

Mechanochemical Kinetics of Transcription Elongation

Lu Bai, Robert M. Fulbright, and Michelle D. Wang*

Department of Physics, Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853, USA
(Received 25 May 2006; published 8 February 2007)

The mechanochemical kinetics of transcription elongation was examined with a combination of theoretical and experimental approaches. The predictive power of a sequence-dependent thermal ratchet model for transcription elongation was tested by establishing model parameters based solely on measurements under chemical perturbations and then directly predicting responses under mechanical perturbations without additional model parameters. Agreement between predicted and measured force-velocity curves provides strong support for a simple mechanochemical coupling mechanism.

DOI: [10.1103/PhysRevLett.98.068103](https://doi.org/10.1103/PhysRevLett.98.068103)

PACS numbers: 87.15.Aa, 87.14.Gg, 87.15.Rn, 87.80.Cc

RNA polymerase (RNAP) is the key enzyme involved in transcription, an essential process in gene expression and its regulation. During transcription, RNAP moves processively along a DNA template, incorporating nucleotides into the 3' end of the nascent transcript RNA. From a mechanical point of view, RNAP functions as a molecular motor, utilizing chemical catalysis to forward translocate along the DNA (for a review, see [1]). One of the most fundamental questions regarding mechanisms of transcription is how RNAP couples chemical catalysis to mechanical translocations during transcription elongation.

The mechanochemical coupling of transcription elongation by *E. coli* RNAP has been modeled with distinct reaction pathways [2–10], which differ in complexity as manifested in the number of translocation states, the number of nucleotide binding sites, etc. Given that each model has been shown to yield reasonable fits to some data from biochemical or single molecule measurements, a central question remains as to how to evaluate these models. This evaluation should not rely on the quality of the fits, since it is well known that a model with a sufficiently large number of parameters can fit almost any set of data. A more stringent test of a model would be based on its predictive power. That is, once model parameters are determined under a set of conditions, the quality of a model could then be assessed by its ability to correctly predict the outcome of new measurements under compatible conditions but far from those used to tune the model parameters.

Elongation kinetics is governed by DNA sequence, NTP incorporation chemistry, and external load applied to RNAP. Previously, we formulated a sequence-dependent thermal ratchet model for *E. coli* RNAP elongation [7]. This initial formulation focused on the sequence-dependent kinetics while simplifying the NTP incorporation chemistry and without consideration of external load. To extend and test the predictive power of our previous model, we first established NTP-specific model parameters based on experimental measurements in response to chemical perturbations. Then this model, without any additional parameters, was used to predict how transcription

would respond to mechanical perturbations. Thus, these direct predictions served as straightforward tests of the model and, specifically, the mechanochemical coupling mechanism on which the model is based.

Experimental data presented here were all acquired from single molecule measurements using an optical trap as previously described [11,12]. RNAP template position was monitored in real time and was used to indicate the length of the RNA transcript. The instantaneous velocity distribution was accumulated by giving equal weight to velocity at each template position [12]. Obviously, both the main (essential) and the branched reaction pathways contribute to the measured velocity. However, branched reactions often lead to transcription pauses [13–15], which had a low probability of occurrence (typically <1 pause per 100 bp) under all of our experimental conditions. Thus, this method of velocity determination reflects primarily the kinetics of the main reaction pathway.

We probed elongation velocity responses to chemical perturbations by varying the template composition and concentration of the nucleoside triphosphates (NTPs) [16]. In Figs. 1 and 2, we present experimental evidence to show that kinetic properties of the NTP incorporation reaction are NTP-type-specific.

In the first experiment, four sets of data were acquired on four different DNA templates, each with a segment lacking one nucleotide (i.e., an *N*-less cassette) while containing a random sequence of the remaining three nucleotides [17]. Each *N*-less cassette was joined to the same upstream DNA template containing all four nucleotides [17]. As shown in Fig. 1(b), the elongation velocities were clearly dependent on the *N*-less cassette type, with the *A*-less cassette trailing and the *U*-less cassette leading, indicating that ATP (UTP) catalyzed faster (slower) than the other three nucleotides. Importantly, the velocities on the template upstream from the *N*-less cassettes at 1 mM NTPs were highly consistent among all four templates [Fig. 1(b)], indicating that the differences in the *N*-less cassette velocities were not due to possible velocity heterogeneity in RNAP molecules [18,19]. These results clearly show that the catalytic rate is NTP-specific.

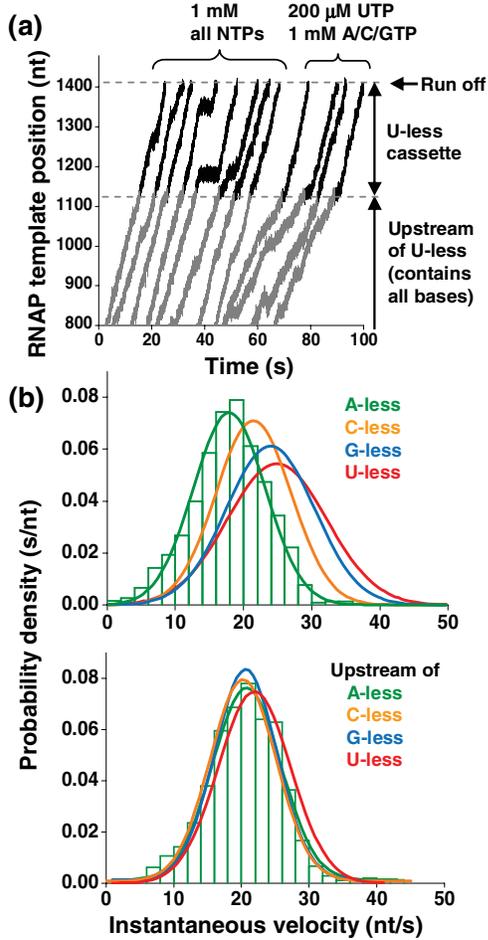


FIG. 1 (color). Elongation velocities on templates containing N -less cassettes. (a) Examples of single molecule traces of RNAP template position (relative to the promoter) vs time on a template containing a U -less cassette at two different NTP concentrations. The U -less cassette (287 bp) was located immediately before the runoff end. As expected, reduced UTP concentration slowed RNAP down on the upstream sequence but not on the U -less cassette. (b) Distributions of instantaneous velocities on all four N -less cassettes (top panel) and the corresponding upstream of the cassettes (bottom panel) at 1 mM NTPs. For each template, an instantaneous velocity histogram was generated from 30 single molecule traces. The means and the standard errors of the means (in nt/s) are 16.6 ± 0.3 (A -less), 20.7 ± 0.3 (C -less), 23.5 ± 0.4 (G -less), 24.2 ± 0.5 (U -less), and 20.6 ± 0.3 (upstream sequence). For clarity, shown are Gaussian fits to histograms. For the A -less cassette, the original histogram is also shown.

In the second experiment, elongation velocity on a DNA template containing all four nucleotides was measured as the concentration of one of the four NTPs was varied while maintaining the others at 1 mM. This type of experiment was performed for each NTP, resulting in four different sets of data. As expected, the velocity increased with higher [NTP] and approached the same plateau at high [NTP]. However, the exact velocity dependence on the varied [NTP] differed for each NTP. For example, when [CTP]

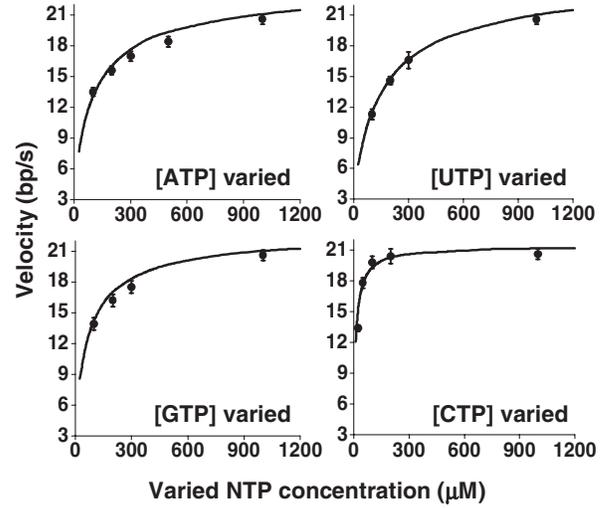
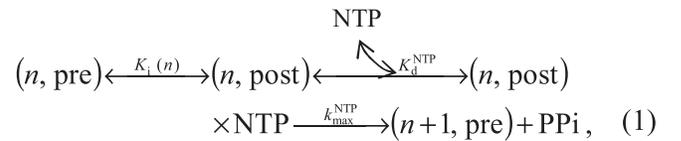


FIG. 2. Elongation velocities at varied A , U , G , and C concentrations, respectively, while maintaining the others at 1 mM. The template contained all four nucleotides. Each data point corresponds to an average of ~ 15 traces from transcription over at least a 400-bp range. The smooth curves are the corresponding fits using the model.

as opposed to [UTP] was varied, the velocity approached a plateau at a much lower concentration. This indicates that the NTP affinity to the active site is also NTP-specific.

We therefore incorporated NTP-specific kinetic parameters into our sequence-dependent elongation model [7]. Briefly, the main reaction pathway is considered as a three-step reaction that includes RNAP translocation, NTP binding, and chemical catalysis resulting in RNA chain elongation by one nucleotide and release of pyrophosphate (PP_i). Both translocation and NTP binding steps are assumed to follow rapid equilibrium [2,5,7,20]:



where (n, pre) and (n, post) represent a transcription complex with a transcript size of n nucleotides in the pre- and post-translocation modes, respectively, K_d^{NTP} is the equilibrium dissociation constant for type-specific NTP binding, and $k_{\text{max}}^{\text{NTP}}$ is the rate of chemical catalysis for that NTP. The equilibrium constant for translocations between (n, pre) and (n, post) is determined solely by their state energies $\Delta G_{(n, \text{pre})}$ and $\Delta G_{(n, \text{post})}$, respectively: $K_i(n) = \exp[(\Delta G_{(n, \text{post})} - \Delta G_{(n, \text{pre})})/k_B T]$, where $k_B T$ is the thermal energy. It is important to note that the free energies are not fit parameters. At each n of a given DNA sequence, they were calculated based on the base pairing energies of the DNA bubble and the DNA-RNA hybrid [7,21,22]. The resulting $\Delta G_{(n, \text{post})} - \Delta G_{(n, \text{pre})}$ varies significantly with sequence.

The overall rate of the three-step reaction in Eq. (1) can be expressed with Michaelis-Menton kinetics [2]:

$$k_{\text{main}}(n) = \frac{k_{\text{max}}^{\text{NTP}}[\text{NTP}]}{K_{\text{d}}^{\text{NTP}}\{1 + K_i(n)\} + [\text{NTP}]} \quad (2)$$

The apparent NTP dissociation constant at a given template location $K_{\text{d,apparent}}^{\text{NTP}} \equiv K_{\text{d}}^{\text{NTP}}\{1 + K_i(n)\}$ is strongly sequence-dependent, consistent with previous findings [23].

To determine the eight model parameters $\{k_{\text{max}}^{\text{NTP}}, K_{\text{d}}^{\text{NTP}}\}$, a global fit of Eq. (2) was performed to data shown in Figs. 1(b) and 2 [16,17]. The data were well fit by the model, and the resulting model parameters from the fit are summarized in Table I. Both $k_{\text{max}}^{\text{NTP}}$ and $K_{\text{d}}^{\text{NTP}}$ differed significantly among the four NTPs. Previous bulk experiments also show that $K_{\text{d,apparent}}^{\text{NTP}}$ is NTP-specific [5,20,23]; however, there have been no comprehensive measurements of $k_{\text{max}}^{\text{NTP}}$ and $K_{\text{d}}^{\text{NTP}}$ for all NTP types. Accurate measurements

TABLE I. NTP-specific parameters.

	ATP	UTP	GTP	CTP
$k_{\text{max}}^{\text{NTP}}$ (s^{-1})	50 ± 6	18 ± 1	36 ± 5	33 ± 6
$K_{\text{d}}^{\text{NTP}}$ (μM)	38 ± 7	24 ± 4	62 ± 18	7 ± 4

of their values using conventional biochemical approaches have been challenging due to the difficulties in separating active elongation from pausing, separating NTP-specific $K_{\text{d}}^{\text{NTP}}$ effects from NTP-specific $k_{\text{max}}^{\text{NTP}}$ contributions, and dealing with the highly sequence-specific nature of the $K_{\text{d,apparent}}^{\text{NTP}}$. Note that the data presented in Figs. 1 and 2 reflect NTP-specific kinetics irrespective of any model.

A simple addition to Eq. (2) will allow it to take into account the effect of a mechanical perturbation on transcription velocity. The application of an external force F on an elongating RNAP tilts the energy landscape between the pre- and post-translocation states and, therefore, alters the elongation velocity:

$$k_{\text{main}}(n) = \frac{k_{\text{max}}^{\text{NTP}}[\text{NTP}]}{K_{\text{d}}^{\text{NTP}}\{1 + \exp[(\Delta G_{(n,\text{post})} - \Delta G_{(n,\text{pre})} - Fd(F))/k_{\text{B}}T]\} + [\text{NTP}]}, \quad (3)$$

where $d(F)$ is extension of 1 bp of DNA at a given force [24] and is ~ 0.34 nm. In this notation, a positive (negative) force assists (resists) elongation, and the effect of force on velocity depends on [NTP]. Given the $k_{\text{max}}^{\text{NTP}}$ and $K_{\text{d}}^{\text{NTP}}$ values in Table I, the effect of force on transcription velocity can be directly predicted without any additional parameters.

To test the predictive ability of Eq. (3), RNAP velocity was measured under a series of constant forces [11] ranging from -10 to $+10$ pN under two different NTP concentrations. The resulting measured force-velocity relations are shown in Fig. 3. Even at 1 mM NTPs, the force-velocity relation shows that there was a statistically significant dependence of velocity on force. This trend is consistent with recent single molecule studies on *E. coli* RNAP [10]. The apparent difference between our results and other previous force-velocity measurements [18,25] is due to different methods of data analysis [17]. The predicted force-velocity relations show good agreement with measurements under both NTP concentrations.

Our model naturally predicts sequence-dependent force-velocity relations. This is demonstrated by a comparison of predicted dwell times at a 20-bp sequence surrounding the $\Delta tR2$ pause sequence with previously measured values [12]. Using the kinetic parameters listed in Table I, the predicted dwell time distributions under both $+4$ and $+15$ pN forces agree well with the corresponding measurements (see Fig. 4).

It is important to note that no other models to date could make such direct force-velocity predictions with model parameters determined solely based on data from chemical perturbations. The ability of this model to do so relied on

the fact that the $\Delta\Delta G \equiv \Delta G_{(n,\text{post})} - \Delta G_{(n,\text{pre})}$ values could be directly computed for a given DNA sequence and, therefore, were not fitting parameters that needed to be determined based on force-velocity relations. For example, on the pRL574 template, $\Delta\Delta G$ has a wide distribution of $+1.6 \pm 1.8k_{\text{B}}T$. The overall positive difference was primarily a result of the loss of one base pairing in the

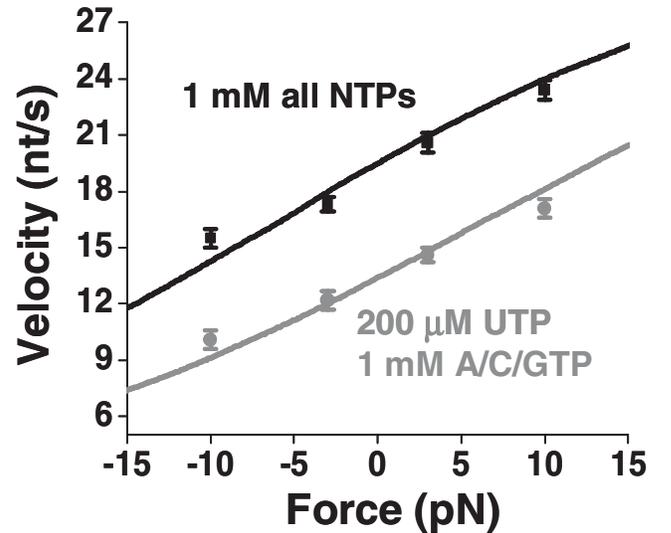


FIG. 3. Force-velocity measurements of transcription elongation and comparison with model predictions. Elongation velocity was measured under various forces at 1 mM all NTPs (black) or 200 μM UTP (arbitrarily chosen) with 1 mM A/C/GTP (gray). Each measured relation (dots with error bars) is compared with a prediction (smooth curve).

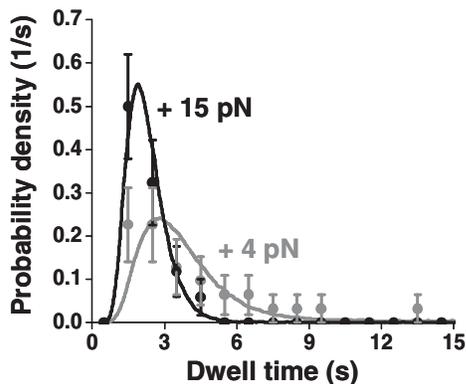


FIG. 4. Experimental (dots with error bars) [12] and predicted (smooth curves) probability density functions of the total dwell times at a 20-bp sequence surrounding the $\Delta tR2$ pause sequence under +4 (gray) and +15 pN (black).

DNA-RNA hybrid in the post-translocation state compared with the pretranslocation state. Estimates of $\Delta\Delta G$ based on fit parameters from a recent study [10] yield values ($0.4\text{--}2.0k_B T$) comparable to those above. Interestingly, studies on T7 RNAP proposed that the post-translocation state is either more stable than the pretranslocation state ($\Delta\Delta G = -1.6k_B T$) [9] or they have similar energies ($\Delta\Delta G \sim 0$) [26]. These differences with our value might be due to different types of RNAP. It would be interesting to see whether our current model could be adapted to other RNAPs.

In this work, the mechanical measurements have provided a stringent test of the mechanochemical coupling of transcription elongation in our model, which is based on a simple thermal ratchet mechanism. Compared with some other models of transcription elongation [5,10,27], this model does not invoke additional NTP binding sites at different translocation states, allosteric NTP binding sites, active/inactive conformational states, etc. The successful predictions of the sequence- [7] and force-dependent kinetics lend strong support to this simple mechanochemical coupling mechanism: In the main reaction pathway, the RNAP rapidly translocates between the pre- and post-translocation states under thermal activation, and on average the RNAP is more likely to be in the pretranslocation state. The binding of NTP stabilizes the post-translocation state, and subsequent NTP incorporation biases the polymerase forward by one base pair, with the binding and incorporation kinetics dependent on the NTP type. Therefore, our work supports a loose-coupling mechanism between chemical catalysis and mechanical translocations during transcription elongation.

We thank members of the Wang lab for critical reading of the manuscript, T.J. Santangelo and C. Simmons for help with RNAP purification, A. Shundrovsky and J. Jin for help with various stages of the project, J. W. Roberts for helpful advice on the DNA template designs, and R. Landick for pRL574. This research was supported by a NSF grant (No. DMR-0517349), a NIH grant (No. R01

GM059849), and the Keck Foundation to M. D. W.

*Corresponding author.

Electronic address: mwang@physics.cornell.edu

- [1] L. Bai, T.J. Santangelo, and M.D. Wang, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 343 (2006).
- [2] R. Guajardo and R. Sousa, *J. Mol. Biol.* **265**, 8 (1997).
- [3] F. Jülicher and R. Bruinsma, *Biophys. J.* **74**, 1169 (1998).
- [4] H. Y. Wang *et al.*, *Biophys. J.* **74**, 1186 (1998).
- [5] J.E. Foster, S.F. Holmes, and D.A. Erie, *Cell* **106**, 243 (2001).
- [6] P.H. von Hippel and Z. Pasman, *Biophys. Chem.* **101–102**, 401 (2002).
- [7] L. Bai, A. Shundrovsky, and M.D. Wang, *J. Mol. Biol.* **344**, 335 (2004).
- [8] G. Bar-Nahum *et al.*, *Cell* **120**, 183 (2005).
- [9] P. Thomen, P.J. Lopez, and F. Heslot, *Phys. Rev. Lett.* **94**, 128102 (2005).
- [10] E.A. Abbondanzieri *et al.*, *Nature (London)* **438**, 460 (2005).
- [11] K. Adelman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13 538 (2002).
- [12] A. Shundrovsky *et al.*, *Biophys. J.* **87**, 3945 (2004).
- [13] N. Komissarova and M. Kashlev, *J. Biol. Chem.* **272**, 15 329 (1997).
- [14] I. Artsimovitch and R. Landick, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7090 (2000).
- [15] J. W. Shaevitz *et al.*, *Nature (London)* **426**, 684 (2003).
- [16] In order to carry out transcription velocity measurements with reasonable accuracy using an optical trap, a small force must be present in order to suppress Brownian motion. Thus, we used +3 pN to approximate a zero-force condition since it produced only a small perturbation ($<10\%$) to the velocity. For accuracy, model parameter values listed in Table I were corrected for this effect, even though values without correction were nearly identical.
- [17] See EPAPS Document No. E-PRLTAO-98-031704 for sequences of the N -less cassettes, global fitting of the NTP-specific parameters, and comparison with previous force-velocity measurements. For more information on EPAPS, see <http://www.aip.org/pubservs/epaps.html>.
- [18] K. C. Neuman *et al.*, *Cell* **115**, 437 (2003).
- [19] S.F. Tolic-Norrelykke *et al.*, *J. Biol. Chem.* **279**, 3292 (2004).
- [20] G. Rhodes and M. J. Chamberlin, *J. Biol. Chem.* **249**, 6675 (1974).
- [21] T.D. Yager and P.H. von Hippel, *Biochemistry* **30**, 1097 (1991).
- [22] The calculated $\Delta G_{(n,\text{post})}$ and $\Delta G_{(n,\text{pre})}$ values depend on the structure of the transcription elongation complex in the model. Here we used the same structure as in Ref. [7]: a DNA bubble size of 12 bp, a DNA-RNA hybrid size of 9 bp, and 1 nt downstream ssDNA.
- [23] J.R. Levin and M.J. Chamberlin, *J. Mol. Biol.* **196**, 61 (1987).
- [24] M.D. Wang *et al.*, *Biophys. J.* **72**, 1335 (1997).
- [25] M.D. Wang *et al.*, *Science* **282**, 902 (1998).
- [26] Q. Guo and R. Sousa, *J. Mol. Biol.* **358**, 241 (2006).
- [27] Y. A. Nedialkov *et al.*, *J. Biol. Chem.* **278**, 18 303 (2003).