Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II

Matthew H. Larso, Jing Zhou, Craig D. Kaplan, Murali Palangat, Roger D. Kornberg, Robert Landick, and Steven M. Block

During transcription, RNA polymerase II (RNAPII) must select the correct nucleotide, catalyze its addition to the growing RNA transcript, and move stepwise along the DNA until a gene is fully transcribed. In all kingdoms of life, transcription must be finely tuned to ensure an appropriate balance between fidelity and speed. Here, we used an optical-trapping assay with high spatiotemporal resolution to probe directly the motion of individual RNAPII molecules as they pass through each of the enzymatic steps of transcript elongation. We report direct evidence that the RNAPII trigger loop, an evolutionarily conserved protein subdomain, serves as a master regulator of transcription, affecting each of the three main phases of elongation, namely: substrate selection, translocation, and catalysis. Global fits to the force-velocity relationships of RNAPII and its trigger loop mutants support a Brownian ratchet model for elongation, where the incoming NTP is able to bind in either the pre- or posttranslocated state, and movement between these two states is governed by the trigger loop. Comparison of the kinetics of pausing by WT and mutant RNAPII under conditions that promote base misincorporation indicate that the trigger loop governs fidelity in substrate selection and mismatch recognition, and thereby controls aspects of both transcriptional accuracy and rate.

The transcription of genetic information stably encoded in DNA into a transient RNA message occurs with remarkable fidelity. RNA polymerase (RNAP) incorporates nucleotides into nascent RNA chains at rates of 10–70 s⁻¹ (1, 2), but only inserts the wrong nucleotide approximately once per 100,000 bases, on average (3). Although the energetics of base-pairing is fundamental to transcription fidelity, the discrimination for correctly inserted nucleotides is kinetically controlled, and accomplished by active site conformational changes centered on matched nucleotide substrates (4). These motions are kinetically controlled, and accomplished by active site conformational changes centered on the trigger loop (TL) (4–7). The TL is an evolutionarily conserved mobile element that stabilizes substrate NTPs in the active site of a variety of multisubunit RNA polymerases, including eukaryotic RNAPI, II, and III, as well as bacterial and archaeal RNAP (8).

In the case of RNAPII, alteration of the TL has important consequences for activity and fidelity. TL residue His1085, for example, interacts with the bound NTP substrate, and is fully conserved among multisubunit RNA polymerases. The importance of this residue is underscored by the lethality of the H1085A substitution in yeast (7) and catalytic defects resulting from substitutions at this position for both the yeast and bacterial enzymes (7, 9–11). Additionally, a substitution for residue Glu1103 of RNAPII has been found not only to promote nucleotide misincorporation (6, 7), but also to increase the elongation rate (6, 7, 12), properties that are jointly consistent with the notion of kinetic proofreading (13). To explore further the interplay between elongation rate and transcriptional fidelity, we determined the effect of E1103G and H1085A substitutions on the nucleotide-addition cycle (NAC) using single-molecule assays.

A key step in the NAC for all polymerases that remains poorly characterized is translocation along the DNA. Traditional ensemble techniques do not observe enzyme motions directly, but instead report the average positions of populations of molecules. This limitation can be problematic, owing to the intrinsic variability in synthesis rates, and for RNAPII, the inherently stochastic nature of transcriptional pausing (1), which can obscure the underlying kinetics of translocation. Here, we describe a single-molecule assay capable of directly observing the motions of individual, elongating RNAPII molecules at subnanometer resolution in real time. Using the assay, we were able to identify and exclude from analysis transcriptional pauses as short as 1 s, and thereby to determine the “pause-free” velocity of an individual molecule, which furnishes an accurate measure of its raw elongation rate. Our results for the pause-free velocity as a function of applied force favor a Brownian ratchet model for transcript elongation, in which the incoming NTP substrate is able to bind in either the pre- or posttranslocated state of RNAPII. By examining the effect of two TL mutant enzymes—containing either E1103G or H1085A/E1105G substitutions—on pause-free velocity, we show that the TL affects NTP binding, translocation and catalysis in the NAC. Furthermore, the effect of the E1103G substitution on transcription provides evidence that the TL promotes fidelity by regulating nucleotide selection and pausing during elongation, which aids in the recognition of misincorporation events.

Results and Discussion

Force-Velocity Relationships Distinguish Between Competing Models for Ordering of Events in the Nucleotide-Addition Cycle. Because every NAC involves translocation along the DNA template, the elongation rates of RNA polymerase molecules are sensitive to modulation by force, which can either assist or hinder translocation, depending upon the experimental geometry of the assay (Fig. 1A). The magnitude of any such modulation, assessed by determining the force-velocity (F-v) relationship, is affected by the step size of RNAPII and by its equilibrium between pre- and posttranslocated states. The F-v relationship is also sensitive to the temporal ordering of the translocation event with respect to NTP binding during the NAC when measured as a function of


The authors declare no conflict of interest.

This Direct Submission article had a prearranged editor.

M.H.L., J.Z., and C.D.K. contributed equally to this work.

Present address: Laboratory of Receptor Biology and Gene Expression, National Cancer Institute, Bethesda, MD 20892.

To whom correspondence should be addressed. E-mail: sblock@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1200939109/-/DCSupplemental.
NTP concentration (14), and this placement has been the subject of some controversy. Kinetic studies have suggested that NTP binding may precede, and possibly promote, translocation itself (15–17). Because most substrate-bound RNAPII crystal structures are in the posttranslocated state, they do not speak to this issue, but the identification of a nonspecific NTP-binding site in some of these structures suggests that substrate binding might precede translocation (18).

We determined the F-v relationships for wild-type (WT) RNAP II and two specific TL mutant enzymes, E1103G (red) and H1085A/E1103G (green). As Models 1 and 2, due to the assumption that NTP binding to the secondary site is energetically equivalent to active site, dies of RNAPII that posit the existence of a secondary NTP binding site, distinct from the primary nucleotide-addition site (18). (We note that although Model 3 has additional states in the reaction pathway, fits involve the same number of free parameters, m = 3, as Models 1 and 2, due to the assumption that NTP binding to the secondary site is energetically equivalent to active site, which leads to identical equilibrium constants for NTP binding pre- and posttranslocation.) Model 4, in which PP_i release can take place from either the pre- or posttranslocated state (Fig. 2D), is inspired by kinetic studies demonstrating that the rate of PP_i release is stimulated by the incoming cognate NTP (26). (Similar to the assumption made for Model 3, PP_i release is assumed to be equally likely from either the pre- or posttranslocated state, resulting in a F-v relationship with the same number of free parameters as Models 1, 2, and 3.)

Global fits were performed against the models for all the force-velocity data acquired with wild-type and mutant enzymes: this

Global fits were performed against the models for all the force-velocity data acquired with wild-type and mutant enzymes: this

Global fits were performed against the models for all the force-velocity data acquired with wild-type and mutant enzymes: this

Global fits were performed against the models for all the force-velocity data acquired with wild-type and mutant enzymes: this

Global fits were performed against the models for all the force-velocity data acquired with wild-type and mutant enzymes: this
procedure implicitly assumes that enzymes function via the same reaction pathway, regardless of mutation, albeit with different rates, and is justified by the goodness of the fit. The best-fit model (Model 3, Fig. 2C) placed the translocation step at the start of the NAC, prior to catalysis, with the incoming NTP able to bind RNAPII in either its pre- or posttranslocated state (reduced NAC, prior to catalysis, with the incoming NTP able to bind (Model 3, Fig. 2C) placed the translocation step at the start of the reaction pathway, regardless of mutation, albeit with different rates, and is justified by the goodness of the fit. The best-fit model

\[ v(F, [NTP]) = \frac{k_{\text{cat}}}{1 + \frac{k_{\text{D}}}{[\text{NTP}]}(1 + K_\delta \exp(-F\delta/k_B T)) + K_\delta \exp(-F\delta/k_B T))}, \]

where \( k_{\text{cat}} \) is the rate of nucleotide condensation followed by PPi release (these steps have been combined into a single rate; see SI Text), \( K_D \) is the NTP dissociation constant, and \( K_\delta \) describes the equilibrium between pre- and posttranslocated states \( (K_\delta = [\text{pre}]/[\text{post}]) \). The effective step size, \( \delta \), is a parameter corresponding to the distance between translocation states; this was fixed to 1 bp (0.34 nm) based on a previous measurement (14). We observed a relatively weak effect of force on the pause-free velocity for the WT RNAPII (Fig. 3A), a dependence that had not been observed previously with RNAPII, presumably due to the limited range of forces that was explored and the larger experimental uncertainties (19). Based on the fit, WT RNAPII is biased toward its posttranslocated state (Table 1), indicating that the translocation step is not rate-limiting for the advance of the WT enzyme. The apparent \( K_D \) for WT RNAPII is somewhat

\[ v(F, [NTP]) = \frac{k_{\text{cat}}}{1 + \frac{k_{\text{D}}}{[\text{NTP}]}(1 + K_\delta \exp(-F\delta/k_B T)) + K_\delta \exp(-F\delta/k_B T))}, \]

where \( k_{\text{cat}} \) is the rate of nucleotide condensation followed by PPi release (these steps have been combined into a single rate; see SI Text), \( K_D \) is the NTP dissociation constant, and \( K_\delta \) describes the equilibrium between pre- and posttranslocated states \( (K_\delta = [\text{pre}]/[\text{post}]) \). The effective step size, \( \delta \), is a parameter corresponding to the distance between translocation states; this was fixed to 1 bp (0.34 nm) based on a previous measurement (14). We observed a relatively weak effect of force on the pause-free velocity for the WT RNAPII (Fig. 3A), a dependence that had not been observed previously with RNAPII, presumably due to the limited range of forces that was explored and the larger experimental uncertainties (19). Based on the fit, WT RNAPII is biased toward its posttranslocated state (Table 1), indicating that the translocation step is not rate-limiting for the advance of the WT enzyme. The apparent \( K_D \) for WT RNAPII is somewhat

\[ v(F, [NTP]) = \frac{k_{\text{cat}}}{1 + \frac{k_{\text{D}}}{[\text{NTP}]}(1 + K_\delta \exp(-F\delta/k_B T)) + K_\delta \exp(-F\delta/k_B T))}, \]

where \( k_{\text{cat}} \) is the rate of nucleotide condensation followed by PPi release (these steps have been combined into a single rate; see SI Text), \( K_D \) is the NTP dissociation constant, and \( K_\delta \) describes the equilibrium between pre- and posttranslocated states \( (K_\delta = [\text{pre}]/[\text{post}]) \). The effective step size, \( \delta \), is a parameter corresponding to the distance between translocation states; this was fixed to 1 bp (0.34 nm) based on a previous measurement (14). We observed a relatively weak effect of force on the pause-free velocity for the WT RNAPII (Fig. 3A), a dependence that had not been observed previously with RNAPII, presumably due to the limited range of forces that was explored and the larger experimental uncertainties (19). Based on the fit, WT RNAPII is biased toward its posttranslocated state (Table 1), indicating that the translocation step is not rate-limiting for the advance of the WT enzyme. The apparent \( K_D \) for WT RNAPII is somewhat

\[ v(F, [NTP]) = \frac{k_{\text{cat}}}{1 + \frac{k_{\text{D}}}{[\text{NTP}]}(1 + K_\delta \exp(-F\delta/k_B T)) + K_\delta \exp(-F\delta/k_B T))}, \]

where \( k_{\text{cat}} \) is the rate of nucleotide condensation followed by PPi release (these steps have been combined into a single rate; see SI Text), \( K_D \) is the NTP dissociation constant, and \( K_\delta \) describes the equilibrium between pre- and posttranslocated states \( (K_\delta = [\text{pre}]/[\text{post}]) \). The effective step size, \( \delta \), is a parameter corresponding to the distance between translocation states; this was fixed to 1 bp (0.34 nm) based on a previous measurement (14). We observed a relatively weak effect of force on the pause-free velocity for the WT RNAPII (Fig. 3A), a dependence that had not been observed previously with RNAPII, presumably due to the limited range of forces that was explored and the larger experimental uncertainties (19). Based on the fit, WT RNAPII is biased toward its posttranslocated state (Table 1), indicating that the translocation step is not rate-limiting for the advance of the WT enzyme. The apparent \( K_D \) for WT RNAPII is somewhat

\[ v(F, [NTP]) = \frac{k_{\text{cat}}}{1 + \frac{k_{\text{D}}}{[\text{NTP}]}(1 + K_\delta \exp(-F\delta/k_B T)) + K_\delta \exp(-F\delta/k_B T))}, \]

where \( k_{\text{cat}} \) is the rate of nucleotide condensation followed by PPi release (these steps have been combined into a single rate; see SI Text), \( K_D \) is the NTP dissociation constant, and \( K_\delta \) describes the equilibrium between pre- and posttranslocated states \( (K_\delta = [\text{pre}]/[\text{post}]) \). The effective step size, \( \delta \), is a parameter corresponding to the distance between translocation states; this was fixed to 1 bp (0.34 nm) based on a previous measurement (14). We observed a relatively weak effect of force on the pause-free velocity for the WT RNAPII (Fig. 3A), a dependence that had not been observed previously with RNAPII, presumably due to the limited range of forces that was explored and the larger experimental uncertainties (19). Based on the fit, WT RNAPII is biased toward its posttranslocated state (Table 1), indicating that the translocation step is not rate-limiting for the advance of the WT enzyme. The apparent \( K_D \) for WT RNAPII is somewhat

\[ v(F, [NTP]) = \frac{k_{\text{cat}}}{1 + \frac{k_{\text{D}}}{[\text{NTP}]}(1 + K_\delta \exp(-F\delta/k_B T)) + K_\delta \exp(-F\delta/k_B T))}, \]

where \( k_{\text{cat}} \) is the rate of nucleotide condensation followed by PPi release (these steps have been combined into a single rate; see SI Text), \( K_D \) is the NTP dissociation constant, and \( K_\delta \) describes the equilibrium between pre- and posttranslocated states \( (K_\delta = [\text{pre}]/[\text{post}]) \). The effective step size, \( \delta \), is a parameter corresponding to the distance between translocation states; this was fixed to 1 bp (0.34 nm) based on a previous measurement (14). We observed a relatively weak effect of force on the pause-free velocity for the WT RNAPII (Fig. 3A), a dependence that had not been observed previously with RNAPII, presumably due to the limited range of forces that was explored and the larger experimental uncertainties (19). Based on the fit, WT RNAPII is biased toward its posttranslocated state (Table 1), indicating that the translocation step is not rate-limiting for the advance of the WT enzyme. The apparent \( K_D \) for WT RNAPII is somewhat
higher than that suggested by some previous reports (6); however, those determinations were carried out for a single NTP species binding to a given template position, whereas the present value corresponds to an average of all four NTPs measured over thousands of template positions.

Role of Ammonium in Transcript Elongation. A stimulatory effect of ammonium on the overall rate of transcription has been previously reported (20–22), but the mechanism for its action is still unknown. To assess the role of ammonium on individual steps in the NAC, we added 100 mM ammonium chloride to the buffer (in addition to the 130 mM KCl already present) for transcription reactions of WT RNAPII and its TL mutants. This additional set of conditions allowed us to discriminate better among the different models for the NAC, as well as to observe transcription by the H1085A/E1103G mutant enzyme at low NTPs, a condition which would otherwise produce rates too slow to observe practically in single-molecule assays due to residual amounts of baseline drift. Ammonium increased the pause-free velocity by factors of 1.5 to 4 for WT and mutant RNAPII (Fig. 3 A–C). Based on fits to Model 3 of the NAC, ammonium ion increases the pause-free elongation rate by increasing the value of $k_{cat}$ for the both WT and mutant RNAPIII, leaving $K_D$ and $K_s$ substantially unchanged (Table 1). Furthermore, the present value of ammonium ion for the WT and E1103G were statistically indistinguishable, suggesting the existence of a common, rate-limiting step under conditions where $k_{cat}$ is rapid and an assisting load is applied. Taken together, the results indicate that ammonium ion primarily alters the rate of the catalytic step of the nucleotide-addition cycle, and acts independently of trigger loop changes to E1103 or H1085.

Substitutions in the TL Affect NTP Binding, Catalysis, and Translocation. The force-velocity dependence of the TL mutant enzymes (Fig. 3 B and C) confirms that the TL is involved in catalysis (5, 9, 10), because the E1103G mutation produced a 1.6-fold increase in $k_{cat}$ over the WT (Table 1). We also observed more than a 2.5-fold decrease in $K_D$ compared to the WT, indicating that the incoming NTP is more tightly bound to the active site. This reduction may be attributable to a preferential closure of the mutant TL around the bound substrate. The increases observed in both NTP-binding affinity and catalysis were accompanied by an inhibitory effect on translocation, because the E1103G mutant enzyme was found to be significantly more biased toward its pretranslocated state than the WT. A pretranslational bias is somewhat counterintuitive, in view of the increased elongation rate for the mutant enzyme. However, it can be understood because the TL substitution promotes a net forward translocation of RNAPII. In the case of E1103G, the combination of tighter NTP binding and a faster catalytic rate (Table 1) lead to an increase in the elongation rate despite an inhibition of the translocation step. It was previously observed that E1103G displayed a bias for the pretranslocated state in the absence of the correct NTP when tested at a single template position, and that this bias shifted toward the posttranslocated state in the presence of the correct NTP (6). By contrast, our best-fit model of the NAC, which supplies uncoupled measures of translocation ($K_s$) and NTP binding ($K_D$) over many different template positions, instead suggests a pretranslocation bias for E1103G, even at high NTP levels.

If the observed properties of the E1103G substitution were attributable to a stabilization of the closed state of the TL (dependent upon a bound, template-matched NTP), which tends to be more open in the WT, we reasoned that a compensatory destabilization of this closed state might reverse some, or all, of the mutant enzyme properties. Guided by structural data for the TL and RNAPII active site (4, 18), we altered a second residue of the TL, His1085, which, through interaction with NTP substrates, promotes catalysis and likely favors the closed conformation of the TL (4, 5, 7, 9–11). H1085A lethality is suppressed by combination with E1103G, but the transcriptional and phenotypic properties of the double substitution suggest mutual suppression (11). As anticipated, based on compromise of the critical substrate-interacting sidechain, the H1085A/E1103G double-mutant enzyme exhibited a significantly slower elongation rate than either the WT or E1103G single-mutant enzymes, and therefore returned a reduced value for $k_{cat}$ (Table 1). However, in several other key respects, the WT characteristics were restored. Notably, the load dependencies of the F-v curves for the H1085A/E1103G and WT enzymes were similarly flat (Fig. 3C). Moreover, the nucleotide binding affinity ($K_D$) of the double-mutant enzyme was statistically indistinguishable (Table 1). Taken together, these results suggest that the TL serves to affect not only the rate of catalysis, as previously found, but also the NTP-binding affinity and the translocation equilibrium. An inhibitory effect of TL closure on translocation was recently proposed, based on structural modeling and computational studies (30), and is consistent with the proposed “two-ratchet model” for RNAP translocation (27). In that proposal, the closed state of the TL, which is dependent

**Table 1.** Trigger loop affects catalysis, NTP binding, and translocation equilibrium. Global fit parameters (errors as SD) obtained from modeling of F-v curves in Fig. 3 to the two-binding site Brownian ratchet model of Fig. 2C.

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>E1103G</th>
<th>H1085A/E1103G</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$</td>
<td>34 ± 2 s$^{-1}$</td>
<td>56 ± 2 s$^{-1}$</td>
<td>6 ± 1 s$^{-1}$</td>
</tr>
<tr>
<td>$k_{cat}$(NH$_4^+$)</td>
<td>77 ± 3 s$^{-1}$</td>
<td>77 ± 3 s$^{-1}$</td>
<td>22 ± 2 s$^{-1}$</td>
</tr>
<tr>
<td>$K_D$</td>
<td>140 ± 16 μM</td>
<td>49 ± 2 μM</td>
<td>94 ± 7 μM</td>
</tr>
<tr>
<td>$K_s$</td>
<td>0.2 ± 0.1</td>
<td>0.8 ± 0.1</td>
<td>0.3 ± 0.1</td>
</tr>
<tr>
<td>$δ$</td>
<td>0.34 nm (fixed)</td>
<td>0.34 nm (fixed)</td>
<td>0.34 nm (fixed)</td>
</tr>
</tbody>
</table>

**Fig. 3.** Force-velocity curves. Average pause-free elongation rates (errors as SEM) measured as functions of force for (A) WT, (B) E1103G, and (C) H1085A/E1103G RNAPII at saturating (1 mM) and subsaturating (0.1 mM) NTPs, as well as in the presence of 100 mM ammonium chloride, which was added to our standard elongation buffer. Global fits to the two-binding site Brownian ratchet model (Model 3, Fig. 2C) are shown (dashed lines).
upon a match between the incoming NTP and the template DNA, precludes translocation. However, as the correct NTP is condensed into the nascent RNA, forward translocation is facilitated by the opening of the TL, induced by thermal motion, likely along with a corresponding alteration in the conformation of the bridge helix (31). The single-molecule data presented here provide experimental evidence consistent with such TL-gated translocation. This gating is analogous to models of translocation in DNA polymerases, where the reversals of domain movements coupled to substrate selection and catalysis have been implicated in the subsequent translocation step, but by and large have been experimentally inaccessible (32).

Pausing Kinetics Imply a Role for the Trigger Loop in Transcriptional Fidelity. Because pausing has been implicated in RNAP proofreading (33, 34), we examined the characteristics of transcriptional fidelity and pausing behavior in WT RNAPII and its TL mutants. Misincorporation of certain bases (such as the GTP analog, ITP) by RNAP results in backtracking along the DNA template, leading to a relatively long-lived, paused state. The backtrack motion displaces the 3′-end of the nascent RNA from the enzyme active site, permitting transcription factors (such as TFIIIS in eukaryotes or Gre factors in bacteria) to activate cleavage of the most recently incorporated bases, thereby removing the error and restoring a new 3′-end of RNA in the active site, allowing transcription to resume (33–35).

The RNAPII pause lifetime distribution could be subdivided into two categories, the majority of which (>95%) were “short,” with an average time of ~1.1 s, and the remainder of which (~5%) were “long” (>15 s) (Fig. S5). WT RNAPII entered these long pauses approximately once every 10 kb at subsaturating NTP concentrations, a rate that corresponds closely to the frequency for the incorporation of an incorrect nucleotide substrate, as measured by ensemble assays in vitro (36). It has been shown previously that E1103G promotes NTP misincorporation (6, 7). Consistent with that finding, we observed a slight, but not statistically significant, increase in the long pause density (LPD, number of long pauses/kb) for the E1103G mutant enzyme over the WT enzyme at 0.1 mM NTPs (Fig. 4A).

To ascertain whether the increase in long pausing observed for the RNAPII mutant was caused by an increased rate of mismatches arising from a deficiency in substrate selection, we added 1 mM additional nucleotide for each NTP in turn to subsaturating NTPs (0.1 mM NTPs), which serves to promote mismatching of the excess nucleotide by mass action. The LPD for E1103G increased nearly fivefold upon the addition of excess ATP and about threefold in the presence of excess CTP or UTP (Fig. 4B). Only a modest increase in LPD was observed in the presence of excess GTP. By contrast, the LPD of wild-type RNAPII was comparatively insensitive to excess NTPs (Fig. 4A) and reached a minimum in the presence of additional GTP. These results imply that the E1103G mutant enzyme differentially undertakes long pauses dependent on the specific excess NTP. In particular, the inferred misincorporation events involving GTP are less frequent, or less able to trigger a long pause, or both. This finding is consistent with a mismatch-specific proofreading mechanism proposed previously (33, 36, 37), and may be attributable to specific interactions between the TL and the misincorporated 3′-end of the RNA, which could trigger long pauses.

Inosine triphosphate (ITP), a GTP analog that forms a weak Watson-Crick pair with cytosine and has a similar binding constant and rate of incorporation into RNA, also inhibits next-nucleotide-addition rates in a manner similar to that of certain mismatched base pairs. This property has been exploited previously to investigate postincorporation nucleotide discrimination and removal by RNAPII (38). At saturating concentration of NTPs (1 mM) in the absence of ITP, E1103G showed a clear increase in LPD over WT (Fig. 4B). For the WT RNAPII, the addition of 200 μM ITP to transcription reactions containing 1 mM NTPs significantly increased the LPD (Fig. 4B), as previously described for E. coli RNAP (34), which is indicative of misincorporation-induced pause enhancement. Perhaps surprisingly, the LPD for E1103G decreased in the presence of ITP, despite the fact that the mutant enzyme is able to incorporate ITP into transcripts to the same extent as the WT (Fig. S4). Taken together, these results suggest that the TL promotes transcriptional fidelity in two distinct ways. First, the TL helps to ensure correct NTP selection from among a pool of mixed substrates. Mutation of the TL (E1103G) generally leads to more mismatches being added to the transcript. Second, the TL functions in the efficient recognition of misincorporation events by RNAPII once a selection error has been made, as evidenced by backtracking (38). The E1103G mutant enzyme also appears to be deficient in this regard, because the misincorporation of ITP fails to be recognized, and therefore fails to produce a corresponding increase in the proportion of long pauses.

Conclusion

The observation of elongation by individual RNAPII molecules has provided additional insights into eukaryotic transcription mechanisms, and it offers the ability to separate effects on catalysis from those on translocation, which is critical for the dissection of the RNAPII mechanism. Based on high-resolution measurements of the load dependence of the elongation rate, we find support for a biochemical pathway where incoming NTPs can associate with RNAPII in either its pre- or posttranslocated state. Both single-molecule and bulk biochemical data, comparing wild-type behavior with substitutions in the trigger loop, suggest that this key subdomain of RNAPII functions not only to regulate the basic elongation rate, via changes to the translocation bias and catalytic rate, but also plays a critical role in both NTP substrate selection and mismatch recognition. In particular, substrate selection defects were characterized under conditions of competing levels of matched and mismatched NTPs. The experiments also revealed significant differences in the recognition of ITP misincorporation by the TL mutant enzyme, implicating the TL in the promotion of pausing associated with proofreading repair. We conclude that the sequence of the trigger loop evolved to strike a critical balance between the opposing constraints of catalytic prowess and transcriptional fidelity.

Materials and Methods

A detailed description of materials and methods is given in SI Text.

Fig. 4. Substrate selection and mismatch recognition. (A) Long pause density (long pauses per 10 kb transcribed; errors as SEM) for the WT (blue) and E1103G mutant enzymes (red) at subsaturating (0.1 mM) NTPs under the conditions indicated. The long pause density increases for the mutant enzyme when a single NTP is present in excess, indicative of a deficiency in substrate selectivity. (B) The inclusion of 200 μM ITP in a saturating NTP buffer (1 mM NTPs) increases the long pause density for the WT (blue), but decreases the density for the E1103G mutant enzyme (red), indicative of a failure in mismatch recognition for ITP incorporation events.
Transcription Initiation. Biotinylated, 12-subunit Saccharomyces cerevisiae RNA Pol II (prepared as described in SI Text) was assembled stepwise onto a DNA-RNA scaffold using a modification of a previous method (39), allowing us to circumvent the need for initiation factors normally required to assemble a transcriptional elongation complex. The 2.7 kb upstream and 4.8 kb downstream DNA “handles” were ligated to the ends of the template DNA by incubating the RNA Pol II scaffolding complexes (5 nM) with these DNA fragments (20 nM each) in elongation buffer (25 mM Hepes-KOH, pH 8.0, 130 mM KCl, 5 mM MgCl2, 1 mM DTT, 0.15 mM EDTA, 5% glycerol and 25 μg acetylated bovine serum albumin/mL) in the presence of ligase (2 K units) and ATP (1 mM) at 12 °C for 1 h.

Data Collection and Analysis. Records of position data were acquired at 2 kHz using custom software (written in Labview), filtered at 1 kHz by an 8-pole lowpass Bessel filter, and analyzed using custom software (written in Igor Pro). The distance between optically trapped beads was converted into the DNA contour length (14) and pseuods were identified and scored as described previously (34). Pause-free elongation velocities of single RNA Pol II molecules were determined by fitting the velocity distribution to a sum of two Gaussianians (one corresponding to pausing, and the other to active elongation) as previously described (1).

Note Added in Proof. Additional information on materials and methods may be found in ref. 42.

ACKNOWLEDGMENTS. We thank D. Koslover and V. Schweikhard for help with data collection, D. Larson for critical reading of the manuscript, and J. Gelles and members of the Block lab for helpful discussions. This work was supported by a National Science Foundation Graduate Research Fellowship (to M.H.L.) and grants from the National Institute of General Medical Sciences (to S.M.B., R.L., and R.D.K.).
**Supporting Information**

**Larson et al. 10.1073/pnas.1200939109**

**SI Results and Discussion**

**Fitting the F-v Data to Different Kinetic Models.** To determine the location of the translocation step within the nucleotide-addition cycle, as well as the biochemical steps that are affected by ammonium (and related) ions, we fit the global data to each of four different Brownian ratchet-type models (Fig. 2). Three of these pathways place the NTP binding event before, after, or either before or after translocation, but prior to pyrophosphate release; the fourth pathway allows pyrophosphate release to occur either before or after translocation. In Model 1 (Fig. 2A), RNA polymerase II (RNAPII) oscillates stochastically between the pre- and posttranslocated states with thermal motion. Assuming a Boltzmann-type behavior for the load-dependent step, this model predicts that the pause-free elongation rate, v, will depend upon the applied force and NTP concentrations as follows:

\[ v(F, [NTP]) = \frac{k_{cat}}{1 + K_D^{NTP}(1 + K_S \cdot \exp[-F \cdot \delta/k_B T])}. \]  

[S1]

where \( K_D \) is the NTP dissociation constant, \( K_S \) describes the equilibrium between the pre- and posttranslocated states (\( K_S = \text{pre}/\text{post} \)), and \( \delta \) is a characteristic distance parameter, which here represents the distance between the pre- and posttranslocated states (fixed to 1 bp = 0.34 nm in these fits). The condensation reaction and PPi release steps are not separately visualized in the single-molecule assay, and are therefore combined here into a single kinetic parameter representing the catalysis event, \( k_{cat} \).

In Model 2 (Fig. 2B), substrate binding occurs prior to translocation, suggesting that the incoming NTP drives translocation (1), and the thermal motion between the pre- and posttranslocated states is rectified by the NTP condensation reaction. Model 2 predicts the following F-v behavior:

\[ v(F, [NTP]) = \frac{k_{cat}}{1 + K_S \cdot \exp[-F \cdot \delta/k_B T](1 + K_D^{NTP})}. \]  

[S2]

In Model 3, the incoming NTP can bind in either the pre- or posttranslocated state (Fig. 2C) (2). We assume that both binding sites are energetically equivalent, which leads to identical equilibrium constants for NTP binding the pre- and posttranslocated states. As such, Model 3 has the same number of free parameters (\( m = 3 \)) as both Models 1 and 2, and predicts the following F-v relationship:

\[ v(F, [NTP]) \]

\[ = \frac{k_{cat}}{1 + K_S \cdot \exp(-F \cdot \delta/k_B T) + K_S \cdot \exp(-F \cdot \delta/k_B T)}. \]  

[S3]

Model 4, in which PPi release can take place from either the pre- or posttranslocated state (Fig. 2D), is inspired by kinetic studies demonstrating that the rate of PPi release is stimulated by the incoming cognate NTP (3). Similar to the assumption made for Model 3, PPi is assumed to be equally likely from either the pre- or posttranslocated state, resulting in a F-v relationship that has the same number of free parameters as Models 1, 2, and 3:

\[ v(F, [NTP]) = \frac{k_p}{1 + K_P^{NTP} (K_S \cdot \exp[-F \cdot \delta/k_B T])}. \]  

[S4]

Global fits to the measured pause-free elongation rates for WT, E1103G, and H1085A/E1103G RNAPII at all our experimental conditions are shown for all four models in Fig. S1.

The experimental data best support the secondary binding-site model (Model 3, Fig. 2C), where the NTP can bind RNAPII in either the pre- or the posttranslocated state. As before, we have assumed that the equilibrium constants for both NTP-addition pathways, \( K_S \) and \( K_D \), are identical (i.e., that NTP binding to either the pre- or the posttranslocated state is energetically equivalent) and that the energetics of the translocation event are unchanged by NTP binding. If we relax that particular assumption, we arrive at the more complex Model 5 (Fig. S24). The F-v relationship predicted by this model is:

\[ v(F, [NTP]) = \frac{k_{cat}}{1 + K_S^{NTP} (1 + K_D \cdot \exp(-F \cdot \delta/k_B T)) \cdot A \cdot K_S \cdot \exp(-F \cdot \delta/k_B T)}. \]  

[S5]

Here, an additional free parameter, \( A \), has been introduced, which represents the energetic penalty associated with having an NTP in the secondary binding site whereas the enzyme remains in its pretranslocated position. A value of \( A = 0 \) would preclude any binding at all in this position, whereas a value of \( A = 1 \) suggests that binding to either the primary or secondary binding site is energetically equivalent (2). Global fits of the F-v curves for WT RNAPII and its trigger loop mutants to various \( A \) values spanning an order of magnitude are shown in Fig. S2B–D. Fitting to a range of \( A \) values still captured the main features of the F-v relationship for the WT RNAPII and its two trigger loop mutants. Moreover, any qualitative differences among the remaining parameters (i.e., changes to \( k_{cat}, K_D \), or \( K_S \)) for WT RNAPII compared to its trigger loop (TL) mutants were unaltered. The fits are therefore not particularly sensitive to the value of \( A \) (Model 5) under the conditions of these experiments, making the simpler version represented by Model 3 more attractive.

**ITP Misincorporation Assay.** The addition of 200 μM ITP to the transcription buffer containing saturating concentrations of NTPs led to a large increase in long pause density (LPD) for WT RNAPII (Fig. 4B), consistent with previous single-molecule studies of ITP misincorporation by *Escherichia coli* RNAP (4). By contrast, the long pause density (LPD) for the E1103G mutant enzyme dropped somewhat in the presence of ITP under identical conditions (Fig. 4B). The inability of ITP to increase the LPD in the case of E1103G could come about in two ways, in principle, that are not mutually exclusive (1). The E1103G mutant enzyme fails to incorporate ITP into the transcript, thereby precluding any need for mismatch recognition and the subsequent triggering of a long pause, or (2) The E1103G mutant enzyme misincorporates ITP, but thereafter fails to recognize the mismatch and trigger a long pause. To distinguish between these alternatives, we monitored transcript elongation in bulk by both the WT and E1103G enzymes in the presence of ITP.

WT and E1103G ECs were assembled using the preformed synthetic scaffold described in the supporting methods below, resulting in ECs with a 14 nt transcript, and both EC preparations...
were incubated with four different NTP mixtures. To correct for differences in the concentration of active ECs, WT, and E1103G ECs were incubated in the presence of *AG mix (0.33 μM α,β-ATP, 1.3 μM each of ATP and GTP) for 2 min at 30 °C, allowing extension along an AG cassette (AGGAAA) until the enzyme stalled at A20 due to CTP deprivation. We found that E1103G complexes were present in a 1.8-fold excess compared to the WT (compare lanes 1 and 5 in Fig. S4), and both ECs were able to extend to the end of the template in the presence of *ACG mix (0.33 μM α,β-ATP, 1.3 μM each of ATP, CTP, and GTP) as evidenced by lanes 2 and 7. In the presence of *AI mix (0.33 μM α,β-ATP, and 6.6 μM ITP), both ECs showed extension to the A20 position (lanes 3 and 7), indicating ITP misincorporation at positions coding for ITP), both ECs showed extension to the A20 position (lanes 3 and 7), indicating ITP misincorporation at positions coding for ITP, respectively. In the absence of ITP, indicating that E1103G is deficient in both substrate selection and mismatch recognition in the case of ITP.

SL Materials and Methods

Materials. All DNA and RNA oligonucleotides were from IDT and were purified by denaturing PAGE before use. FPLC-purified NTPs were from Promega Corp. The α,β,γ-ATP was from Perkin-Elmer. T4 Polynucleotide Kinase and T4 DNA Ligase were from New England Biolabs.

Purification of Yeast RNA Polymerase II. pBS1479 (5), a plasmid designed for C-terminal tagging proteins in yeast with the tandem-affinity purification (TAP)-epitope tag, was derivatized to additionally encode a high-affinity peptide [the Avitag (6)] BirA biotinylase. This biotinylase catalyzes the biotinylation of proteins and is specifically encoded a high-affinity peptide [the Avitag (6)] BirA biotinylase. This biotinylase catalyzes the biotinylation of proteins and is specifically targeted for fusion of Avitag-TAP encoding RPB1, RPB2, and RPB3 subunits of Pol II. RNAPII was purified before use.

Transcript elongation by A20 ECs. After ligation of the DNA fragments to the labeled A20 ECs, the complexes were incubated with ATP, CTP, and GTP (100 μM each) for 5 min at 30 °C to form G76 ECs, which were then elongated in the presence of all 4 NTPs (1 mM each) for 5 min at 30 °C.

template strand (100 nM) oligonucleotide 5′-GAGTTGGTGCTGCGTGGTTTCTC-3′ and 5′-CCCAAGCGACACCACGGG, and incubated for another 10 min at 30 °C. The resulting ECs were then elongated in the presence of all 4 NTPs (1 mM each) for 5 min at 30 °C.

Fig. S1. Global fits to Brownian ratchet models of elongation. The single-molecule, pause-free elongation rates (error bars as SEM) measured as a function of external load for WT, E1103G, and H1085A/E1103G RNAPII at saturating (1 mM) and subsaturating (0.1 mM) NTPs, as well as in the presence of 100 mM ammonium, color-coded as indicated by the legends (Inset). The experimental datasets were fit globally to each the four Brownian ratchet reaction pathways shown in Fig. 2 (fits to Eqs. S1–S4, represented by dashed lines). There were 12 independent parameters fit for each model: specifically, the four kinetic parameters $k_{cat}$, $k_{cat}(NH_4^+)$, $K_i$, and $K_D$ for each of three RNAPII species (WT, E1103G, and H1085A/E1103G).

Model 1: NTP binds after translocation ($\chi^2 = 2.55$; $v = 33; N = 45; p < 0.005$)

Model 2: NTP binds before translocation ($\chi^2 = 2.57$; $v = 33; N = 45; p < 0.005$)

Model 3: NTP binds before or after translocation ($\chi^2 = 0.52; v = 33; N = 45; p = 0.99$)

Model 4: PPI release before or after translocation ($\chi^2 = 3.82; v = 33; N = 45; p < 0.005$)
Fig. S2. Alternate fitting of the F-v data. (A) Model 5 is identical to Model 3, but with the addition of another free parameter, $A$, that allows for different rates of binding to the pre- and posttranslocated states. The F-v data were fit to Model 5 with fixed $A$ values spanning an order of magnitude ($A = 10$, pink solid line; $A = 0.1$, green solid line). Dashed lines represent fits to Model 5 with $A = 1$, which causes this model to reduce to Model 3, depicted in Fig. 2C.

Fig. S3. Pause lifetime distribution for WT RNAPII. Histogram of the measured pause durations ($N = 98$, error bars as SD) from records of WT RNAPII transcribing at 10 pN assisting load and 1 mM NTPs. Dashed line represents a single-exponential fit to the data ($\tau = 1.1$ s).
Fig. S4. ITP misincorporation assay. WT and E1103G ECs were incubated with four different NTP mixtures (see SI Results and Discussion), and the radiolabeled transcripts were separated on a 12.5% denaturing polyacrylamide gel. The positions of the A20 halt site and run-off product (75-mer) are indicated.