



## Supplementary Materials for

### **A Pause Sequence Enriched at Translation Start Sites Drives Transcription Dynamics In Vivo**

Matthew H. Larson, Rachel A. Mooney, Jason M. Peters, Tricia Windgassen, Dhananjaya  
Nayak, Carol A. Gross, Steven M. Block, William J. Greenleaf,\* Robert Landick,\*  
Jonathan S. Weissman\*

\*Corresponding author. E-mail: wjg@stanford.edu (W.J.G.); landick@biochem.wisc.edu (R.L.);  
weissman@cmp.ucsf.edu (J.S.W.)

Published 1 May 2014 on *Science Express*  
DOI: 10.1126/science.1251871

**This PDF file includes:**

Materials and Methods  
Supplementary Text  
Figs. S1 to S14  
Tables S1 to S4  
References

## Materials and Methods

### Strains

Strains used in this study are listed in table S1. To construct the *E. coli* RNAP::3XFLAG strain, plasmid pRM678 was transformed into RL324 and recombination was selected by growth on LB-Kan plates (20). Plasmid loss from single-colony isolates was confirmed by loss of ampicillin-resistance. Replacement of *rpoC* with the *rpoC*::3XFLAG fusion was confirmed by western blot analysis of cell lysates separated by PAGE, using an antibody to the FLAG peptide. P1<sub>vir</sub> lysate grown on this strain was used to transduce RL1655 to give strain RL2081. Deletions of *greA*::*kan* and *greB*::*kan* were obtained from the Keio collection (21). Kan markers were excised with FLP recombinase to yield strains with the unmarked deletions (RL2849 and RL2850). The *rpoC*::3XFLAG allele was moved from RL2081 into these strains to yield strains RL2851 and RL2852. Mutation of the *lacZ* RBS (AGGAAA) to an orthogonal RBS sequence (o-RBS: ATCCCT) was performed using oligo-mediated recombineering (22). Briefly, RL2081 was transformed with the pSIM6 plasmid, which contains the  $\lambda$ -derived Red recombination proteins and an Amp selection marker. Mutagenesis was directed by a ssDNA oligo (*lacZ*-oRBS, table S1) with 42 nt of homology flanking either side of the mutation site. Colonies were blue-white screened for the absence of LacZ expression on X-gal/ITPG LB-Amp/Kan plates. Individual white colonies were restreaked and reselected, and the mutation verified using colony PCR to yield the final strain (RL2937).

The *B. subtilis* RNAP::3XFLAG strain was constructed in multiple steps. First, a *kan* marker with 1 kb flanking homology to the last 1 kb of *rpoC* and 1 kb of sequence immediately downstream of *rpoC* was generated by overlap extension PCR (23). The *kan* marker was then transformed into CAG17168; double crossover recombinants were selected on LB-Kan plates. Next, the *kan* marker and 1 kb downstream homology was amplified from the chromosome using a primer that contained a *NotI* restriction site and the coding sequence for 3XFLAG. In addition, the last 1 kb of the *rpoC* gene was amplified using a primer that contained a *NotI* site. The two PCR products were digested with *NotI*, ligated together, and then the ligation product was amplified again by PCR. The final product was transformed into CAG74168, selecting for kanamycin resistance, which generated the *rpoC*::3XFLAG strain, CAG74145.

### Cell growth and harvesting

For each sample, a monoclonal culture was grown at 37 °C from an OD (600 nm) 0.05 to early log-phase (OD 0.45 ± 0.05) in 500 mL of either MOPS EZ rich defined medium (Teknova) for *E. coli*, or LB medium for *B. subtilis*. Cells were harvested by filtration over 0.22  $\mu$ m nitrocellulose filters (GE) and frozen in liquid nitrogen to simultaneously halt all transcriptional progress. Frozen cells (100  $\mu$ g) were pulverized on a Qiagen TissueLyser II mixer mill 6 times at 15 Hz for 3 min in the presence of 500  $\mu$ L frozen lysis buffer (20 mM Tris pH 8, 0.4% Triton X-100, 0.1% NP-40, 100 mM NH<sub>4</sub>Cl, 50 U/mL SUPERase•In; Ambion) and 1X protease inhibitor cocktail (Complete, EDTA-free, Roche), supplemented with 10 mM MnCl<sub>2</sub>. The lysate was resuspended on ice by pipetting. RQ1 DNase I (110 U total, Promega) was added and incubated for 20 min on ice. The reaction was quenched with EDTA (25 mM final), which releases polysomes from the transcript (fig. S2) and reduces contamination from ribosomal RNA and

ribosome-associated tRNAs without affecting elongation complex stability (24). The lysate was clarified at 4 °C by centrifugation at 20,000 g for 10 min. The lysate was loaded onto a PD MiniTrap G-25 column (GE Healthcare) and eluted with lysis buffer supplemented with 1 mM EDTA.

For the measurement of start codon pausing in *lacZ* (fig. S12), both the WT (RL2081) and mutant *lacZ* (RL2937) strains were grown in the presence of IPTG (1 mM) and bicyclomycin (20 µg/mL; kind gift of Max Gottesman at Columbia University), which was necessary to prevent premature termination by Rho in the absence of translation.

#### Total RNA purification

Total RNA was purified from the clarified lysate using the miRNeasy kit (Qiagen). 1 µg of RNA in 20 µL of 10 mM Tris pH 7 was mixed with an equal volume of 2X alkaline fragmentation solution (2 mM EDTA, 10 mM Na<sub>2</sub>CO<sub>3</sub>, 90 mM NaHCO<sub>3</sub>, pH 9.3) and incubated for ~25 min at 95 °C to generate fragments ranging from 30-100 nt. The fragmentation reaction was stopped by adding 0.56 mL of ice-cold precipitation solution (300 mM NaOAc pH 5.5 plus GlycoBlue; Ambion), and the RNA was purified by a standard isopropanol precipitation. The fragmented mRNA was then dephosphorylated in a 50 µL reaction with 25 U T4 PNK (NEB) in 1X PNK buffer (without ATP) plus 0.5 U SUPERase•In, and precipitated with GlycoBlue via standard isopropanol precipitation methods.

#### Nascent RNA purification

For nascent RNA purification, the clarified lysate was added to 0.5 mL anti-FLAG M2 affinity gel (Sigma Aldrich) as described previously (7). The affinity gel was washed twice with lysis buffer supplemented with 1 mM EDTA before incubation with the clarified lysate at 4 °C for 2.5 h with nutation. The immunoprecipitation was washed 4 × 10 mL with lysis buffer supplemented with 300 mM KCl, and bound RNAP was eluted twice with lysis buffer supplemented with 1 mM EDTA and 2 mg/mL 3XFLAG peptide (Sigma Aldrich). Nascent RNA was purified from the eluate using the miRNeasy kit (Qiagen) and converted to DNA using a previously established library generation protocol (7).

#### DNA library preparation and DNA sequencing

The DNA library was sequencing on an Illumina HiSeq 2000. Reads were processed using the HTSeq Python package and other custom software written in Python 2.7. The 3' end of the sequenced transcript was aligned to the reference genome using Bowtie (<http://bowtie-bio.sourceforge.net/>) (25) and the RNAP profiles generated in MochiView (<http://johnsonlab.ucsf.edu/mochi.html>) (26).

#### NET-seq data analysis

Data analysis was performed using custom scripts written in Python 2.7. For pause detection, positions with counts greater than four standard deviations above the mean were scored as pauses. A list of pause positions identified using this criteria is included in the accompanying Gene Expression Omnibus submission (accession number GSE56720). Pause detection was largely independent of threshold choice: using a threshold value of two to four standard deviations above the gene average changed pause statistics only slightly, and did not affect the overall conclusions. To generate the pause consensus sequence logo, pause sequences were aligned at their 3' end and submitted to [weblogo.berkeley.edu](http://weblogo.berkeley.edu) (27).

### Single-molecule assay

We used a previously described, high-resolution optical-trapping assay to follow single *E. coli* RNAP molecules under assisting load as they transcribed a 1450 bp section of the *E. coli rpoB* gene (Fig. 2D) (10, 11). Briefly, elongation complexes (ECs) were stalled 29 bp after a T7A1 promoter on a digoxigenin-labeled template derived from the pRL777 plasmid. These ECs were incubated with 600 nm avidin-coated beads along with 730 nm antidigoxigenin-coated beads to form a bead:DNA:RNAP:bead dumbbell as previously described (10). Transcription was restarted by adding transcription buffer containing a rate-limiting concentration of GTP (2.5  $\mu$ M), along with a saturating concentration (1 mM) of the other nucleotide species. The nucleotide reaction mixture induced RNAP to pause at each template position that required GTP addition. Transcription data, taken at an assisting load of 18 pN using an optical force-clamp (28), were collected at 2 kHz, filtered at 1 kHz, then smoothed and decimated to a 100 Hz data rate.

### Single-molecule data analysis

To align transcription records absolutely against the known DNA template, the pattern of pauses was autocorrelated against a mask of expected pause positions (corresponding to limiting GTP) generated from the template sequence (Fig. 2D), as described previously (11). The automated alignment algorithm was allowed to stretch (by up to 10%) and shift (by up to 2 nm) individual 20-nm sections of transcription records in order to maximize their overlap with the expected pause patterns. These stretching and shifting transformations allowed small variations in rise-per-base of dsDNA, as well as long-term drift of the system, to be compensated. The shifted traces were checked by eye after alignment, and in some cases, portions of the record were re-aligned by hand. Finally, pause durations at expected base positions were measured using a semi-automated algorithm: an initial guess to identify the pause was made by marking the points at which the polymerase entered and exited a  $\pm 1.7$  Å window around the expected pause position. These automatically chosen pause demarcation points were then checked by eye, and modified if transient noise on the trace caused misidentification of the start or end of the pause. Noisy sections of certain traces were discarded. Because the sequential addition of 2 or more rate-limiting GTPs results in small spatial separations of pauses, these events were most challenging to separate into clearly resolved pause events. However, to avoid biasing data collection, pauses at sequential template positions were included in the analysis wherever feasible. We attempted to correct for possible misincorporation events, by excluding outlier pauses significantly longer than average dwell time measured at the sequence location.

### Calculation of pause energetics

The natural logarithm of the average dwell time was taken as a measure of the relative energetic barrier to elongation at each sequence position. For the single-molecule data, the dwell time was measured as described in the previous section. For the NET-seq data, the relative dwell time was determined from the total number of counts at a given position divided by the average over the gene. For both data sets, we used the same approach to analyze the sequence-specific contributions to next-nucleotide addition at the *i*th base under the footprint of RNAP. We partitioned our relative energetic barrier data into four subsets containing template positions that had each of the four possible species of nucleotide at the *i*th position under the RNAP footprint. The effective energy barriers for each of these sub-populations were then averaged, and the average effective contribution at each base position was then normalized, such that the sum over

all base possibilities of the relative energetic contributions at each position was zero. Errors ( $\pm$  SD) were estimated by measuring the variability of energetic contributions at other base positions that are not expected to make contact with the polymerase.

To determine the total energy at each position, we summed the energetic contribution of each base from position -1 to -11 upstream of the pause. The resulting energies for both pause and non-pause sequences are plotted in Fig. 2C. The separation of the cumulative probability curves along the x-axis indicates that pause-associated sequences have, on average, a higher energetic barrier for next-nucleotide addition. The slope of the curve is also shallower for pause sequences, suggesting that there may be greater variation in next-nucleotide energetics for pause sequences compared to non-pause sequences.

### RNA polymerases for *in vitro* transcription assays

The single-molecule experiments were performed with biotin-labeled RNAP, modified by addition of biotin carboxyl carrier protein to the C-terminus of the  $\beta'$  subunit, and purified as described previously (29). Ensemble transcription experiments were performed with wild-type *E. coli* RNAP overexpressed from plasmid pRM756 (table S1) and purified by Ni-NTA agarose affinity chromatography followed by chromatography on a HiTrap heparin column, as described previously (30). *Bta*RNAPII was purified from calf thymus by PEI precipitation, anion-exchange chromatography on High Q Sepharose, 8WG16 antibody affinity chromatography, and anion-exchange chromatography, on Uno-Q as described previously (31). *Rsp*RNAP and *Mbo*RNAP were kind gifts from Chris Lennon and Agata Czyz, respectively (University of Wisconsin). *Tth*RNAP was purified as described previously (32). NusG, GreA, and GreB were purified as described previously (33, 34). RfaH NTD was a kind gift from Pyae Hein (University of Wisconsin). TFIIS used here was TFIIS $\Delta$ 1-130, a derivative of TFIIS lacking the N-terminal mediator-interaction domain (35), and was a kind gift of C. Kane (University of California-Berkeley).

### Ensemble *in vitro* transcription assays with bacterial RNAPs

All DNA and RNA oligonucleotides were from IDT. FPLC-purified NTPs were from Promega Corp. Radionucleotides ( $[\alpha\text{-}^{32}\text{P}]\text{GTP}$  or  $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ ) were from Perkin-Elmer. Elongation complexes were assembled using reconstitution essentially as described elsewhere (30) but with 50 nM scaffold and 200 nM core RNAPs, or after initiation from a promoter (Fig. S6). For scaffold-based pause assays except for *Tth*RNAP, complexes were reconstituted 2 nt upstream from the pause site, incubated with 10  $\mu\text{M}$   $[\alpha\text{-}^{32}\text{P}]\text{GTP}$  (20 Ci/mmol) for 5 min at 37  $^{\circ}\text{C}$  to form complexes halted 1 nt before the pause site. For scaffold-based assays on the anti-consensus sequence except for *Tth*RNAP, complexes were similarly reconstituted 2 nt upstream from the pause site, but incubated with 10  $\mu\text{M}$   $[\alpha\text{-}^{32}\text{P}]\text{UTP}$  (20 Ci/mmol) for 5 min at 37  $^{\circ}\text{C}$  to form complexes halted 1 nt before the site corresponding to the pause site on the consensus scaffold. Transcription was restarted by addition of 100  $\mu\text{M}$  GTP and CTP to allow transcription through the pause signal, or 100  $\mu\text{M}$  GTP, CTP, and UTP for the anti-consensus. For *Tth*RNAP, complexes were reconstituted 14 nt upstream from the pause site (G16, fig S10), incubated with 10  $\mu\text{M}$   $[\alpha\text{-}^{32}\text{P}]\text{CTP}$  (20 Ci/mmol) for 5 min at 37  $^{\circ}\text{C}$  to form C17 complexes halted 13 nt before the pause site. Transcription was restarted by addition of 100  $\mu\text{M}$  GTP, CTP, ATP, and UTP and incubated at 37  $^{\circ}\text{C}$  (fig. S10B) or 50  $^{\circ}\text{C}$  (fig. S10C). Samples were removed at different time points and quenched by the addition of an equal volume of 2X urea STOP buffer (7 M urea, 50 mM EDTA, 90 mM Tris-borate buffer pH 8.3, 0.02% bromophenol blue, and 0.02% xylene

cyanol). Samples were heated for 2 min at 95 °C and separated by electrophoresis through 20% denaturing polyacrylamide (19:1) gels in 0.5X TBE (45 mM Tris-borate pH 8.3, 1.25 mM EDTA). Rates were determined by performing single- or double-exponential fits of the fraction of RNA at the pause position relative to the total amount of RNA at this position and beyond as a function of time. Relative pause strengths ( $\tau_0$ ) shown in Fig. 3C and fig. S7 correspond to the area under the pause RNA fraction *vs.* time, and were calculated as the product of pause efficiency,  $E$ , and pause dwell time,  $\tau$ , ( $E \cdot \tau$  or  $E/k_e$ , where  $k_e$  is the pause escape rate (36)) and reported as the average pause strength relative to wild type  $\pm$  SD from  $\geq 3$  experimental replicates. The fraction of RNA species present in the reaction was normalized to 1 for each time point and the change in pause RNA fraction over time was then fit to a single- or double-exponential decay equation,  $E_1 \cdot e^{-k_1 \cdot t} + E_2 \cdot e^{-k_2 \cdot t}$ , where the slower rate represented a minor fraction of complexes ( $E_2$  and  $k_2$ ) that may enter a backtracked state or was omitted for single-exponential fits when the slow phase was negligible ( $E_2=0$ ). The total pause efficiency,  $E$ , was  $E_1 + E_2$ . Pause escape rates for regulators (Fig. 3C) were determined by adding the regulators to reactions after the incorporation labeling step at concentrations indicated in the legend. GreA and GreB gave similar results; the GreA/B pause escape rate shown in Fig. 3C is the average of 3 replicates with GreA and one replicate with GreB.

To determine consensus pause kinetic parameters ( $V_{\max}$  and  $K_{\text{NTP}}$ ; Fig. 3B), pause escape rates were measured over a range of GTP concentrations (10  $\mu\text{M}$  – 20 mM) using a KinTek quench-flow apparatus, as described previously (30), except that a weighted average of slow and fast rates (based on relative  $E$ ) was used for GTP concentrations at which two rates were evident (typically  $\leq 100$   $\mu\text{M}$  GTP). Briefly, 1.2  $\mu\text{M}$  RNA (8342), 1  $\mu\text{M}$  template DNA (8334), and 2  $\mu\text{M}$  non-template DNA (8333) were annealed to form a nucleic-acid scaffold and then reconstituted with RNAP (250 nM RNAP, 200 nM scaffold) in 1X MB (20 mM Tris-HCl, pH 7.9, 20 mM NaCl, 0.125 mM EDTA, 15 mM MgCl<sub>2</sub>, 5% glycerol, 25 mg acetylated BSA/mL, and 2 mM  $\beta$ -mercaptoethanol) for 15 min at room temperature. Complexes were diluted 2 fold in 1X MB, supplemented with MgCl<sub>2</sub> to 5 mM greater than the NTP concentration, and then incorporation labeled with 2  $\mu\text{M}$  [ $\alpha$ -<sup>32</sup>P]GTP. Labeled ECs (20  $\mu\text{L}$ ) were injected into one sample loop and NTPs (20  $\mu\text{L}$ ; varying [GTP]) and 2 mM CTP in transcription buffer were injected into the other loop. Reactions were performed for predetermined times at 37 °C. The samples were quenched with 2 M HCl (40  $\mu\text{L}$ ) and immediately neutralized to pH 8.0 by addition of 3 M Tris base (40  $\mu\text{L}$ ). The RNA was phenol extracted, ethanol precipitated, suspended in formamide loading dye, separated on a denaturing polyacrylamide (17.5%; 8 M urea) 0.5X TBE gel, and quantified using a Phosphorimager. Pause escape rates for each GTP concentration were determined in triplicate by using KaleidaGraph (Synergy Software) to fit pause decay to a single- or double-exponential. These rates were used to determine the pause escape  $V_{\max}$  of  $0.5 \pm 0.1 \text{ s}^{-1}$  and  $K_{\text{GTP}}$  of  $2.5 \pm 1 \text{ mM}$  (Fig. 3B). Kinetic parameters for the *his* pause shown for comparison in Fig. 3B were determined by the same methods and were reported previously (37).

### Ensemble *in vitro* transcription assays with mammalian RNAPII

Elongation complexes were assembled using reconstitution essentially as described previously (38), but with 150 nM scaffold and 20 nM *Bta*RNAPII. Complexes were reconstituted 2 nt upstream from the pause site, incubated with 10  $\mu\text{M}$  [ $\alpha$ -<sup>32</sup>P]GTP (20 Ci/mmol) for 3 min at 30 °C to form complexes halted 1 nt before the pause site. If present, TFIIS was then added to 300 nM and allowed to bind for 3 minutes at 30 °C. Transcription was restarted by addition of 10  $\mu\text{M}$  CTP to allow transcription through the pause signal. Samples were removed at

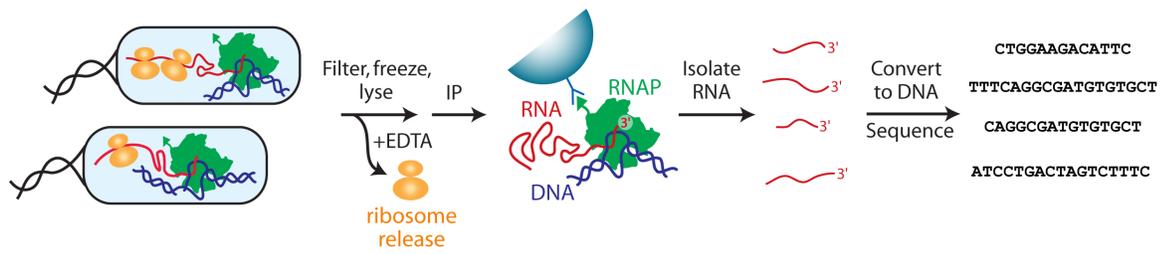
different time points and quenched by the addition of an equal volume of 2X urea STOP buffer. Samples were heated for 2 min at 95 °C and separated by electrophoresis through 20% denaturing polyacrylamide (19:1) gels in 0.5X TBE. Rates were determined as described above.

## Supplementary Text

### Determination of consensus sequence contributions to pausing *in vitro*

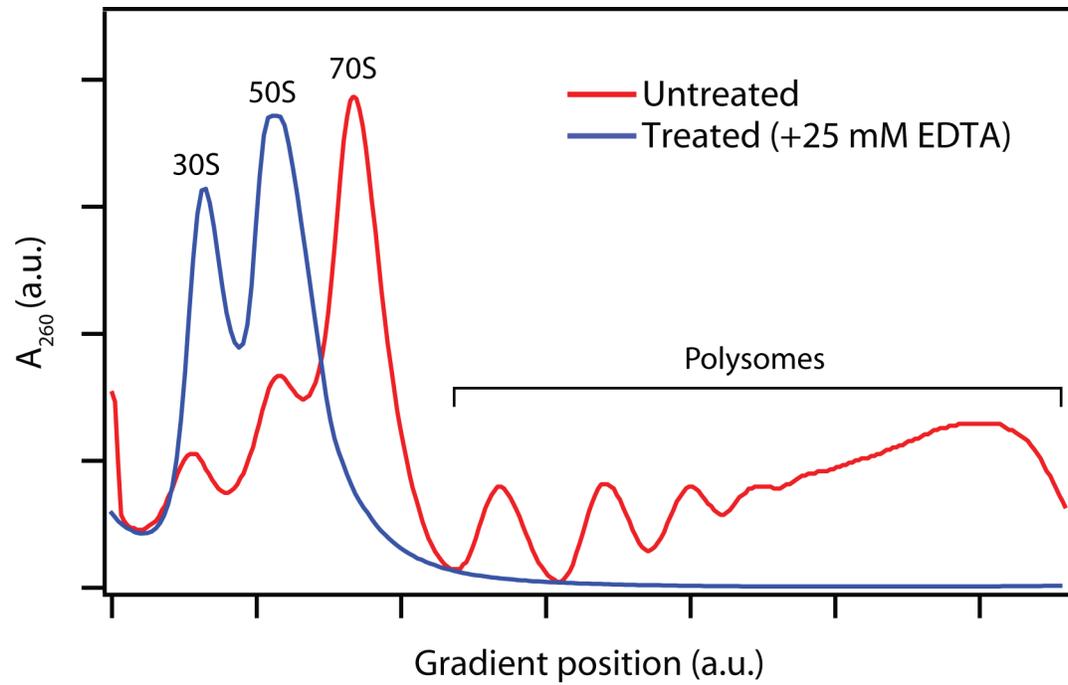
To determine how changes to the consensus sequence affected pausing by RNAP, we measured the rates at which RNAP added the next nucleotide after the pause on a series of scaffolds with single or multiple substitutions at different locations of the consensus scaffold (Fig 3C and fig. S7). To aid reconstitution, the consensus scaffold contained non-complementary bases in the non-template strand opposite the 8-nt core RNA:DNA hybrid, positions for which DNA-DNA base-pairing should have little or no effect on pausing. The effects of different substitutions on consensus pausing are attributable to discrete, previously described interactions of RNAP with separate parts of the scaffold that combine additively to create a multipartite pause signal that traps RNAP in an off-line state (13, 39, 40). Substitutions at -10 and -11 reveal that RNA:DNA base-pairing in an overextended hybrid rather than G in the unpaired RNA contributes to pausing, consistent with prior observations (41) and Gilbert's first proposal of a pause mechanism in which difficulty in unwinding the RNA:DNA hybrid inhibits nucleotide addition (42). Substitutions in the RNA:DNA hybrid (positions -2, -3, and -5) are consistent with a modest contribution to pausing from interactions between RNAP and the hybrid (13, 39, 40, 43, 44). Effects of substitutions at -1 and +1 are consistent with the consensus pause sequence (Fig. 2) and previous reports that pausing is favored when pyrimidines at the RNA 3' end react with incoming purine NTPs (13, 15, 45, 46). Abrogation of pausing by removal of the nontemplate G at +1 could suggest interaction of the nontemplate G makes an additive contribution to consensus pausing of magnitude similar to hybrid overextension and downstream DNA interaction, consistent with the proposal by Ebright and co-workers (47), but also is consistent with a positive contribution of base-pairing between the nontemplate strand G and template strand C that could inhibit translocation. The effects of downstream substitutions are consistent with prior suggestions that interactions of duplex DNA near +2 to +4 and +5 to +8 with the downstream DNA entry channel of RNAP also facilitate pausing (13, 40, 46, 48, 49).

Although we found evidence for backtracking at the consensus pause site (fig. S8), our results strongly suggest the major barrier to pause escape occurs in the non-backtracked elemental pause state (12, 14, 37), especially at *in vivo* concentrations of NTP. We note, however, that variations in the sequence might change the propensity for backtracking and that we have not yet systematically explored the effect of sequence changes on the extent of backtracking, including for the substitutions shown in Fig. 3C and fig. S7.



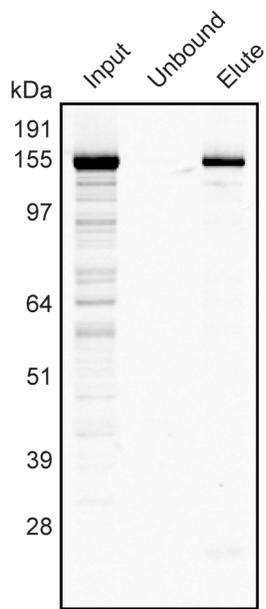
**Fig. S1. Schematic of the NET-seq protocol.**

Actively transcribed RNA is isolated from bacteria via immunoprecipitation of FLAG-tagged RNAP molecules and converted to a DNA library sequenced with deep coverage.



**Fig. S2. Polysome profiles of *E. coli* cell lysates.**

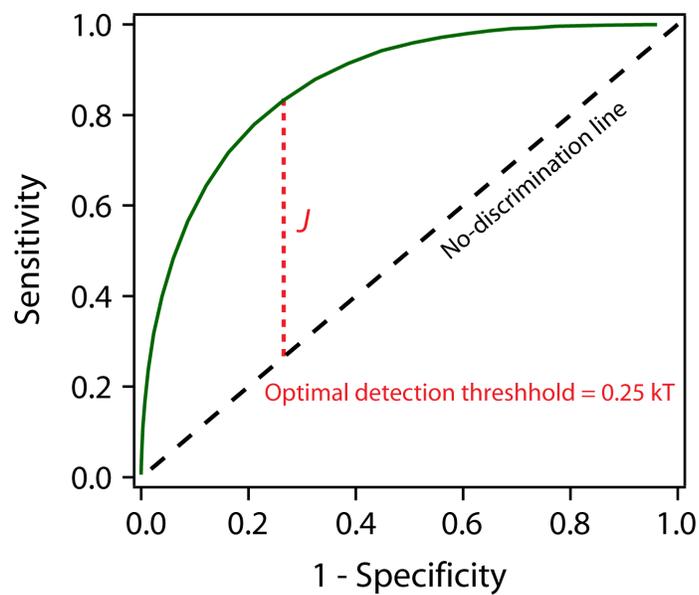
Treatment of lysate with 25 mM EDTA chelates  $Mg^{2+}$  and leads to release of cotranslating ribosomes, reducing rRNA and tRNA contamination.



**Fig. S3. Immunoprecipitation of RNAP.**

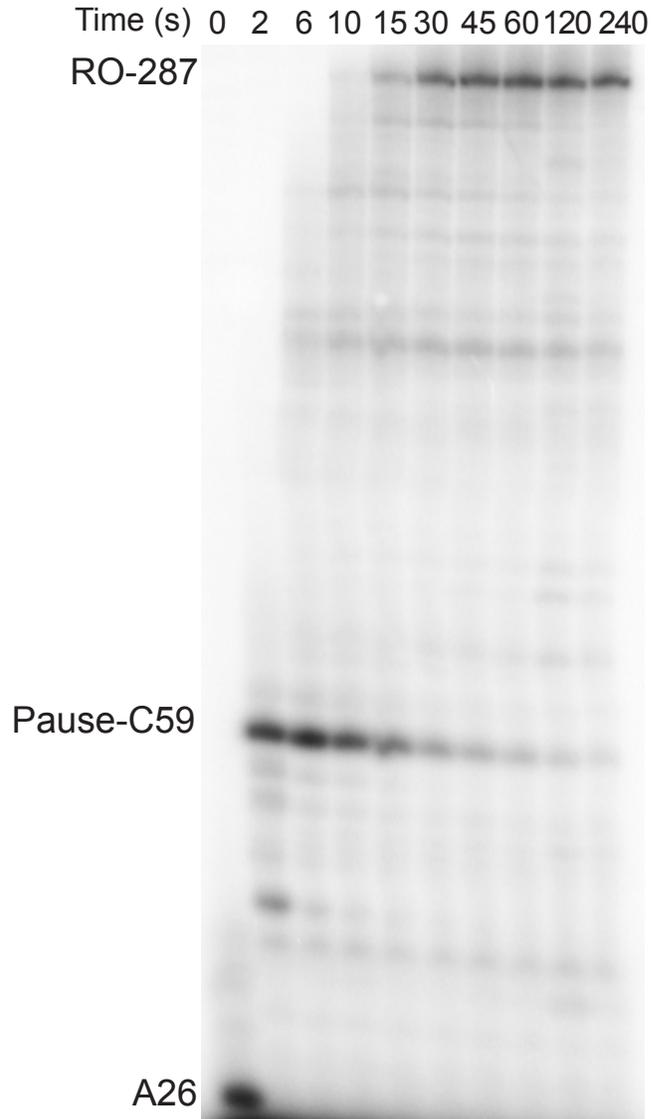
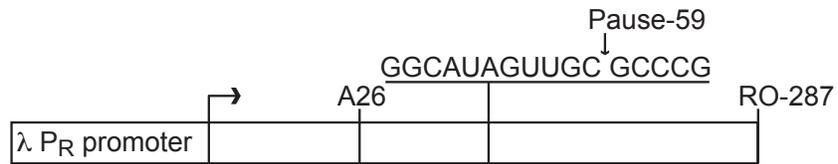
Western blot detecting FLAG-labeled RpoC in immunoprecipitation samples corresponding to input lysate, unbound lysate, and eluted protein.





**Fig. S5. Receiver-operating characteristic (ROC) analysis for the prediction of pauses using next-nucleotide energetics.**

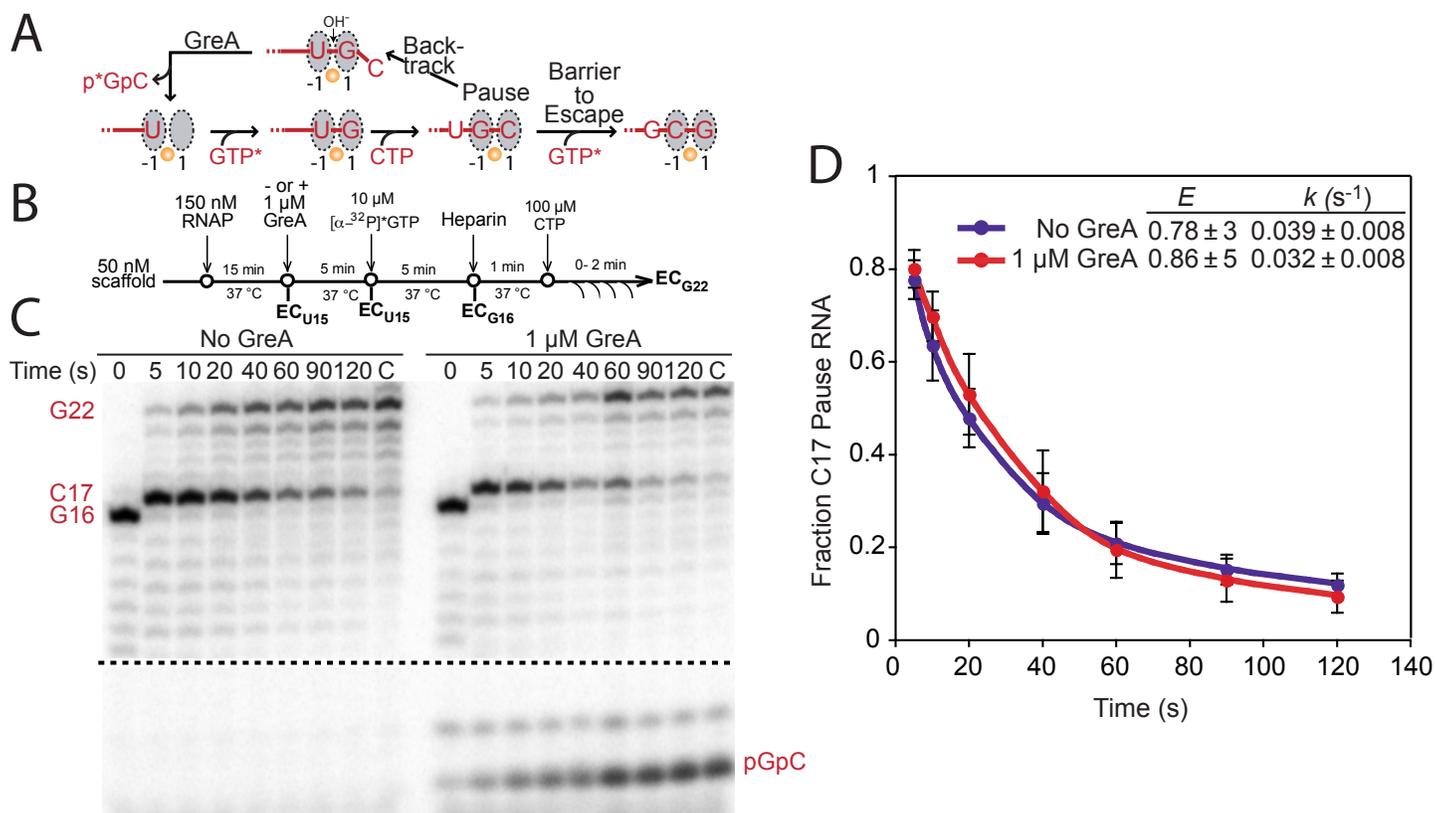
The optimal energy threshold was determined by maximizing the Youden index ( $J$ ).



**Fig. S6. Transcription of the consensus pause sequence initiated from a  $\lambda$   $P_R$  promoter shows a strong transcriptional pause.**

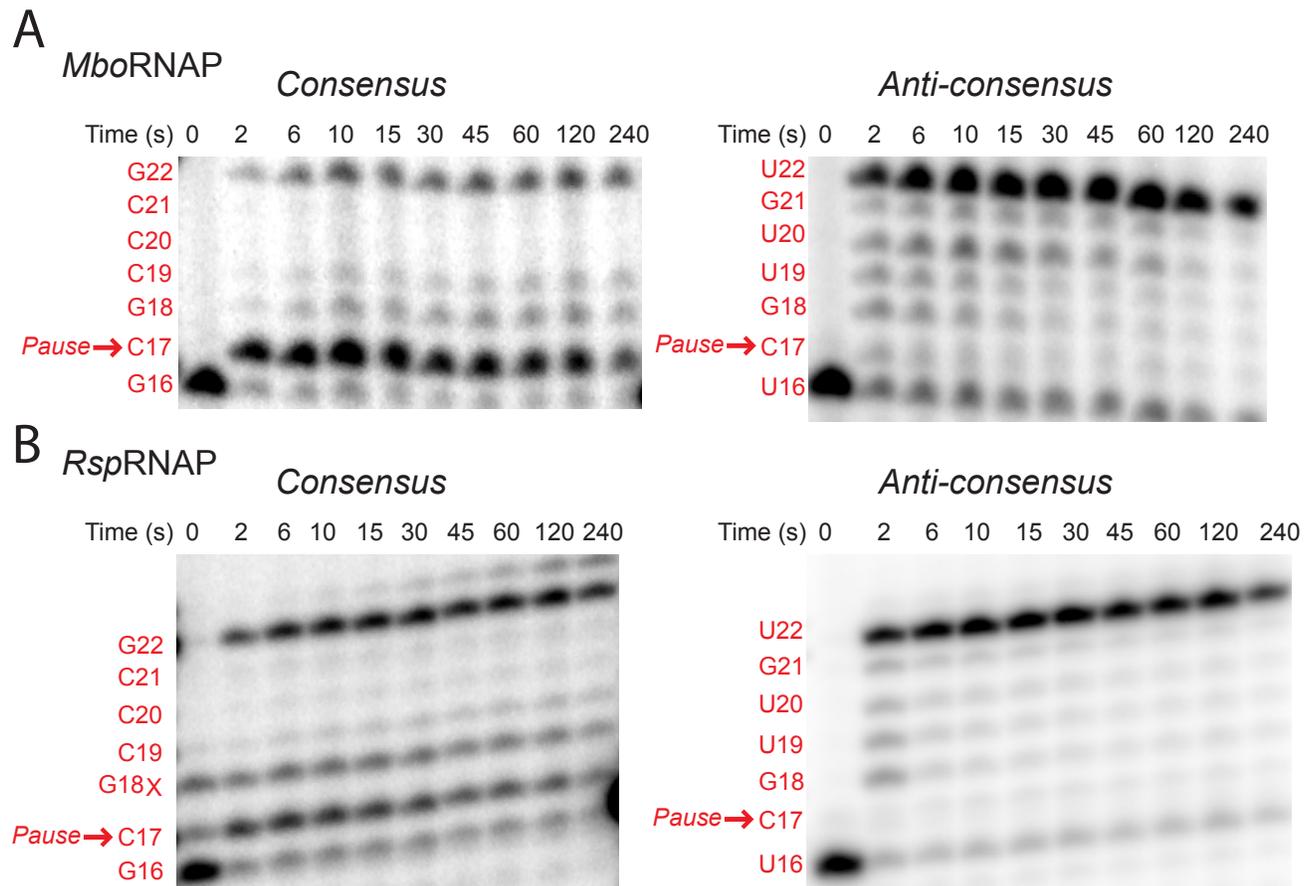
Halted complexes (A26) were formed by incubating 40 nM RNAP with 25 nM linear DNA template (prepared by PCR from pRM1002; table S1; methods), 150  $\mu$ M ApU, 2.5  $\mu$ M ATP and UTP, and 1  $\mu$ M [ $\alpha$ - $^{32}$ P]-GTP for 15 minutes at 37  $^{\circ}$ C. Transcription was resumed by addition of 100  $\mu$ M of all 4 NTPs plus 150  $\mu$ g rifampicin/mL. Samples were removed at 0, 2, 6, 10, 15, 30, 45, 60, and 120 seconds. Pause escape rate =  $0.09 \pm 0.007$ .





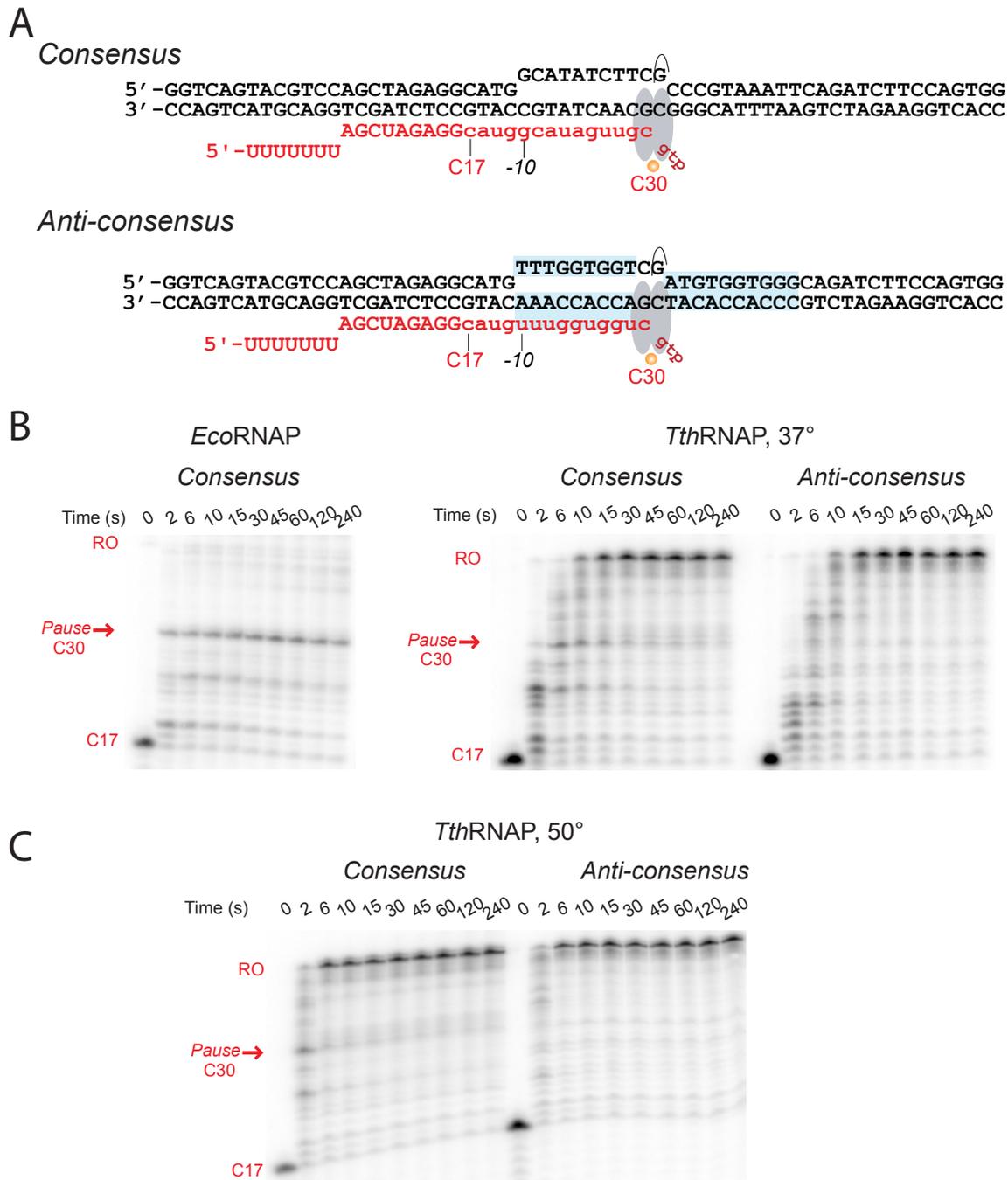
**Fig. S8. Effect of transcript cleavage factor GreA on consensus sequence pausing.**

(A) Schematic indicating the reactions taken during the pause assay shown in panel C conducted on the consensus scaffold shown in fig. S7A. Constant specific activity [ $\alpha$ -<sup>32</sup>P]GTP was used during the labeling and extension parts of the assay to ensure that the cleavage and re-extension of the transcript did not alter its specific activity. (B) Schematic of the experiment shown in panel C. (C) Pause assay on consensus scaffold without (left panel) or with (right panel) 1  $\mu$ M GreA. The bands are labeled with numbering shown in fig. S7A. The appearance of the cleavage product pGpC in the presence of GreA suggests that GreA can induce backtracking of a consensus paused complex that is not backtracked in the absence of GreA (Fig. 3D), consistent with prior suggestion that GreA can stabilize backtracking by RNAP (54). (D) Quantitation of the pause RNA (C17) as a function of time from the data shown in panel C ( $\pm$  SD from 3 replicates). Although GreA may slightly increase the efficiency of pausing, consistent with its interaction in the secondary channel of RNAP, the absence of overall effect of GreA on consensus pausing establishes that the barrier to consensus pause escape occurs at the step of GTP binding to or reaction with the non-backtracked C18 complex, rather than entry into a backtracked state from which it can be rescued by GreA.



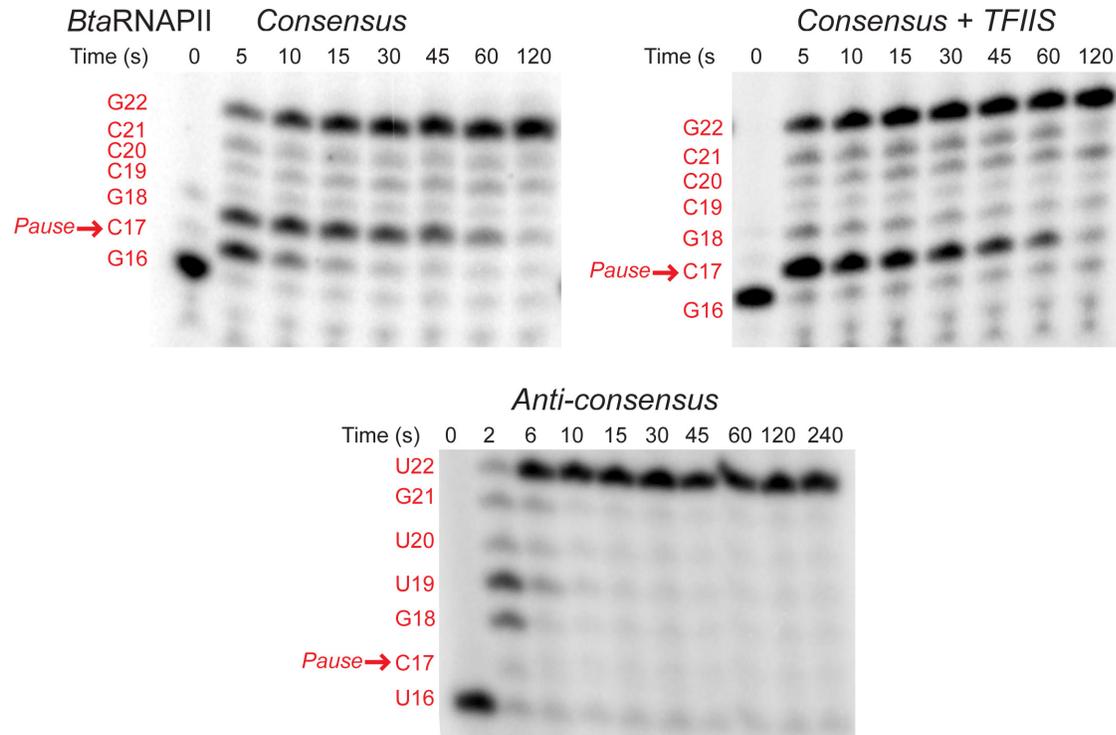
**Fig. S9. Consensus pause sequence also leads to pausing by RNAPs from diverse lineages.**

Pausing on the consensus and anti-consensus scaffold by RNAPs from *Mycobacteria bovis* (A) and *Rhodobacter pseudomonas* (B). Pause assays were performed using the scaffolds shown in fig. S7A and S7B and as described in supplementary methods. Pause strength at C17 was quantified as described in the supplementary methods and is depicted in Fig. 3C with SD from 3 replicates.



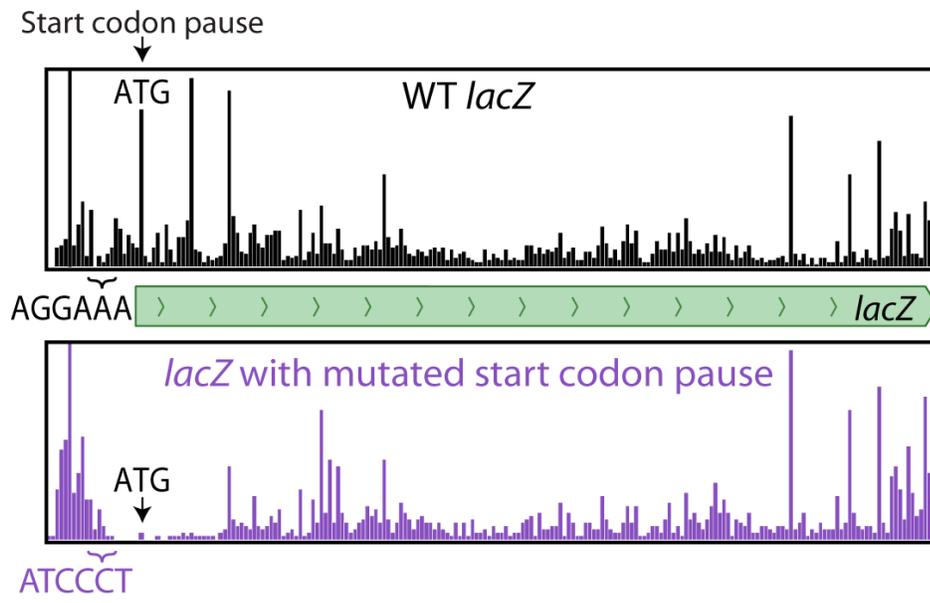
**Fig. S10. Pausing by *Tth*RNAP on consensus and anti-consensus scaffolds.**

(A) Representative pause assays were performed using *Eco* or *Tth* RNAPs reconstituted on the scaffolds shown in which longer DNA duplexes were used to ensure annealing when higher temperatures were used for *Tth*RNAP (e.g., 50 °C). Portions of the sequence highlighted in blue denote the sequences that are changed between the scaffolds. The sequence of the starting RNA is capitalized whereas nucleotides that are added by incorporation are shown in lower-case. Nascent RNAs were 3'-end labeled by incorporation of [ $\alpha$ -<sup>32</sup>P]CMP to form C17 complexes (time 0). Transcription was then restarted by addition of 100  $\mu$ M ATP, GTP, CTP, and UTP and samples were removed at the times indicated. (B) Transcription with *Eco* or *Tth* RNAPs at 37 °C. (C) Transcription by *Tth*RNAP at 50 °C. Pausing behavior of *Tth*RNAP relative to that exhibited by *Eco*RNAP is consistent with previous work (14, 15).



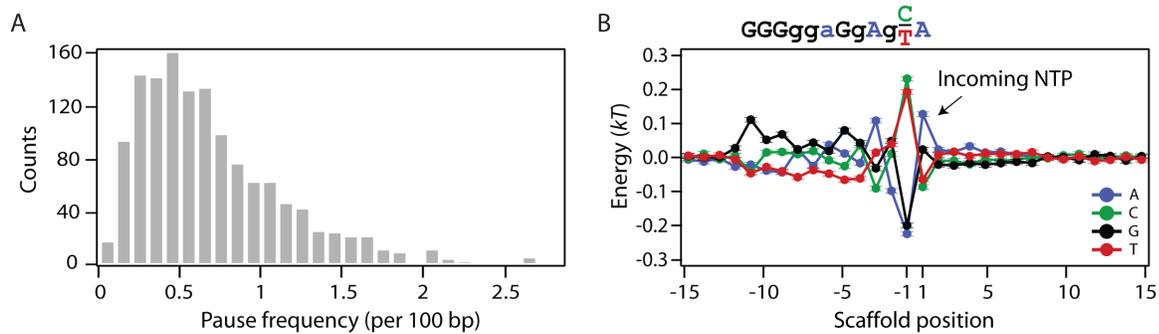
**Fig. S11. Pausing by *BtaRNAPII* on the consensus and anti-consensus scaffold.**

Representative pause assays were performed using calf thymus RNAPII (*BtaRNAPII*) reconstituted on the consensus (*top two panels*) or anti-consensus (*lower panel*) pause scaffolds as shown in fig. S7 and described in the supplementary methods. Nascent RNAs were 3'-end labeled by incorporation of [ $\alpha$ - $^{32}$ P]GMP (consensus) or [ $\alpha$ - $^{32}$ P]UMP (anticonsensus) one nt upstream from the pause (time 0), and then samples were removed at the times indicated after addition of CTP to 100  $\mu$ M, GTP to 10  $\mu$ M, and, for the anticonsensus, UTP to 100  $\mu$ M. For the consensus pause assay +TFIIS, human TFIIS lacking the N-terminal domain ( $\Delta$ 1-130; see supplementary methods) was added to 300 nM to the G16 complexes prior to the addition of CTP. Samples were removed from the reactions at the times indicated, then processed and separated on 20% polyacrylamide denaturing gels as described in the supplementary methods. The RNAs were quantified by phosphorimaging. Triplicate data for the consensus scaffold + TFIIS and the anti-consensus scaffolds were used to calculate pause strengths shown in Fig. 3C.



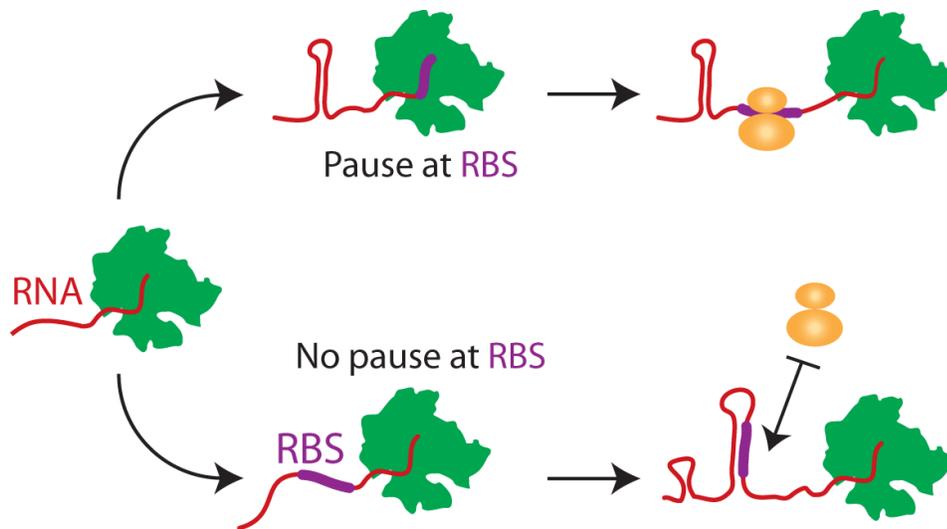
**Fig. S12. *LacZ* RBS is required for start codon pausing.**

A strong start codon pause is observed for *lacZ* in the WT strain (black). Mutation of upstream elements of the pause consensus sequence, located within the RBS, abolishes the start codon pause in the mutated strain (purple).



**Fig. S13. Bacterial NET-seq results for *B. subtilis*.**

(A) Histogram of pause frequency for highly transcribed genes ( $n = 1437$ , gene average  $>1$  read/bp) within the protein coding sequence. (B) The pause energetics indicate that pausing is strongly associated with ATP addition.



**Fig. S14. Model of the role of transcriptional pausing in translation initiation.**

RNAP pausing at the start codon could direct folding of the 5'-UTR into structures that preserve accessibility of the RBS once transcription resumes. Upper pathway: RNAP pauses at the start codon, allowing formation of RNA structure upstream of the RBS. The RBS is left unstructured when transcription resumes, promoting translation initiation. Lower pathway: RNAP does not pause at the start codon, leading to the formation of RBS-occluding structures in the nascent transcript that inhibit ribosome binding.

**Table S1. Bacterial strains, plasmids, and oligonucleotides used in this study.**

Stock #	Name	Description	Source or Note
<b>Strains</b>			
RL1655	MG1655	<i>Escherichia coli</i> K12; F- $\lambda$ - <i>ilvG</i> - <i>rfb</i> -50 <i>rph</i> -1	(55)
RL324	JC7623	$\lambda$ F <sup>-</sup> <i>recB21 recC22 sbcB15 thr-1 leuB6 thi-1 lacY1 galK2 ara-14 xyl-5 mtl-1 proA2 his4 argE3 rpsL31(Str) tsx-33 supE44</i>	(20)
RL2081		MG1655 <i>rpoC</i> ::3XFLAG	This work
RL2688		$\Delta$ <i>greA</i> ::kan; Keio collection	(21)
RL2690		$\Delta$ <i>greB</i> ::kan; Keio collection	(21)
RL2849		MG1655 $\Delta$ <i>greA</i>	This work
RL2850		MG1655 $\Delta$ <i>greB</i>	This work
RL2851		MG1655 $\Delta$ <i>greA rpoC</i> ::3XFLAG	This work
RL2852		MG1655 $\Delta$ <i>greB rpoC</i> ::3XFLAG	This work
RL1293	CLT252	$\Delta$ <i>greA</i> ::Cm	C. Turnbough
RL2858		MG1655 $\Delta$ <i>greA</i> ::Cm $\Delta$ <i>greB rpoC</i> ::3XFLAG	This work
RL2937		MG1655 o-RBS <i>lacZ rpoC</i> ::3XFLAG	This work
CAG74168	168	<i>Bacillus subtilis subtilis</i> 168; <i>trpC2</i>	(56)
CAG74145	JMP27	168; <i>rpoC</i> ::3XFLAG-kan	This work
<b>Plasmids</b>			
2956	pRM756	Expresses wild-type <i>E. coli</i> RNAP ( $\alpha_2\beta\beta'\omega$ ) with His <sub>10</sub> tag on the $\beta'$ C-terminus	(30)
1777	pRL777	T7 A1 <i>rpoB'</i> DNA template plasmid for single-molecule transcription	(57)
2878	pRM678	<i>rpoC</i> ::3X FLAG	This work
5302	pRM1002	$\lambda$ PR-consensus pause	This work
<b>Oligonucleotides (5'→3')</b>			
8333	NT, consensus pause	GGTCAGTACGTCCGGTCGATCTTCGCCCGTAAATTCAGATCTTCCAGTGG	
8334	T, consensus pause	CCACTGGAAGATCTGAATTTACGGGCGCAACTATGCCGGACGTACTGACC	
8342	R, consensus pause	UUUUUUGGCAUAGUU	
8953	NT, anti-consensus	GGTCAGTACGTCCTTGCCAGCTGCGTTGTAGTGGGCAGATCTTCCAGTGG	
8954	T, anti-consensus	CCACTGGAAGATCTGCCCACTACAACGACCACCAAAGGACGTACTGACC	

8952	R, anti-consensus	UUUUUUUUUUGGUGG
8609	NT, -1 C to A mutant	GGTCAGTACGTCCGGTCGATCTTAGCCCGTAAATTCAGATCTTCCAGTGG
8610	T, -1 C to A mutant	CCACTGGAAGATCTGAATTTACGGGCTCAACTATGCCGGACGTACTGACC
9074	NT, -1 C TO T mutant	GGTCAGTACGTCCGGTCGATCTTTGCCCGTAAATTCAGATCTTCCAGTGG
9075	T, -1 C to U mutant	CCACTGGAAGATCTGAATTTACGGGCACA ACTATGCCGGACGTACTGACC
8607	T, -2 mutant	CCACTGGAAGATCTGAATTTACGGGCGTAACTATGCCGGACGTACTGACC
8873	T, -3 mutant	CCACTGGAAGATCTGAATTTACGGGCGCGACTATGCCGGACGTACTGACC
8875	R, -3 mutant	UUUUUUGGCAUAGUC
8874	T, -5 mutant	CCACTGGAAGATCTGAATTTACGGGCGCAAATATGCCGGACGTACTGACC
8876	R, -5 mutant	UUUUUUGGCAUAUUU
8604	NT, -10mismatch	GGTCAGTACGTCCGTTCGATCTTCGCCCGTAAATTCAGATCTTCCAGTGG
8605	T, -10 mismatch	CCACTGGAAGATCTGAATTTACGGGCGCAACTATGACGGACGTACTGACC
8601	NT, -11 mismatch	GGTCAGTACGTCCGTTCGATCTTCGCCCGTAAATTCAGATCTTCCAGTGG
8602	T, -11 mismatch	CCACTGGAAGATCTGAATTTACGGGCGCAACTATGCAGGACGTACTGACC
8969	NT, -1 abasic	GGTCAGTACGTCCGGTCGATCTT-GCCCGTAAATTCAGATCTTCCAGTGG
8965	NT, +1 abasic	GGTCAGTACGTCCGGTCGATCTTC-CCCGTAAATTCAGATCTTCCAGTGG
8979	NT, +2 to +11 mutant	GGTCAGTACGTCCGGTCGATCTTCGATGTGGTGGGCAGATCTTCCAGTGG
8980	T, +2 to +11 mutant	CCACTGGAAGATCTGCCCACCACATCGCAACTATGCCGGACGTACTGACC
8975	NT, +2 to +4 mutant	GGTCAGTACGTCCGGTCGATCTTCGATGGTAAATTCAGATCTTCCAGTGG
8976	T, +2 to +4 mutant	CCACTGGAAGATCTGAATTTACCATCGCAACTATGCCGGACGTACTGACC
8977	NT, +5 to +8	GGTCAGTACGTCCGGTCGATCTTCGCCCTAGTATTCAGATCTTCCAGTGG
8978	T, +5 to +8	CCACTGGAAGATCTGAATACTAGGGCGCAACTATGCCGGACGTACTGACC
lacZ-oRBS		T*C*GTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACATCCCTCAGCTATGACCATGATTACGGATTCACTGCCGTCGTTTT*A*C (*denotes phosphorothioate bonds)

**Table S2. Alignment statistics.**

The total number of aligned reads and the number of transcripts followed (each with an average density of at least 1 read/bp) for each sample.

<b>Sample</b>	<b>Aligned reads</b>	<b>Transcripts followed</b>
<i>E. coli</i> , WT (RL2081)	33,630,893	1,984
<i>E. coli</i> , $\Delta greA$ (RL2851)	49,046,813	2,077
<i>E. coli</i> , $\Delta greB$ (RL2852)	30,130,047	2,095
<i>E. coli</i> , $\Delta greA / \Delta greB$ (RL2858)	38,792,824	1,557
<i>E. coli</i> , WT (RL2081) + 20 ug/mL BCM	34,119,727	1,923
<i>E. coli</i> , <i>lacZ</i> -oRBS (RL2937) + 20 ug/mL BCM	47,676,729	1,945
<i>B. subtilis</i> , WT (CAG74145)	17,217,456	1,564

**Table S3. Demonstration of specificity in the purification of nascent transcripts.**

Two IPs were performed using the protocol detailed above. IP1 used a mixed lysate of two strains: (1) a strain endogenously expressing FLAG-labeled RNAP and (2) a strain expressing unlabeled RNAP and GFP. IP2 was performed on lysate from a strain expressing both FLAG-labeled RNAP and GFP. If free RNA does not reassociate with RNAP after cell lysis, then qPCR of GFP transcripts (normalized based on the levels of an endogenous gene, *ompA*) should not detect GFP transcripts, while the second IP should detect a significant population of GFP transcripts due to their transcription by FLAG-labeled RNAP. qPCR on the RNA that copurified from each IP quantified the *ompA*:GFP ratio, which is summarized in the table. These results show that nascent transcripts expressed in the same cells as a FLAG-labeled RpoC are purified at least 41-fold more than mature messages.

<b>Sample</b>	<b>IP 1</b>	<b>IP 2</b>
GFP (a.u.)	0.4	15.5
<i>ompA</i> (a.u.)	22.7	21.3
<i>ompA</i> /gFP	56.8	1.4
IP 1 / IP 2	<b>41 fold enrichment</b>	

**Table S4. List of previously characterized regulatory pauses observed by NET-seq.**  
The nucleotide associated with the peak in RNAP density is underlined, while the next-nucleotide to be added to the transcript 3' end is shown in lowercase.

<b>Pause location</b>	<b>Pause sequence</b>
<i>hisL</i>	CAGGCGATGTGTGC <u>T</u> gGAAG
<i>ilvL</i>	TGCGGGGCTGCACT <u>T</u> gGACG
<i>ivbL</i>	CGTGCGTGTGGTGG <u>T</u> gGTCC
<i>leuL</i>	GCGCGGTAGACGAG <u>T</u> gAGCG
<i>mgtL</i>	GTAAGGCTTCGCCA <u>C</u> gCCTG
<i>rfaQ (ops)</i>	CTGGGGCGGTAGCG <u>T</u> gCTTT
<i>pheL</i>	GAGGCGTTTCGTCG <u>T</u> gTGAA
<i>pheM</i>	AGGAGGCTAGCGCG <u>T</u> gAGAA
<i>pyrL</i>	ATTTTGTCTTACCG <u>C</u> gTCTG
<i>rnpB</i>	GGGGGGAAACCCA <u>C</u> gACCA
<i>thrL</i> (1)	ACGGTGCGGGCTGA <u>C</u> gCGTA
<i>thrL</i> (2)	GGTGCGGGCTGACG <u>C</u> gTACA
<i>tnaC</i>	CACCGCCCTTGATT <u>T</u> gCCCT
<i>trpL</i>	CAGTGTATTCACCA <u>T</u> gCGTA

## References and Notes

1. T. Pan, T. Sosnick, RNA folding during transcription. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 161–175 (2006). [Medline doi:10.1146/annurev.biophys.35.040405.102053](#)
2. I. Artsimovitch, R. Landick, The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* **109**, 193–203 (2002). [Medline doi:10.1016/S0092-8674\(02\)00724-9](#)
3. I. Gusarov, E. Nudler, The mechanism of intrinsic transcription termination. *Mol. Cell* **3**, 495–504 (1999). [Medline doi:10.1016/S1097-2765\(00\)80477-3](#)
4. R. Landick, J. Carey, C. Yanofsky, Detection of transcription-pausing in vivo in the trp operon leader region. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1507–1511 (1987). [Medline doi:10.1073/pnas.84.6.1507](#)
5. S. Proshkin, A. R. Rahmouni, A. Mironov, E. Nudler, Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504–508 (2010). [Medline doi:10.1126/science.1184939](#)
6. R. Landick, The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**, 1062–1066 (2006). [Medline doi:10.1042/BST0341062](#)
7. L. S. Churchman, J. S. Weissman, Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368–373 (2011). [Medline doi:10.1038/nature09652](#)
8. D. G. Vassylyev, M. N. Vassylyeva, A. Perederina, T. H. Tahirov, I. Artsimovitch, Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157–162 (2007). [Medline doi:10.1038/nature05932](#)
9. M. Palangat, T. I. Meier, R. G. Keene, R. Landick, Transcriptional pausing at +62 of the HIV-1 nascent RNA modulates formation of the TAR RNA structure. *Mol. Cell* **1**, 1033–1042 (1998). [Medline doi:10.1016/S1097-2765\(00\)80103-3](#)
10. E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, S. M. Block, Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460–465 (2005). [Medline doi:10.1038/nature04268](#)
11. W. J. Greenleaf, S. M. Block, Single-molecule, motion-based DNA sequencing using RNA polymerase. *Science* **313**, 801 (2006). [Medline doi:10.1126/science.1130105](#)
12. K. M. Herbert, A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick, S. M. Block, Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell* **125**, 1083–1094 (2006). [Medline doi:10.1016/j.cell.2006.04.032](#)
13. C. L. Chan, D. Wang, R. Landick, Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *J. Mol. Biol.* **268**, 54–68 (1997). [Medline doi:10.1006/jmbi.1997.0935](#)
14. A. Weixlbaumer, K. Leon, R. Landick, S. A. Darst, Structural basis of transcriptional pausing in bacteria. *Cell* **152**, 431–441 (2013). [Medline doi:10.1016/j.cell.2012.12.020](#)
15. M. L. Kireeva, M. Kashlev, Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8900–8905 (2009). [Medline doi:10.1073/pnas.0900407106](#)

16. J. Starmer, A. Stomp, M. Vouk, D. Bitzer, Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Comput. Biol.* **2**, e57 (2006). [Medline](#)  
[doi:10.1371/journal.pcbi.0020057](https://doi.org/10.1371/journal.pcbi.0020057)
17. R. A. Mooney, S. E. Davis, J. M. Peters, J. L. Rowland, A. Z. Ansari, R. Landick, Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell* **33**, 97–108 (2009). [Medline](#)  
[doi:10.1016/j.molcel.2008.12.021](https://doi.org/10.1016/j.molcel.2008.12.021)
18. D. B. Goodman, G. M. Church, S. Kosuri, Causes and effects of N-terminal codon bias in bacterial genes. *Science* **342**, 475–479 (2013). [Medline](#) [doi:10.1126/science.1241934](https://doi.org/10.1126/science.1241934)
19. G. Kudla, A. W. Murray, D. Tollervey, J. B. Plotkin, Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009). [Medline](#)  
[doi:10.1126/science.1170160](https://doi.org/10.1126/science.1170160)
20. Z. Horii, A. J. Clark, Genetic analysis of the recF pathway to genetic recombination in *Escherichia coli* K12: Isolation and characterization of mutants. *J. Mol. Biol.* **80**, 327–344 (1973). [Medline](#)  
[doi:10.1016/0022-2836\(73\)90176-9](https://doi.org/10.1016/0022-2836(73)90176-9)
21. T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2**, 0008 (2006). [Medline](#)  
[doi:10.1038/msb4100050](https://doi.org/10.1038/msb4100050)
22. S. K. Sharan, L. C. Thomason, S. G. Kuznetsov, D. L. Court, Recombineering: A homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223 (2009). [Medline](#)  
[doi:10.1038/nprot.2008.227](https://doi.org/10.1038/nprot.2008.227)
23. R. Higuchi, B. Krummel, R. K. Saiki, A general method of in vitro preparation and specific mutagenesis of DNA fragments: Study of protein and DNA interactions. *Nucleic Acids Res.* **16**, 7351–7367 (1988). [Medline](#) [doi:10.1093/nar/16.15.7351](https://doi.org/10.1093/nar/16.15.7351)
24. K. M. Arndt, M. J. Chamberlin, RNA chain elongation by *Escherichia coli* RNA polymerase. Factors affecting the stability of elongating ternary complexes. *J. Mol. Biol.* **213**, 79–108 (1990). [Medline](#) [doi:10.1016/S0022-2836\(05\)80123-8](https://doi.org/10.1016/S0022-2836(05)80123-8)
25. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009). [Medline](#)  
[doi:10.1186/gb-2009-10-3-r25](https://doi.org/10.1186/gb-2009-10-3-r25)
26. O. R. Homann, A. D. Johnson, MochiView: Versatile software for genome browsing and DNA motif analysis. *BMC Biol.* **8**, 49 (2010). [Medline](#) [doi:10.1186/1741-7007-8-49](https://doi.org/10.1186/1741-7007-8-49)
27. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004). [Medline](#) [doi:10.1101/gr.849004](https://doi.org/10.1101/gr.849004)
28. W. J. Greenleaf, M. T. Woodside, E. A. Abbondanzieri, S. M. Block, Passive all-optical force clamp for high-resolution laser trapping. *Phys. Rev. Lett.* **95**, 208102 (2005). [Medline](#)  
[doi:10.1103/PhysRevLett.95.208102](https://doi.org/10.1103/PhysRevLett.95.208102)
29. S. F. Tolić-Nørrelykke, A. M. Engh, R. Landick, J. Gelles, Diversity in the rates of transcript elongation by single RNA polymerase molecules. *J. Biol. Chem.* **279**, 3292–3299 (2004). [Medline](#) [doi:10.1074/jbc.M310290200](https://doi.org/10.1074/jbc.M310290200)

30. D. Nayak, M. Voss, T. Windgassen, R. A. Mooney, R. Landick, Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol. Cell* **50**, 882–893 (2013). [Medline doi:10.1016/j.molcel.2013.05.015](#)
31. X. Hu, S. Malik, C. C. Negroiu, K. Hubbard, C. N. Velalar, B. Hampton, D. Grosu, J. Catalano, R. G. Roeder, A. Gnatt, A Mediator-responsive form of metazoan RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9506–9511 (2006). [Medline doi:10.1073/pnas.0603702103](#)
32. M. N. Vassylyeva, J. Lee, S. I. Sekine, O. Laptenko, S. Kuramitsu, T. Shibata, Y. Inoue, S. Borukhov, D. G. Vassylyev, S. Yokoyama, Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from *Thermus thermophilus*. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1497–1500 (2002). [Medline doi:10.1107/S0907444902011770](#)
33. R. A. Mooney, K. Schweimer, P. Rösch, M. Gottesman, R. Landick, Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* **391**, 341–358 (2009). [Medline doi:10.1016/j.jmb.2009.05.078](#)
34. G. H. Feng, D. N. Lee, D. Wang, C. L. Chan, R. Landick, GreA-induced transcript cleavage in transcription complexes containing *Escherichia coli* RNA polymerase is controlled by multiple factors, including nascent transcript location and structure. *J. Biol. Chem.* **269**, 22282–22294 (1994). [Medline](#)
35. G. Cipres-Palacin, C. M. Kane, Alanine-scanning mutagenesis of human transcript elongation factor TFIIS. *Biochemistry* **34**, 15375–15380 (1995). [Medline doi:10.1021/bi00046a046](#)
36. R. Landick, D. Wang, C. L. Chan, Quantitative analysis of transcriptional pausing by *Escherichia coli* RNA polymerase: His leader pause site as paradigm. *Methods Enzymol.* **274**, 334–353 (1996). [Medline doi:10.1016/S0076-6879\(96\)74029-6](#)
37. I. Touloukhonov, J. Zhang, M. Palangat, R. Landick, A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell* **27**, 406–419 (2007). [Medline doi:10.1016/j.molcel.2007.06.008](#)
38. M. Palangat, M. H. Larson, X. Hu, A. Gnatt, S. M. Block, R. Landick, Efficient reconstitution of transcription elongation complexes for single-molecule studies of eukaryotic RNA polymerase II. *Transcription* **3**, 146–153 (2012). [Medline doi:10.4161/trns.20269](#)
39. C. L. Chan, R. Landick, Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. *J. Mol. Biol.* **233**, 25–42 (1993). [Medline doi:10.1006/jmbi.1993.1482](#)
40. D. Wang, T. I. Meier, C. L. Chan, G. Feng, D. N. Lee, R. Landick, Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell* **81**, 341–350 (1995). [Medline doi:10.1016/0092-8674\(95\)90387-9](#)
41. S. Kyzer, K. S. Ha, R. Landick, M. Palangat, Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J. Biol. Chem.* **282**, 19020–19028 (2007). [Medline doi:10.1074/jbc.M701483200](#)
42. W. Gilbert, N. Maizels, A. Maxam, Sequences of controlling regions of the lactose operon. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 845–855 (1974). [Medline doi:10.1101/SQB.1974.038.01.087](#)

43. M. Palangat, R. Landick, Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. *J. Mol. Biol.* **311**, 265–282 (2001). [Medline doi:10.1006/jmbi.2001.4842](#)
44. A. Bochkareva, Y. Yuzenkova, V. R. Tadigotla, N. Zenkin, Factor-independent transcription pausing caused by recognition of the RNA-DNA hybrid sequence. *EMBO J.* **31**, 630–639 (2012). [Medline doi:10.1038/emboj.2011.432](#)
45. V. A. Aivazashvili, R. Sh. Bibilashvili, R. M. Vartikian, T. A. Kutateladze, Effect of the primary structure of RNA on the pulse character of RNA elongation in vitro by *Escherichia coli* RNA polymerase: A model *Mol. Biol. (Mosk.)* **15**, 915–929 (1981). [Article in Russian] [Medline](#)
46. D. N. Lee, L. Phung, J. Stewart, R. Landick, Transcription pausing by *Escherichia coli* RNA polymerase is modulated by downstream DNA sequences. *J. Biol. Chem.* **265**, 15145–15153 (1990). [Medline](#)
47. Y. Zhang, Y. Feng, S. Chatterjee, S. Tuske, M. X. Ho, E. Arnold, R. H. Ebright, Structural basis of transcription initiation. *Science* **338**, 1076–1080 (2012). [Medline doi:10.1126/science.1227786](#)
48. S. E. LaFlamme, F. R. Kramer, D. R. Mills, Comparison of pausing during transcription and replication. *Nucleic Acids Res.* **13**, 8425–8440 (1985). [Medline doi:10.1093/nar/13.23.8425](#)
49. M. Palangat, C. T. Hittinger, R. Landick, Downstream DNA selectively affects a paused conformation of human RNA polymerase II. *J. Mol. Biol.* **341**, 429–442 (2004). [Medline doi:10.1016/j.jmb.2004.06.009](#)
50. M. T. Yang, J. F. Gardner, Isolation and footprint analysis of the *Escherichia coli* thr leader paused transcription complex. *Nucleic Acids Res.* **19**, 1671–1680 (1991). [Medline doi:10.1093/nar/19.7.1671](#)
51. J. M. Bartkus, B. Tyler, J. M. Calvo, Transcription attenuation-mediated control of leu operon expression: Influence of the number of Leu control codons. *J. Bacteriol.* **173**, 1634–1641 (1991). [Medline](#)
52. C. A. Hauser, J. A. Sharp, L. K. Hatfield, G. W. Hatfield, Pausing of RNA polymerase during in vitro transcription through the ilvB and ilvGEDA attenuator regions of *Escherichia coli* K12. *J. Biol. Chem.* **260**, 1765–1770 (1985). [Medline](#)
53. T. N. Wong, T. R. Sosnick, T. Pan, Folding of noncoding RNAs during transcription facilitated by pausing-induced nonnative structures. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 17995–18000 (2007). [Medline doi:10.1073/pnas.0705038104](#)
54. O. Laptenko, J. Lee, I. Lomakin, S. Borukhov, Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. *EMBO J.* **22**, 6322–6334 (2003). [Medline doi:10.1093/emboj/cdg610](#)
55. F. R. Blattner, G. Plunkett 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, Y. Shao, The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997). [Medline doi:10.1126/science.277.5331.1453](#)
56. P. R. Burkholder, N. H. Giles Jr., Induced biochemical mutations in *Bacillus subtilis*. *Am. J. Bot.* **34**, 345–348 (1947). [Medline doi:10.2307/2437147](#)

57. K. C. Neuman, E. A. Abbondanzieri, R. Landick, J. Gelles, S. M. Block, Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* **115**, 437–447 (2003). [Medline](#)  
[doi:10.1016/S0092-8674\(03\)00845-6](https://doi.org/10.1016/S0092-8674(03)00845-6)