A Single-Molecule Technique to Study Sequence-Dependent Transcription Pausing

Alla Shundrovsky,* Thomas J. Santangelo,† Jeffrey W. Roberts,† and Michelle D. Wang*
*Department of Physics, Laboratory of Atomic and Solid State Physics, and †Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York

ABSTRACT We present a technique that allows sequence-dependent analysis of transcription elongation using single-molecule optical trapping techniques. Observation of individual molecules of RNA polymerase (RNAP) allows determination of elongation kinetics that are difficult or impossible to accurately obtain from bulk studies, and provides high temporal resolution of the RNAP motion under a calibrated mechanical load. One limitation of previous single molecule studies was the difficulty in correlating the observed motion of RNAP with its actual position on the DNA template to better than ~100 bp. In this work, we improved the spatial precision of optical trapping studies of transcription to ~5 bp by using runoff transcription as an unambiguous marker of RNAP template position. This runoff method was sufficient to unequivocally locate and study a single known pause sequence (ΔtR2). By applying various loads to assist RNAP forward translocation, we specifically investigated elongation kinetics within this pause region and found that the dwell time at the pause sequence decreased with increasing assisting load. This observation is consistent with bulk biochemical studies that suggest RNAP reverse translocates, or "backtracks," at the ΔtR2 pause sequence.

INTRODUCTION

Transcription elongation by RNA polymerase (RNAP) is neither uniform nor continuous. Specific DNA sequences, called pause sites, temporarily halt the progress of RNAP. Pausing is thought to occur because of misalignment of the RNA 3′ end with the RNAP active site due to RNAP backtracking or RNA hairpin formation (Komissarova and Kashlev, 1997a; Artsimovitch and Landick, 2000). Numerous pause sequences have been shown, or are suspected, to provide regulatory functions such as allowing transcription factors to bind and thereby modify gene expression (Uptain et al., 1997; Roberts et al., 1998). Other pause sequences that have been detected in vitro have no known biological function, but nonetheless reflect the intrinsic sequence-dependence of RNAP motion.

Experimental study of the mechanism of pausing is nontrivial, partially due to the difficulty in measuring kinetics during continuous elongation. Traditional bulk experiments are only capable of detecting the overall elongation behavior of a large population of molecules that may not be homogeneous. In many bulk transcription experiments, the RNAP population is first halted at a specific template position via nucleotide starvation, and the kinetics at a subsequent pause sequence are observed after elongation has been restarted by the addition of nucleotides. However, after transcription restart the RNAP population becomes asynchronous very rapidly, and thus different RNAP molecules arrive at the pause sequence at different times. Furthermore, although nucleotide-starved transcription complexes do not allow further nucleotide incorporation, RNAP may slide backward along the DNA template (Komissarova and Kashlev, 1997b; Nudler et al., 1997), resulting in an RNAP population that will resume elongation from different translocation states upon nucleotide addition. The combination of spatial variations in the starting population and asynchrony during elongation results in a spatially and temporally heterogeneous population, making it difficult to precisely determine the kinetics at a distant pause sequence.

Observation of individual molecules of RNAP offers the possibility to probe the kinetics of transcription pausing without the complications of population heterogeneity. If motions of each RNAP molecule along the DNA template could be monitored at any given sequence, the pause duration could then be directly obtained from single-molecule data, avoiding the necessity of synchronizing the RNAP population. Sequence-dependent pause kinetics could then be obtained by polling elongation data from many RNAP molecules. To accurately determine the pause kinetics and draw statistically meaningful conclusions, a large dataset of individual single molecule traces must be acquired, and this can be time-consuming. However, current and future attempts to automate data acquisition will likely broaden the feasibility of using such an approach.

In addition, single-molecule mechanical techniques make it possible to probe the mechanism of pausing. The application of an external load to the transcription elongation complex (TEC), to either assist or hinder transcription, may bias the translocation motion of RNAP. This may alter the measured elongation kinetics, especially for transcription pauses where translocation is rate-limiting, for example when pauses are caused by RNAP noncatalytically backtracking (i.e., reverse translocating) along the DNA template.

Submitted April 4, 2004, and accepted for publication September 1, 2004.

Address reprint requests to Michelle D. Wang, E-mail: mdw17@cornell.edu.

0006-3495/04/12/3945/09 $2.00

doi: 10.1529/biophysj.104.044081

To probe sequence-dependent transcription kinetics, single-molecule techniques must have sufficient precision to locate individual pause sequences encountered by the RNAP during elongation. This has not been attainable in previous single-molecule mechanical studies. Although it is well known that these techniques can resolve a displacement of RNAP to a few bps or better, the precision of the location of RNAP on the DNA template during elongation so far has been limited to ~100 bp (Yin et al., 1995; Wang et al., 1998; Davenport et al., 2000; Adelman et al., 2002; Forde et al., 2002; Neuman et al., 2003; Shaevitz et al., 2003). Within this measurement uncertainty, multiple pause sequences may exist, and therefore transcription pausing has been examined without specific reference to the corresponding DNA sequences (Adelman et al., 2002; Forde et al., 2002; Neuman et al., 2003; Shaevitz et al., 2003). The lack of sequence-dependent information inevitably complicates the interpretation of transcription-pausing data, since different types of pauses could exhibit kinetics that respond differently to chemical and mechanical perturbations. For example, an assisting load is expected to discourage RNAP entry into and dwell within a backtracked pause, whereas the same force is expected to promote possible pausing due to hypertranslocation. The elucidation of the mechanism of pausing therefore requires accurate location of the transcribed pause sequence.

In this work, we significantly improved the precision of single-molecule optical trapping studies of transcription so that kinetics of an individual pause sequence could be probed. This improvement was achieved by using runoff transcription as a well-defined position marker for alignment of the measured RNAP position on DNA template. The precision of this method was determined by using a well-defined pause sequence ΔtR2 (Yarnell and Roberts, 1999; Ryder and Roberts, 2003). We found that pausing within the ΔtR2 sequence positioned near the end of the DNA template could be accurately located with a ~5 bp precision under our experimental conditions, representing a ~20-fold enhancement over previous methods of RNAP position detection. We further studied pausing within the ΔtR2 sequence and demonstrated that the pausing kinetics were significantly altered by the application of an external load. Our results show a force-dependent behavior that is consistent with possible backtracking within this sequence.

MATERIALS AND METHODS

DNA templates

Three DNA templates were constructed for the single molecule experiments (Fig. 1 *A*). The three single-molecule templates were prepared by PCR with the forward primer biotinylated to provide a single biotin tag ~2 kbp upstream of a T7A1 promoter and different reverse primers. Following the T7A1 promoter, each DNA template contained a 1.1 kbp fragment of the *Escherichia coli rpoB* gene (derived from pRL574; Schafer et al., 1991), a 51 bp T-less region (region I), a known pause sequence ΔtR2, a second T-less region (region II), and, depending on the template, a length of DNA near the runoff end. The ΔtR2 pause sequence (Yarnell and Roberts, 1999; Ryder and Roberts, 2003) was constructed from the intrinsic tR2 terminator by disrupting the upstream half of its hairpin to prevent termination while leaving the T-rich pause-inducing sequence intact. The design of the templates sandwiched the ΔtR2 pause sequence between two T-less regions to reduce pausing immediately before and after the ΔtR2 pause sequence under low UTP concentration. Template 1 contained a 30 bp of the T-less region II and no end region, so that the ΔtR2 pause sequence was located 30 bp from the runoff end. In templates 2 and 3, the T-less region II was 53 bp long and the ΔtR2 pause sequence was located 105 bp and 226 bp from the runoff end, respectively.

Experimental configuration

Transcription was initiated by incubating 25 nM *E. coli* RNAP bearing a hemagglutinin (HA) epitope tag on the C-terminus of its α-subunit, 5 nM DNA template containing a T7A1 promoter, 250 μM ApU initiating dinucleotide, and 50 μM ATP/CTP/GTP in transcription buffer (25 mM Tris·Cl, pH 8.0, 100 mM KCl, 4 mM $MgCl_2$, 1 mM DTT, 3% (v/v) glycerol, 0.15 mg/mL acetylated BSA) for 20 min at 37°C. Stalled TECs prepared in this way contained the RNAP, DNA and a 20 nt nascent RNA and were then attached to an anti-HA antibody-coated glass coverslip surface as previously described (Adelman et al., 2002) (Fig. 1 *B*). The upstream end of the DNA template was attached via biotin/streptavidin linkage to a 0.5 μm polystyrene microsphere held by an optical trap. This configuration allows application of loads in the direction assisting RNAP motion during elongation. Transcription was resumed by flowing in a solution containing all four NTPs (1 mM ATP, CTP, and GTP, 50 μM UTP at 23°C) in transcription buffer plus 0.2 mg/mL heparin; the concentration of UTP was lowered to a subsaturating level to increase the probability of pausing at the ΔtR2 sequence.

A computer-controlled feedback loop was used to apply a constant force to the TEC. This force clamp mode was achieved by using a 1-D piezoelectric stage (Physik Instrumente, Waldbronn, Germany) to modulate the position of the coverslip relative to the trapped microsphere, which was effectively held at a fixed position relative to the trap center (Brower-Toland et al., 2002). The position of the trapped microsphere relative to the trap center and the trapping force was determined by measuring the deflection and power of transmitted laser light using a quadrant photodiode detector (Hamamatsu, Bridgewater, NJ). The detector signals, as well as the piezo stage sensor were low-pass filtered at 5 kHz, digitized at 13 kHz using a 16-bit DAQ board (National Instruments, Austin, TX), and averaged to 130 Hz. The data were then smoothed using a Gaussian weight function with a standard deviation of 1.0 s to generate instantaneous velocity (Adelman et al., 2002). This Gaussian low-pass filtering was performed to reduce the noise in the data but also resulted in smearing short pauses with durations <1 s. The smoothed data of RNAP template position versus time were also used to calculate dwell time versus template position by summing the total time that the RNAP spent at a given bp.

Five different assisting forces were used to determine the force-dependence of RNAP motion: 4, 6, 8, 12, and 15 pN, corresponding to 31, 26, 27, 28, and 34 single molecule traces, respectively. Different forces were produced by setting the relative distance between the microsphere and the trap center to preset fixed positions, while keeping the laser intensity (trap stiffness) constant. RNAP molecules that paused for >60 s anywhere on the template were considered arrested and their traces were excluded from further data analysis.

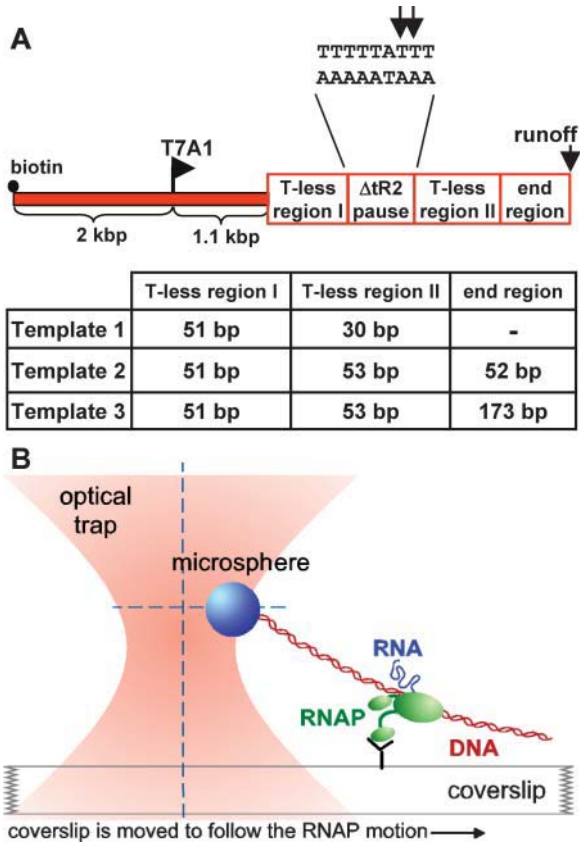
Determination of DNA tether length

To correctly analyze transcription data, it was necessary to determine the number of base pairs between the RNAP and the trapped microsphere, i.e., the DNA tether length in bp (see Fig. 1 *B*). This was achieved using an approach adapted from Wang et al. (1997, 1998) and is summarized below.

First, both the force and extension of the DNA tether between the RNAP and the microsphere needed to be obtained according to the experimental geometry (Fig. 1 *B*). All our experiments were performed with the microscope objective focused onto the surface of the coverslip. At this focus, the trap center was designed to locate ~650 nm above the surface and this trap height was constant for all experiments described in this work. Before the start of a transcription measurement on a given tether, the lateral location of the RNAP relative to the trap center was determined as follows. The tethered microsphere was manually positioned to the trap center laterally and the DNA tether was then stretched by moving the stage piezoelectrically along one horizontal axis in both directions. The symmetry point of the resultant position detector signal versus piezo position curve corresponded to the piezo stage position at which the RNAP was located directly below the trap center. Thus, in subsequent experiments the horizontal location of the RNAP relative to the trap center could be determined via the piezo position.

During a transcription experiment, the tethered microsphere was displaced both vertically and horizontally from the trap center due to the force exerted by the RNAP. Since only the horizontal displacement and force could be detected, the vertical displacement and force were computed based on balancing the forces on the microsphere while satisfying the geometric constraints. The vertical displacement from the trap center in a typical transcription experiment was ~100 nm. This ensured that the microsphere did not touch the surface of the glass coverslip. The net force and extension of the DNA tether were then computed by performing a vector sum of their components along the horizontal and vertical directions.

Second, the force and extension were converted to DNA tether length (in bp) by using a modified Marko-Siggia worm-like-chain model of DNA elasticity with the following DNA elasticity parameters: 0.338 nm contour length per base pair, 41.0 nm persistence length, and 1277 pN stretch modulus. These values were determined by stretching DNA of known sizes.



	T-less region I	T-less region II	end region
Template 1	51 bp	30 bp	-
Template 2	51 bp	53 bp	52 bp
Template 3	51 bp	53 bp	173 bp

FIGURE 1 Experimental configuration. (*A*) Schematic of the three DNA templates used in the single molecule experiments. The upstream end of each DNA contained a single biotin tag for attachment to a streptavidin-coated microsphere. Transcription was initiated on the T7A1 promoter and continued until the RNAP reached the runoff end of the DNA. Each template contained a known T-rich pause sequence ΔtR2 (intrinsic terminator tR2 with its hairpin disrupted to prevent termination) flanked by two T-less cassettes. The two major pause positions within the ΔtR2 sequence are indicated by arrows. The major difference in the three DNA templates was the location of the ΔtR2 sequence relative to the runoff end. (*B*) A stalled TEC was specifically attached to a coverslip surface via interaction between HA-tagged RNAP and an anti-HA antibody nonspecifically adsorbed to the surface. The upstream end of the DNA molecule was attached via biotin/streptavidin linkage to a polystyrene microsphere held by the optical trap. The coverslip was mounted on a piezoelectric stage and constant assisting force was maintained on the transcribing RNAP by modulating the coverslip position relative to the trapped microsphere. When the RNAP reached the end of the DNA template, it dissociated and produced a characteristic runoff signal (see text).

RESULTS AND DISCUSSION

RNAP position uncertainty in conventional optical trapping transcription experiments

Fig. 1 *B* shows our experimental configuration to monitor the motions of individual RNAP molecules under constant applied load using an optical trap. The motion of an RNAP was determined by measuring the time course of its DNA tether length between the RNAP and the trapped microsphere, since this tether length was directly related to the position of the RNAP on the DNA template. The RNAP template position is defined here as the distance (in bp) that the RNAP has moved from the transcription start site. This is identical to RNA transcript size when the RNAP does not noncatalytically backtrack or forward translocate along the DNA template. Simultaneous detection of force (in pN) and extension (in nm) of the DNA tether allowed computation of tether length (in bp) using a modified Marko-Siggia worm-like-chain model of DNA elasticity (Wang et al., 1997; also see Materials and Methods). The RNAP template position was obtained from the measured tether length by subtracting the known length of the DNA between the trapped microsphere and the transcription start site (Wang et al., 1998). The precision of the tether length measurement thus determined the precision of locating specific sequences transcribed by the RNAP during elongation.

We performed the following experiments to investigate the precision of the RNAP template position determination under our experimental configuration. These experiments used TECs stalled after transcribing the first 20 bp as described in Materials and Methods, each with 1958 bp of DNA between the trapped microsphere and the transcription start site. Each experiment consisted of two basic steps: 1), the microscope objective was focused onto the surface of the glass coverslip. The optical trap was also centered laterally on a microsphere tethered via a stalled RNAP (see Fig. 1 *B*).

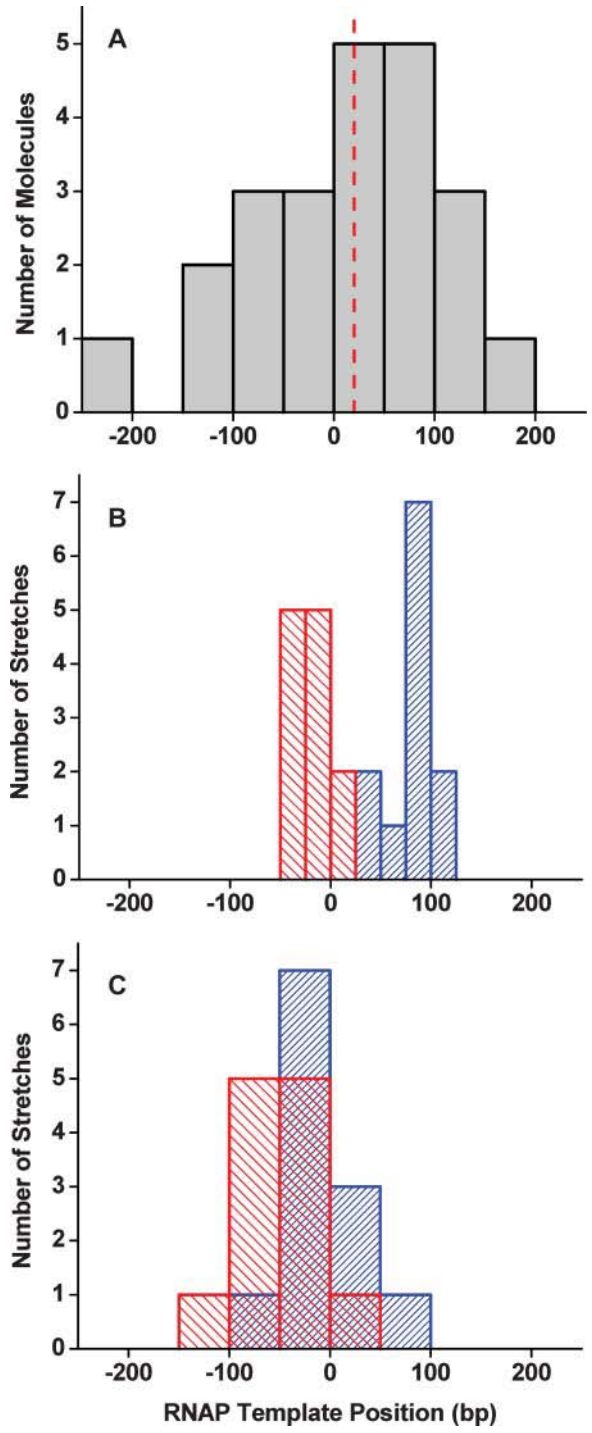


FIGURE 2 RNAP template position uncertainty. Distributions of RNAP template positions were measured for stalled transcription complexes. The expected RNAP template position of the stalled RNAP was +20 bp (counted from the transcription start site) and is indicated by a red dashed line. (*A*) A histogram of RNAP template position from measurements of 23 individual tethers. The measured template position was +6 ± 94 bp (mean ± SD). (*B*) Histograms of RNAP template position, each from multiple measurements of a single DNA tether. Data from two different tethers are shown in different colors. After initial objective focusing and optical trap positioning, each DNA tether was repetitively stretched twelve times without refocusing or repositioning. The histograms show RNAP template positions of −16 ± 19 bp and +78 ± 24 bp (mean ± SD) (*red* and *blue histograms*, respectively). (*C*) Histograms of RNAP template position, each from multiple measurements of a single DNA tether. Data from two different tethers are shown in different colors. Each DNA tether was repetitively stretched 12 times with refocusing and repositioning preceding each stretch. The histograms show RNAP template positions of −51 ± 42 bp and −6 ± 44 bp (mean ± SD) (*red* and *blue histograms*, respectively).

This step was performed manually. 2), Subsequently, the tether was stretched by moving the stage piezoelectrically along one horizontal axis and the force and extension from this measurement were used to determine the DNA tether length and thereby the RNAP template position (see Materials and Methods).

Fig. 2 *A* shows the distribution of the measured RNAP template positions obtained by pooling data from a number of DNA tethers. This distribution has a mean of +6 bp, which is close to the expected stalled RNAP template position of +20 bp. Its standard deviation of 94 bp gives the best case estimate of our precision in determining the time-dependent RNAP position on the DNA template, which is comparable to those obtained in previous similar single-molecule experiments (Davenport et al., 2000; Adelman et al., 2002; Forde et al., 2002; Neuman et al., 2003; Shaevitz et al., 2003). Thus the measured mean RNAP template position was consistent with the expected value, but a rather large uncertainty exists from tether to tether.

We performed two types of experiments to further determine the sources of this RNAP template position uncertainty. In the first type of experiment, a single DNA tether was stretched repetitively without manual objective refocusing and optical trap recentering (i.e., without step 1 above). Fig. 2 *B* shows examples of RNAP template position distributions from this type of measurement, each obtained from a single DNA tether. These distributions have a standard deviation of ~22 bp, which must be due to uncertainties introduced by the optical trap, the detection mechanism, and the piezo stage (but not objective focusing or optical trap centering). In the second type of experiment, a single DNA tether was again stretched repetitively, but each time, step 1 was repeated. Fig. 2 *C* shows examples of RNAP template position distributions from this type of measurement, each obtained from a single DNA tether. These distributions have a standard deviation of ~43 bp, which must also include uncertainties introduced by manual objective focusing and optical trap centering. Under the assumption that uncertainties introduced by various sources are independent, these measurements indicate that uncertainties introduced by objective focusing and optical trap centering should be ~37 bp, and uncertainties introduced by sources other than those mentioned above should be ~83 bp. Thus these other sources of uncertainty are the dominant ones. Indeed, this conclusion is in good agreement with a recent transcription study that employed automated three-dimensional focusing

and positioning and obtained absolute position uncertainty of ~75 bp (Neuman et al., 2003).

In our experiments, it is unlikely that the dominant uncertainties came from variations of DNA template sizes, since a single well-defined length PCR product was observed by gel electrophoresis (data not shown). It is also unlikely that they were due to extensive RNAP backtracking, which can result in complexes becoming arrested and unable to resume elongation unaided (Komissarova and Kashlev, 1997b). Since we obtained a distribution similar to that of Fig. 2 *A* from TECs that continued transcription after the addition of NTPs (data not shown), this argues against stalled complexes backtracking for more than a few bp. Possible sources of this uncertainty may include variations in microsphere size, nonspecific sticking of the DNA to the trapped microsphere, and spontaneous nicks in the DNA tether. Regardless of the source, the presence of this large uncertainty in the RNAP template position determination makes it impossible to accurately correlate the RNAP motion with specific transcribed DNA sequences.

The runoff method and its precision

We developed a runoff method to improve the precision of RNAP template position determination in single molecule experiments. We used runoff transcription as a well-defined marker to relate the measured final length of the DNA tether to the known RNAP runoff position on the DNA template, thus circumventing the large uncertainty of tether length measurement. In this runoff method, elongating RNAP is monitored until it reaches the end of the DNA template and dissociates, producing an immediate and characteristic jump in the apparent RNAP position (the runoff signal). Although RNAP is highly processive, spontaneous dissociation of the TEC did rarely occur (~1% per 100 bp). Thus only when the dissociation signal occurred near the runoff end to within the uncertainty of the RNAP template position measurement was it considered a genuine runoff signal and therefore served as an alignment marker for the preceding data. The detected RNAP template position was subsequently converted to the actual template position by adding a constant offset to align it with the runoff end.

To determine the precision of the runoff method, a specific sequence was engineered to contain a known pause sequence (ΔtR2) flanked by two T-less regions. Three DNA templates were used that contained the ΔtR2 sequence located at three different distances relative to the runoff end (Fig. 1 *A*; see also Materials and Methods). Under our experimental conditions (1 mM ATP, CTP, and GTP; 50 μM UTP) the RNAP is expected to move at optimal rate through the two T-less regions, which serve to accentuate the pause signal.

For all templates, the major pause positions within the ΔtR2 sequence should occur at two adjacent sites at positions +1182 and +1183, based on our bulk experiments (data not shown), consistent with previous studies (Gusarov and Nudler, 1999). Additionally, other T's within the pause sequence should also result in some pausing due to the lowered UTP concentration. Therefore, we expected that pausing within the ΔtR2 sequence should occur in the position range of +1176 to +1184 bp.

An example of data obtained from template 1 and analyzed using the runoff method is shown in Fig. 3. The RNAP template position versus time curve (Fig. 3 *A*) was smoothed, and the corresponding instantaneous velocity versus time curve was computed (Adelman et al., 2002; also see Materials and Methods). These data were then used to generate both the dwell time and instantaneous velocity versus RNAP template position curves (Fig. 3, *B* and *C*). As shown in Fig. 3,

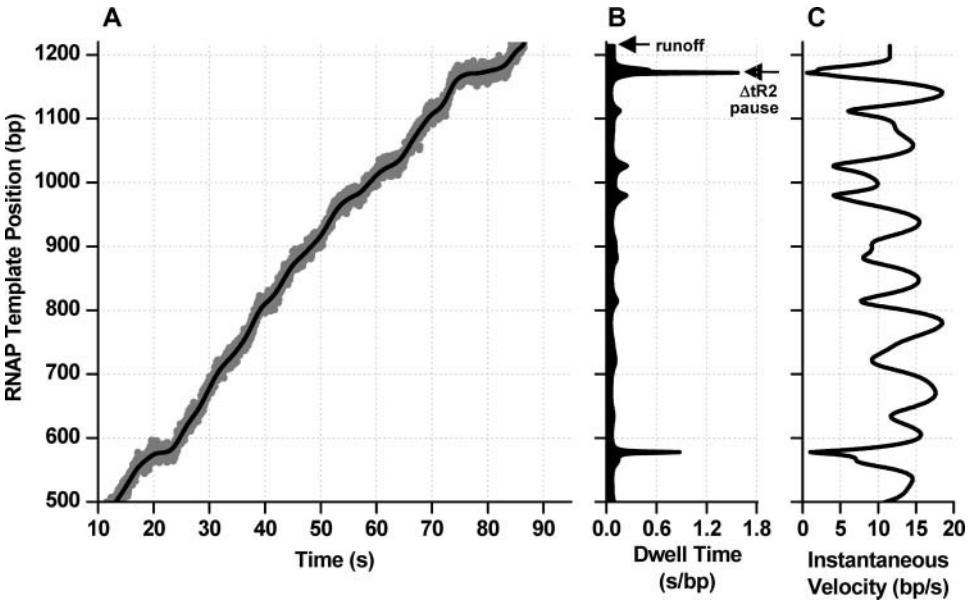


FIGURE 3 Analysis of movement of a single RNAP molecule. (*A*) Time course of the motion of a single molecule of RNAP taken on template 1 under 4 pN of assisting force and aligned using the runoff method. Both raw data (*shaded area*) and filtered data (*black line*) are shown. (*B*) RNAP dwell time at each template position over the corresponding template range. (*C*) Corresponding instantaneous velocity plotted as a function of template position.

steady RNAP molecule motion was interrupted by transcription pauses, which are reflected as flat steps in Fig. 3 *A*, distinct peaks in the corresponding dwell time in Fig. 3 *B*, and drops in the instantaneous velocity curve in Fig. 3 *C*. As expected, we observed pausing around the predicted positions of the ΔtR2 pause sequence.

Dwell time data such as those shown in Fig. 3 *B* were compared with the expected ΔtR2 pause position range. Additional examples are shown in Fig. 4, with each curve obtained from a single RNAP molecule. In the majority of these curves, a single pause peak located near the expected ΔtR2 pause sequence was flanked by two regions of low dwell time (see Fig. 4, *A–E*, as examples) corresponding to transcription through the two T-less regions. The pause peaks had an average standard deviation of ~3 bp. Some curves did not show any detectable dwell time peak within the expected sequence (see Fig. 4 *F* as an example).

We determined the precision of locating the ΔtR2 pause sequence using the runoff method by plotting a histogram of the pause dwell time peak positions (for an example, see Fig. 5 *A*). RNAP molecules that did not pause for at least 1 s did not produce detectable dwell time peaks (see Materials and Methods) and were excluded from this analysis. This histogram shows that the pausing was centered at template position 1179 ± 5 bp (mean ± SD). The mean is entirely consistent with the expected pausing range of the ΔtR2 pause sequence, proving that our method can be used to accurately locate individual pausing sequences near the runoff end. The distribution was broadened by other factors in addition to the intrinsic measurement uncertainty and the possibility that for a given transcription size RNAP may backtrack or forward track. For example, RNAP does not pause at a unique template position within the 9 bp ΔtR2 pause sequence as discussed above. Thus the 5 bp uncertainty for sequences near the runoff end is a very conservative estimate for the precision of the runoff method.

It is worth noting that there are two major configurations in single molecule mechanical measurements, each with different considerations for positional precision.

In one configuration, only a small internal change in the molecular structure needs to be determined. In this case, the molecule of interest is suspended between a surface and a microsphere. Some examples include titin unfolding (Tskhovrebova et al., 1997), RNA unfolding (Liphardt et al., 2001),

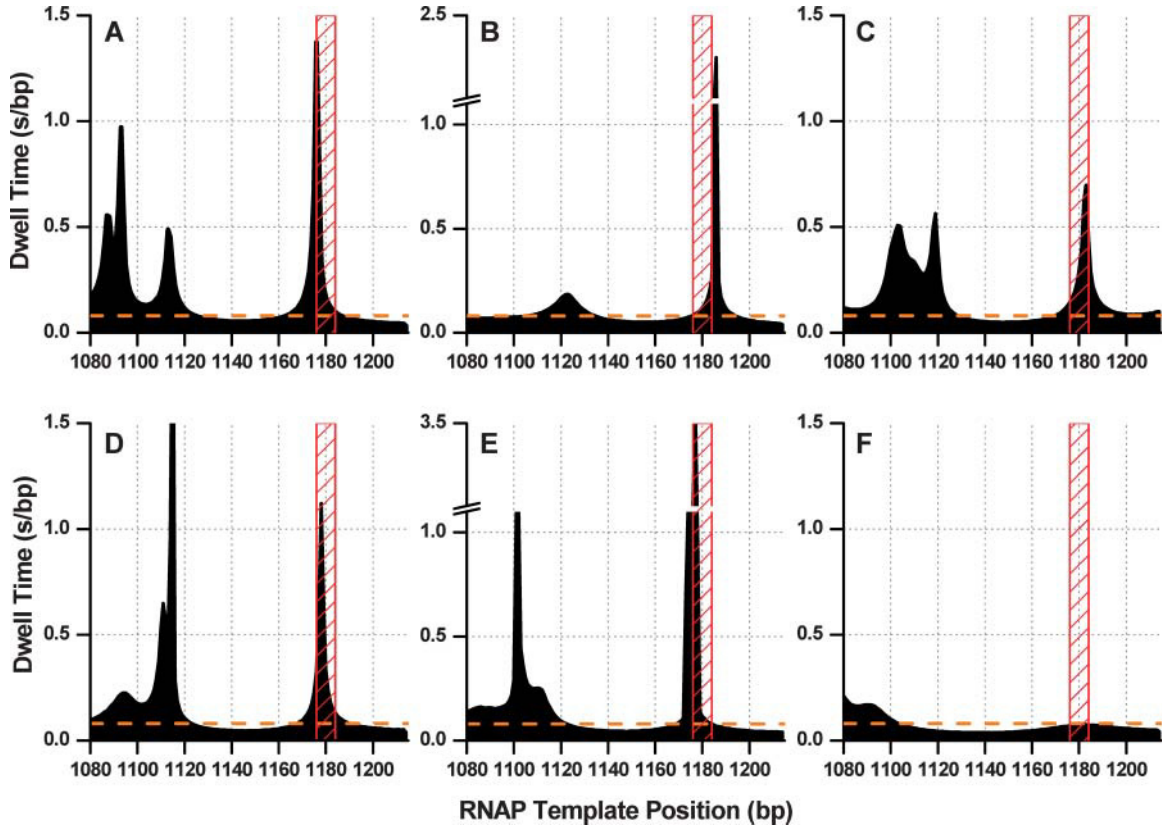


FIGURE 4 RNAP dwell time near the template end. Examples of RNAP dwell time as a function of template position after alignment using the runoff method. Data were taken on DNA template 1 under 4 pN of assisting force and only data near the runoff end are shown. The expected location of the ΔtR2 pause sequence is marked by a red bar in each graph. The most likely dwell time excluding the last 100 bp near the runoff end was computed by averaging dwell time data from different RNAP molecules. Its value, 0.08 s/bp, is indicated as a horizontal orange dashed line in each graph to serve as a reference. Notice that for most of the traces the long dwell time near the expected pause range is flanked by regions of faster than average elongation.

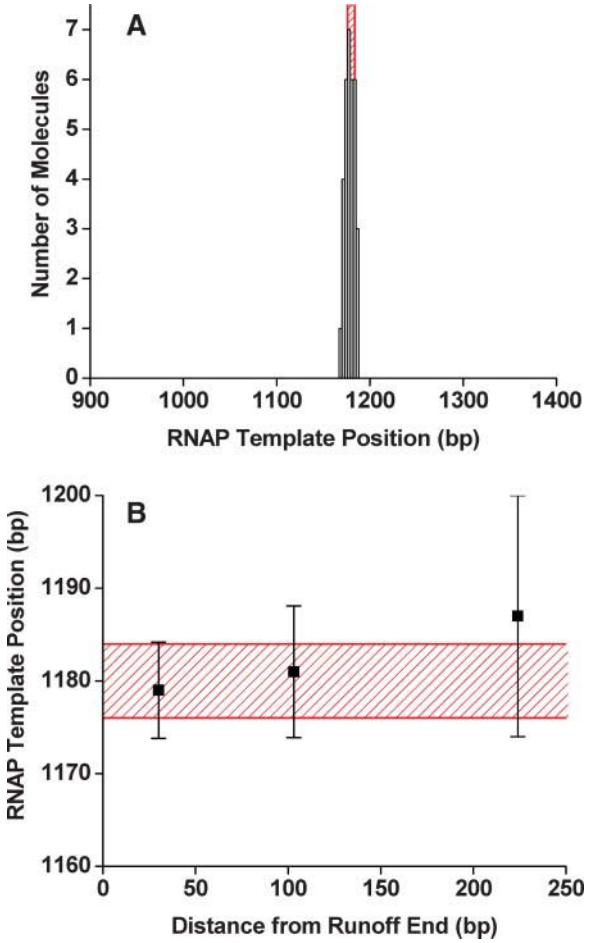


FIGURE 5 Precision of the runoff method based on the ΔtR2 pause position. (*A*) The RNAP template positions of the dwell time peaks within the ΔtR2 sequence from different molecules on template 1 (see Fig. 4) were used to generate this histogram with 3-bp bin size. Only those traces with dwell times long enough to produce peaks greater than twice the most likely dwell time (0.16 s/bp threshold) were included for analysis (21 out of 31 molecules). The expected location of the ΔtR2 pause sequence is marked by a red bar. Note that the range of the horizontal axis for this histogram (500 bp) is identical to that of Fig. 2. (*B*) Precision of the runoff method as a function of distance from the template end. Mean ΔtR2 pause positions are plotted for single molecule DNA templates of three different lengths. The vertical error bars correspond to the standard deviations of the position histograms. Thus a mean represents the accuracy and a standard deviation represents the precision of the RNAP template position determination using the runoff method. The number of measurements is 21, 16, and 17 for templates 1, 2, and 3, respectively.

and single nucleosome disruption (Brower-Toland et al., 2002). The mechanical force is applied directly to the molecule of interest and a small change in length, rather than the absolute length, is relevant. Since a change in tether length is subject to roughly the same percent uncertainty as the total tether length, a typical change of tens of nm has an uncertainty on the order of a few nm. This precision is normally sufficient to draw important conclusions regarding molecular mechanisms.

In the other configuration, the absolute location of a molecule of interest relative to its track molecule needs to be determined. This is the case for transcription experiments where the location of the RNAP relative to its DNA template is determined by measuring the DNA tether length. Although the percent uncertainty in the tether length might be rather small (e.g., a few percent), the resulting uncertainty in the absolute location of the RNAP on the DNA template is typically ~100 base pairs (~30 nm in DNA contour length) due to the relatively long DNA tethers used to optimize experimental geometry (typically over 1 μm in contour length). In essence, the runoff method achieves its high precision by measuring a small change in tether length from a well defined reference point (the runoff end) and in this way is similar to the configuration described in the previous paragraph.

Precision versus distance from the runoff end

To test the general applicability of the runoff method for pause sequences located further away from the runoff end, we used two additional single molecule templates containing 105 bp (template 2) and 226 bp (template 3) after the ΔtR2 pause sequence. Fig. 5 *B* shows that each mean measured pause position is consistent with the expected pausing range of the ΔtR2 sequence within its standard error of the mean (not shown, but can be readily computed based on the standard deviation and the number of measurements). The precision, which is represented by the standard deviation in the figure, decreased with the distance of the pause sequence from the runoff end.

In principle, the precision of the runoff method should be limited only by the Brownian motion of the tethered microsphere (~1 bp at 1 Hz bandwidth). However, in our and other experimental configurations, low-frequency drift of the instrument typically contributes significantly to the position uncertainty (Adelman et al., 2002; Forde et al., 2002; Neuman et al., 2003). This drift was observed to be ~1 bp/s and bidirectional in our configuration. Thus, when the runoff method was used to locate the pause position, the cumulative errors increased with the distance from the runoff end largely due to the increased time required for the RNAP to reach the runoff. Consequently, the derived RNAP template position is expected to be most precise near the runoff end, with the precision decreasing for distant sequences, likely due to instrument drift. We observed that the precision of the measured ΔtR2 pause position for template 1 was 5 bp, which was comparable to the expected pausing range of the ΔtR2 sequence. The precision decreased to 7 bp and 12 bp for templates 2 and 3, respectively.

In summary, the runoff method precision near the runoff end represents a minimum of ~20-fold enhancement over previous approaches for measuring sequence-specific locations of RNAP. Although the precision of the method decreases with distance from the runoff end, the decrease is

not dramatic, which allows for considerable flexibility in the placement of a sequence of interest.

Probing the mechanism of pausing within the ΔtR2 sequence

Previous bulk studies have shown that RNAP can backtrack along the DNA template without changing the nascent RNA size, and backtracking has been suggested to be primarily responsible for pausing within T-rich sequences where the increased dwell time likely reflects the additional time it takes for the TEC to return to the active configuration (Nudler et al., 1997; Komissarova and Kashlev, 1997a). A number of previous studies (Komissarova and Kashlev, 1997a; Nudler et al., 1997; Gusarov and Nudler, 1999) as well as our bulk experiments (data not shown) showed that the ΔtR2 sequence caused RNAP to backtrack at positions +1182 and +1183 if stalled at the pause site by nucleotide starvation. Furthermore, RNAP dwell time within the ΔtR2 region was sensitive to GreB (data not shown). Transcription factor GreB assists backtracked complexes to resume active elongation by stimulating transcript cleavage at the RNAP active site (Borukhov et al., 1993, 2001; Opalka et al., 2003). The observed reduction in ΔtR2 pause duration in the presence of GreB suggests that polymerase tends to transiently backtrack within the pause region.

The runoff method allowed precise and unambiguous location of the ΔtR2 pause sequence and made it possible to study its pausing kinetics. Using the runoff method, we investigated whether it was possible to probe if RNAP backtracks within the ΔtR2 pause sequence by examining the load-dependence of pausing. An assisting force applied to the RNAP is expected to prevent it from back-translocation as well as to reduce the time spent in backtracked states. Therefore, the dwell time at the ΔtR2 sequence is predicted to decrease with increasing assisting force.

Fig. 6 shows the load dependence of RNAP dwell time at the ΔtR2 sequence. To ensure that all pausing within the sequence was included in the analysis, the RNAP dwell time was summed over a 20 bp window centered about the average pause position +1179 bp. This window size was large enough to include the 5 bp RNAP template position uncertainty of the runoff method as well as the ~3 bp standard deviation of individual pauses. Comparison of the dwell time histogram at 4 and 15 pN assisting loads shows that a larger assisting force shifted the dwell time distributions toward shorter times (Fig. 6 *A*). Also, long-lived (>5 s) pauses that were present in the 4 pN distribution disappeared when the force was increased to 15 pN. The dwell time within the ΔtR2 sequence exhibits a much stronger load-dependence compared with that of active elongation (Fig. 6 *B*). These results show that the RNAP dwell time within the ΔtR2 sequence could be significantly reduced by an applied assisting load. This observation is consistent with RNAP backtracking within this sequence.

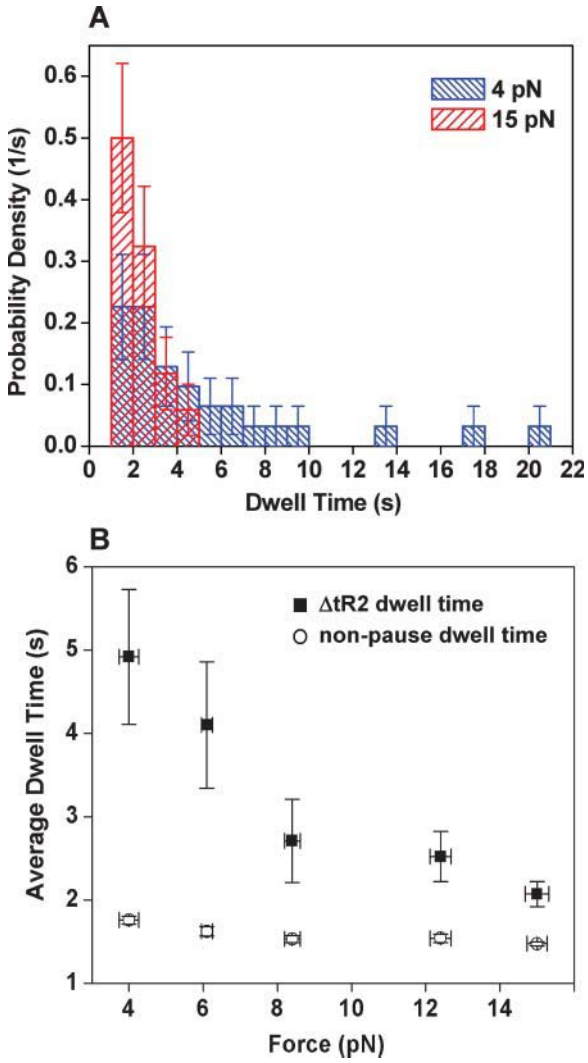


FIGURE 6 Load-dependence of pausing within the ΔtR2 sequence. (*A*) Pause kinetics of the ΔtR2 pause sequence under 4 pN (*blue*) and 15 pN (*red*) of assisting loads. The dwell time within the ΔtR2 region was measured by computing the total dwell time within a 20 bp window centered at the average pause template position of +1179 bp (see text). The dwell times from all single molecule traces were pooled and the resulting probability density function was normalized. (*B*) Average dwell time within the ΔtR2 region as a function of assisting force (■) with error bars representing standard error of the mean. For comparison, we also measured the average time it took for RNAP to actively elongate 20 bp without pausing (○), by dividing 20 bp by the mean active (nonpause) elongation velocity (in bp/s) at a given load.

CONCLUSIONS

The method we have developed adds the capability of sequence-dependent studies to single molecule transcription experiments. Using a known position marker (the runoff end) makes it possible to locate preceding transcribed sequences with high accuracy and precision. The method is of general applicability because any sequence of interest can readily be engineered into a template close to a runoff end

through conventional molecular biological techniques. For instance, a single RNAP molecule transcribing through multiple successive pause sequences can be studied simply by creating DNA templates containing several pause sequences 10–20 bp apart placed reasonably close to the runoff end. For distances far away from the runoff end precision can be improved further by minimizing instrument drift, for example by removing the TEC attachment to the sample chamber (Shaevitz et al., 2003).

The runoff method is not restricted to single molecule studies of RNAP and should find broader applications in the study of other DNA-based motors, such as DNA polymerases, helicases, exonucleases, etc., whose motions are likely or known to be sequence-dependent.

We are grateful to L. Bai and Dr. R. M. Fulbright for helpful discussions and critical comments on the manuscript. We thank Dr. S. J. Koch and R. C. Yeh for participation in the construction of the optical trapping setup, Dr. K. Adelman for participation in the purification of HA-tagged *E. coli* RNAP, and Dr. A. La Porta for helpful discussions on data analysis. We also thank Dr. R. Landick for the gift of pRL574 plasmid.

M.D.W. has been supported by grants from the National Institutes of Health, the Beckman Young Investigator Award, the Alfred P. Sloan Research Fellow Award, and the Keck Foundation's Distinguished Young Scholar Award. A.S. has also been supported by Cornell University Molecular Biophysics Training Grant from National Institutes of Health.

REFERENCES

Adelman, K., A. La Porta, T. J. Santangelo, J. T. Lis, J. W. Roberts, and M. D. Wang. 2002. Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc. Natl. Acad. Sci. USA.* 99:13538–13543.

Artsimovitch, I., and R. Landick. 2000. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. USA.* 97:7090–7095.

Borukhov, S., V. Sagitov, and A. Goldfarb. 1993. Transcript cleavage factors from *Escherichia coli*. *Cell.* 72:459–466.

Borukhov, S., O. Laptenko, and J. Lee. 2001. Escherichia coli transcript cleavage factors GreA and GreB: functions and mechanisms of action. *Methods Enzymol.* 342:64–76.

Brower-Toland, B. D., C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, and M. D. Wang. 2002. Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proc. Natl. Acad. Sci. USA.* 99:1960–1965.

Davenport, R. J., G. J. Wuite, R. Landick, and C. Bustamante. 2000. Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase. *Science.* 287:2497–2500.

Forde, N. R., D. Izhaky, G. R. Woodcock, G. J. L. Wuite, and C. Bustamante. 2002. Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. USA.* 99:11682–11687.

Gusarov, I., and E. Nudler. 1999. The mechanism of intrinsic transcription termination. *Mol. Cell.* 3:495–504.

Komissarova, N., and M. Kashlev. 1997a. RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and RNA. *J. Biol. Chem.* 272:15329–15338.

Komissarova, N., and M. Kashlev. 1997b. Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3′ end of the RNA intact and extruded. *Proc. Natl. Acad. Sci. USA.* 94:1755–1760.

Liphardt, J., B. Onoa, S. B. Smith, I. Tinoco, Jr., and C. Bustamante. 2001. Reversible unfolding of single RNA molecules by mechanical force. *Science.* 292:733–737.

Neuman, K. C., E. A. Abbondanzieri, R. Landick, J. Gelles, and S. M. Block. 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell.* 115:437–447.

Nudler, E., A. Mustaev, E. Lukhtanov, and A. Goldfarb. 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell.* 89:33–41.

Opalka, N., M. Chlenov, P. Chacon, W. J. Rice, W. Wriggers, and S. A. Darst. 2003. Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell.* 114:335–345.

Roberts, J. W., W. Yarnell, E. Bartlett, J. Guo, M. Marr, D. C. Ko, H. Sun, and C. W. Roberts. 1998. Antitermination by bacteriophage λ Q protein. *Cold Spring Harb. Symp. Quant. Biol.* 63:319–325.

Ryder, A. M., and J. W. Roberts. 2003. Role of the non-template strand of the elongation bubble in intrinsic transcription termination. *J. Mol. Biol.* 334:205–213.

Schafer, D. A., J. Gelles, M. P. Sheetz, and R. Landick. 1991. Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature.* 352:444–448.

Shaevitz, J. W., E. A. Abbondanzieri, R. Landick, and S. M. Block. 2003. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature.* 426:684–687.

Tskhovrebova, L., J. Trinick, J. A. Sleep, and R. M. Simmons. 1997. Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature.* 387:308–312.

Uptain, S. M., C. M. Kane, and M. J. Chamberlin. 1997. Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* 66:117–172.

Wang, M. D., M. J. Schnitzer, H. Yin, R. Landick, J. Gelles, and S. M. Block. 1998. Force and velocity measured for single molecules of RNA polymerase. *Science.* 282:902–907.

Wang, M. D., H. Yin, R. Landick, J. Gelles, and S. M. Block. 1997. Stretching DNA with optical tweezers. *Biophys. J.* 72:1335–1346.

Yarnell, W. S., and J. W. Roberts. 1999. Mechanism of intrinsic transcription termination and antitermination. *Science.* 284:611–615.

Yin, H., M. D. Wang, K. Svoboda, R. Landick, S. M. Block, and J. Gelles. 1995. Transcription against an applied force. *Science.* 270:1653–1657.