, breadth and position of these transients are determined by the rates of synthesis and elongation of each RNA length, which are determined by concentrations of both NTP and the initiating complexes as well as the kinetic constants (k_{cat} , K_{m} , $k_{\text{cat}}/K_{\text{m}}$) for these enzyme-catalyzed reactions. Global fitting of these FL and short RNA profiles, in the next section of Results, yields these kinetic constants and allows the height, breadth and position of these transients to be predicted.

At all NTP concentrations investigated (e.g. those of Figs. 1 and 2), FL RNA synthesis exhibits an initial lag of several seconds, followed by a first order (single-exponential) approach of the amount of FL RNA to a plateau (or near-plateau) value. For both low (10 μ M) GTP conditions investigated (Fig 1 A, B), the lag in FL synthesis is about 2.5 s and the rate constant for the approach to the plateau is $\sim 0.11 \text{ s}^{-1}$. For the low (10 μ M) UTP conditions of Fig. 1 C, the lag is longer (3.6 s) and the rate constant for FL RNA synthesis is smaller ($\sim 0.06 \text{ s}^{-1}$). At the low GTP sets of NTP concentrations studied in panels A and B of Figs 1–2, FL RNA synthesis is complete in ~ 20 –40 s. At the low UTP NTP concentration in panel C of Figs 1–2, FL RNA synthesis is complete in 90 s, as compared to 10–20 s at the low-UTP set of NTP concentrations investigated previously at 37 °C⁵.

For the low-UTP set of NTP concentrations investigated, the kinetics of accumulation of 7- and 8-mer RNAs (plotted in Fig. 3) also exhibit first order approaches to small but significant plateau values ($\sim 0.014 \text{ RNA/OC}$) with similar rate constants of $\sim 0.015 \text{ s}^{-1}$ after brief (1–2 s) lags. These kinetics for 7- and 8-mer at this low-UTP condition, judged to be mostly if not entirely from the first round of synthesis by non-productive complexes, have the same functional form and similar rate constants (though shorter lag times) as for FL RNA synthesis by productive complexes. Because 7- and 8-mer synthesis by nonproductive complexes, as well as FL RNA synthesis by productive complexes, exhibit first-order approaches to plateau values after brief initial lags, we assume this behavior is general and use it to model the kinetics of first-round synthesis of each short RNA length by non-productive complexes (see Fig. S2–S4). In this way we are able to predict the time course of synthesis of the first short RNA of each length by nonproductive complexes. At each time point in FL RNA synthesis, subtraction of the amount of each RNA synthesized by nonproductive complexes from the observed total amount of that RNA yields the amount of each intermediate RNA length present as a RNA-DNA hybrid in a productive complex (on the pathway to FL RNA synthesis) as a function of time. Details of this analysis are described in SI text, SI Figs S2–S4 and SI Tables S4–S6.

Quantifying the Buildup and Decay of Short RNA Intermediates in FL RNA Synthesis by Productive Initiation Complexes

Time courses of formation and decay of each short RNA transient intermediate (3-mer to 10-mer) on the pathway to FL RNA synthesis by productive initiation complexes at 19 °C at the low (10 μ M) GTP and low (10 μ M) UTP conditions of Figures 1 and 2, calculated from the data of Figs. 1–2 and SI Figs. S2–S4 as described above, are shown in Fig. 4 and tabulated in SI Tables S4–S6. In all cases, a progression from the shortest (3-mer) to the longest (10-mer) RNA-DNA hybrid with increasing time in the process of promoter escape and FL RNA synthesis is observed.

In experiments performed at low (10 μ M) GTP, high (75 or 200 μ M) ATP and UTP (Fig. 4 transient concentrations of 5-, 9- and 10-mer RNA intermediates in FL RNA synthesis are large because the next base incorporated is G, and the rate of incorporation of G is reduced at low GTP concentration. However, transient concentrations of 3- and 6-mer (next base U) in these low-G experiments are also large, indicating that the following steps in FL RNA synthesis (making 4- and 7-mer) must be intrinsically slow and that these steps have small

overall rate constants. Large transient concentrations of all these short RNAs are also observed in experiments at lower (1, 3 μM) GTP and high (200 μM) ATP and UTP concentrations (Figure S6).

Concentration profiles of transient intermediate RNAs on the pathway to FL RNA synthesis that are detected in initiation kinetic experiments at low (10 μM) UTP, high (200 μM) ATP and GTP are shown in Figure 4C. Only three transient intermediate short RNAs (3-mer, 6-mer, 10-mer) are observed; concentrations of other RNA intermediates are too small to detect for this choice of NTP concentrations. Accumulation of 3-mer and 6-mer is expected because positions 4 and 7 of the ITR sequence are U, but accumulation of 10-mer (next base G) cannot be explained this way and must be the result of an intrinsically small overall rate constant for synthesis of 11-mer from 10-mer (see below).

Tables S4–S6 list concentrations of each small RNA and of FL RNA present in productive complexes at each time up to completion of FL RNA synthesis. Experimental values are given for 3-mer RNA and all larger RNAs. For 2-mer, experimental data are lacking but concentrations can be predicted by the quantitative analysis developed in the next section. Tables S4–S6 show that within experimental uncertainty of ± 10 –20% the sum of concentrations of all RNA lengths is equal to the concentration of productive complexes, showing that the analysis of the data of Figs. 1–2 and SI Figs. S2–S4 to obtain these amounts is correct and internally consistent.

A key result of this and previous sections for the analysis that follows is the observation of significant transient concentrations of 3-, 6- and 10-mer in both low-GTP and low-UTP experiments (Figs. 2 and 3). These results cannot be simply explained by appealing to rate reductions from the low concentration of the next NTP and indicate that overall rate constants for the subsequent step(s) of NTP incorporation are small. The following quantitative analysis confirms this and indicates that these small rate constants arise from large translocation stresses and disruption of RNAP-promoter interactions in these steps.

Dependence of Rate Constants for RNA-DNA Hybrid Extension on Hybrid Length

Results of Figure 3 for the time course of each step ($i > 2$) of NTP incorporation and hybrid extension in initiation are interpreted in terms of the mechanism shown as Scheme 2 that includes reversible translocation, reversible NTP binding and catalysis. This mechanism was previously used to interpret single-molecule kinetic data for elongation of a transcript by pol II RNAP²⁰. Steady-state analysis⁴³ of Scheme 2 yields an analytical expression for the velocity as a function of NTP concentration (Eq. 1)²⁰, which is as applicable to initiation as elongation.

Expressed per open complex, the velocity v_i of step i (Scheme 2) as a function of $[\text{NTP}_i]$ has the standard hyperbolic functional form of noncooperative enzyme-catalyzed reactions:

$$v_i = k_i[\text{NTP}_i] \left(1 + \frac{[\text{NTP}_i]}{K_{m,i}^{\text{obs}}} \right)^{-1} \quad \text{Eq. 1}$$

In Eq. 1, k_i is the overall second order rate constant for NTP incorporation and hybrid extension:

$$k_i = f_{i-1}^{\text{post-tr}} \left(\frac{k_{\text{cat},i}}{K_{m,i}^{\text{NTP}}} \right) \quad \text{Eq. 2}$$

In Eq. 2, $k_{\text{cat},i}$ is the rate constant of the catalytic step in Scheme 2 and $K_{m,i}^{\text{NTP}}$ is the intrinsic Michaelis constant for the interaction of NTP_i with the post-translocated state ($\text{IC}_{i-1}^{\text{post-tr}}$):

$$K_{m,i}^{\text{NTP}} = \frac{(k_{\text{NTP},i}^- + k_{\text{cat},i})}{k_{\text{NTP},i}^+} = K_{d,i}^{\text{NTP}} \left(1 + \frac{k_{\text{cat},i}}{k_{\text{NTP},i}^+} \right) \quad \text{Eq. 3}$$

where $k_{\text{NTP},i}^+$ and $k_{\text{NTP},i}^-$ are rate constants for binding of NTP_i to and dissociation from $\text{IC}_{i-1}^{\text{post-tr}}$ and $K_{d,i}^{\text{NTP}} = k_{\text{NTP},i}^- / k_{\text{NTP},i}^+$. Also, in Eq. 2 $f_{i-1}^{\text{post-tr}}$ is the fraction of initiating complexes with a hybrid length of $i-1$ bases present as $\text{IC}_{i-1}^{\text{post-tr}}$ prior to binding of NTP_i :

$$f_{i-1}^{\text{post-tr}} = \left(\frac{\text{IC}_{i-1}^{\text{post-tr}}}{[\text{IC}_{i-1}^{\text{pre-tr}}] + [\text{IC}_{i-1}^{\text{post-tr}}]} \right) = \left(\frac{K_{i-1}^{\text{tr}}}{1 + K_{i-1}^{\text{tr}}} \right) \quad \text{Eq. 4}$$

In Eq. 4 the equilibrium constant K_{i-1}^{tr} is the quotient of forward ($k_{i-1}^{+, \text{tr}}$) and back ($k_{i-1}^{-, \text{tr}}$) translocation rate constants (Scheme 2):

$$K_{i-1}^{\text{tr}} = \left(\frac{f_{i-1}^{\text{post-tr}}}{1 - f_{i-1}^{\text{post-tr}}} \right) = \left(\frac{k_{i-1}^{+, \text{tr}}}{k_{i-1}^{-, \text{tr}}} \right) \quad \text{Eq. 5}$$

Also in Eq. 1, $K_{m,i}^{\text{obs}}$ is the observed Michaelis constant, related to the intrinsic Michaelis constant $K_{m,i}^{\text{NTP}}$ (Eq. 3) by:

$$K_{m,i}^{\text{obs}} = K_{m,i}^{\text{NTP}} (f_{i-1}^{\text{post-tr}})^{-1} \left(1 + \frac{k_{\text{cat},i}}{k_{i-1}^{+, \text{tr}}} \right)^{-1} = K_{d,i}^{\text{NTP}} (f_{i-1}^{\text{post-tr}})^{-1} \quad \text{Eq. 6}$$

$K_{m,i}^{\text{obs}}$ and k_i generalize the intrinsic quantities $K_{m,i}^{\text{NTP}}$ and $k_{\text{cat},i} / K_{m,i}^{\text{NTP}}$ to include the consequences of reversible translocation on the velocity v_i of the enzyme-catalyzed reaction (Eq. 1).

A global fit of amounts of transient short RNA species and FL RNA (Fig. 3) to the irreversible eleven-step mechanism of RNA-DNA hybrid elongation and FL RNA synthesis (Scheme 1) yields composite second order rate constants k_i (Eq. 4) for all steps of initiation. Rate constants for all steps involving translocation are plotted in Fig. 5 (see also Table S8). The fits to the transient populations of each RNA species vs time obtained using these rate constants are shown in Figure 4. Composite fits to the total transient population of short RNA as a function of time are shown in Figure S7. These fits accurately reproduce many aspects of the time evolution of the short RNA intermediates and the synthesis of FL RNA. Some deviations in the decay of individual short RNA species are observed which may result from competition between correct and incorrect NTP, especially at steps of G-incorporation at low GTP and of U-incorporation at low UTP, and/or from deviations from the one parameter (k_i) fit used in Scheme 1 for $i > 2$.

Synthesis of the initial RNA dinucleotide (here pppApU) is the only step of initiation without translocation. Notably, we find from our fitting that this is the only step of the mechanism for which separation of reversible NTP binding from catalysis is necessary in order to obtain a good fit to the initiation kinetic data at the different low-G and low-U conditions investigated. The fitting analysis indicates that binding of UTP is in rapid equilibrium on the time scale of catalysis of pppApU synthesis, with a lower bound on the UTP dissociation rate constant ($k_{\text{NTP},2}^- > 72 \text{ s}^{-1}$) which is significantly larger than the catalytic rate constant for pppApU synthesis ($k_{\text{cat},2} = 17 \pm 1 \text{ s}^{-1}$). The fitted dissociation equilibrium constant for UTP is $K_{\text{d},2}^{\text{UTP}} = (3.6 \pm 0.5) \times 10^{-5} \text{ M}$. From these results we calculate an overall second order rate constant $k_2 = k_{\text{cat},2}/K_{\text{m},2}^{\text{UTP}} = (4.4 \pm 0.6) \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ from the average of upper-bound and lower-bound estimates of $K_{\text{m},2}^{\text{UTP}}$ from the above $K_{\text{d},2}^{\text{UTP}}$ and $k_{\text{cat},2}$ values.

The kinetics of each subsequent step of initiation (including the sub-steps of reversible translocation, reversible NTP binding, and catalysis as in Scheme 2) are found to be well-fit with a single composite rate constant k_i (Eq. 1–2). Indeed, for the range of [NTP] examined ($< 200 \mu\text{M}$), unique fits cannot be obtained for these steps ($i > 2$) if NTP binding is separated from catalysis in an attempt to obtain both k_i and $K_{\text{m},i}^{\text{obs}}$ from the fit. This shows that velocities of these steps are well described by $v_i \approx k_i[\text{NTP}_i]$, the limiting form of Eq. 1 valid if $K_{\text{m},i}^{\text{obs}} > [\text{NTP}_i]$. In Discussion, we conclude that these $K_{\text{m},i}^{\text{obs}}$ values are much larger than Michaelis constants for UTP in pppApU synthesis ($K_{\text{m},2}^{\text{UTP}} \approx K_{\text{d},2}^{\text{UTP}} = 3.6 \times 10^{-5} \text{ M}$; see above) or for elongation ($K_{\text{m},\text{elong}}^{\text{NTP}} \approx 9 \times 10^{-5} \text{ M}$, $K_{\text{m},\text{elong}}^{\text{obs}} \approx 8.0 \times 10^{-5} \text{ M}$; ²⁰) because of the additional free energy cost of translocation of the hybrid into the cleft in initiation, generating translocation stress and disrupting RNAP-promoter contacts prior to RNAP to escape. Consistent with this picture, values of k_i for all these steps (Table S8, Fig. 4) are significantly smaller than k_2 , by factors ranging from 1/3 to 1/40.

Rate constants for the extension of 2- to 3-mer ($1.5 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$), of 5- to 6-mer ($1.6 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$), and of 9- to 10-mer ($1.3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$) are about one-third as large as k_2 .

Subsequent steps exhibit much smaller rate constants. The smallest rate constants, indicating the most difficult translocation steps, are for extension of 3- to 4-mer ($1.2 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$), for extension of 6- to 7-mer ($1.6 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$) and for promoter escape (10-mer to FL RNA; $2.7 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$). Rate constants for extension of 4- to 5-mer, of 7- to 8-mer, and of 8- to 9-mer are also small. The significant transient populations of 3-, 6- and 10-mer observed at all NTP conditions investigated result from the combination of large rate constants for forming these species and small rate constants for subsequent steps. Most dramatically, the combination of small rate constants for conversions of 3- to 4-mer and 6- to 7-mer with the need to incorporate U as the 4th and 7th base in the RNA results in strong transient accumulation of 3- and 6-mer as the two major intermediates on the time-course of FL RNA synthesis at low UTP (Fig 3 Panel C). This parallels the observation that, at low UTP, the main stall-points for nonproductive complexes are after during extension of 3- and 6-mer (Fig. S5).

No correlation is observed between rate constants k_i for steps of initiation and the NTP incorporated. Rate constants k_i are relatively large for four steps of initiation, one incorporating UTP and three incorporating GTP (Figure 5). Rate constants are much smaller for the other six steps of initiation, three incorporating ATP, two incorporating UTP and one incorporating GTP.

Inclusion in the global analysis of experiments performed at 1 μM and 3 μM GTP resulted in systematic deviations from observed transients at steps where G was the base incorporated. These deviations were resolved by replacing these actual GTP concentrations by effective GTP concentrations (parameterizations for misincorporation^{31, 38}) of 2 μM at 1 μM actual GTP and 5 μM at 3 μM GTP. Fits using these effective GTP concentrations and the k_i values reported here are shown in SI Fig S6.

DISCUSSION

Interpretation of Reductions in Initiation Rate Constants (k_i) in Terms of Translocation Stress and Disruption of RNAP-Promoter Interactions

To escape from the promoter and transition to the elongation phase of transcription, the interactions between RNAP and the promoter responsible for stability and specificity of the open complex must be disrupted. Translocation of the RNA-DNA hybrid into the cleft provides the steric^{4, 21, 22} and scrunching^{23, 24} stress necessary to accomplish this. We deduce that translocation in initiation is rapidly reversible and unfavorable, as is the case for elongation^{20, 29, 33}, and that each step of translocation is driven by NTP binding to the post-translocated state. Translocation stresses in initiation are predicted to reduce $f_{i-1}^{\text{post-tr}}$ and K_{i-1}^{tr} for reversible translocation, relative to their values in the absence of translocation stress or in elongation, and thereby increase observed $K_{m,i}^{\text{obs}}$ value for NTP binding (Eq. 6) and reduce NTP incorporation rate constants k_i .

A range of scenarios are possible for how translocation stress drives promoter escape. Stress could build up steadily in the region of the initiation bubble until sufficient to disrupt all

specific RNAP-promoter interactions at a critical hybrid length, resulting in promoter escape. Alternatively, stress could be minimized at each step of hybrid extension by disruption of RNAP-promoter interactions contiguous to the initiation bubble, thereby distributing translocation stress over a larger region of promoter DNA.

The striking pattern of k_i values in Figure 4 indicates that the actual situation does not correspond to either of the above extremes. Instead, the k_i values indicate that stress buildup and disruption of RNAP-promoter interactions occurs in three stages, ending in escape of RNAP. From a comparison of OC stabilities and escape points for λ PR and T7A1 promoters, we previously deduced that each step of translocation of the hybrid into the cleft in initiation destabilized the initiating complex by ~ 1 kcal/mol⁵. Single-molecule analysis of pre-scrunching energetics yielded a similar result²⁵. Escape of RNAP from the λ PR promoter occurs after 9 translocation steps and therefore should require ~ 9 kcal of driving free energy from these translocation steps. In what follows, we analyze initiation k_i values to test this result and dissect the 9 kcal into contributions from disruption of three structurally-distinct groups of RNAP-promoter interactions.

For elongation by pol II, translocation is reversible and the pre-translocated state is strongly favored ($f_{\text{elong}}^{\text{post-tr}} = 0.13$) in the absence of NTP binding²⁰, presumably because translocation in elongation involves a net disruption of 1 base pair (bp) of RNA-DNA hybrid as well as formation (upstream) and disruption (downstream) of one DNA-DNA bp. We propose that a similar intrinsic bias toward the pre-translocated state is present in the steps of initiation, because translocation in initiation involves opening one DNA-DNA bp at the downstream end of the bubble with no compensating effect from hybrid or upstream DNA base pairing. We define the reference quantity $f_{\text{-trs}}^{\text{post-tr}}$, the (hypothetical) fraction of complexes of a given hybrid length in the post-translocated state in the absence of initiation-specific translocation stresses (steric, scrunching, disruption of RNAP-promoter interactions). We assume that $f_{\text{-trs}}^{\text{post-tr}}$ is small ($f_{\text{-trs}}^{\text{post-tr}} \ll 1$, probably comparable to $f_{\text{elong}}^{\text{post-tr}} = 0.13$), because of the cost of base pair opening. Initiation-specific translocation stresses are expected to increase the bias toward the pre-translocated state significantly, causing $f_{i-1}^{\text{post-tr}}$ to be less than $f_{\text{-trs}}^{\text{post-tr}}$ and therefore reducing k_i (Eq. 2) relative to its hypothetical value without translocation stress ($k_{\text{-trs}}$).

The reductions in $f_{i-1}^{\text{post-tr}}$ in steps of initiation result in increases in $K_{m,i}^{\text{obs}}$ (Eq. 6) relative to its value ($K_{m,-\text{trs}}^{\text{obs}}$) for the hypothetical situation where translocation in initiation occurs without stress-buildup and concomitant disruption of RNAP-promoter contacts: $K_{m,i}^{\text{obs}} > K_{m,-\text{trs}}^{\text{obs}} > K_{m,i}^{\text{NTP}}$. If $f_{i-1}^{\text{post-tr}}$ were not much less than $f_{\text{-trs}}^{\text{post-tr}}$, then $K_{m,i}^{\text{obs}}$ would not exceed the maximum NTP concentration investigated here (200 μM) and composite rate constants k_i would not be sufficient to fit kinetic data for translocation steps in Scheme 1 (see Results).

From Eqs. 2–4, for the situation $f_{i-1}^{\text{post-tr}} < f_{\text{-trs}}^{\text{post-tr}} \ll 1$ and $K_{i-1}^{\text{tr}} < K_{\text{-trs}}^{\text{tr}} \ll 1$ applicable in initiation and discussed above:

$$\ln\left(\frac{k_i}{k_{-tr}}\right) \approx \ln\left(\frac{f_{i-1}^{post-tr}}{f_{-tr}^{post-tr}}\right) \approx \ln\left(\frac{K_{i-1}^{tr}}{K_{-tr}^{tr}}\right) = -\left(\frac{\Delta G_{i-1}^{0,tr} - \Delta G_{-tr}^{0,tr}}{RT}\right) = -\frac{\Delta G_{tr,i}^0}{RT} \quad \text{Eq. 7}$$

Equation 7 relates the ratio k_i/k_{-tr} to the free energy change $\Delta G_{tr,i}^0$, from buildup of steric and scrunching stress and disruption of RNAP-promoter interactions in the i -th step.

Setting the sum of $\Delta G_{tr,i}^0$ values for the nine translocation steps in Fig. 4 ($3 \leq i \leq 10$) equal to 9 kcal/mol, as deduced previously⁵, yields a predicted $k_{-tr} \cong 2.7 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ from Eq. 7. This k_{-tr} value is eminently reasonable, since it falls in the relatively narrow (three-fold) range between the lower bound for k_{-tr} provided by the largest k_i values for steps of initiation involving translocation ($k_6 = 1.6 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $k_3 = 1.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) and plausible upper bounds provided by rate constants for two steps of transcription without steric and scrunching contributions to translocation stress ($k_2 \cong 4.4 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for pppApU synthesis (Fig. 4); $k_{elong} \cong 4.4 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for pol II elongation²⁰).

Values of $\Delta G_{tr,i}^0$ calculated from Eq. 7 using $k_{-tr} \cong 2.7 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ are plotted in Figure 6A for the steps of initiation at λ PR. This analysis reveals that translocation stress (steric, scrunching) makes these steps unfavorable by amounts ranging from + 0.4 kcal/mol to +1.8 kcal/mol. The nine steps divide into three cycles, as observed for rate constants k_i in Fig. 5. $\Delta G_{tr,i}^0$ values in the steps synthesizing 3-, 6- and 10-mer RNAs are modest (0.3 – 0.4 kcal/mol for $i = 3, 6, 10$). But the steps immediately following these have much larger $\Delta G_{tr,i}^0$ values (in the range 1.3 – 1.8 kcal/mol) and for the first two cycles are followed by one or two steps with moderately large $\Delta G_{tr,i}^0$ values (0.8 – 1.2 kcal/mol).

Do NTP-specific effects on intrinsic NTP K_D or k_{cat} values in Scheme 2 contribute to the observed dependence of k_i on the step of initiation in Fig. 5? Evidence that any such NTP-specific effects are small in comparison to those observed here and attributed to large differences in K_{tr} for different translocation steps comes from single-molecule elongation kinetic studies with *E. coli* RNAP at room temperature²⁹. For steps of elongation, both K_D and k_{cat} were found to be somewhat NTP-specific but differences were modest and corresponding values of k_{cat}/K_D (an approximation to k_{cat}/K_m) differed by only two-fold for the NTP incorporated here, being about twice as large for incorporation of ATP ($\sim 1.3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$) as for incorporation of GTP ($\sim 7.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) or UTP ($\sim 5.8 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$). Effects of NTP identity on these elongation rate constants are far smaller than the observed variations in initiation rate constants k_i (Fig. 4; Table S8). Initiation k_i values for steps involving translocation are smaller than for elongation and do not correlate with the NTP incorporated. The three steps in which ATP is incorporated are among the six with relatively small k_i values. Of the three steps of initiation incorporating UTP, one has a large k_i value and two have small k_i values; of the four steps incorporating GTP, three have large k_i values and one has a small k_i value. If NTP-specific effects on the k_i values like those identified in elongation were also present in initiation, correction for them would in many cases increase

the differences between the larger and smaller rate constants that we attribute to differences in the translocation equilibrium constant K_{tr} .

The simplest explanation for these patterns is that stress build-up at the beginning of each cycle is used to disrupt RNAP-promoter interactions in the later steps of that cycle. Once a group of interactions are disrupted, additional promoter DNA becomes available absorb the stress of the next translation and $\Delta G_{trs,i}^0$ becomes much less unfavorable. With another translocation, stress builds up and the cycle repeats. After three cycles, all specific RNAP-promoter interactions are disrupted and escape occurs. Sums of $\Delta G_{trs,i}^0$ values for the three cycles are approximately 3, 4, and 2 kcal, respectively, averaging to 1 kcal/mol per step, the same as the overall average.

Relating the Pattern of Rate Constants for Hybrid Extension to Disruption of RNAP-Promoter Interactions; Comparison of *E. coli* and T7 RNAP

Figure 6B proposes a structural basis for the semi-periodic pattern of the rate constants k_i (Fig. 5) and corresponding free energy changes $\Delta G_{trs,i}^0$ (Eq. 7, Fig. 6A) of steps of initiation leading to escape of RNAP from the λ PR promoter. In this proposal, the three cycles of k_i and $\Delta G_{trs,i}^0$ (Fig 6A) values arise from sequential (downstream to upstream) disruption of specific contacts between RNAP and three promoter sequence elements. Figure 6B proposes that interactions in the active site cleft involving the strands of the discriminator element and σ^{70} region 1.2 and/or the core recognition element (CRE)⁴⁴ are disrupted first, as the RNA-DNA hybrid translocates into the cleft in synthesis of the 5-mer RNA from the 2-mer and scrunches the discriminator strands. Next, specific interactions of the nt and t strands of the -10 element with σ^{70} regions 2.3 and 3.0 are broken as the hybrid translocates further into the cleft in synthesis of the 9-mer RNA from the 5-mer, interacting with σ^{70} region 3.2 and scrunching the DNA strands of the bubble. Disruption of these -10 interactions presumably results in closing of the upstream end of the bubble. Lastly, specific upstream interactions between the -35 element and σ^{70} region 4 and between the UP element and the α -CTD break from the steric and scrunching stress generated as the hybrid translocates further in synthesis of the 11-mer RNA from the 9-mer.

Support for this proposal is provided by previous structural and kinetic findings, summarized below, that translocation of the 6-mer hybrid in the second of the three high-stress regions in Figures 5 and 6A is slow and that this step initiates disruption of the strong interactions with the -10 element. In addition, this proposal for *E. coli* RNAP is fundamentally analogous to the previously-proposed pathway of sequential (downstream to upstream) disruption of specific contacts in initiation by and escape of T7 RNAP from a consensus promoter, based on structural and kinetic evidence (Figure 7A, B)^{9, 45}. For both RNAPs, disruption of specific interactions by downstream-applied translocation stress proceeds from downstream to upstream.

In vitro transcription assays and structural studies revealed conformational changes in the *coli* initiation complex that occur when the RNA-DNA hybrid reaches a length of 5–6 nucleotides. At this length, the upstream (5') end of the RNA is near $\sigma_{3.2}$, positioned in the

RNA exit channel²¹. Unfavorable steric and electrostatic interactions develop as the hybrid grows and translocates into the cleft, pushing against $\sigma_{3.2}$ and removing it from the exit channel^{21, 22, 46, 47}. These conformational changes are proposed to break contacts between the nontemplate -10 region and $\sigma_{2.3}$ and between the template strand and σ_3 (Fig 6B, blue structures)²².

Aspects of initiation and escape of RNAP from stable open complexes at other promoters with 6 bp discriminators and near-consensus or consensus -10 and -35 elements are relevant here. *In vivo*, pauses at transcript lengths between +6 to +12 were observed in initiation from two synthetic promoters with consensus -10 and -35 elements⁴⁸. Studies with -10 variants of λ PR found that the overall rate of initiation and escape of RNAP decreased as the -10 sequence approached consensus². Single-molecule kinetic studies on initiation from the lacCONS promoter with consensus -10 and -35 elements found an extended pause ($t_{1/2} \sim 20$ s) in initiation after synthesis of a 6 bp hybrid⁴. This step may be slower than the corresponding step for λ PR because of the greater strength of RNAP contacts with the consensus -10 element of lacCONS than those with the near-consensus -10 element of λ PR.

We therefore deduce that the slow kinetics of hybrid elongation from 6 bp to 9 bp in the second of the three cycles in Fig 5, and the large, positive sum of $\Delta G_{\text{trs},i}^0$ values of these four steps (4 kcal, 44% of the total; Fig. 6A) result from disruption of interactions between *E. coli* RNAP with the -10 region of the λ PR promoter. Hence interactions that involve the discriminator strands with $\sigma_{1.2}$ and/or the CRE (Fig. 6B yellow structure)^{26, 44, 49} in the active site cleft must be disrupted first in the cycle as the hybrid translocates into the cleft in synthesis of the 5-mer from the 2-mer. The sum of $\Delta G_{\text{trs},i}^0$ values for these steps is about 3 kcal, about 33% of the total for promoter escape. Scrunching distortions in the strands of the open region are presumably sufficient to disrupt these in-cleft interactions. Interactions between the -35 region and σ_4 (Fig. 6B green structure) must be broken last to allow promoter escape. For lacCONS, a significant pause ($t_{1/2} \sim 8$ s) in the FRET signal was observed in the step extending the 11-mer in promoter escape, proposed to be from an analogous conformational change⁸. For λ PR, the sum of $\Delta G_{\text{trs},i}^0$ values for these steps is about 2 kcal, about 22% of the total for promoter escape.

Figure 6 and the above discussion propose a model for step-wise disruption of σ^{70} - λ PR promoter interactions with extension of the RNA-DNA hybrid, leading to escape of RNAP from the promoter in the conversion of 10-mer to 11-mer RNA. Our results provide no direct evidence for when the σ^{70} subunit is released from core polymerase. Photo-crosslinking studies of initiation at the λP_R promoter revealed that σ^{70} was no longer present at RNA lengths greater than a 13-mer³⁷. Perhaps relevant to this, 13-mer is the primary post-escape RNA that transiently accumulates in FL RNA synthesis (Fig. 1), indicating that the rate constant for incorporation of GTP to convert 13-mer to 14-mer is smaller than for incorporation of GTP to convert 12- to 13-mer.

Tang '09 obtained results analogous to Fig 5 for initiation by T7 RNAP in studies at a single NTP concentration (500 μ M of each NTP)⁵⁰, much higher than the upper end of the range of

NTP concentrations used to obtain the k_i values in Figure 5 and high enough so it may be comparable to the average $K_{m,i}^{\text{obs}}$ value for initiation with translocation stress. Even so, because the concentrations of all NTPs are the same, the reported first order rate constants for steps of initiation by T7 RNAP can be analyzed by Eq. 7 to obtain a consistent set of lower-bound estimates of corresponding free energy inputs $\Delta G_{\text{trs},i}^0$ for these steps. For this lower-bound calculation, we assume that the largest rate constant in the series corresponds to a situation without translocation stress. This is probably an underestimate; if the hypothetical baseline rate constant $k_{\text{-trs}}$ is two-fold larger, all $\Delta G_{\text{trs},i}^0$ values would increase by 0.4 kcal. If $K_{m,i}^{\text{obs}}$ values for these steps of initiation are not large in comparison to $[\text{NTP}_i] = 500 \mu\text{M}$, that will further increase all $\Delta G_{\text{trs},i}^0$ values.

Lower-bound values of $\Delta G_{\text{trs},i}^0$ for T7 RNAP initiation at a consensus promoter sequence, calculated as above, are plotted vs the step of initiation in Figure 7A. The average $\Delta G_{\text{trs},i}^0$ for T7 RNAP is 0.7 kcal/mol, about 30% less than obtained above for *E. coli* RNAP. Increasing the baseline rate constant $k_{\text{-trs}}$ two-fold (similar to the *E. coli* RNAP situation; see above) would yield average $\Delta G_{\text{trs},i}^0$ for T7 RNAP of 1.1 kcal/mol, almost the same as obtained for *E. coli* RNAP. Similar to initiation by *E. coli* RNAP at λ PR promoter (Figs 5, 6A & 6B), three regions (or more) are observed for T7RNAP in Figure 7A: 2- to 5-mer, 5- to 9-mer, and 9- to 13-mer. Figure 7B shows a structural interpretation of these effects.

The largest free energy input for T7 RNAP occurs for the extension of 5- to a 9-mer RNA, analogous to the disruption of -10 contacts for *E. coli* RNA. Footprinting assays, along with single molecule FRET and structural studies indicated that interactions between the T7 RNAP specificity loop (Fig. 7B, blue structure) and the -7 to -11 region of DNA, along with the valine loop that acts as an intercalating hairpin to open and stabilize the complex, are disrupted during these steps. The sum of lower-bound $\Delta G_{\text{trs},i}^0$ values for conversion of 5- to 9-mer is 4.2 kcal, about 60% of the total for promoter escape and similar to that required to break the -10 contacts between *E. coli* RNAP and λ PR.

Translocation of the hybrid was previously found to disrupt T7 RNAP-promoter interactions sequentially from downstream to upstream^{9, 45}. The non-monotonic progression of rate constants and of lower-bound $\Delta G_{\text{trs},i}^0$ values (Figure 7A)⁵⁰ for synthesis of 5-mer from 2-mer by T7 RNAP indicate that scrunching of in-cleft DNA and disruption of in-cleft interactions downstream of the -7 to -11 element of the T7 promoter must occur in these first steps of hybrid translocation⁹. The sum of lower-bound $\Delta G_{\text{trs},i}^0$ values for conversion of 2-mer RNA to 5-mer is about 1.5 kcal, about 20% of the total for promoter escape.

Single molecule FRET experiments⁹ and studies crosslinking the AT-rich recognition loop (Figure 6B green structure) with the upstream DNA⁵¹ indicate that the final promoter escape step is associated with disruption of the upstream interactions. The sum of lower-bound $\Delta G_{\text{trs},i}^0$ values for conversion of 9-mer RNA to 13-mer is about 1.4 kcal, about 20% of the

total for promoter escape. This process is immediately followed by a large conformational change in subdomain H to form the RNA exit channel^{9, 52, 53}.

The kinetics and energetics of the first step of hybrid synthesis at the promoters investigated differ greatly for *E. coli* and T7 RNAP. For *E. coli* RNAP, the rate constant k_2 of this step is larger than those of all subsequent steps of initiation because the open region of the binary OC includes +2 and, with a 6 bp discriminator, the +2 template strand base is presumably in the active site already before binding of UTP⁴⁰. Therefore k_2 is not reduced by the unfavorable free energy costs of base pair opening and stress buildup from hybrid translocation in initiation that affect subsequent steps. In contrast, the open region in the T7 RNAP-consensus T7 promoter binary OC extends from -4 to -1, and both +1 and +2 base pairs are opened in initiation, driven by binding the complementary NTP^{28, 54}. Consequently, this step is slower than all subsequent steps of initiation, with a lower-bound estimate of its rate constant which is only 5% as large as the *E. coli* RNAP k_2 .

The above kinetic and free energy analyses of translocation steps in initiation by *E. coli* and T7 RNAP are consistent with an average free energy increase from translocation stress ($\Sigma \Delta G_{trs,i}^0 / N$, where N is the number of steps) of ~ 1 kcal/mol per translocation step, in agreement with single-molecule force measurements²⁵. Our kinetic-mechanistic analysis of initiation by *E. coli* RNAP indicates that NTP binding shifts the population distribution from the pre- to the post-translocated state and thereby overcomes these positive $\Delta G_{trs,i}^0$ values.

Individual values of $\Delta G_{trs,i}^0$ vary by about an order of magnitude for different steps of initiation as RNAP-promoter contacts are disrupted to relieve this translocation stress. The strength of the specific interactions of the -10, -35, UP-element and discriminator regions of the promoter with RNAP determine both the rate of escape at the NTP concentration examined^{2, 26, 28, 55}, and the RNA-DNA hybrid length for escape⁵. Stabilizing interactions involving these promoter elements typically increase the overall rate of OC formation, at least at low RNAP concentration, by increasing the closed-complex binding constant, but also reduce the rate of escape and increase the length of the hybrid required for escape. The strength of promoter interactions may also affect the percent of OCs that can productively initiate, or affect the branch point in the pathway of productive and nonproductive complexes^{2, 7, 8}. Hence, as discussed previously, the relationship between homology-to-consensus and promoter strength (rate of productive initiation) is not straightforward. Systematic studies of effects of promoter sequence and RNAP and NTP concentrations on the overall kinetics of promoter binding, OC formation, initiation by productive OC and escape of RNAP from the promoter are ongoing in our laboratory. These are needed to understand the generality of the results reported here and their relevance for regulation of promoter strength by promoter sequence and the many other relevant variables.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We gratefully acknowledge discussions with Bob Landick, Eric Galburt, other colleagues, and the reviewers. We thank Laura Vanderploeg for her assistance in generating publication quality figures.

Funding Sources

This work was supported by NIH Grants GM118100 to M.T.R. and GM122303 to K.L.H.

REFERENCES

- [1]. Ruff EF, Record MT Jr., and Artsimovitch I (2015) Initial events in bacterial transcription initiation, *Biomolecules* 5, 1035–1062. [PubMed: 26023916]
- [2]. Ko J, and Heyduk T (2014) Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse, *Biochemistry Journal* 463, 135–144.
- [3]. Winkelman JT, Vvedenskaya IO, Zhang Y, Zhang Y, Bird JG, Taylor DM, Gourse RL, Ebright RH, and Nickels BE (2016) Multiplexed protein-DNA cross-linking: Scrunching in transcription start site selection, *Science* 351, 1090–1093. [PubMed: 26941320]
- [4]. Duchi D, Bauer DLV, Fernandez L, Evans G, Robb N, Hwang LC, Gryte K, Tomescu A, Zawadzki P, Morichaud Z, Brodolin K, and Kapanidis AN (2016) RNA polymerase pausing during initial transcription, *Molecular Cell* 63, 939–950. [PubMed: 27618490]
- [5]. Henderson KL, Felth LC, Molzahn CM, Shkel IA, Wang S, Chhabra M, Ruff EF, Bieter L, Kraft JE, and Record MTJ (2017) Mechanism of transcription initiation and promoter escape by *E. coli* RNA polymerase, *Proceedings of the National Academy of Sciences* 114, E3032–E3040.
- [6]. Lerner E, Chung S, Allen BL, Wang S, Lee J, Lu SW, Grimaud LW, Ingargiola A, Michalet X, Alhadid Y, Borukhov S, Strick TR, Taatjes DJ, and Weiss S (2016) Backtracked and paused transcription initiation intermediate of *E. coli* RNA polymerase, *Proceedings of the National Academy of Sciences* 113, E6562–E6571.
- [7]. Heyduk E, and Heyduk T (2018) DNA template sequence control of bacterial RNA polymerase escape from the promoter, *Nucleic Acids Research* 46, 4469–4486. [PubMed: 29546317]
- [8]. Dulin D, Bauer DLV, Malinen AM, Bakermans JJW, Kaller M, Morichaud Z, Petushkov I, Depken M, Brodolin K, Kulbachinskiy A, and Kapanidis AN (2018) Pausing controls branching between productive and non-productive pathways during initial transcription in bacteria, *Nature Communications* 9, 1478–1490.
- [9]. Koh HR, Roy R, Sorokina M, Tang G-Q, Nandakumar D, Patel SS, and Ha T (2018) Correlating transcription initiation and conformational changes by a single-subunit RNA polymerase with near base-pair resolution, *Molecular Cell* 70, 695–706. [PubMed: 29775583]
- [10]. Flentie K, Garner A, and Stallings C (2016) Mycobacterium tuberculosis transcription machinery: Ready to respond to host attacks, *Journal of Bacteriology* 198, 1360–1373. [PubMed: 26883824]
- [11]. Abascal-Palacios G, Ramsay EP, Beuron F, Morris E, and Vannini A (2018) Structural basis of RNA polymerase III transcription initiation, *Nature* 553, 301–306. [PubMed: 29345637]
- [12]. He Y, Yan C, Fang J, Inouye C, Tjian R, Ivanov I, and Nogales E (2016) Near-atomic resolution visualization of human transcription promoter opening, *Nature* 533, 359. [PubMed: 27193682]
- [13]. Engel C, Gubbey T, Neyer S, Sainsbury S, Oberthuer C, Baejen C, Bernecky C, and Cramer P (2017) Structural basis of RNA polymerase I transcription initiation, *Cell* 169, 120–131.e122. [PubMed: 28340337]
- [14]. Plaschka C, Hantsche M, Dienemann C, Burzinski C, Plitzko J, and Cramer P (2016) Transcription initiation complex structures elucidate DNA opening, *Nature* 533, 353–358. [PubMed: 27193681]
- [15]. Feklistov A, Bae B, Hauver J, Lass-Napiorkowska A, Kalesse M, Glaus F, Altmann K-H, Heyduk T, Landick R, and Darst SA (2017) RNA polymerase motions during promoter melting, *Science* 356, 863–866. [PubMed: 28546214]

- [16]. Pupov D, Kuzin I, Bass I, and Kulbachinskiy A (2014) Distinct functions of the RNA polymerase subunit region 3.2 in RNA priming and promoter escape, *Nucleic Acids Research* 42, 4494–4504. [PubMed: 24452800]
- [17]. pupov D, Petushkov I, Esyunina D, Murakami KS, and Kulbachinskiy A (2018) Region 3.2 of the sigma factor controls the stability of rRNA promoter complexes and potentiates their repression by DksA, *Nucleic Acids Research* 46, 11477–11487. [PubMed: 30321408]
- [18]. Rammohan J, Manzano AR, Garner AL, Stallings CL, and Galburt EA (2015) CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism, *Nucleic Acids Research* 43, 3272–3285. [PubMed: 25697505]
- [19]. Kubori T, and Shimamoto N (1996) A branched pathway in the early stages of transcription by *E. coli* RNA polymerase, *Journal of Molecular Biology* 256, 449–457. [PubMed: 8604130]
- [20]. Dangkulwanich M, Ishibashi T, Liu S, Kireeva ML, Lubkowska L, Kashlev M, and Bustamante CJ (2013) Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism, *eLIFE*
- [21]. Murakami KS, Masuda S, and Darst SA (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution, *Science* 296, 1280–1284. [PubMed: 12016306]
- [22]. Zuo Y, and Steitz TA (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Molecular Cell* 58, 534–540. [PubMed: 25866247]
- [23]. Kapanidis AN, Margeat E, Ho SO, Kortkhonja E, Weiss S, and Ebright RH (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism, *Science* 314, 1144–1147. [PubMed: 17110578]
- [24]. Revyakin A, Liu C, Ebright RH, and Strick TR (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching, *Science* 314, 1139–1143. [PubMed: 17110577]
- [25]. Yu L, Winkelman JT, Pukhrambam C, Strick TR, Nickels BE, and Ebright RH (2017) The mechanism of variability in transcription start site selection, *eLIFE* 6.
- [26]. Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, and Gourse RL (2006), rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase, *Cell* 125, 1069–1082. [PubMed: 16777598]
- [27]. Fenton MS, Lee SJ, and Gralla JD (2000) *E. coli* promoter opening and –10 recognition: mutational analysis of σ^{70} , *The EMBO Journal* 19, 1130–1137. [PubMed: 10698953]
- [28]. Martin CT, Esposito EA, Theis K, and Gong P (2005) Structure and function in promoter escape by T7 RNA polymerase, *Progress in Nucleic Acid Research and Molecular Biology* 80, 323–347. [PubMed: 16164978]
- [29]. Bai L, Fullbright RM, and Wang MD (2007) Mechanochemical kinetics of transcription elongation, *Physical Review Letters* 98, 06803;06801–06804.
- [30]. Applng FD, Lucius AL, and Schneider DA (2015) Transient-state kinetic analysis of the RNA polymerase I nucleotide incorporation mechanism, *Biophysical Journal* 109, 2382–2393. [PubMed: 26636949]
- [31]. Greive SJ, and von Hippel PH (2005) Thinking quantitatively about transcription regulation, *Nature Reviews* 6.
- [32]. Erie DA, Yager TD, and von Hippel PH (1992) The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective, *Annual Reviews in Biophysics and Biomolecular Structure* 21, 379–415.
- [33]. Guajardo R, and Sousa R (1997) A model for the mechanism of polymerase translocation, *Journal of Molecular Biology* 265, 8–19. [PubMed: 8995520]
- [34]. Svetlov V, and Artsimovitch I (2015) Purification of bacterial RNA polymerase: tools and protocols, *Methods in Molecular Biology* 1276, 13–29. [PubMed: 25665556]
- [35]. Borukhov S, and Goldfarb A (1993) Recombinant *E. coli* RNA polymerase: Purification of individually overexpressed subunits and in vitro assembly, *Protein Expression and Purification* 4, 503–511. [PubMed: 8286946]
- [36]. Roe J-H, Burgess RR, and Record MTJ (1984) Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λ PR promoter, *Journal of Molecular Biology* 176, 495–521. [PubMed: 6235375]

- [37]. Stackhouse TM, Telesnitsky AP, and Meares CF (1989) Release of the σ subunit from *E. coli* RNA polymerase transcription complexes is dependent on the promoter sequence, *Biochemistry* 28, 7781–7788. [PubMed: 2482069]
- [38]. Erie DA, Hajiseyedi O, Young MC, and von Hippel PH (1993) Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription, *Science* 262, 867–873. [PubMed: 8235608]
- [39]. McClure WR, Cech CL, and Johnston DE (1978) A steady state assay for the RNA polymerase initiation reaction, *Journal of Biological Chemistry* 253, 8941–8948. [PubMed: 363712]
- [40]. Basu RS, Warner BA, Molodtsov V, Pupov D, Eshunina D, Fernandez-Tornero C, Kulbachinsky A, and Murakami KS (2014) Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme, *Journal of Biological Chemistry* 289, 24549–24559. [PubMed: 24973216]
- [41]. Straney DC, and Crothers DM (1987) A stressed intermediate in the formation of stably initiated RNA chains at the *E. coli lac* UV5 promoter, *Journal of Molecular Biology* 193, 267–278. [PubMed: 2439694]
- [42]. Bloomfield VA, Crothers DM, and Tinoco IJ (2000) *Nucleic Acids: Structures, Properties, and Functions*, University Science Books, Sausalito CA.
- [43]. Cleland WW (1975) Partition analysis and concept of net rate constants as tools in enzyme kinetics, *Biochemistry* 14, 3220–3224. [PubMed: 1148201]
- [44]. Vvedenskaya IO, Vahedian-Movahed H, Bird JG, Knoblauch JG, Goldman SR, Zhang Y, Ebright RH, and Nickels BE (2014) Interactions between RNA polymerase and the core recognition element counteract pausing *Science* 344, 1285–1289. [PubMed: 24926020]
- [45]. Bandwar RP, Tang G-Q, and Patel SS (2006) Sequential release of promoter contacts during transcription initiation to elongation transition, *Journal of Molecular Biology* 360, 466–483. [PubMed: 16780876]
- [46]. Kulbachinsky A, and Mustaev A (2006) Region 3.2 of the σ subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation, *Journal of Biological Chemistry* 281, 18273–18276. [PubMed: 16690607]
- [47]. Samanta S, and Martin CT (2013) Insights into the mechanism of initial transcription in *E. coli* RNA polymerase, *Journal of Biological Chemistry* 288, 31993–32003. [PubMed: 24047893]
- [48]. Ellinger T, Behnke D, Bujard H, and Gralla JD (1994) Stalling of *E. coli* RNA polymerase in the +6 to +12 region in vivo is associated with tight binding to consensus promoter elements, *Journal of Molecular Biology* 239, 455–465. [PubMed: 8006961]
- [49]. Haugen SP, Ross W, Manrique M, and Gourse RL (2008) Fine structure of the promoter- σ region 1.2 interaction, *Proceedings of the National Academy of Sciences* 105, 3292–3297.
- [50]. Tang G-Q, Roy R, Bandwar RP, Ha T, and Patel SS (2009) Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase, *Proceedings of the National Academy of Sciences* 106, 22175–22180.
- [51]. Esposito EA, and Martin CT (2004) Cross-linking of promoter DNA to T7 RNA polymerase does not prevent formation of a stable elongation complex, *Journal of Biological Chemistry* 279, 44270–44276. [PubMed: 15304497]
- [52]. Durniak KJ, Bailey S, and Steitz TA (2008) The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation, *Science* 322, 553–557. [PubMed: 18948533]
- [53]. Theis K, Gong P, and Martin CT (2004) Topological and conformational analysis of the initiation and elongation complex of T7 RNA polymerase suggest a new twist, *Biochemistry* 43, 12709–12715. [PubMed: 15461442]
- [54]. Stano NM, Levin MK, and Patel SS (2002) The +2 NTP binding drives open complex formation in T7 RNA polymerase, *Journal of Biological Chemistry* 277, 37292–37300. [PubMed: 12151383]
- [55]. Galburt EA (2018) The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression, *Proceedings of the National Academy of Sciences* 115, E11604–E11613.

Figure 1.

Time Course of Transcription Initiation by *E. coli* RNA Polymerase at the λ PR Promoter. Panels A-C show representative gels of RNA products synthesized at 0.5 s to 90 or 150 s after mixing preformed RNAP- λ PR open complexes with NTPs and heparin (50 μ g/mL) at 19 °C. The transcript sequence is pppA⁺¹pUpGpUpApG⁺⁶pUpApApGpG⁺¹¹pApGpGpUpU⁺¹⁶pC... Nucleotide concentrations are listed above the figure panels. The escape point is at the 10-mer to 11-mer step^{5,37}. CTP is omitted to stop transcription of this modified λ PR sequence after making a 16-mer, but misincorporation allows slow readthrough to the next stop after a 31-mer. Labeling of RNAs is with ~17 nM α -³²P-GTP in Panels A and B and ~17 nM α -³²P-UTP in Panel C. Lengths of major RNA species are indicated at the right of each gel image.

Figure 2.

Amounts of Full-length (FL) and Selected (3-, 6-, 10-mer) Short RNAs Present as a Function of Time in Faster, Slower Phases of Initiation at Different NTP Concentrations. Columns A, B, and C summarize kinetic data for FL and 3-, 6-, and 10-mer short RNAs (from averages of 2–4 independent experiments like Fig. 1) at the NTP concentrations of Fig. 1 A, B and C respectively. In all cases, normalized amounts of RNA per initiating complex are plotted. In each column, Panel 1 plots kinetic data for the synthesis of FL RNA by productive complexes and shows the fit of these data to a single exponential approach to a plateau value after a brief lag (black line). Amounts of 3-mer (panels 2), 6-mer (panels 3), and 10-mer (panels 4) RNAs each increase to a maximum and then decay to a non-zero value on the timescale of FL RNA synthesis (the faster phase of initiation). Amounts of 3-mer (panels 2) subsequently increase linearly in the slower (abortive) phase of initiation,

while amounts of 6-mer and 10-mer remain constant, demonstrating the absence of abortive synthesis of 6- and 10-mer in this slower phase.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Examples of the Kinetics of Synthesis of the First Short RNA by Nonproductive Complexes in the Faster Kinetic Phase of Initiation.

Amounts of 7-mer (panel A) and 8-mer (panel B) RNA quantified from gels of experiments at 10 μ M UTP, 200 μ M ATP and GTP (Fig 1C) are plotted from 0 s to 90 s. After a short lag (1 s for 7-mer; 1.5 s for 8-mer), both approach plateau amounts (~ 0.014 RNA/OC) with similar first order rate constants (~ 0.015 s $^{-1}$) in this faster phase of initiation where FL RNA is synthesized (see Fig. 2C). No additional 7-mer or 8-mer is made in the slower (abortive) kinetic phase of initiation, consistent with observations for 6- and 10-mer in Fig. 2.

Figure 4.

Time Evolution of the Population of Short RNA Intermediates in FL RNA Synthesis.

Amounts of each detectable short RNA intermediate, normalized per open complex (OC) are plotted vs. time for initiation experiments at 19 °C for the three sets of NTP concentrations of Fig. 1. Points in Panel A are averages and ranges for two experiments; points in Panels B and C are averages $\pm$ 1 SD for four experiments. Smooth curves are predicted from rate constants (Fig. 5 below) obtained from the fit of these data to the mechanism of Scheme 1 as described in Results.

Figure 5.

Rate Constants for Steps of RNA-DNA Hybrid Growth in Initiation. Composite second order rate constants (k_i values, Eq. 1) describing each step of RNA-DNA hybrid extension (including reversible translocation, reversible NTP binding and catalysis; Scheme 2) up to the escape point are plotted vs the step of initiation (bottom axis) and the nucleobase incorporated (top axis).

Figure 6.

Free Energy and Structural Analysis of Disruption of σ^{70} -PR Promoter Interactions in Initiation and Promoter Escape by *E. coli* RNAP. A) Free energy profile of the stages of buildup and release of translocation stress ($\Delta G_{tr,i}^o$) as the RNA-DNA hybrid grows and disrupts subsets of RNAP-promoter interactions. $\Delta G_{tr,i}^o$ values are calculated from the corresponding rate-constant reductions (Eq. 7). High $\Delta G_{tr,i}^o$ values in the three regions (yellow, blue, green) of this plot are proposed to indicate disruption of the three regions of RNAP-promoter interactions, relieving translocation stress in the next step. B) Structural proposal that in-cleft interactions of λ PR promoter DNA with $\sigma_{1,2}$ (yellow) and other in-cleft interactions are broken as the 3-mer hybrid extends to a 5-mer. Interactions of $\sigma_{2,3}$ and σ_3 (blue) with the -10 region are proposed to break as the 6-mer extends to a 9-mer. Interactions of the -35 region with σ_4 (green) and other upstream interactions are proposed

to break in extension of the 10-mer to an 11-mer, resulting in promoter escape. Structural model was made using 4YLP.pdb.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 7.

Free Energy and Structural Analysis of Disruption of T7 RNAP-Consensus Promoter Interactions in Initiation and Promoter Escape. A) Free energy profile of the stages of buildup and release of translocation stress ($\Delta G_{tr,i}^o$) as the RNA-DNA hybrid grows, and disrupts subsets of RNAP-promoter interactions. Estimates of $\Delta G_{tr,i}^o$ values are obtained from the corresponding rate-constant reductions (Eq. 7), using the results of Tang et al⁴⁸. Increases in $\Delta G_{tr,i}^o$ values occur in three regions (yellow, blue, green). B) Proposed relationship between $\Delta G_{tr,i}^o$ values for steps of initiation by T7 RNAP and previous structural proposals. DNA bending and scrunching with unfavorable $\Delta G_{tr,i}^o$ (yellow in Fig 7A), are proposed to occur in forming the 5-mer RNA-DNA hybrid⁹. Contacts of the specificity loop and intercalating hairpin (blue) with the -7 to -11 region of the promoter

are disrupted with concomitant bubble collapse during extension of the 5-mer to the 9-mer⁵⁴. Interactions between the A-T-recognition region (green) and A-T rich upstream promoter element are disrupted in extension of the 9-mer to the 13-mer^{9, 28}. Structural model was made using 1QLN.pdb.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Scheme 1.

Mechanism of Transcription Initiation by *E. coli* RNAP at λ PR Promoter. The minimal 11-step mechanism needed to fit initiation kinetic data as a function of NTP concentrations is shown. Initiation begins with reversible binding of ATP (step 1) to the open complex (OC) to form the first initiation complex (IC1); ATP binding is quantified by a dissociation equilibrium constant $K_{d,1} = k_{NTP,1}^-/k_{NTP,1}^+$. In the second step, UTP binds reversibly (step 2a, for which $K_{d,2} = k_{NTP,2}^-/k_{NTP,2}^+$), followed by irreversible catalysis to form IC2 (pppApU; step 2b, with rate constant $k_{cat,2}$). Subsequent steps of initiation all start with reversible translocation to create the NTP binding site, followed by reversible NTP binding and catalysis. For the NTP concentrations investigated these sub-steps cannot be separated in the fit, yielding composite 2nd order rate constants k_i for each step of hybrid extension up to the escape point and synthesis of full-length RNA (k_{11}).

Scheme 2.

Mechanism of Steps of Initiation that Begin with Translocation ($3 \leq i \leq 10$) Used to Interpret Rate Constants in Scheme 1. In Scheme 2, IC_{i-1}^{pre-tr} and $IC_{i-1}^{post-tr}$ are pre-translocated and post-translocated conformations of the initiation complex (IC) with a (i-1)-mer RNA, $k_{i-1}^{tr,+}$ and $k_{i-1}^{tr,-}$ are the forward and back rate constants and the equilibrium constant for the translocation step (equilibrium constant $K_{i-1}^{tr} = k_{i-1}^{tr,+}/k_{i-1}^{tr,-}$), $k_{NTP,i}^{+}$ and $k_{NTP,i}^{-}$ are rate constants for binding and dissociation of NTP_i , and $k_{cat,i}^{NTP}$ is the catalytic rate constant. The catalytic step forming IC_i^{pre-tr} from $NTP_i IC_{i-1}^{post-tr}$ is irreversible because the PP_i concentration is negligibly small.