Biophysical Journal Volume 109 December 2015 2382–2393

Article

Transient-State Kinetic Analysis of the RNA Polymerase I Nucleotide Incorporation Mechanism

Francis D. Appling,[1] Aaron L. Lucius,[2,*] and David A. Schneider[1,*]
[1]Department of Biochemistry and Molecular Genetics and [2]Department of Chemistry, University of Alabama at Birmingham, Birmingham, Alabama

ABSTRACT Eukaryotes express three or more multisubunit nuclear RNA polymerases (Pols) referred to as Pols I, II, and III, each of which synthesizes a specific subset of RNAs. Consistent with the diversity of their target genes, eukaryotic cells have evolved divergent cohorts of transcription factors and enzymatic properties for each RNA polymerase system. Over the years, many *trans*-acting factors that orchestrate transcription by the individual Pols have been described; however, little effort has been devoted to characterizing the molecular mechanisms of Pol I activity. To begin to address this gap in our understanding of eukaryotic gene expression, here we establish transient-state kinetic approaches to characterize the nucleotide incorporation mechanism of Pol I. We collected time courses for single turnover nucleotide incorporation reactions over a range of substrate ATP concentrations that provide information on both Pol I's nucleotide addition and nuclease activities. The data were analyzed by model-independent and model-dependent approaches, resulting in, to our knowledge, the first minimal model for the nucleotide addition pathway for Pol I. Using a grid searching approach we provide rigorous bounds on estimated values of the individual elementary rate constants within the proposed model. This work reports the most detailed analysis of Pol I mechanism to date. Furthermore, in addition to their use in transient state kinetic analyses, the computational approaches applied here are broadly applicable to global optimization problems.

INTRODUCTION

Eukaryotic cells express at least three multisubunit nuclear RNA polymerases referred to as Pols I, II, and III. These enzymes are composed of 14, 12, and 17 subunits, respectively. Nuclear RNA synthesis is divided between the Pols with Pol I being responsible for production of the majority of the ribosomal RNA, Pol II being responsible for the production of messenger RNA and many small RNAs, and Pol III being responsible for production of transfer RNAs and the smallest ribosomal RNA.

RNA expression is the first critical step in gene expression, thus understanding the enzymatic features and regulatory pathways that control polymerase function is critical for establishing a thorough understanding of cell biology. Although the three nuclear RNA polymerases share considerable homology, we and others have recently demonstrated critical structural and/or functional differences between Pols I and II (1,2). Thus, to develop a detailed understanding of eukaryotic gene expression, genetic, biochemical, and biophysical studies focused on each of the three Pols must be conducted. To date, mechanistic studies of multisubunit RNA polymerase transcription have focused on prokaryotic RNA polymerases or eukaryotic RNA polymerase II. This study uses transient-state kinetic approaches to probe the mechanism of transcription by Pol I.

Although Pol I transcribes a single target gene, it is responsible for the majority of RNA synthesis in the nucleus (3). Ribosome biosynthesis, and by association Pol I transcription, are tightly linked to cell proliferation and this link has recently gained attention from the chemotherapeutic field (4). Indeed, a small number of inhibitors that potentially target Pol I activity have been identified (5). However, the molecular mechanisms of inhibition by any of these compounds are unknown.

Due to its central role in cell growth and its potential as a chemotherapeutic target, there is a need to understand mechanistic details of transcription by Pol I. The aim of this study was to determine a minimal kinetic model that describes the nucleotide addition cycle of Pol I. Such a model provides quantitative estimates of elementary rate constants governing the nucleotide addition cycle, ultimately yielding insight into the evolutionary forces that have given rise to the partitioning of eukaryotic transcriptional responsibilities among three or more polymerases.

To develop this kinetic model, we used a transient state kinetic approach to monitor single nucleotide addition and nuclease reactions catalyzed by Pol I. We developed and applied, to our knowledge, a novel optimization method that enabled us to fit our data to a single model describing the Pol I nucleotide addition and nuclease cycles. Confidence intervals for each estimated elementary rate constant were calculated using a rigorous grid search-based approach. To our knowledge, this is the first study of

Submitted August 13, 2015, and accepted for publication October 28, 2015.
*Correspondence: allucius@uab.edu or dschneid@uab.edu
Editor: Timothy Lohman.

0006-3495/15/12/2382/12
http://dx.doi.org/10.1016/j.bpj.2015.10.037

its kind for Pol I and represents the first set of quantitative mechanistic data concerning this important enzyme. In addition, the computational strategies developed in this work can be broadly applied to diverse transient-state kinetic data sets, and more generally, to global optimization problems in which many parameters are being simultaneously determined.

MATERIALS AND METHODS

Buffers

Buffers were prepared using water that was twice deionized and filtered using a Millipore ultrafiltration system (Billerica, MA) ($R \geq 18$ MΩ) and reagent grade chemicals. A list of all buffers is available in the Supporting Materials and Methods.

Proteins

Pol I was purified according to Appling and Schneider (6). Pol I is stored in a storage buffer at −20°C.

Heparin and bovine serum albumin preparation

Heparin and BSA were dialyzed extensively before use. A detailed description of heparin and BSA preparation is available in the Supporting Materials and Methods.

Nucleic acids

All nucleic acids were commercially synthesized by Integrated DNA Technologies (Coralville, IA). All nucleic acids were purified and dialyzed before use. A detailed description of nucleic acid preparation is available in the Supporting Materials and Methods.

All data presented here were collected using elongation complexes (ECs) assembled with the following nucleic acid sequences:

RNA(5′-AUCGAGAGG)
DNA template strand (DNA-t) (5′ ACCAGCAGGCCGATTGGGATGG GTATTCCCTCCTGCCTCTCGATGGCTGTAAGTATCCTATAGG)
DNA nontemplate strand (DNA-nt) (5′ CCTATAGGATACTTACAGC CATCGAGAGGCAGGAGGGAATACCCATCCCAATCGGCCTGC TGGT).

EC assembly

ECs were assembled in a stepwise manner from purified protein and nucleic acid components. The EC assembly protocol can be found in the Supporting Materials and Methods.

ATP substrates

ATP substrates were dialyzed before use. Preparation of ATP substrates is described in the Supporting Materials and Methods.

Quenched flow time courses

Nucleotide incorporation time courses were collected using a chemical quenched flow instrument. The chemical quenched flow conditions and protocol are described in the main text and in the Supporting Material.

Nucleotide addition reaction electrophoresis

Reactant and product RNAs were separated by high resolution denaturing polyacrylamide gel electrophoresis (PAGE). A description of PAGE conditions is available in the Supporting Materials and Methods.

Data analysis

Data analysis was performed using MATLAB (The MathWorks, Natick, MA) and KaleidaGraph (Synergy Software, Reading, PA). Matrix manipulations were performed in MATLAB. Optimization of the parameters shown in Schemes 1 and 2 was accomplished using a custom-built genetic algorithm. A detailed description of all data analyses is available in the Supporting Materials and Methods.

RESULTS

Single turnover nucleotide addition reactions

The transcription cycle can be divided into three phases: initiation, elongation, and termination (Fig. 1 *A*). The elongation phase is composed of a repetition of the nucleotide addition cycle. At a minimum the nucleotide addition cycle must consist of nucleotide binding, phosphodiester bond formation, pyrophosphate release, and translocation as schematized in Fig. 1 *A*. The goal of this work is to establish a minimal kinetic mechanism describing a single round of the Pol I nucleotide addition cycle.

The functional form of Pol I during the elongation phase of transcription is referred to as the EC. The minimal components of an EC are polymerase, template and nontemplate DNA, and RNA. In this work we have reconstituted Pol I ECs using the scaffold template approach, which involves directly building ECs from purified protein and nucleic acid components (7,8). To label RNA for visualization, once the EC was assembled, we used the polymerase to incorporate a single ^{32}P-labeled CMP (cytosine monophosphate) into the 3′-end of a 9 nucleotide RNA as depicted in Fig. 1 *B*. This was accomplished by mixing the assembled EC with Mg^{2+} and α-^{32}P-CTP (cytosine triphosphate) (see Fig. 1 *B* and the Supporting Materials and Methods). After 10 min, reactions were stopped by mixing ECs with EDTA and an excess (over radiolabeled CTP) of unlabeled CTP (Fig. 1 *B*). Chelation of solution divalent cation by EDTA leaves Pol I reversibly inactivated because divalent cation is required for both Pol I nucleotide addition and nuclease activities (9,10).

A chemical quenched flow instrument was used, as schematized, in Fig. 1 *C* to acquire time courses for single nucleotide extension from the radiolabeled RNA. Reactions were initiated by rapidly mixing the radiolabeled Pol I ECs illustrated in Fig. 1 *B* with a solution containing ATP (adenosine triphosphate), Mg^{2+}, and heparin. After rapid mixing of the two reagents and a delay time, the reaction was stopped by mixing with 1 M HCl. Nucleotide addition time courses were reconstructed by varying the quenched flow delay time between 0.005 and 10 s (see the Supporting Materials and Methods).

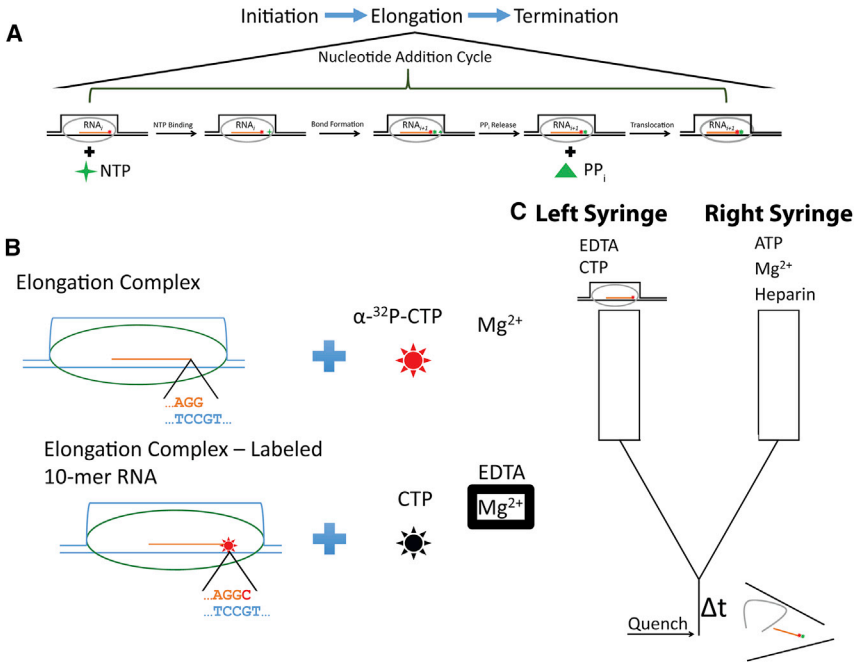


FIGURE 1 Schematic representations of the Pol I nucleotide addition cycle, EC-catalyzed RNA labeling and the quenched flow setup used to collect nucleotide addition time courses. (*A*) The transcription cycle is composed of three distinct phases corresponding to initiation, elongation, and termination. The elongation phase is composed of a repetition of the nucleotide addition cycle. The expanded schematic represents the minimal steps that must occur during the nucleotide addition cycle. These steps include substrate NTP binding, phosphodiester bond formation, pyrophosphate release, and translocation. (*Stars*) NMPs; (*diamond*) NTPs; (*triangle*) pyrophosphate. (*B*) ECs are formed using the scaffold template approach. DNA template and nontemplate strands and hybridized RNA bound by oval Pol I are illustrated. The expansion underneath the schematized EC displays the sequence of the 3′-end of the RNA and its cognate DNA template sequence. To visualize RNA, the polymerase is forced to incorporate a single labeled CMP into the 3′-end of the RNA. This labeling reaction is accomplished by adding labeled CTP and Mg^{2+}. (*Stars*) Labeled CTP and CMP; (*black star*) unlabeled CTP. Mg^{2+} inside a box below "EDTA" denotes that solution Mg^{2+} has been chelated by EDTA. (*C*) Chemical quenched flow setup used to collect nucleotide incorporation time courses. To initiate each reaction the contents of the left and right syringes are rapidly mixed. (Δ*t*) Line indicates that the reaction is allowed to proceed for a specified amount of time. After this specified reaction time the reaction mixture is rapidly mixed with 1 M HCl (as indicated by the *arrow* labeled *quench*). (*Disordered lines* to the *right* of the *quench arrow*) Quenching reaction inactivates and denatures the EC. The left syringe contains the EC mixture depicted at the bottom of (*B*). The right syringe contains ATP, Mg^{2+}, and heparin. To see this figure in color, go online.

Reactions collected in the quenched flow were subjected to high resolution denaturing PAGE to separate reactants and products. ^{32}P-labeled RNA was visualized by phosphorimaging; a representative gel is shown in Fig. 2 *A*. Five distinguishable ^{32}P-labeled species corresponding to 10- and 11-mer RNA, CA, and GC dinucleotide RNA, and unincorporated cytosine nucleotide were observed on imaged gels (Fig. 2 *A*—unincorporated nucleotide was cut off). The identities of these individual species were confirmed by comparison to commercial standards and control reactions (see Fig. S1). Upon quantification we found that [GC] does not exhibit time dependence during nucleotide incorporation time courses (see Fig. S2). Visualization of these RNA species enabled analysis of nucleotide addition and nuclease activities of Pol I.

Model-independent analysis of nucleotide addition and nuclease reactions

Fig. 2, *B* and *C*, displays 11-mer and CA time courses, respectively. The time courses were reconstructed from quantification of the 11-mer and CA bands on the gel displayed in Fig. 2 *A* using Eq. S1 in the Supporting Materials and Methods. The nucleotide addition reaction analyzed in Fig. 2 was collected at 10 μM ATP.

Fig. 2 *B* shows that the fraction of RNA in the 11-mer state increases between ~0 and 200 ms before decaying slowly for the remainder of the time course (note log timescale). In contrast, Fig. 2 *C* shows that the fraction of CA RNA rises in a single phase throughout the time course. The quantifications in Fig. 2, *B* and *C*, indicate that the fall of [11-mer] and the rise of [CA] occur over approximately the same timescale. The time courses in Fig. 2, *B* and *C*, suggest that 11-mer is directly converted to labeled CA and unlabeled 9-mer.

To quantify the time courses, we applied nonlinear least-squares (NLLS) analysis. The time courses appear to be exponential, thus the data were fit to a sum of exponential terms (see Eq. S2). The experiments were performed three times and each time course was subjected to NLLS analysis. The three determinations of the parameters obtained from this approach were averaged and the standard deviation was determined.

The CA time course was adequately described by a single exponential function while the 11-mer data required two exponentials (*solid red lines* in Fig. 2, *B* and *C*, display representative fits). The 11-mer time course yielded a fast and a slow observed rate constant with values of $(15 \pm 1)\ s^{-1}$ and $(0.29 \pm 0.07)\ s^{-1}$, respectively, whereas CA accumulated with a rate constant of $(0.35 \pm 0.05)\ s^{-1}$.

If CA production is the result of cleavage of the 11-mer then the rate constant for formation of CA would be expected to be the same as the rate constant for disappearance of the 11-mer. Consistently, the rate constant for the disappearance of the 11-mer, $(0.29 \pm 0.07)\ s^{-1}$, is within error of the observed rate constant for CA formation $(0.35 \pm 0.05)\ s^{-1}$.

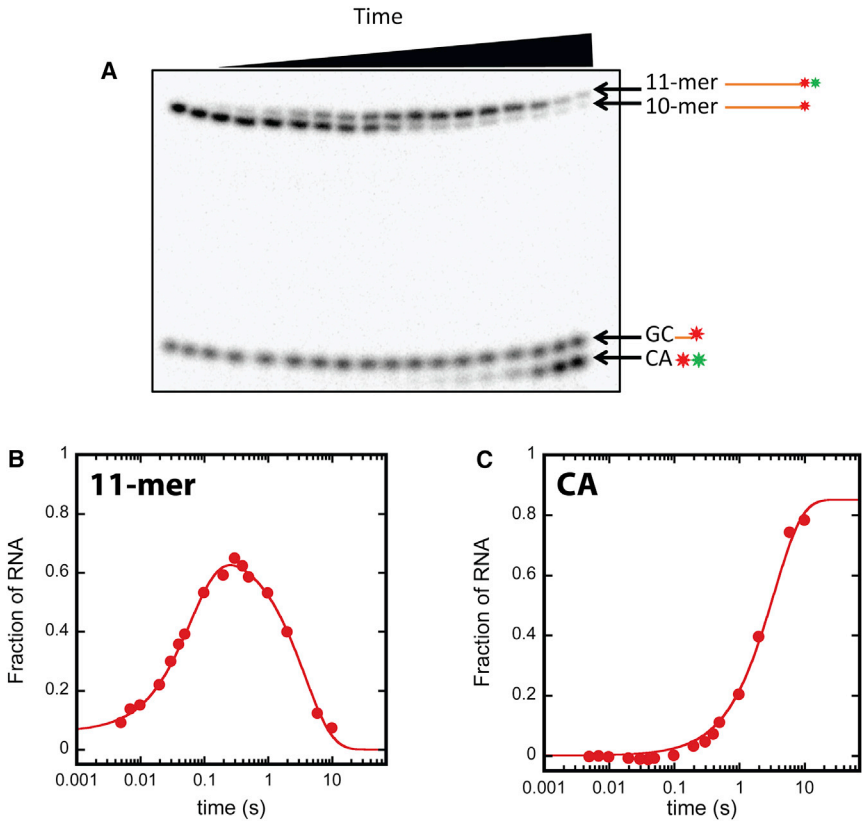


FIGURE 2 AMP incorporation time course collected at 10 μM substrate ATP. Representative gel and quantifications of the gel. (*A*) Denaturing PAGE separation of reactants and products from a nucleotide addition time course. (*Leftmost two lanes*) $t = 0$ points. Quenched flow delay time increases from left to right. The species in the gel are labeled to the right of the gel. (*B*) Quantified 11-mer time course. (*Circles*) Fraction of RNA in the 11-mer state as a function of time quantified according to Eq. S1. (*Solid traces*) NLLS fit of the data to Eq. S2 ($j = 2$). (*C*) Quantified CA time course. (*Circles*) Fraction of RNA in the CA state as a function of time quantified according to Eq. S1 . (*Solid traces*) NLLS fit of the data to Eq. S2 ($j = 1$). To see this figure in color, go online.

This observation indicates that these two rate constants directly report on Pol I's nuclease activity.

To obtain additional information about the nucleotide addition mechanism and further probe the linkage between 11-mer decay and CA accumulation, the system must be perturbed and the kinetic response(s) measured. Taking into account the minimal steps that must occur during the nucleotide addition cycle, Fig. 1 *A*, a simple way to perturb the system is to vary the [ATP]. At sufficiently low ATP concentrations the ATP binding step will become rate limiting. In contrast, at sufficiently high ATP concentrations, the time courses will be independent of ATP concentration. Thus, the behavior of the 11-mer and CA time courses as a function of [ATP] will contain information about the nucleotide addition mechanism.

Fig. 3 *A* shows 11-mer time courses collected between 10 μM and 1 mM ATP. As described above and observed in Fig. 2 *B*, all of the 11-mer time courses are biphasic. As seen in Fig. 3 *A*, the two phases clearly respond differently to changing [ATP]. The rising phases of the 11-mer time courses (first $\approx$100 ms) appear to accelerate as [ATP] is increased (traces shift *left* in Fig. 3 *A*). In sharp contrast, the decay phases of the 11-mer time courses virtually overlay indicating no dependence on [ATP].

Fig. 3 *B* shows the corresponding CA time courses collected between 10 μM and 1 mM ATP. As described above and observed in Fig. 2 *C*, [CA] rises continuously throughout the time course. In contrast to the time courses for 11-mer, the time courses for CA are invariant across the experimental [ATP]. This observation indicates that this reaction is independent of [ATP].

To quantify the effects of [ATP] on the 11-mer and CA time courses, each time course was subjected to NLLS-weighted nonlinear least-squares (WNLLS) analysis using a sum of

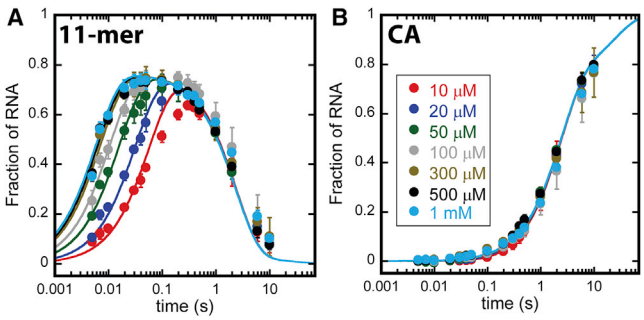


FIGURE 3 Nucleotide addition time courses collected as a function of [ATP] (*A*) 11-mer time courses collected as a function of [ATP]. (*Circles*) Average value of three independent measurements quantified according to Eq. S2. Uncertainty bars represent standard deviation about the average. (*Solid traces*) Global WNLLS fit of the data to Scheme 2. Colors denoting [ATP] are the same as those displayed in the legend shown in (*B*). (*B*) CA time courses collected as a function of [ATP]. (*Circles*) Average value of three independent measurements quantified according to Eq. S2. (*Uncertainty bars*) Standard deviation about the average. (*Solid traces*) Global WNLLS fit of the data to Scheme 2. (*Inset legend*) Experimental substrate [ATP].

exponential terms as described above for the 10 μM ATP data (Fig. 2, *B* and *C*) (see Eq. S2). WNLLS analysis of the 11-mer time courses yielded two observed rate constants which are plotted as a function of [ATP] in Fig. 4, *A* and *B* (*circles*). It is clear from Fig. 4, *A* and *B*, that the two 11-mer observed rate constants (11-mer k_{obs1} and 11-mer k_{obs2}, respectively) differ in both magnitude and response to [ATP]. Specifically, 11-mer k_{obs1} (Fig. 4 *A*) exhibits a hyperbolic dependence on [ATP] whereas the 11-mer k_{obs2} (Fig. 4 *B*) appears to be independent of [ATP].

The 11-mer k_{obs1} versus [ATP] data were subjected to weighted nonlinear least-squares (WNLLS) analysis using the equation for a rectangular hyperbola given by Eq. S3. The analysis yielded $k_{max} = (270 \pm 30)\ s^{-1}$, and $K_{1/2} = (170 \pm 30)\ \mu M$. The 11-mer k_{obs2} versus [ATP] data in Fig. 4 *B* were also subjected to WNLLS analysis to a constant (see Eq. S4). From this analysis a constant corresponding to 11-mer $k_{obs2} = (0.27 \pm 0.02)\ s^{-1}$ was determined.

CA time courses collected at each [ATP] were similarly subjected to WNLLS analysis. As observed for the 10 μM ATP data set (Fig. 2 *C*), CA time courses at each [ATP] were adequately described by a single exponential function (see Eq. S2). In Fig. 4 *C*, CA observed rate constants (CA k_{obs}) are plotted against [ATP]. As described above for 11-mer k_{obs2}, CA k_{obs} versus [ATP] data were fit to a constant (see Eq. S4) to yield a value of $(0.40 \pm 0.02)\ s^{-1}$.

The analysis described above leads to three conclusions: 1) The hyperbolic dependence of 11-mer k_{obs1} on [ATP] indicates that ATP binding is best described as a rapid equilibrium process with respect to the step following ATP binding. 2) The lack of a lag in the 11-mer time courses at any [ATP] along with the hyperbolic dependence of 11-mer k_{obs1} on [ATP] indicate that the step reported on by 11-mer k_{obs1} is directly linked to ATP binding (11). 3) The correspondence between 11-mer k_{obs2} and CA k_{obs} values and [ATP] independence indicates that these two rate constants are reporting on the same process, most likely the nuclease activity of Pol I.

Model-dependent analysis of single turnover nucleotide addition and nuclease reactions

Scheme 1 of Fig. 5 shows the minimal steps that must be present in the nucleotide addition reaction. In this scheme EC_n refers to EC with RNA of length n, PP_i refers to pyrophosphate, and *CA* refers to the dinucleotide fragment liberated from the 3′-end of the RNA. The system of coupled ordinary differential equations (ODEs) defined by Scheme 1 was numerically integrated to globally fit 11-mer and CA time courses collected at each [ATP] by WNLLS analysis using a custom built genetic algorithm (see the Supporting Materials and Methods). Fig. 6, *A*

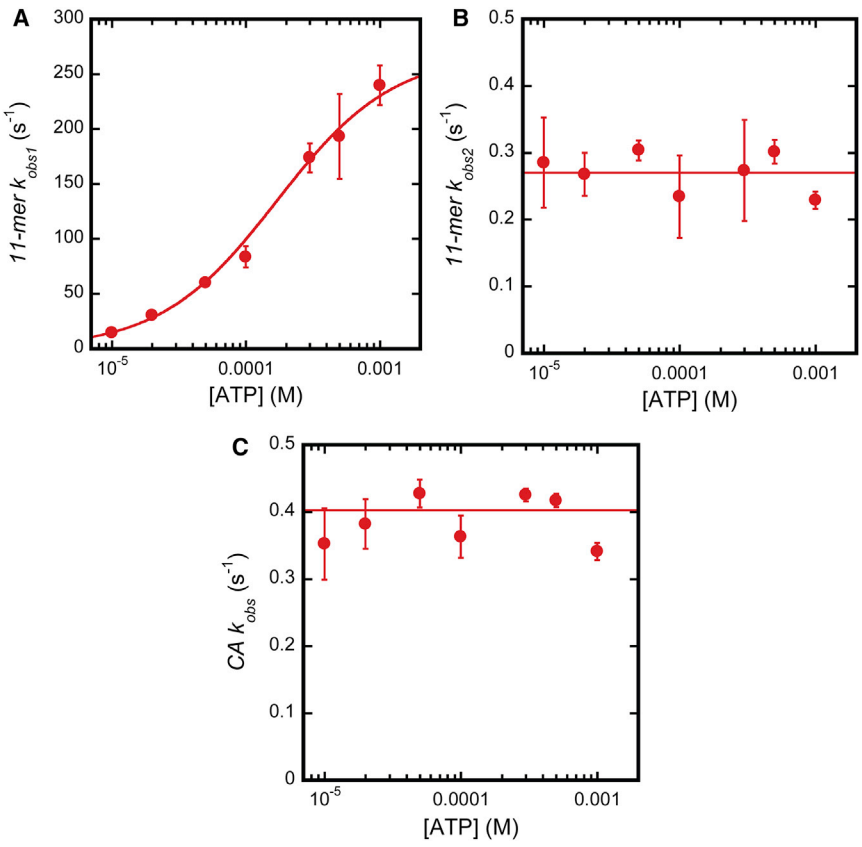


FIGURE 4 k_{obs} versus [ATP]. (*A*) 11-mer k_{obs1} plotted as a function of [ATP]. (*Circles*) Average k_{obs} value determined from analysis of three independent time courses according to Eq. S2 ($j = 2$). (*Uncertainty bars*) Standard deviation about the average. (*Solid traces*) WNLLS fit of the data to Eq. S3. (*B*) 11-mer k_{obs2} plotted as a function of [ATP]. (*Circles*) Average k_{obs} value determined from analysis of three independent time courses according to Eq. S1 ($j = 2$). Uncertainty bars represent standard deviation about the average. (*Solid traces*) WNLLS fit of the data to Eq. S4. (*C*) CA k_{obs} plotted as a function of [ATP]. (*Circles*) Average k_{obs} value determined from analysis of three independent time courses according to Eq. S2 ($j = 1$). Uncertainty bars represent standard deviation about the average. (*Solid traces*) WNLLS fit of the data to Eq. S4. To see this figure in color, go online.

$$EC_{10} + ATP \underset{k_2}{\overset{1{\cdot}10^8\,M^{-1}s^{-1}}{\rightleftarrows}} EC_{10} \bullet ATP \underset{k_4}{\overset{k_3}{\rightleftarrows}} EC_{11} \bullet PP_i \xrightarrow{k_5} EC_{11} \xrightarrow{k_6} EC_9 + CA$$

$$EC^*_{10} \xrightarrow{k_7} EC_{10} + ATP \underset{k_2}{\overset{1{\cdot}10^8\,M^{-1}s^{-1}}{\rightleftarrows}} EC_{10} \bullet ATP \underset{k_4}{\overset{k_3}{\rightleftarrows}} EC_{11} \bullet PP_i \xrightarrow{k_5} EC_{11} \xrightarrow{k_6} EC_9 + CA$$

FIGURE 5 Reaction schemes. (*Scheme 1*, *top*) EC_n, EC with RNA of length n; PP_i, pyrophosphate; *CA*, dinucleotide fragment removed from 3′-terminus of RNA. (*Scheme 2*, *bottom*) All species are the same as in Scheme 1, with the exception of EC^*_{10}, which corresponds to a subset of the EC_{10} population that is reversibly inactivated.

and *B*, displays the 11-mer and CA time courses, respectively, collected at each [ATP] (*circles*) overlaid with simulations produced from the best fit parameter values using Scheme 1 (*solid lines*).

From inspection of Fig. 6, *A* and *B*, it is clear that there are systematic deviations between the fit (*solid lines*) and data (*circles*). The largest deviation occurs in the 11-mer time courses between ~0.1 and 1 s (see Fig. 6 *A*). The best-fit lines according to Scheme 1 predict 11-mer peak heights that are consistently greater than those observed in the experimental data (see Fig. 6 *A* residuals between 0.1 and 1 s). Specifically, Scheme 1 predicts 11-mer peak heights ≈ 1 whereas the experimental data do not exhibit peak heights above ≈ 0.75 (see *light-blue solid line* and *circles* for comparison of fit and data at 1 mM ATP). The inability of Scheme 1 to describe the data indicates that a step or steps in addition to the minimal number of steps included in Scheme 1 must be present in the Pol I nucleotide addition mechanism. Furthermore, this step(s) must regulate the peak height of the 11-mer.

In the nucleotide addition time courses that we have collected, the 11-mer RNA species can be considered an intermediate and CA can be considered the reaction product. The relative rates of influx and efflux through an intermediate in a reaction pathway regulates the peak height of the intermediate. As an example a generic three state system of the form given by

$$A \overset{k_1}{\rightarrow} B \overset{k_2}{\rightarrow} C \tag{1}$$

can be used to explore how the relative magnitudes of flux through an intermediate can regulate the peak height of

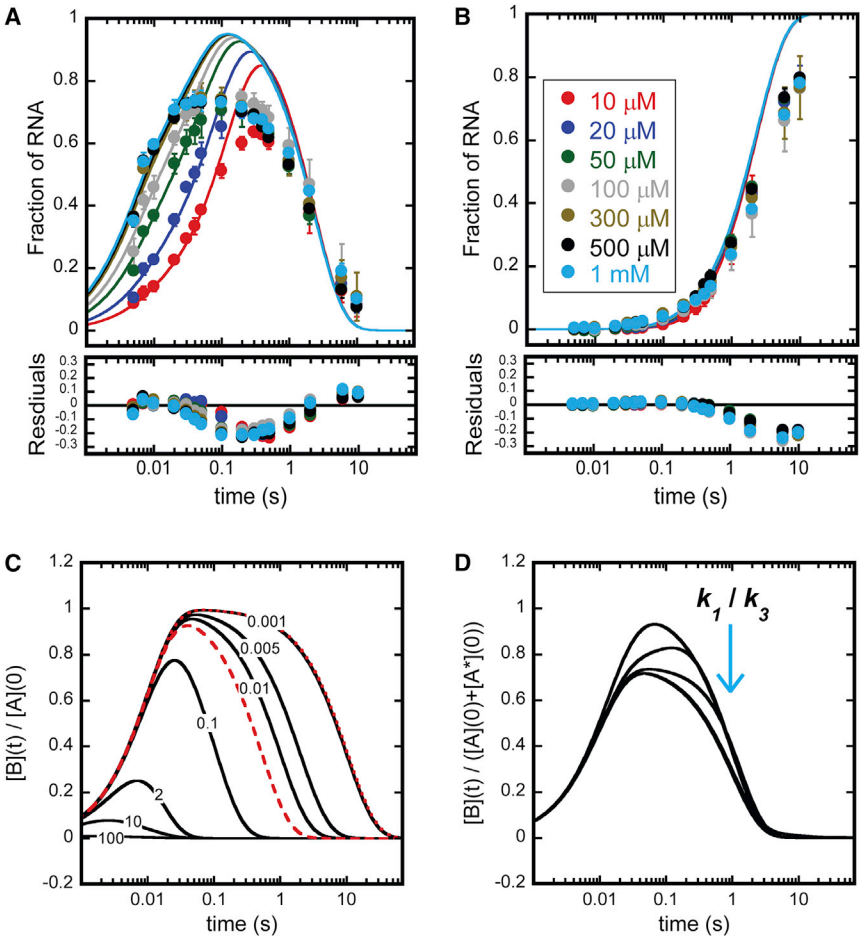


FIGURE 6 Global fit of Scheme 1. Simulations of Eqs. 2 and 4. (*A*) Global WNLLS fit of 11-mer time course to Scheme 1. (*Circles*) Experimental data as described in Fig. 3 *A*. (*Solid lines*) Scheme 1 simulated at best-fit parameter values. $k_1 = 1 \times 10^8\ s^{-1}$, $k_2 = 5534\ s^{-1}$, $k_3 = 142.8\ s^{-1}$, $k_4 = 84.6\ s^{-1}$, $k_5 = 52.3\ s^{-1}$, and $k_6 = 0.43\ s^{-1}$. (*B*) Global WNLLS fit of CA time courses to Scheme 1. (*Circles*) Experimental data as described in Fig. 3 *B*. (*Solid lines*) Scheme 1 simulated at best-fit parameter values listed in the description in (*A*). (*C*) Evaluation of Eq. 2 at a range of k_2/k_1. Each trace is labeled with its respective k_2/k_1 ratio. For each evaluation, $k_1 = 100\ s^{-1}$. To generate solid lines, k_2 was set to the values {10,000, 1000, 200, 10, 1, 0.5, 0.1} s^{-1}, to result in the k_2/k_1 ratio {100, 10, 2, 0.1, 0.01, 0.005, 0.001}, respectively. To generate broken and dashed red lines, k_2 was set to the values {0.0955, 1.92} s^{-1} to result in the k_1/k_2 ratios {0.00096, 0.019}, respectively. (*D*) Evaluation of Eq. 4 at a range of k_1/k_3 ratios. For each evaluation, $k_2 = 100\ s^{-1}$, $k_3 = 1\ s^{-1}$, $A(0) = 0.75$, $A^*(0) = 0.25$. To generate solid lines, k_1 was set to the values {50, 10, 2, 0.1, 0.01, 0.005, 0.001} s^{-1}, to result in the k_1/k_3 ratios {50, 10, 2, 0.1, 0.01, 0.005, 0.001}, respectively. Note that below a k_1/k_3 ratio of 0.1, all curves overlay.

the intermediate. According to Eq. 1, $[B](t)$ normalized to the initial concentration of reactant ($[A](0)$) is given by

$$\frac{[B](t)}{[A](0)} = \left(\frac{1}{1-\frac{k_2}{k_1}}\right)\left(e^{-k_2 t} - e^{-k_1 t}\right). \tag{2}$$

Inspection of Eq. 2 indicates that $[B](t)$ is scaled by a factor that is a function of the ratio of rate constants governing the formation and decay of B. In addition to this scaling factor, two time–dependent exponential terms of opposite sign modulate both the height of the $[B]$ peak and its position with respect to the time axis.

Due to the time dependence of Eq. 2, simple inferences involving the k_2/k_1 ratio regarding the peak height of B are difficult to draw. However, Eq. 2 can be evaluated at a range of k_2/k_1 ratios and the behavior of $[B]$ peak heights monitored. Fig. 6 *C* displays time courses simulated from Eq. 2 at seven k_2/k_1 ratios (100, 10, 2, 0.1, 0.01, 0.005, and 0.001, *solid black lines*) and plotted over the experimental time range. In addition, Fig. 6 *C* displays Eq. 2 evaluated at the k_2/k_1 ratios obtained from 11-mer time courses collected at 1 mM and 10 μM ATP (11-mer k_{obs1}/11-mer k_{obs2} ratios) plotted as dotted and broken red lines, respectively. From these simulations, it is clear that there is no combination of experimental 11-mer influx and efflux rates (11-mer k_{obs1} and 11-mer k_{obs2} values) that result in an 11-mer peak height below ≈ 1. In fact, an analysis of Fig. 6 *C* indicates that a k_2/k_1 ratio ≈ 0.1 is required to result in a peak height ≈ 0.8. According to the 11-mer k_{obs2} values plotted in Fig. 4 *B*, this ratio would correspond to an 11-mer k_{obs1} value of $\approx$2–3 s^{-1}. This value is much smaller than any experimentally observed 11-mer k_{obs1} value (Fig. 4 *A*). In terms of the nucleotide addition mechanism these analyses indicate that there is not a step that can follow ATP binding that could modulate the 11-mer peak heights while preserving the 11-mer influx and efflux rates (see Fig. 4, *A* and *B*, for 11-mer flux rates reported as 11-mer k_{obs1} and 11-mer k_{obs2}).

The simplest model that preserves the 11-mer k_{obs1} and 11-mer k_{obs2} values while reducing the 11-mer peak heights is one that includes an EC subpopulation that undergoes an inactive to active transition. If this inactive to active transition were governed by a rate constant less than that governing CA production, the 11-mer peak heights would respond to the fraction of the EC population in the active form when the reaction begins. A model of this form can be approximated by the generic four state system given by

$$A^* \xrightarrow{k_1} A \xrightarrow{k_2} B \xrightarrow{k_3} C \tag{3}$$

where $A*$ denotes the subpopulation that must undergo an inactive to active transition. B[B](t) normalized to the initial total $[A]$ ($[A^*](0)+[A](0)$) can be obtained and takes the form of Eq. 4, where $A(0)$ and $A^*(0)$ denote the initial concentrations of active and inactive species, respectively:

$$\begin{aligned}\frac{B(t)}{A^*(0)+A(0)} = k_2\Bigg(&\frac{A^*(0)e^{-k_1 t}k_1}{(k_1-k_2)(k_1-k_3)} \\ &+ \frac{e^{-k_2 t}((A(0)+A^*(0))k_1 - A(0)k_2)}{(-k_1+k_2)(k_2-k_3)} \\ &+ \frac{e^{-k_3 t}((A(0)+A^*(0))k_1 - A(0)k_3)}{(-k_1+k_3)(-k_2+k_3)}\Bigg).\end{aligned} \tag{4}$$

Equation 4 demonstrates that the interplay of rate constants that governs an intermediate's peak height quickly grows in complexity as a function of the number of steps in a reaction pathway. In the limit that k_1 approaches zero, Eq. 4 becomes

$$\lim_{k_1 \to 0} \frac{B(t)}{A^*(0)+A(0)} = \left(\frac{A(0)}{A(0)+A^*(0)}\right)\frac{(-e^{-k_2 t}+e^{-k_3 t})k_2}{(k_2-k_3)}. \tag{5}$$

Although this is the trivial case in which Eq. 3 collapses to Eq. 1 with an inactive reactant population, this limit demonstrates that as k_1 becomes very small the entire function is scaled by the fraction of reactant A in the active form. In fact, inspection and rearrangement of Eq. 5 yields Eq. 6, where $frac_{active}$ is given by Eq. 7, below:

$$\lim_{k_1 \to 0} \frac{B(t)}{A^*(0)+A(0)} = frac_{active}\left(\frac{1}{1-\frac{k_3}{k_2}}\right)\left(-e^{-k_2 t}+e^{-k_3 t}\right), \tag{6}$$

$$frac_{active} = \frac{A(0)}{A(0)+A^*(0)}. \tag{7}$$

Comparison between Eqs. 6 and 2 reveals that Eq. 6 is simply Eq. 2 scaled by a constant. This constant scaling factor is the fraction of total reactant population in the active state at $t = 0$.

Fig. 6 *D* displays Eq. 4 evaluated at seven k_1/k_3 ratios (50, 10, 2, 0.1, 0.01, 0.005, and 0.001) and plotted over the experimental time range. Importantly, the scaling factor given by Eq. 15 was fixed to 0.75 in all evaluations plotted in Fig. 6 *D*. It is clear from Fig. 6 *D* that as the k_1/k_3 ratio decreases, the peak height quickly approaches that given by Eq. 7 (note that below a k_1/k_3 ratio of 0.1 the curves overlay giving the appearance that only four curves are plotted). Fig. 6 *D* indicates that Eq. 4 approaches Eq. 6 before the limit in Eq. 5 is achieved.

Scheme 2 places the inactive to active transition model into the context of the nucleotide addition cycle. In Scheme

2 there are initially two EC_{10} populations before rapid mixing with ATP to initiate single nucleotide addition; only one of which is initially active. As described above for Scheme 1, the ODEs derived from Scheme 2 were numerically integrated to globally fit all time course data. Fig. 3, *A* and *B*, display 11-mer and CA time course data, respectively (*circles*) overlaid with simulations of Scheme 2 evaluated at the parameter values obtained from the global fit (*solid lines*, Table 1). There is clear agreement between the time courses predicted by Scheme 2 and the experimental data.

To further test Scheme 2's ability to describe our experimental data we obtained the eigenvalues from the coefficient matrix of the system of ODEs describing Scheme 2. These eigenvalues are functions describing the negative observable rate constants that Scheme 2 would produce. Scheme 2 gives rise to four unique nonzero observable rate constant functions. These functions were evaluated at the parameter values obtained from globally fitting 11-mer and CA time courses at each [ATP] to Scheme 2 and are plotted (*solid lines*) along with experimental k_{obs} versus [ATP] data (*circles*) in Fig. 7. Two of these observable rate constants are functions of [ATP] whereas two are constants. It is clear from Fig. 7 (see comparison to the reciprocal of a 2 ms quenched-flow dead time, *broken line*) that only one of the [ATP]-dependent observable rate constants would actually be observed under our experimental conditions. This leaves three observable rate constant functions that give rise to values detectable under our experimental conditions. Two of these three observable rate constants correspond well to experimental k_{obs} versus [ATP] data (Fig. 7).

Scheme 2 is clearly capable of describing both time course data (Fig. 3, *A* and *B*) and k_{obs} versus [ATP] data (Fig. 7). Thus, we used Scheme 2 to provide estimates of elementary rate constant values describing the nucleotide addition cycle.

Assignment of parameter uncertainties

It is well established that χ^2 functions resulting from systems of equations derived from models such as Scheme 2 display correlated parameters (12,13). These correlations make optimization of the parameter values and calculation of uncertainties on parameter values challenging. For example, the on and off rate constants of a bimolecular reaction are highly correlated. Taking this correlation into account, k_1 was fixed in all global fits performed in this work.

TABLE 1 Scheme 2 fitted parameter values and parameter value limits

Parameter	Fitted Value[a]	Lower Bound	Upper Bound
k_1	1×10^8 M^{-1} s^{-1}	—	—
k_2	9800 s^{-1}	8100 s^{-1}	12,100 s^{-1}
k_3	180 s^{-1}	160 s^{-1}	220 s^{-1}
k_4	≈ 0 s^1	—	—
k_5	>300 s^{-1}	300 s^{-1}	—
k_6	0.42 s^{-1}	0.40 s^{-1}	0.44 s^{-1}
k_7	0.04 s^{-1}	0.01 s^{-1}	0.07 s^{-1}
$[EC_{10}]/[EC_{total}]$	0.77	0.74	0.79

[a] k_1 and k_4 values were not optimized in the fitting routine.

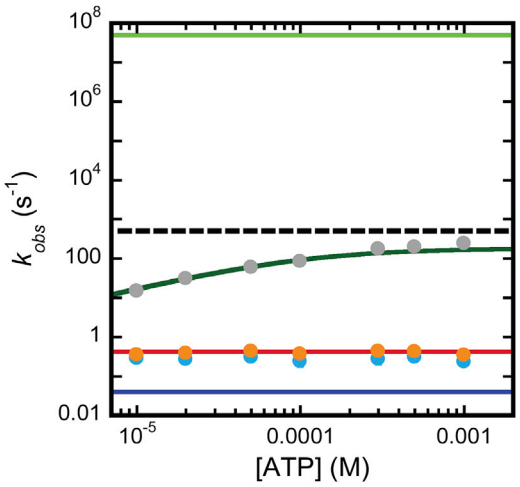


FIGURE 7 Eigenvalue decomposition of the ODEs describing Scheme 2. (*Solid lines*) Negative eigenvalues evaluated at the best-fit parameter values obtained from global WNLLS analysis of the 11-mer and CA time courses according to Scheme 2 (*red*, $-\lambda_1$; blue, $-\lambda_2$; *dark green*, $-\lambda_3$; *light green*, $-\lambda_4$). (*Circles*) Experimental k_{obs} values (*gray*, 11-mer k_{obs1}; *blue*, 11-mer k_{obs2}; *orange*, CA k_{obs}). (*Broken line*) Upper limit of a rate constant detectable in our quenched flow instrument according to a dead time of 0.002 s. To see this figure in color, go online.

To provide bounds on the elementary rate constants we obtained by globally fitting time course data to Scheme 2, we employed a grid searching strategy. In this approach one parameter at a time was chosen and its value scanned across a specified range (13). At each value the scanned parameter was held constant while all other parameters were optimized to minimize χ^2. This procedure enabled calculation of a parameter value versus *F*-statistic data set. This data set was analyzed in the context of an *F*-critical value and parameter bounds were calculated (see the Supporting Materials and Methods for a detailed description) (14). Parameter value versus *F*-statistic data sets for each parameter that floated during global fitting of Scheme 2 are presented in Fig. 8.

Grid searching revealed that not all parameters in Scheme 2 were equally constrained (see relative width of each data set at the *F*-critical value (*broken lines*) in Fig. 8). Importantly, we found that k_3, k_4, and k_5 values could not be simultaneously determined. This observation required that we constrain the fit by fixing one of the parameters and determining bounds for the remaining parameter values. Our 11-mer k_{obs1} versus [ATP] data indicate that the reverse rate of bond formation is very near zero (see Eq. 8). For this reason we chose to constrain k_4 to a value of zero and optimize the remainder of the parameters to yield the values shown in Table 1. The grid searches displayed in Fig. 8

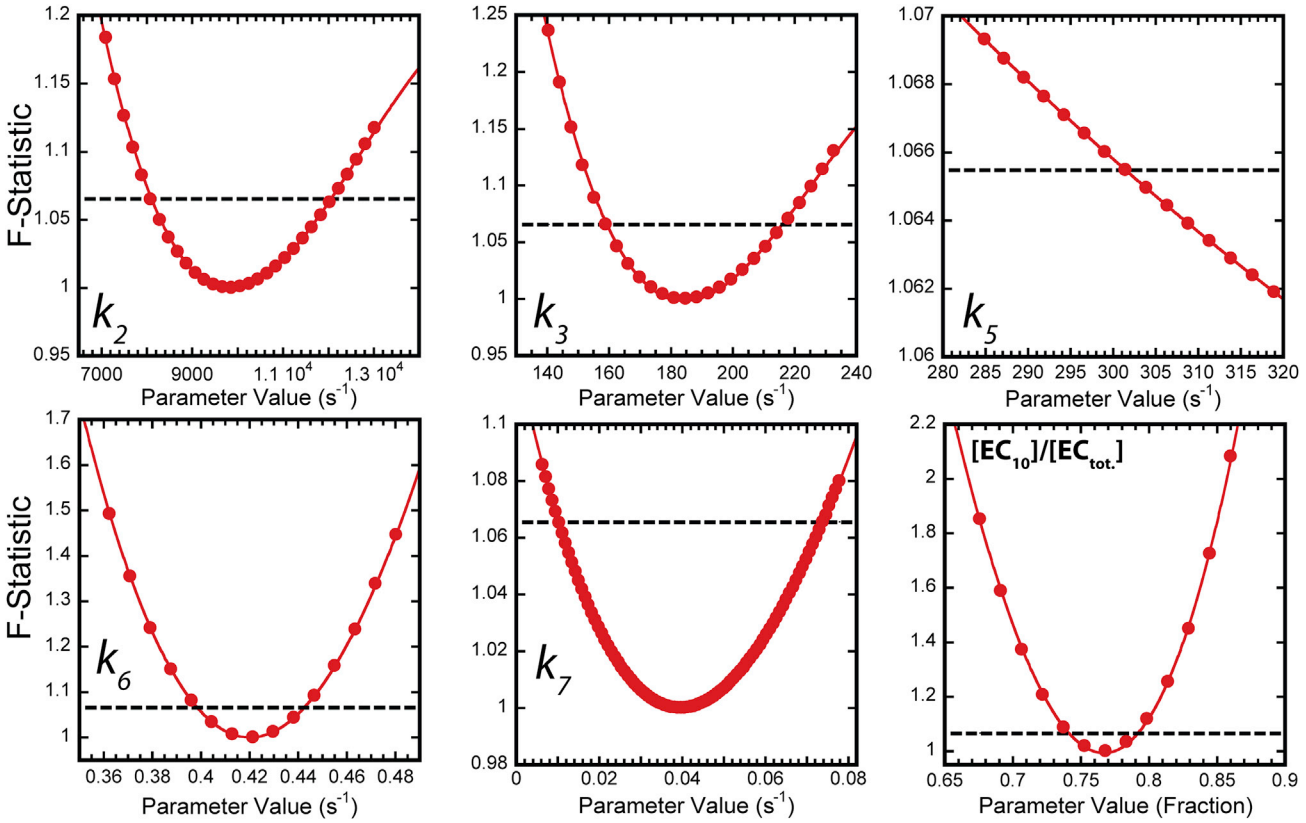


FIGURE 8 Calculation of parameter uncertainties. Each panel represents the resultant *F*-statistic versus parameter value data set from a grid search performed as described in Materials and Methods. Each panel is labeled with the specific fixed parameter. (*Circles*) Individual *F*-statistic values; (*solid lines*) fits to a third-order polynomial (with the exception of the k_5 data set, which was fit to a second-order polynomial). (*Broken line* in each panel) *F*-critical value calculated as described in Materials and Methods. To see this figure in color, go online.

were performed based on the parameter values optimized in the global fit and the calculated uncertainties are reported in Table 1. Global fitting of time course data combined with grid searching provides rigorous determination of parameter values and bounds on each parameter value.

DISCUSSION

Information content of nucleotide addition time courses

The minimal steps that must occur in the Pol I nucleotide addition cycle are displayed in Fig. 1 *A*. Using the methods described above, we have generated the first model describing the Pol I single nucleotide addition mechanism. It is important to note that our experimental approach may or may not be sensitive to all the steps in the nucleotide addition cycle. There are two simple scenarios that could each render a step invisible to our measurements. First, an event governed by a very large rate constant would reach completion in the dead time of the quenched flow instrument. Furthermore, an internal step governed by a relatively large rate constant (in comparison to flanking steps) will not contribute to the shape of the observed time course. Second, our experimental approach may leave our measurements insensitive to a specific step. A noteworthy example of a step that may fall into both categories is translocation of the polymerase along the nucleic acid. Translocation could either be relatively fast (with respect to other steps in the nucleotide addition cycle) or not be required before nucleotide addition. We do not know the translocation state of the EC with respect to the 3′-terminus of the RNA following our EC assembly protocol. Therefore, before nucleotide addition Pol I may or may not have to take a translocation step. Taking this uncertainty into account, we have not assigned any of the kinetic steps proposed in Scheme 2 to translocation.

In addition to the minimal steps depicted in Fig. 1 *A*, our experimental setup may render our measurements sensitive to steps that do not occur frequently during processive elongation. Upon conducting single nucleotide addition reactions we observed robust nuclease activity (Figs. 1 *D* and 2 *B*) that liberated a dinucleotide fragment from the 3′-end of the RNA. This observation required that we include a dinucleotide production step in any model describing the Pol I nucleotide addition pathway. With these information content considerations in mind we proceed with a synopsis of the Pol I single nucleotide addition mechanism.

ATP binding

Our data indicate that ATP binding is best described as a rapid equilibrium process with respect to the rest of the pathway. There are two key pieces of evidence that support this interpretation. First, the model-independent analyses of the time courses displayed in Fig. 3 *C* reveal that there is a step leading to 11-mer formation that is governed by an [ATP]-dependent rate constant that saturates at $k_{max} = (270 \pm 30)\ s^{-1}$ and is characterized by a half-maximal [ATP] of $K_{1/2} = (170 \pm 30)\ \mu M$. Second, the time courses shown in Fig. 3 *A* do not display a lag at any [ATP]. The lack of a lag in the time courses and the hyperbolic dependence of 11-mer k_{obs1} on [ATP] indicate that the step that is being reported on by 11-mer k_{obs1} is directly linked to ATP binding and occurs immediately after ATP binding. Assuming that ATP binding is a diffusion-limited process occurring with a rate constant of $\approx 1 \times 10^8\ M^{-1}\ s^{-1}$ and that $K_{1/2}$ is approximating the ATP dissociation equilibrium constant, ATP release must be governed by a rate constant $\approx 2 \times 10^4\ s^{-1}$. A rate constant of this magnitude places the ATP binding step well within the rapid equilibrium regime with respect to the subsequent step in the pathway (i.e., $k_{reverse} >> k_{forward}$). Rapid equilibrium ATP binding governed by a dissociation equilibrium constant $\approx 100\ \mu M$ (obtained from global fit; see Table 1, k_2/k_1) are consistent with other RNA polymerases and ATPase motor enzymes (15,16).

Work from the Erie group has indicated that *Escherichia coli* RNA polymerase possesses an allosteric NTP binding site (15,17). They observed that occupation of this site leads to alterations in nucleotide incorporation rates. Our data contain no indications of such allostery. Specifically, 11-mer k_{obs1} versus [ATP] data are described well by a rectangular hyperbola (see Fig. 4 *A*). The difference between our results and those from the Erie group highlights the need to carefully investigate each RNA polymerase system, rather than assuming conservation of mechanism. As discussed in the introduction, eukaryotes have evolved to require at least three nuclear RNA polymerases. It is unlikely that each of these enzymes is governed by the same molecular mechanisms.

Phosphodiester bond formation

We have interpreted the step immediately after ATP binding as phosphodiester bond formation or a conformational change immediately preceding rapid bond formation. In Scheme 2, k_3 is similar to k_{max} but not identical. Taking into account Scheme 2 and the fitted parameter values, k_3 should equal k_{max}. To probe this discrepancy we attempted to fit individual 11-mer time courses collected at 1 mM ATP (roughly saturating [ATP]) to a sum of two exponential terms in which one of the rate constants was constrained to equal k_3. The fits with or without this constraint were virtually indistinguishable (data not shown). In time courses collected at 1 mM ATP there are very few data points that fall in the region of the curve defining the faster of the two rate constants. This region of low data density leads to increased uncertainty in this parameter's fitted value. The observation of this discrepancy highlights the strengths of simultaneously fitting a data set collected as a function of two or more variables (in our case time and [ATP]) to constrain parameter values. Together, these analyses indicate that k_{max} is reporting on k_3.

Considering the analyses discussed in the above two paragraphs and the relative values of the fitted parameters from Table 1, 11-mer k_{obs1} can be approximated by Eq. 8 (11) in which k_n are defined in Scheme 2 and $K_{1/2}$ refers to the [ATP] at $k_3/2$,

$$11 - mer\ k_{obs1} \approx \frac{k_3[ATP]}{K_{1/2} + [ATP]} + k_4. \quad (8)$$

In the limit that [ATP] in Eq. 8 goes to zero, 11-mer k_{obs1} would equal k_4. Inspection of Fig. 4 *A* indicates that in this [ATP] limit, 11-mer $k_{obs1} \approx 0$. Satisfying the simple relationship described by Eq. 8 requires that the remaining steps of the pathway leading to signal change be rapid with respect to k_3 and k_4 (11). Although our data do not contain direct information on k_4 or k_5, the above discussion and the k_5 grid search results presented in Fig. 8 and table 1 indicate that k_4 must be very small and/or k_5 must be very large. Constraint on k_4 and k_5 will require a direct measure of one of these steps.

The Belogurov group has recently reported direct measurements of pyrophosphate release kinetics for *E. coli* RNA polymerase (18). Each nucleotide incorporation was monitored (AMP, CMP, GMP, UMP) at a single [NTP]. The Belogurov group reported rate constant values $\approx 100\ s^{-1}$ for pyrophosphate release following each incorporation event. Considering the difference in enzyme and analytical approaches between our study and theirs, this value is in reasonable agreement with the lower limit of $300\ s^{-1}$ we have placed on Pol I pyrophosphate release following an AMP incorporation (see Table 1, k_5). Further refinement of the Pol I pyrophosphate release rate constant value will require direct measurement of this step.

Nuclease activity of Pol I as a possible fidelity mechanism

Pol I-catalyzed dinucleotide production is governed by a rate constant much slower than all other steps in the pathway with the exception of the active-inactive transition (Table 1). This observation indicates that a nucleolytic event would be rare during processive elongation. It has been reported that polymerase elongation rate decreases immediately following a misincorporation event (19). Under these conditions, the probability of a nucleolytic event by Pol I would greatly increase. Thus, Pol I's nuclease activity would serve

a role in increasing transcriptional fidelity. This hypothesis is the subject of ongoing investigation.

Both Pols I and III possess robust nuclease activities, whereas purified Pol II and *E. coli* RNAP require additional *trans*-acting factors to stimulate cleavage of the nascent RNA (TFIIS and GreA/B, respectively (10)). There are two potential explanations for this disparity. First, it is possible that the cleavage assisting factors saturate elongating Pol II and RNAP in vivo. If these polymerases are continuously occupied by their cognate cleavage assisting factor in vivo the argument is an artifact of enzyme purification. Alternatively, eukaryotic evolution may have selected for enhanced fidelity of synthesis for the stable RNAs involved in translation (ribosomal RNA and transfer RNA), resulting in the observed nuclease activities in Pols I and III.

Inactive to active EC interconversion

The physical origins of a pre-ATP binding step and the observation of two polymerase populations are unclear (see EC^* and EC_{10} in Scheme 2). Heterogeneity in motor enzyme systems is not uncommon. Data obtained both at the single molecule level as well as in the ensemble point to the presence of both static (distinct enzyme populations) and dynamic (single enzymes converting between distinct conformers) heterogeneity (18,20–22). In the past decade there has been significant interest in the mechanisms governing RNA polymerase pausing (23–26). Pausing is often discussed as an on-pathway or off-pathway event in which the entire polymerase population experiences the pause or only a subset, respectively (25,26). The off-pathway events could be due to either static (only a subset of the polymerase molecules are sensitive to the conditions leading to a pause) or dynamic (stochastic recognition of the conditions leading to a pause by a given polymerase molecule) heterogeneity. As of this writing, we are monitoring multi-NTP incorporation reactions to better understand the inactive to active conversion we have observed (see EC^* and EC_{10} in Scheme 2). Specifically, we are working to determine if the polymerase population repartitions after each incorporation event or if once the inactive population has entered the active polymerase pool it remains there during a processive elongation event.

Measuring the time dependence of each intermediate RNA length in a processive elongation reaction will reveal whether or not a transiently inactivated population is generated after each incorporation event. Although these ensemble measurements will not be able to distinguish between static and dynamic heterogeneity, they will be able to distinguish between on-pathway and off-pathway events. An off-pathway event will be experienced by only a subset of the polymerase population and will be reported on as a fraction of the polymerase population entering an inactivated state after the nucleotide incorporation event. Experimentally, these events will give rise to multiphasic decays of the paused RNA intermediate. It is important to recognize that an on-pathway pause observed in the ensemble is a bit of a misnomer. Instead, this event should be viewed as a difference in relative nucleotide incorporation rates between the paused base and the flanking incorporations. An on-pathway event would be characterized by monophasic decay of the paused RNA intermediate. For now, we can say that under our conditions there is a Pol I population that must undergo a slow transition to achieve catalytic competency during a single nucleotide incorporation event.

CONCLUSION

We have developed a minimal kinetic model describing Pol I-catalyzed nucleotide addition and nuclease reactions. Using reconstituted ECs we monitored single nucleotide addition and nuclease reactions under single turnover conditions. Using global fitting techniques and a novel fitting algorithm, we provide quantitative estimates and rigorous limits of the elementary rate constants governing Pol I-catalyzed nucleotide addition and phosphodiester bond hydrolysis. Using the experimental and analytical methods described herein, a deeper understanding of the mechanistic consequences of inhibitory compounds, *trans*-acting factors, and mutations in Pol I is within reach.

SUPPORTING MATERIAL

Supporting Materials and Methods and three figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)01113-3.

AUTHOR CONTRIBUTIONS

F.D.A. designed research, performed research, contributed analytical tools, analyzed data, and wrote the article; A.L.L. designed research, contributed analytical tools, analyzed data, and wrote the article; and D.A.S. designed research, analyzed data, and wrote the article.

ACKNOWLEDGMENTS

We thank members of the Lucius and Schneider labs for critical evaluation of this work. We thank the University of Alabama at Birmingham fermentation facility for growth and harvest of yeast strains used in the RNA polymerase purification. In addition, we thank David Hilton for advisement on genetic algorithms and general MATLAB strategies.

This work was funded by a grant from the National Institutes of Health to D.A.S. (No. GM084946).

REFERENCES

1. Vannini, A., and P. Cramer. 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell.* 45:439–446.

2. Viktorovskaya, O. V., K. L. Engel, …, D. A. Schneider. 2013. Divergent contributions of conserved active site residues to

transcription by eukaryotic RNA polymerases I and II. *Cell Reports.* 4:974–984.

3. Warner, J. R. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* 24:437–440.
4. Drygin, D., W. G. Rice, and I. Grummt. 2010. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu. Rev. Pharmacol. Toxicol.* 50:131–156.
5. Bywater, M. J., R. B. Pearson, …, R. D. Hannan. 2013. Dysregulation of the basal RNA polymerase transcription apparatus in cancer. *Nat. Rev. Cancer.* 13:299–314.
6. Appling, F. D., and D. A. Schneider. 2015. Purification of active RNA polymerase I from yeast. *Methods Mol. Biol.* 1276:281–289.
7. Daube, S. S., and P. H. von Hippel. 1992. Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Science.* 258:1320–1324.
8. Sidorenkov, I., N. Komissarova, and M. Kashlev. 1998. Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol. Cell.* 2:55–64.
9. Roeder, R. G., and W. J. Rutter. 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature.* 224:234–237.
10. Yuzenkova, Y., M. Roghanian, and N. Zenkin. 2012. Multiple active centers of multi-subunit RNA polymerases. *Transcription.* 3:115–118.
11. Johnson, K. A. 1992. Transient-state kinetic analysis of enzyme reaction pathways. *Enzymes.* 20:1–61.
12. Magar, M. E. 1972. Data Analysis in Biochemistry and Biophysics. Academic Press, New York.
13. Johnson, K. A., Z. B. Simpson, and T. Blom. 2009. FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal. Biochem.* 387:30–41.
14. Correia, J. J., and W. F. Stafford. 2009. Chapter 15, extracting equilibrium constants from kinetically limited reacting systems. *Methods Enzymol.* 455:419–446.
15. Foster, J. E., S. F. Holmes, and D. A. Erie. 2001. Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation. *Cell.* 106:243–252.
16. Lucius, A. L., and T. M. Lohman. 2004. Effects of temperature and ATP on the kinetic mechanism and kinetic step-size for *E. coli* RecBCD helicase-catalyzed DNA unwinding. *J. Mol. Biol.* 339:751–771.
17. Kennedy, S. R., and D. A. Erie. 2011. Templated nucleoside triphosphate binding to a noncatalytic site on RNA polymerase regulates transcription. *Proc. Natl. Acad. Sci. USA.* 108:6079–6084.
18. Malinen, A. M., M. Turtola, …, G. A. Belogurov. 2012. Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Res.* 40:7442–7451.
19. Sydow, J. F., and P. Cramer. 2009. RNA polymerase fidelity and transcriptional proofreading. *Curr. Opin. Struct. Biol.* 19:732–739.
20. Lucius, A. L., A. Vindigni, …, T. M. Lohman. 2002. DNA unwinding step-size of *E. coli* RecBCD helicase determined from single turnover chemical quenched-flow kinetic studies. *J. Mol. Biol.* 324:409–428.
21. Liu, B., R. J. Baskin, and S. C. Kowalczykowski. 2013. DNA unwinding heterogeneity by RecBCD results from static molecules able to equilibrate. *Nature.* 500:482–485.
22. Abbondanzieri, E. A., W. J. Greenleaf, …, S. M. Block. 2005. Direct observation of base-pair stepping by RNA polymerase. *Nature.* 438:460–465.
23. Landick, R. 2006. The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* 34:1062–1066.
24. Kireeva, M. L., and M. Kashlev. 2009. Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. USA.* 106:8900–8905.
25. Bochkareva, A., Y. Yuzenkova, …, N. Zenkin. 2012. Factor-independent transcription pausing caused by recognition of the RNA-DNA hybrid sequence. *EMBO J.* 31:630–639.
26. Landick, R. 2009. Transcriptional pausing without backtracking. *Proc. Natl. Acad. Sci. USA.* 106:8797–8798.

Biophysical Journal

Supporting Material

Transient-State Kinetic Analysis of the RNA Polymerase I Nucleotide Incorporation Mechanism

Francis D. Appling,[1] Aaron L. Lucius,[2,*] and David A. Schneider[1,*]

[1]Department of Biochemistry and Molecular Genetics and [2]Department of Chemistry, University of Alabama at Birmingham, Birmingham, Alabama

Supplemental Methods

Buffers

With the exception of electrophoresis components, all buffers were filtered through 0.22 μm Millipore express plus vacuum-driven filters (Millipore) unless specified otherwise.

Buffer A: (40 mM KCl, 20 mM Tris-OAc, pH 7.9 at 25 °C); Buffer B: (40 mM KCl, 20 mM Tris-Acetate (OAc) pH 7.9 at 25 °C, 2 mM dithiothreitol, 0.2 $mg \cdot ml^{-1}$ bovine serum albumin (BSA)); Protein measurement buffer: (6 M guanidine-HCl, 0.05 M K_2HPO_4, pH adjusted to 7.0 with phosphoric acid); Storage buffer: (0.55 M K-OAc, 10 mM K-HEPES, 0.5 mM $MgCl_2$, 45% (v/v) glycerol pH 7.8); Gel loading buffer: (90% formamide, 25 mM EDTA-Na pH 8.54, 0.025 $mgml^{-1}$ bromophenol blue); Extraction buffer: (10 mM Tris-Cl pH 7.9, 1 mM EDTA-NA pH 8.54); Digestion buffer: (Tris-Cl pH 9.2, 3 mM $MgCl_2$); NA measurement buffer: (100 mM NaCl, 50 mM K_2HPO_4, pH adjusted to 7.0 with phosphoric acid); Neutralization buffer: (2.67 M Tris-Cl pH 8.4)

Proteins

Pol I was purified according to (1). Pol I is stored in storage buffer at -20 °C.

Heparin and bovine serum albumin (BSA) preparation

BSA stocks were prepared by dissolving lyophilized BSA (Fisher Scientific) in buffer A to a concentration of ≈40 $mg \cdot ml^{-1}$ and passing through a 0.2 μm nylon syringe filter (Fisher Scientific). This solution was dialyzed against three 1:500 exchanges of buffer A at 4 °C using 3,500-5,000 MWCO cellulose ester dialysis tubing (Spectra/Por). The concentration of BSA was determined by a spectroscopic assay in protein measurement buffer by taking absorbance scans between 360 nm and 220 nm. Absorbance values at 280 nm were corrected by subtracting absorbance due to scattering effects at 320 nm and an extinction coefficient of 48,150 $M^{-1} \cdot cm^{-1}$ was used to calculate concentration as described (2).

Heparin stocks were prepared by dissolving Na-heparin salt (Acros Organics) in buffer A to achieve ≈40 $mg \cdot ml^{-1}$ and filtering through a 0.2 μm nylon syringe filter (Fisher Scientific). This solution was dialyzed against three 1:500 exchanges of buffer A at 4 °C using 3,500-5,000 MWCO cellulose ester dialysis tubing (Spectra/Por). The concentration of heparin stocks were determined by a colorimetric assay (3). Absorbance at 505 nm of a 5:1:1 mixture of heparin, 0.06 M Barbital buffer (Sigma Aldrich), and 0.09 $mg \cdot ml^{-1}$ azure A was measured and concentration was quantified by comparison to a standard curve.

Nucleic Acids

RNA oligonucleotides were commercially synthesized and purified by HPLC by the manufacture (Integrated DNA Technologies; IDT). RNA oligonucleotides were dialyzed against three 1:5000 exchanges of buffer A at 4 °C using 100-500 MWCO cellulose ester dialysis tubing (Spectra/Por) and were stored at -20 °C.

DNA oligonucleotides were synthesized and de-salted by IDT. DNA oligonucleotides were further purified by denaturing polyacrylamide gel electrophoresis (PAGE) using a method based on that described by Ellington and Pollard (4). Specifically:

Lyophylized oligonucleotide was dissolved in ultrapure water at room temperature for at least 30 min to an estimated concentration of >200 μM. Oligonucleotide solution is diluted into gel loading buffer to achieve a concentration of 10-20 μM oligonucleotide. This solution was loaded onto 20 well (16 cm wide) 0.75 mm thick 18% polyacrylamide gels (19:1 acrylamide:bisacrylamide, 7 M urea, TBE). Resolution falls with increasing amounts of oligonucleotide loaded per lane and we have found that 150-300 pmole·lane^{-1} provides an acceptable balance between purity and yield. Material was electrophoresed at constant temperature (53 °C) in TBE which usually results in ≈1,500 V. Gels are laid on top of thin layer chromatography plates with UV-shadowing capability (Millipore) and exposed to a 254 nm UV lamp. Oligonucleotides were located by UV shadowing and excised with a stainless steel razor blade. Gel slices are crushed, mixed with extraction buffer to achieve 11.11 ml·g^{-1} (buffer/gel), and frozen at -80 °C. Samples were rapidly thawed by incubation in a 65 °C water bath and were incubated at 95 °C after thawing for 5 min. Samples were shaken for 16 hrs in an air shaker at 37 °C. Gel fragments were pelleted by centrifugation for 2 min at 1,000 times g, room temperature and solution was removed and filtered through 0.2 μm nylon syringe filters (Fisher Scientific). Nucleic acid was concentrated by extracting water with butanol. Oligonucleotide was precipitated by adding 0.3 volumes of 3 M Na-OAc pH 5.2 and 1.5 volumes (total volume after addition of Na-OAc) of 100% ethanol and incubated >30min at -20 °C. To calculate volumes of 3 M Na-OAc pH 5.2 and ethanol to add to oligonucleotide solutions, volume of oligonucleotide solution was estimated by weight assuming a density of 1 g·ml^{-1}. Precipitated nucleic acid was pelleted by centrifugation for 10min at 17,000 times g, 4 °C. Supernatant was removed and the precipitate was dried 15 min at room temperature. Precipitate was dissolved with 105 μL buffer A for >15 min at room temperature.

DNA oligonucleotides were dialyzed against three 1:5000 exchanges of buffer A at 4 °C using 3,500-5,000 MWCO cellulose ester dialysis tubing (Spectra/Por) and were stored at -20 °C. We did not observe degradation after long-term storage of RNA or DNA oligonucleotides under these conditions. Purity and integrity of nucleic acids were assayed by 5'-end labeling a small sample with ^{32}P and standard T4 polynucleotide kinase approaches. Labeled samples were separated on denaturing PAGE gels as described for nucleic acid purification, exposed to phosphoimager plates, and scanned with a typhoon imager for analysis.

The concentrations of oligonucleotides were determined by a spectroscopic assay. Nucleic acids were enzymatically degraded to nucleoside monophosphates by incubation with phosphodiesterase I (PDEI) (Worthington Biochemical Corp.)(5). PDEI was dissolved and stored in digestion buffer supplemented with 40% glycerol (v/v) at -20 °C. Typically 5 μL of dialyzed nucleic acid was digested in a 50 μL reaction. Digestion reactions were carried out in digestion buffer with ≈12 u·ml^{-1} PDEI at room temperature >20 min. Absorbance scans were collected between 360 nm and 220 nm in NA measurement buffer. Absorbance values at 260 nm were corrected by subtracting absorbance due to scattering effects at 320 nm. Oligonucleotide extinction coefficients were calculated by summing the extinction coefficients of each NMP multiplied by their stochiometric constant.

Extinction coefficients were obtained from Gray, Hung, and Johnson (6).

All data presented here were collected using ECs assembled with the following nucleic acid sequences:

RNA(5'-AUCGAGAGG)

DNA template strand (DNA-t) (5'
ACCAGCAGGCCGATTGGGATGGGTATTCCCTCCTGCCTCTCGATGGCTGTAAGTATC
CTATAGG)

DNA non-template strand (DNA-nt) (5'
CCTATAGGATACTTACAGCCATCGAGAGGCAGGAGGGAATACCCATCCCAATCGGC
CTGCTGGT)

RNA – DNA-t annealing reactions were carried out as follows: a solution of 10 μM RNA, 3.33 μM DNA-t in buffer A was heated to 95 °C for 5 min then cooled at a rate of 0.1 $°C \cdot s^{-1}$ to a final temperature of 4 °C.

EC Assembly

ECs were formed in buffer B at ambient temperature (≈20-23 °C). BSA and DTT aliquots were stored at 10X concentration at -20 °C and were thawed and used once. Pol I stock was diluted 1:40 into elongation complex assembly reactions.

Due to the sequential nature of the EC formation procedure the concentrations of reactants change during assembly due to dilution. The following list of concentrations corresponds to final reactant concentrations. This list will be followed by an assembly protocol in which each added component is presented as a fraction of the total reaction volume. The bold numbers on the far left of the reagent list denote the order of addition.

We found that under our conditions single nucleotide incorporation time courses are very sensitive to the presence of DNA-nt. In our experiments we used a DNA-nt / DNA-t ratio of 3 / 1. We found that beyond a DNA-nt / DNA-t ratio of 2 / 1 single nucleotide incorporation time courses are insensitive to further increase in [DNA-nt].

(≈16 nM Pol I, 162.75 nM RNA, 54.26 nM DNA-t, 162.75 nM DNA-nt, ≈5 nM α-^{32}P-CTP, 100 μM $Mg(OAc)_2$, 1.1 mM ethylenediamine teraacetic acid (EDTA) – K_3, 5 μM unlabeled CTP)

1) 0.07675 10X concentrated buffer A
2) 0.1 2 $mg \cdot ml^{-1}$ BSA
3) 0.1 20 mM DTT
4) 0.5657 H_2O
5) 0.025 Pol I
6) 0.016275 RNA – DNA-t (20min incubation)
7) 0.016275 DNA-nt (40min incubation)
8) 0.05 labeling mixture (2 mM $Mg(OAc)_2$, 100 nM α-^{32}P-CTP; in buffer A) (10min incubation)

9) 0.05 EDTA / unlabeled CTP chase mixture (22 mM EDTA-K_3, 100 μM unlabeled CTP; in buffer A)

ATP Substrates

Adenosine triphosphate (ATP) substrates were purchased from Sigma Aldrich as lyophilized sodium salts. ATP substrate stocks were prepared by dissolving in buffer A, filtering through 0.2 μm nylon syringe filters (Fisher Scientific), and dialyzing extensively into buffer A at 4 °C using 100-500 MWCO cellulose ester dialysis tubing (Spectra/Por).

ATP substrate mixtures for polymerization/nuclease reactions are prepared in buffer B and supplemented with 1.1 mM EDTA-K_3. In ATP substrate mixtures ATP, heparin, and $Mg(OAc)_2$ are adjusted to 2X concentration (with respect to reaction concentrations) to result in 1X concentration following a one to one mixing event.

Quenched Flow Time Courses

Substrate ATP mixtures were prepared in buffer B supplemented with $Mg(OAc)^2$ and ATP substrate to the desired concentration and heparin to a concentration of 0.05 $mg \cdot ml^{-1}$. Drive syringes were filled with ultrapure ddH_2O and quench syringe was filled with 1 M HCl. Each reaction was initiated by rapidly mixing 20 μL EC solution with 20 μL ATP substrate solution. Reactions were quenched by delivery of 1 M HCl. The volume of quench solution delivered is dependent upon the reaction loop and excess 1M HCl was added to bring final quench solution volume to 200 μL. Quenched reactions were neutralized by the addition of 115 μL of neutralization buffer. An aliquot of neutralized reaction was mixed in a 1:2 ratio with gel loading buffer and placed on ice for the remainder of time course collection. Samples were stored at -20 °C. We have not observed degradation of reactant or product nucleic acids during storage.

Quenched flow time courses were thermo stated at 25 °C by a circulating water bath. The order of time point collection was randomized by Microsoft Excel's random number generation function.

Nucleotide Addition Reaction Electrophoresis

Samples were heated to 95 °C >5 min and 15 μL of sample per well were loaded onto gels of the dimensions described for nucleic acid purification. Gels for analysis of nucleotide incorporation reactions were 28 % acrylamide (19:1 acrylamide:bisacrylamide), 7 M urea, TBE. A gradient buffer system was employed in which the top tank contains TBE and the bottom tank contains TBE supplemented with 0.3 M Na-OAc. We have found that it is important to pre-run gradient buffer gels ≈1 hr at 53 °C to establish the optimal state of the gradient. During nucleic acid separation gels are run at 53 °C (≈900 V). Gels were exposed and imaged as described for analysis of nucleic acid purity.

Data Analysis

Exposed gels were quantified according to EQ. S1,

$$Frac_i(t) = \frac{\frac{[RNA_i](t)}{\sum_i [RNA_i](t)} - \frac{[RNA_i](0)}{\sum_i [RNA_i](0)}}{1 - \frac{[RNA_i](0)}{\sum_i [RNA_i](0)}} \quad \text{S1}$$

where $Frac_i(t)$ is the fraction of RNA in the state i at time t and $[RNA_i](t)$ is the pixel counts of RNA corresponding to RNA in state i at time t. The summations in EQ. 2 included pixel counts corresponding to 10-mer, 11-mer, and CA. Quenched flow nucleotide addition time courses were collected three times and the time course data points plotted in figure 3 of the text represent the average data point. The uncertainty bars plotted in figure 3 of the text represent the standard deviation about this average.

The majority of data analysis was performed using custom built MATLAB functions. All code can be made available upon request.

Fitting to Analytical Functions

11-mer and CA time courses were fit individually by EQ. S2

$$Frac_i(t) = \sum_{j=1}^{n} \alpha_j e^{-k_{obs} t} \quad \text{S2}$$

to obtain observed rate constants k_{obs}. For 11-mer time courses $j = 2$ and for CA time courses $j = 1$. The observed rate constants obtained from each experimental replicate were averaged. This analysis yielded three observed rate constants corresponding to $11-mer\ k_{obs1}$, $11-mer\ k_{obs2}$, and $CA\ k_{obs}$. The uncertainty associated with the observed rate constant values obtained from fitting time course data to EQ. S2 were assumed to be equal to the standard deviation about the average of the three observed rate constants obtained from fitting experimental replicates. Observed rate constants were plotted as a function of [ATP]. *11-mer* k_{obs1} was fit using EQ. S4

$$11-mer\ k_{obs1}([ATP]) = \frac{k_{\max}[ATP]}{K_{1/2} + [ATP]} \quad \text{S3}$$

where $11-mer\ k_{obs1}([ATP])$ is the observed rate constant as a function of [ATP], $k_{\max}$ is the extrapolated observed rate constant at infinite [ATP], and $K_{1/2}$ is the [ATP] at which $11-mer\ k_{obs1}([ATP]) = 0.5k_{\max}$. $11-mer\ k_{obs2}([ATP])$ and $CA\ k_{obs}([ATP])$ were fit using EQ. S5

$$k_{obs}([ATP]) = k_{obs} + 0[ATP] \quad \text{S4}$$

which is a constant function. Analysis using EQ.s S2-S4 was performed in MATLAB using the MATLAB function Lsqnonlin. The data were fit to EQ.s S4 and S5 by minimizing the weighted sum of squared residuals (WSSR) in which each data point was weighted by the standard deviation calculated for each observed rate constant value. Uncertainty on parameter values obtained from fitting the data to EQs S3 and S4was calculated by grid searching (see below).

Fitting by Numerical Integration

11-mer and CA time courses were globally fit using the ordinary differential equations (ODEs) derived from each reaction scheme. Specifically, WSSR was calculated by summing weighted residuals of each time course (both 11-mer and CA time courses collected at each [ATP]). Numerical integration was performed in MATLAB using the Ode23s numerical integration algorithm. During numerical integration the jacobian matrix (EQ. S5) defined by the ODEs was passed to the numerical integration algorithm.

$$\frac{\partial\left\{\frac{ds_i}{dt}\cdots\frac{ds_n}{dt}\right\}}{\partial\{s_i\cdots s_n\}}=\begin{pmatrix}\frac{\partial\frac{ds_i}{dt}}{\partial s_i} & \cdots & \frac{\partial\frac{ds_i}{dt}}{\partial s_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial\frac{ds_n}{dt}}{\partial s_i} & \cdots & \frac{\partial\frac{ds_n}{dt}}{\partial s_n}\end{pmatrix} \qquad \text{S5}$$

In EQ. 6 s_n denotes the concentration of an individual species in a given reaction pathway. When fitting by numerical integration of ODEs WSSR was minimized using a custom built algorithm in MATLAB. This algorithm is based on a genetic algorithm but uses a gradient-based non-linear least squares algorithm (GBNLLS) to further minimize WSSR at each iteration. The structure of the algorithm is briefly described below.

A population of parameter sets is defined by randomly selecting individual parameter values from a Gaussian distribution of defined mean and standard deviation. WSSR is calculated for each individual parameter set in the population by the above described numerical integration procedure. The population is ranked according to WSSR. Like standard genetic algorithms; a subset of the population advances to the next iteration (generation) unaltered, a subset is subjected to mutation at a specified probability, and a subset is subjected to crossover at a specified probability (7). In our algorithm a subset of the population is further optimized by the GBNLLS algorithm Lsqnonlin. The number of individuals and the identity of individuals subjected to GBNLLS is adjustable. All fitting by numerical integration in this study was accomplished by subjecting four individuals to GBNLLS at each iteration of the algorithm. Three individuals are the top three based on WSSR and one is chosen randomly from the population. After a specified number of iterations the parameter values and WSSR is reported.

In globally fitting Schemes 1 and 2 we observed that the rate constants governing the steps corresponding to pyrophosphate release and dinucleotide production (k_5 and k_6) were interchangeable with little effect on the fit. This lack of constraint is due to the fact that only one of these steps must be rate limiting (with a rate constant of $\approx$0.4 s^{-1}) to describe the time course

data. Both published (8) and unpublished data (data not shown) indicate that during processive elongation Pol I catalyzes nucleotide incorporation at ≈20-50 nucleotides s^{-1}. These rates are not consistent with a rate constant governing pyrophosphate release (a step that must occur after ever nucleotide addition event) of ≈0.4 s^{-1}. We therefore constrained the rate constant governing pyrophosphate release (k_5) in Schemes 1 and 2 to a lower limit of 10 s^{-1} in all global fitting routines.

$k_{\mathbf{obs}}\left(\left[\mathbf{ATP}\right]\right)$ *Derivation and Simulation*

$k_{obs}\left(\left[ATP\right]\right)$ are the negative eigenvalues of the coefficient matrix defined by the DEs. Specifically, DEs derived from reaction pathways studied in this work are of the form EQ. S6

$$\frac{d}{dt}\mathbf{c} = \mathbf{M}\mathbf{c} \qquad \text{S6}$$

where **c** is the solution vector of concentrations of species in a given reaction pathway and **M** is the coefficient matrix. Importantly, for the sake of derivation, these equations were derived under pseudo first order conditions, i.e. [ATP]=constant. This does not affect the comparison of simulated $k_{obs}\left(\left[ATP\right]\right)$ functions to experimental data because all reactions were performed under pseudo first order conditions with respect to ATP, i.e. [EC]<<[ATP]. The negative analytical eigenvalues of **M** were found using MATLAB. The parameter values obtained from fitting DEs by numerical integration were substituted into the analytical eigenvalues. k_{obs} values were calculated over a range of [ATP] and the real component of the k_{obs} values were compared to experimental data.

Grid Searching

Grid searching was accomplished by systematically varying a single parameter's value and subsequently minimizing WSSR by allowing all other parameter values to be optimized. Each WSSR minimization was accomplished by Lsqnonlin. F-statistics at each parameter value were calculated according to EQ. S7

$$F-stat\left(param\ value\right) = \frac{WSSR_{grid}\left(param\ value\right)}{WSSR_{fit}} \qquad \text{S7}$$

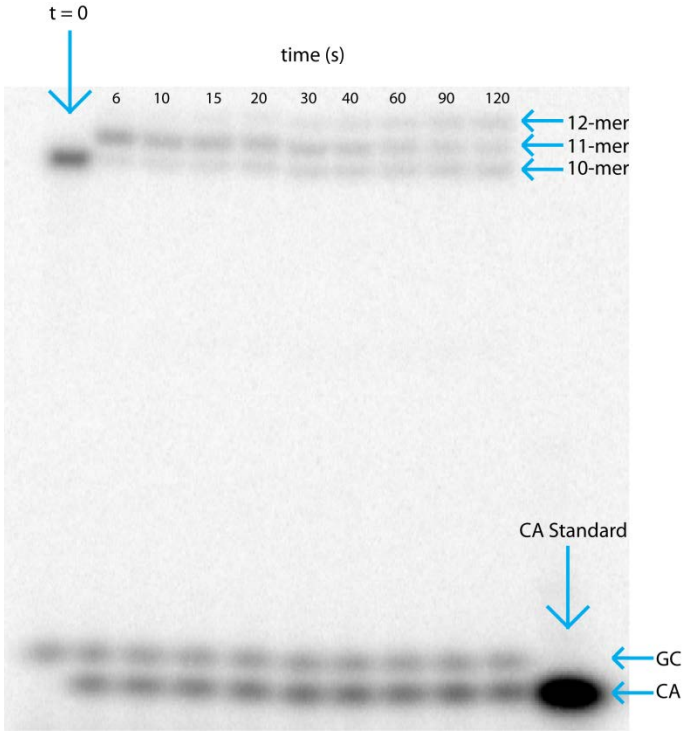
where the WSSR subscript refers to the best-fit WSSR value obtained from the original fit (*fit*) or that obtained at a fixed parameter value during a given grid search iteration (*grid*); note: $WSSR_{grid}$, and subsequently $F-stat$, are functions of the parameter values of the parameter being scanned. An F-critical ($F-crit$) value at the 68% confidence interval was calculated using Microsoft Excel's FINV function. The relationship between each parameter and F-statistics ($F-stat\left(param\ value\right)$) was defined by polynomial interpolation of F-statistic vs parameter value data in the vicinity of $F-stat = F-crit$. Polynomial fitting was performed using Kaleidagraph. *param value* at $F-crit$ was calculated using MATLAB by finding the roots of the polynomial used to fit the F-stat vs parameter value data. The roots defining the

parameter bounds of a given parameter were chosen based on inspection of $F-stat$ vs $param\ value$ plots. Simply stated, the parameter value bounds are defined by where the interpolating polynomial defining $F-stat\left(param\ value\right)$ crosses the F-crit threshold.

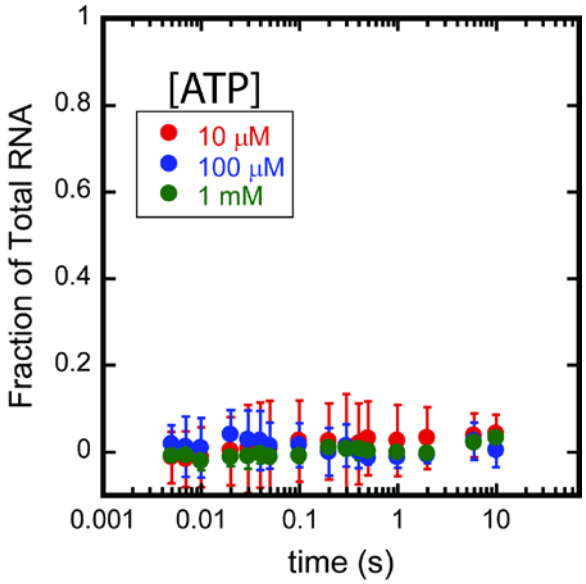
Supplemental References

1. Appling, F. D., and D. A. Schneider. 2015. Purification of active RNA polymerase I from yeast. Methods in molecular biology 1276:281-289.
2. Gill, S. C., and P. H. v. Hippel. 1989. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. Analytical Biochemistry:319.
3. Jaques, L. B. 1977. Determination of Heparin and Related Sulfated Mucopolysaccharides. Methods of Biochemical Analysis 24:203-312.
4. Ellington, A., and J. Jack D. Pollard. 1998. Purification of Oligonucleotides Using Denaturing Polyacrylamide Gel Electrophoresis. Current Protocols in Molecular Biology:2.12.11-12.12.17.
5. Lucius, A. L., A. Vindigni, R. Gregorian, J. A. Ali, A. F. Taylor, G. R. Smith, and T. M. Lohman. 2002. DNA Unwinding Step-size of E.coli RecBCD Helicase Determined from Single Turnover Chemical Quenched-flow Kinetic Studies. Journal of molecular biology 324:409-428.
6. Gray, D. M., S. H. Hung, and K. H. Johnson. 1995. Absorption and Circular Dichroism Spectroscopy of Nucleic Acid Duplexes and Triplexes. Methods in Enzymology 246:19-34.
7. Back, T., U. Hammel, and H.-P. Schwefel. 1997. Evolutionary Computation: Comments on the History and Current State. IEEE Transactions on Evolutionary Computation 1:3-17.
8. Viktorovskaya, O. V., K. L. Engel, S. L. French, P. Cui, P. J. Vandeventer, E. M. Pavlovic, A. L. Beyer, C. D. Kaplan, and D. A. Schneider. 2013. Divergent contributions of conserved active site residues to transcription by eukaryotic RNA polymerases I and II. Cell reports 4:974-984.

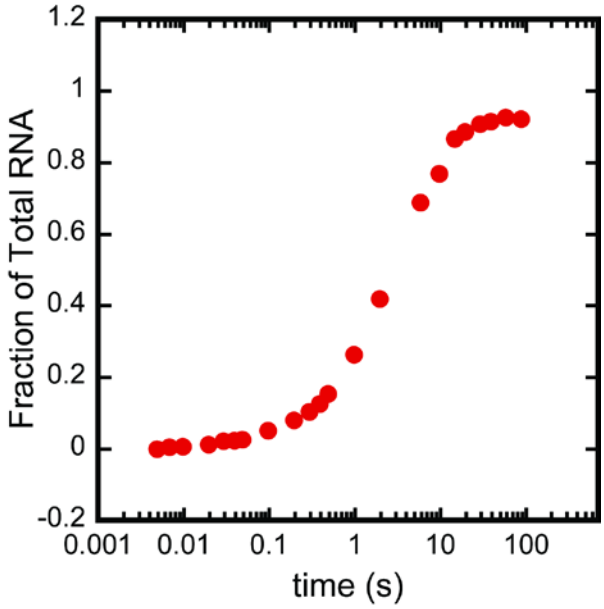
Supplemental Figures



Supplemental Figure S1. Identification of RNA species. This is a representative gel of an AMP incorporation time course collected by hand. The reactions were carried out as described in the text and materials and methods sections for quenched flow time courses except all volumes were reduced by half and time points were collected by hand. As opposed to the short time scale monitored using the quenched flow instrument, on this time scale 12-mer RNA is observed. We do not know the sequence of this RNA but it is likely due to misincorporation or trace contamination of NTP other than ATP. Its presence on these long time scales does not affect the interpretation of our data. The first lane on the left represents t = 0 material. For this lane labeled elongation complexes were denatured and electrophoresed. In this lane two bands are visible corresponding to 10-mer and GC RNA. The 10-mer RNA is the expected species following incorporation of a single ^{32}P-labeled CMP into the 9-mer substrate RNA converting unlabeled 9-mer RNA into labeled 10-mer RNA. The GC dinucleotide is the expected product from a nuclease reaction removing the incorporated ^{32}P-labeled CMP and the terminal GMP from the synthetic RNA. Beginning with the second lane from the left and moving from left to right each lane represents an increased delay time (denoted by numbers above each lane) between addition of ATP substrate and addition of HCl quench. The appearance of a slower migrating species is consistent with a single AMP incorporation extending the 10-mer RNA to the 11-mer state. The appearance of a species migrating slightly faster than GC is consistent with a dinucleotide production event immediately following incorporation of an AMP. This species co migrates with a CA standard.



Supplemental Figure S2. [GC] vs. time at 10 μM, 20 μM, and 1mM ATP. The fraction of RNA in the GC state was quantified according to EQ S1 (supplemental materials and methods). The circles represent the average of three measurements and the error bars represent the standard deviation about this average.



Supplemental Figure S3. CA time course collected at 300 μM ATP. This time course demonstrates that the ≈ 0.75 peak height values of 11-mer discussed in the text are not due to enzyme inactivation or RNA degradation.