Supporting Information

SI Appendix

A. Additional Methods

DNA Preparations

Two DNA templates were prepared for the single-molecule experiments. A 'downstream' DNA template was used for real-time single-molecule transcription assays to measure the downstream stall torque; while a 'calibration' DNA template was used in the control experiment to measure the drift velocity. Both DNA templates contain a T7A1 promoter, and the preparation of these two templates was reported previously (1). For the 'downstream' DNA template, a 4.7-kb DNA segment was amplified by PCR from the plasmid pMDW2 (sequence available upon request) and digested via Basl-HF at 37 °C for 1 hour to generate a 4-bp 5'overhang downstream of the T7A1 promoter. The digested fragment was then ligated to a 66bp DNA torsional-constraint anchor with 6 biotin linkers evenly spaced for attachment to the streptavidin-coated bottom surface of a quartz cylinder at 16 °C overnight. The 'calibration' template was the same as the 'downstream' template except that the upstream end of the T7A1 promoter was also digested to generate a 4-bp 5'-overhang and ligated to a 6-digoxigenin linker with dig tags evenly spaced for attachment to an anti-dig coated coverslip surface. The sequences of these two templates are available upon request.

Purification of E. coli RNAP

E. coli RNAP containing one hemagglutinin (HA) epitope tag on the C-terminus of each α-

subunit was expressed and purified as previously described (2, 3). Briefly, RNAP was expressed at low levels in DH-5 α -competent *E. coli* transformed with the plasmid pKA1 in Superbroth with 100 µg/ml ampicillin at 37 °C until OD_{600nm} reached 2.1. IPTG was added to a final concentration of 1 mM, and the cells were grown for an additional 4 h, harvested by centrifugation and stored at -20 °C. The transformed cells were re-suspended at 3 ml/g in RNAP Lysis Buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM EDTA, 1 mM DTT, and 5% v/v glycerol). Lysozyme was added to 300 μ g/ml, and the cells were incubated for 20 min on ice. Further lysis was achieved by sonication via a Branson Sonifier 250 with 60% duty cycle in small aliquots (< 20 ml). The cell debris was pelleted by centrifugation at 16,300 g for 45 min at 4 °C and the supernatant containing DNA and DNA-bound proteins was collected. Cleared 5% (w/v) polyethyleneimine (PEI) was added dropwise, with slow stirring, to a final concentration of 0.4% (w/v) in order to precipitate nucleic acids and their bound proteins. The DNA and associated RNAP were pelleted from solution and after five washes in a buffer containing 350 mM NaCl, RNAP was eluted from PEI and DNA with a buffer containing 1 M NaCl. The eluted RNAP was purified to homogeneity through chromatography on 3 columns: first on a HiPrep Heparin FF 16/10 column (GE), followed by a HiPrep 26/60 Sephacryl S-300 HR column (GE), and last on a Ni-NTA Superflow column (QIAGEN). Fractions containing pure RNAP were pooled, concentrated and dialyzed against RNAP Storage Buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 50% v/v glycerol) and stored at -20 $^{\circ}$ C.

Purification of GreB

Plasmid pES3, encoding GreB-6xHis in pET-28b(+) (4), was transformed into BL21(DE3) cells for protein overexpression. Cells were grown in LB media supplemented with 50 μ g/ml of

kanamycin at 37 °C until the OD_{600nm} reached 0.6-0.8, and then were induced with 1 mM IPTG. After induction at 37 °C for 3 hours, cells were harvested and stored at -80 °C. The protocol for the purification of GreB proteins was similar to those previous described (4, 5). In brief, the cells were thawed on ice and re-suspended in GreB Lysis Buffer (50 mM Tris-HCl pH 6.9, 500 mM NaCl, and 5% v/v glycerol) with EDTA-free protease-inhibitor cocktail (Roche) and lysozyme (300 µg/ml). The suspension was incubated on ice for 1 hour and followed by a brief sonication to further lyse the cells. The extract was cleared by centrifugation (24,000 g, 20 min at 4 °C) and passed twice through a 0.45-µm filter. The proteins were isolated by a Ni-NTA agarose column and eluted in GreB Lysis Buffer with 200 mM imidazole. The elute was further purified on a Superdex 200 column (GE Healthcare) with Elution Buffer (10 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT, 1 mM EDTA, and 5% v/v glycerol). The purified proteins were dialyzed into GreB storage buffer (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM EDTA, and 50% v/v glycerol), frozen in liquid nitrogen and stored at -80 °C.

Paused Transcription Complexes Assembly

The transcription elongation complexes were formed in bulk and paused at +20 position via nucleotide depletion (1, 3). In brief, 25 nM HA-tagged *E. coli* RNAP was incubated with 5 nM 'Downstream' DNA templates, 250 µM ApU (initiating dinucleotide), and 50 µM GTP/ATP/CTP in the Transcription Buffer (25 mM Tris-HCl pH 8.0, 100 mM KCl, 4 mM MgCl₂, 1 mM DTT, and 3% v/v glycerol) with 1 unit/µl SUPERase-InTM RNase inhibitor and 0.15 mg/ml acetylated BSA at 37 °C for 40 minutes. 0.2 µg/µl heparin was added to the paused transcription complexes (PTCs) in some preparations. PTCs were kept at 4 °C before being used in single molecule assays. Under our single molecule experimental conditions, RNAP molecules not associated with the surface

of the sample chamber were removed prior to the introduction of NTPs. This effectively minimized re-initiation from the same promoter even in the absence of heparin (2).

DNA Supercoiling Characterization

We performed a similar DNA supercoiling property characterization as described before in order to extract desired information (such as stall torque, RNAP template position, transcription velocity, etc.) from measured data (such as force, DNA extension, etc.) (1). These were conducted using the 'calibration' DNA template in the transcription buffer (SI Appendix, Fig. S6).

Instrument Drift Estimate

A control experiment that mimics the real-time single-molecule transcription assay was employed to estimate instrument drift (SI Appendix, Fig. S1). In brief, the 'calibration' DNA had one end tonsionally anchored on a coverslip surface coated with anti-dig while the other end was torsionally constrained to the bottom surface of a quartz cylinder via multiple biotin/streptavidin linkages. The DNA was first stretched and mechanically wound by 12 turns under the force clamp of 0.3 pN to generate the (+) plectoneme. Then the force clamp was turned off. The trap position was held constant for 60 s to mimic the stall torque measurement step in the real-time transcription assay and the extension change was recorded. The drift velocity was obtained by the linear-fitting of the extension change.

Real-Time Single-Molecule Transcription Assay

The angular-optical-trap-based single-molecule transcription assays were performed as previously described (1). Briefly, an RNAP in the PTC was torsionally anchored on a coverslip

surface coated with anti-HA through two HA/anti-HA interactions; while the downstream end of DNA was torsionally constrained to the bottom surface of a quartz cylinder via multiple biotin/streptavidin linkages (Fig. 1A). Transcription was resumed by the addition of the transcription buffer supplemented by 1 mM NTPs as well as 2.5 mM protocatechuic acid (PCA), and 20 nM protocatechuate-3,4-dioxygenase (PCD) to reduce photo-damage (6) as indicated. During an experiment, the trapping laser power entering the objective was kept at 15 mW. DNA was mechanically unwound (4 turns/s) by no more than 44 turns under a force clamp of 0.3 pN until a (-) plectoneme of 300 nm in extension was formed. (Fig. 1B). Subsequent translocation of RNAP neutralized the (-) plectoneme and led to the (+) plectoneme formation. The force clamp was then turned off. Subsequent RNAP translocation increased the resisting force and the corresponding torque until reaching a stall. For experiments in the presence of GreB factor, 1 μ M of GreB was mixed with 1 mM NTPs to restart transcription elongation. These singlemolecule transcription assays were carried out in an environmentally controlled room at 23.2 ± 0.5 °C.

For the transcription resumption assay, after RNAP was stalled for ~60s, the DNA was unwound by 12 turns at 3 turns/s to release the torsion in DNA and then the optical trap was turned off for ~ 18 s, during which RNAP's activity was not monitored. Subsequently the same process as shown in Fig. 1B was repeated. The resumption time is defined as the time duration between torque relaxation and subsequent detection of RNAP forward translocation.

Data Acquisition and Analysis

During a transcription experiment, RNAP buckled the DNA and was allowed to stall under

an increasing torque (Fig. 1A). The DNA contour length and torque as a function of time were obtained from the measured force, DNA extension, and cylinder orientation as previously derived (1). Raw data for measured force (*F*), DNA extension (*z*), and cylinder orientation were low-pass filtered to 1 kHz and digitized at 2 kHz. The RNAP template position relative to the transcription start site was thus determined from the DNA contour length and the known location of the transcription start site (TSS). The stall torque was taken as the torque value averaged over 2 s upon stalling.

For pause-free velocity analysis, the converted RNAP template position data were smoothed using a Gaussian weight function with a standard deviation of 0.1 s and pauses were identified as previously reported (1). The cutoff duration was set to be 0.2 s (i.e., an RNAP was considered to have paused if the time it spent at a single nucleotide position was greater than 0.2 s). The 0.2 s was found to be 4 times the most likely dwell time which represented the baseline. Pause-free velocity was determined for the data segments between pauses.

To obtain RNAP backtracking velocity at a stall versus torque (Fig. 3B), we took into account the decrease in torque as RNAP backtracked in each trace (see an example in Fig. 1C). Each trace was first divided into 20 s segments, and for each segment, the RNAP velocity and the mean torque was determined. RNAP velocities from different segments of all traces were then binned according to their torque values before being averaged.

For the variance analysis shown in Fig. 3C, we estimate the contribution from the Brownian motion to be ~ 10.7 bp² after taking into account of the low pass filtering of the data. This variance corresponds to an essentially negligible offset to the overall variance.

B. Model of Backtracking as a 1D Diffusion Process Biased by Torque

When RNAP stalls and backtracks in the absence of GreB, we treated the backtracking process as a 1D diffusion process that is biased by the applied torque (τ). We assume a uniform backtracking energy landscape where hopping to an adjacent backtracked state has the same rate k_0 in the absence of torque. This treatment does not take into account any sequence-dependence of the hopping rate. Thus k_0 may only be considered as an 'effective' rate. We further assume that the activation barrier is located at the mid-point between two adjacent states. Thus backtracking velocity is

$$v_{-} = k_0 d \left[e^{-\frac{\omega_0 d}{2k_{\rm B}T}\tau} - e^{+\frac{\omega_0 d}{2k_{\rm B}T}\tau} \right],$$
[S1]

where $\omega_0 = \frac{2\pi}{3.55 \text{ nm}} = 1.77 \text{ nm}^{-1}$ is a conversion factor from translocation distance to rotational angle, d = 0.338 nm is the distance between two adjacent barriers of hopping, and $k_{\text{B}}T$ is the thermal energy. Fitting Eq. S1 to data from Figure 3B (top) yielded $k_0 = 0.26 \text{ s}^{-1}$.

C. Two-State Model of Backtracking Rescue by GreB

In the presence of GreB, GreB may rescue a backtracked RNAP and bring it into active elongation (Fig. 3D). Thus the stalling process consists of transitions from two states: backtracking and active elongation. The stochastic transitions between the two states may be described as reversible first order reactions with reaction rates k_{-} and k_{+} , corresponding to the transition rate from elongation state to backtracking state and the transition rate from backtracking state to elongation state, respectively. During backtracking, RNAP moves at a velocity v_{-} ; during active elongation, it moves at a velocity v_{+} . We sought to determine k_{-} and k_{+} in terms of measured RNAP positions and measured velocities v_{-} and v_{+} .

C1. Markov model

We first considered these two-state transitions during an infinitely small time interval Δt (SI Appendix, Fig. S7). During this time, the probabilities of transitioning from '-' to a '+' state and from '+' to '-' state are $\alpha = k_+ \cdot \Delta t$ and $\beta = k_- \cdot \Delta t$ respectively. These transitions follow the Markov process and the probabilities being in the '+' and '-' states evolve as

$$\begin{pmatrix} P_{-}^{(i+1)} \\ P_{+}^{(i+1)} \end{pmatrix} = \begin{pmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{pmatrix} \begin{pmatrix} P_{-}^{(i)} \\ P_{+}^{(i)} \end{pmatrix}$$
[S2]

with the transition matrix:

$$T = \begin{pmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{pmatrix},$$
 [S3]

C2. The eigenvalues and eigenvectors of transition matrix *T*

The eigenvalues and eigenvectors of the transition matrix can be calculated as follows

$$0 = T - \lambda I = \begin{pmatrix} 1 - \alpha - \lambda & \beta \\ \alpha & 1 - \beta - \lambda \end{pmatrix}$$
 [S4]

The two eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 1 - \alpha - \beta$. It is worth noting that $1 - \alpha - \beta \subset$ (-1,1).

The eigenvectors satisfy:

$$T\boldsymbol{V}_k = \lambda_k \boldsymbol{V}_k$$
[S5]

where V_k is the k-th eigenvector (k = 1,2) that corresponds to λ_k . We find the eigenvectors by solving Eq. S5:

$$V_1 = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}, V_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
 [S6]

We define:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0\\ 0 & \lambda_2 \end{pmatrix}, \ V = (V_1 \quad V_2)$$
[S7]

$$TV = T(V_1 \quad V_2) = (\lambda_1 V_1 \quad \lambda_2 V_2) = (V_1 \quad V_2)\Lambda = V\Lambda$$
[S8]

Therefore

$$T = V\Lambda V^{-1}$$
[S9]

and

$$T^N = V\Lambda^N V^{-1}.$$
 [S10]

At time t, $N = \frac{t}{\Delta t}$ number of Δt steps will be have been taken. The *N*-step transition matrix can be written as:

$$T^{N} = VA^{N}V^{-1} = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta & 1\\ \alpha & -1 \end{pmatrix} \begin{pmatrix} 1 & 0\\ 0 & (1-\alpha-\beta)^{N} \end{pmatrix} \begin{pmatrix} 1 & 1\\ \alpha & -\beta \end{pmatrix} = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta+s\alpha & \beta-s\beta\\ \alpha-s\alpha & \alpha+s\beta \end{pmatrix},$$
[S11]

where $s = (\alpha + \beta - 1)^N$. When $N \to \infty$, $s \to 0$,

$$T^{\infty} = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & \beta \\ \alpha & \alpha \end{pmatrix}.$$
 [S12]

which gives the probabilities when the two states are under equilibrium and is independent of the initial states.

C3. Velocity correlation

Although we ultimately wished to find the solution of position correlation, it was mathematically more conducive to start by finding the velocity correlation. We assumed that at the starting point of measurement, the system is in equilibrium. The velocity correlation between the initial state and the *N*th-step can be written as $\langle v_0 v_N \rangle$, which equals to the average of all possible combinations of correlation of stochastic velocity variables, e.g.,

$$\langle v_0 v_N \rangle = v_- P_-^{\infty} \cdot (v_- v_+) \begin{pmatrix} P_{(-|-)}^N \\ P_{(+|-)}^N \end{pmatrix} + v_+ P_+^{\infty} \cdot (v_- v_+) \begin{pmatrix} P_{(-|+)}^N \\ P_{(+|+)}^N \end{pmatrix} = (v_- v_+) T^N \begin{pmatrix} v_- P_-^{\infty} \\ v_+ P_+^{\infty} \end{pmatrix},$$
[S13]

where P_{-}^{∞} and P_{+}^{∞} are the probabilities of state '-' and '+' at 0th-step respectively. $P_{(i|j)}^{N}$ is the conditional probability that the state is j at 0th step and the state becomes i after N steps. Plugging T^{N} , P_{-}^{∞} , and P_{+}^{∞} into Eq. S13, we have

$$\langle v_0 v_N \rangle = \frac{1}{\alpha + \beta} \begin{pmatrix} v_- & v_+ \end{pmatrix} \begin{pmatrix} \beta + s\alpha & \beta - s\beta \\ \alpha - s\alpha & \alpha + s\beta \end{pmatrix} \begin{pmatrix} v_-\beta/(\alpha + \beta) \\ v_+\alpha/(\alpha + \beta) \end{pmatrix}$$
[S14]

To covert this expression to a continuous form, we take the limit that the time step interval $\Delta t \rightarrow 0$. We have thus obtained the velocity correlation as a function of time:

$$\langle v_0 v_N \rangle = A + B e^{-kt}, \tag{S15}$$

where
$$A = \left(\frac{k_-v_- + k_+v_+}{k_- + k_+}\right)^2$$
, $B = k_-k_+ \left(\frac{v_- - v_+}{k_- + k_+}\right)^2$, and $k = k_- + k_+$

C4. Position correlation

We assume that all position trajectories, x(t), are the samples from the same ensemble, therefore:

$$\left\langle \left(x(t) - x(0)\right)^2 \right\rangle = \left\langle \left(\int_0^t v(t_1)dt_1\right)^2 \right\rangle = \left\langle \int_0^t v(t_1)dt_1 \int_0^t v(t_2)dt_2 \right\rangle = \int_0^t \int_0^t \langle v(t_1)v(t_2) \rangle dt_1 dt_2$$

$$= \int_0^t \int_0^t \langle v(0)v(|t_2 - t_1|) \rangle dt_1 dt_2 = \int_0^t \int_0^t [A + Be^{-k|t_2 - t_1|}] dt_1 dt_2$$

$$= 2\int_0^t dt_2 \int_0^{t_2} [A + Be^{-k(t_2 - t_1)}] dt_1$$

$$= At^2 + \frac{2B}{k}t - \frac{2B}{k^2}[1 - e^{-kt}]$$

[S1	61
L 🗠 🖚	~1

It is worth noting that line 2 of Eq. S16 assumes that the system has already reached equilibrium at the time t = 0, so $\langle v(t_1)v(t_2) \rangle = \langle v(0)v(|t_2 - t_1|) \rangle$. This is true in practice since at t = 0, we only know the probabilities that RNAP is in either an elongation state or backtracking state.

In Eq. S16, the first term is a linear drift of the mean positon:

$$\bar{x}(t) = \frac{k_{-}v_{-} + k_{+}v_{+}}{k_{-} + k_{+}}t$$
[S17]

Therefore, variance of the position is

$$\sigma_x^2(t) = \left\langle \left(x(t) - x(0) \right)^2 \right\rangle - \langle x(t) \rangle^2 = \frac{2B}{k} t - \frac{2B}{k^2} [1 - e^{-kt}]$$
[S18]

When $kt \gg 1$, the variance evolves with time as follows

$$\sigma_{\chi}^{2}(t) = \frac{2B}{k}t = \frac{2k_{-}k_{+}}{(k_{-}+k_{+})^{3}}(\nu_{-}-\nu_{+})^{2}t$$
[S19]



Figure S1. Instrument drift velocity. (A) Experimental steps for a control experiment performed to assess the AOT instrument's drift velocity during a stall. (B) An example trace of the extension of a (+) buckled DNA molecule held under an initial force of 0.3 pN. The slight change of the extension over time for this trace was ~ -0.022 nm/s based on a linear fit (black line) as a result of instrument drift. This corresponds to a drift velocity of ~0.05 bp/s at a stall during a transcription experiment. (C) The histogram of the measured instrument drift velocity. The mean obtained from a Gaussian fitting is 0.06 ± 0.01 bp/s (mean ± SEM).



Figure S2. Pause-free active elongation velocity. The pause-free velocities were obtained from traces with a stalling torque between 16-24 pN·nm which is close to the mean stall torque with GreB (18.5 pN·nm) indicated by the dashed line. For each trace, active elongation velocity was determined between pauses immediately prior to stalling (see Methods). All velocities were binned according to their torque values and averaged. The error bars for torque are SDs and the error bars for active elongation velocity are SEMs. At 18.5 pN·nm, pause-free elongation velocities in the absence and presence of GreB have similar values, ~ 3.8 bp/s.

The pause-free active velocity represents active translocation along the main reaction pathway of transcription (7). The decrease with an increase in the resisting torque should be due to hindrance by the torque to translocation from the pre- to the post-translocated state.



Figure S3. RNAP position versus time at a stall in the presence of GreB. The gray curve shows the averaged trajectory of RNAP upon stalling and the black solid line is its linear fit passing through the origin. The fit yielded a mean velocity of 0.05 bp/s.



Figure S4. Rate to enter backtracking (k_{-}) and GreB rescue rate (k_{+}) obtained from variance analysis as a function of backtracking rate $(v_{-}, top row)$ and forward active elongation rate $(v_{+}, bottom row)$. Dashed lines indicate values used in the manuscript.



Figure S5. Transcription resumption in the presence of GreB, either in the presence of absence of PCA/PCD. Error bars indicate standard error of means (SEM). Shown are the fractions of RNAPs that resumed transcription after torque release as a function of the time after torque relaxation. Because the transcription resumption rates with and without PCA/PCD were similar, these two sets of data were combined in Fig. 4.



Figure S6. Characterization of extension and torque during (+) DNA supercoiling. Measurements were carried out as previously described (1). (A) and (B) show examples of the measured extension and torque of DNA, respectively, as a DNA molecule was (+) supercoiled under a constant force of 1.0 pN. Note that once DNA was buckled, torque value remained constant. The dashed vertical line indicates the onset of a bucking or melting transition. We also characterized (C) post-buckling extension per turn Δz_b and (D) post-buckling torque τ_b as a function of force F. These data are in agreement with previous measurements (1). In addition, we fit post-buckling extension per turn to an exponential function, yielding $\Delta z_b = 28.5 + 42.9e^{-1.37}$. Post-buckling torque was fit to a previously derived expression (1): $\tau_b =$



parameters were measured: intrinsic twist persistence length C = 100 nm (8) and bending persistent length $L_p = 43$ nm (9). The fit yielded P = 27 nm.



Figure S7. A cartoon to graphically illustrate the two-state Markov model.

SI References

- 1. Ma J, Bai L, & Wang MD (2013) Transcription Under Torsion. *Science* 340(6140):1580-1583.
- 2. Le TT, et al. (2018) Mfd Dynamically Regulates Transcription via a Release and Catch-Up Mechanism. *Cell* 172(1-2):344-357 e315.
- 3. Adelman K, et al. (2002) Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc Natl Acad Sci USA* 99(21):13538-13543.
- 4. Strobel EJ & Roberts JW (2014) Regulation of promoter-proximal transcription elongation: enhanced DNA scrunching drives lambda Q antiterminator-dependent escape from a sigma 70dependent pause. *Nucleic Acids Res* 42(8):5097-5108.
- 5. Perederina AA, *et al.* (2006) Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription elongation factors GreB from *Escherichia coli* and Gfh1 from Thermus thermophilus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 62:44-46.
- 6. Aitken CE, Marshall RA, & Puglisi JD (2008) An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys J* 94(5):1826-1835.
- 7. Bai L, Fulbright RM, & Wang MD (2007) Mechanochemical kinetics of transcription elongation. *Phys Rev Lett* 98(6):068103.
- 8. Forth S, *et al.* (2008) Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. *Phys Rev Lett* 100(14).
- 9. Wang MD, Yin H, Landick R, Gelles J, & Block SM (1997) Stretching DNA with optical tweezers. *Biophys J* 72(3):1335-1346.