Supporting Online Material for

Initial Transcription by RNA Polymerase Proceeds Through a DNA-Scrunching Mechanism

Achillefs N. Kapanidis,* Emmanuel Margeat, Sam On Ho, Ekaterine Kortkhonjia, Shimon Weiss,* Richard H. Ebright*

*To whom correspondence should be addressed. E-mail: a.kapanidis1@physics.ox.ac.uk., sweiss@chem.ucla.edu, ebright@waksman.rutgers.edu

Published 17 November 2006, *Science* **314**, 1144 (2006)
DOI: 10.1126/science.1131399

This PDF file includes

Supplement: Materials and Methods

Labeled proteins and DNA. σ^{70} derivatives labeled with tetramethylrhodamine at residue 366, 396, 569, or 596, and DNA fragments end-labelled with Cy5, were prepared as in *S1* and *S2*. DNA fragments internally labeled with Cy3B and/or Alexa647 were prepared essentially as in *S1* and *S2*, but using primers labeled internally on amino-dT residues [prepared using Cy3B NHS ester (GE Healthcare, Inc.) and Alexa647 NHS-ester (Molecular Probes, Inc.); *S3*]. Control experiments established that labeled proteins and DNA fragments were functional in open-complex formation and promoter escape. Sequences of DNA fragments are shown in Fig. S1.

Transcription complexes. Reaction mixtures for preparation of transcription complexes contained (30 µl): 60 nM RNAP holoenzyme (Epicentre, Inc.), or 100 nM RNAP core (Epicentre, Inc.) and 80 nM labeled σ^{70} derivative, and 0 or 250 nM rifampicin in transcription buffer (TB; 50 mM Tris-HCl, pH 8, 100 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, 100 µg/ml bovine serum albumin, and 5% glycerol). Samples were incubated 20 min at 30˚C, 0.6 µl of 1 µM labeled DNA fragment was added, and samples were further incubated 15 min at 37˚C. Heparin-Sepharose (GE Healthcare, Inc.; 0.8 µl of 100 mg/ml suspension) was added to disrupt non-specific RNAP-promoter complexes and to remove free RNAP (*S1*), and, after 1 min at 37˚C, samples were centrifuged, and 9.5 µl aliquots were transferred to tubes containing 0.5 µl 10 mM ApA at 37°C (for RP_o) or 0.5 µl 10 mM ApA, 0.5 mM UTP, and 0.5 mM GTP at 37°C (for $RP_{itc,\leq 7}$) at 37°C. [Experiments with RP_o were performed in the presence of the initiating dinucleotide, ApA. The initiating dinucleotide increases stability of complexes (*S4, S5*) and reduces dissociation of complexes during data collection (*S6, S7*). Representative experiments with RP_o in the absence of the initiating dinucleotide yield equivalent results, but inferior signal-to-noise ratio (not shown)].

Single-molecule fluorescence microscopy. Sample preparation, alternating-laser-excitation microscopy, data acquisition, and data analysis were as in *S6, S8*. Transcription complexes were observed for 15-30 min at 37°C at a final concentration of 100 pM in buffer KG7 (40 mM HEPES-NaOH, pH 7, 100 mM potassium glutamate, 10 mM $MgCl_2$, 1 mM DTT, 100 µg/ml BSA, 5% glycerol, 1 mM mercaptoethylamine) containing 0.5 mM ApA at 37°C (for RP_o) or 0.5 mM ApA, 25 µM UTP, and 25 µM GTP at 37°C (for $RP_{itc,\leq 7}$) and, where indicated, also containing 1 nM rifampicin. Excitation intensities were 150-300 µW for 514-nm excitation (D_{exc}), and 50-80 µW for 638-nm excitation (A_{exc}) (measured in the continuous-wave mode). Photons detected at the donor and acceptor emission channels were assigned to 514-excitation or nm or 638-nm excitation based on arrival time, and were used to generate streams $f_{D_{exc}}^{A_{em}}$, $f_{D_{exc}}^{D_{em}}$, $f_{A_{exc}}^{A_{em}}$, and $f_{A_{exc}}^{D_{em}}$ (where $f_{X_{exc}}^{Y_{em}}$ stands for counts per integration period, in the primary spectral range for detecting fluorophore Y, resulting from excitation that primarily excites fluorophore X). $f_{A_{exc}}^{A_{em}}$ thresholds of 7-9 photons per 500 µs and 15-30 photons per burst were used to identify acceptor-containing molecules of appreciable photon count (thus reducing the statistical noise inherent in single-molecule measurements). Signal-to-background ratios for detection of single molecules depended on the fluorophore used, the location of the fluorophore on the labeled biomolecule, and the instrument alignment. In all cases, signal-to-background ratios for detection of single molecules were >15:1. Typically, rates of fluorescence from detected single molecules were 30-400 KHz (photons/s) at the donor-emission channel and 15-200 KHz at the acceptor-emission channel, and rates of background at the respective wavelength ranges were <2 KHz and <1 KHz.
The apparent donor-acceptor stoichiometry parameter, *S*, was calculated as (*S6, S8*):

$$S = \left(F_{D_{exc}}^{A_{em}} + \gamma F_{D_{exc}}^{D_{em}}\right) \Big/ \left(F_{D_{exc}}^{A_{em}} + \gamma F_{D_{exc}}^{D_{em}} + F_{A_{exc}}^{A_{em}}\right) \qquad (1)$$

where γ is a detection correction factor (*8*; 0.23-1.00 in this work, measured using the methods described in *S8*).

The apparent donor-acceptor energy-transfer-efficiency parameter, E^*, was calculated as (*S6, S8*):

$$E^* = F_{D_{exc}}^{A_{em}} \Big/ \left(F_{D_{exc}}^{A_{em}} + F_{D_{exc}}^{D_{em}} \right) \qquad (2)$$

E^*-S histograms (Fig. 1B, right panel) permit identification of species. The parameter S permits identification of molecules containing both donor and acceptor (S = 0.4-0.9; desired species), molecules containing only a donor (S >0.9; undesired species, arising from the presence of free σ^{70} molecules and buffer impurities), and molecules containing only an acceptor (S <0.4; undesired species, arising from the dissociation of non-specific complexes upon heparin challenge during preparation of RP_o; see Materials and Methods, Transcription Complexes).

Values of E^* for donor-acceptor species were constant within error during the period of data acquisition (15-30 min).

In experiments showing changes in measured distances (Figs. 2A, 4A, S2, S3, S7, S8), E^* distributions are broader for $RP_{itc,\leq 7}$ than for RP_o (e.g., SD = 0.127 vs. SD = 0.093 in Fig. 2A). In part this reflects higher statistical noise for E^* distributions having mean E^* closer to 0.5 (see *S9*). In part this also reflects the presence in $RP_{itc,\leq 7}$ of minor populations of complexes containing RNA products 2-6 nt in length as well as the major population of complexes containing RNA products 7 nt in length (see *S10*).

Calculation of accurate energy-transfer efficiencies and corresponding distances. Apparent donor-acceptor energy-transfer efficiencies, E^*, were converted to absolute donor-acceptor energy-transfer efficiencies, E, with precision of ~±0.01, as in *S8* (Table S1). Distances were obtained as $R = R_0 \left[\left(1/E \right) - 1 \right]^{1/6}$, using R_o = 58-65, 58-65, and 68 Å for, respectively, the TMR-Cy5, TMR-Alexa647, and Cy3B-Alexa647 donor-acceptor pairs (determined for each complex using procedures as in *S8, S11*; Table S1).

Supplement: References

S1. J. Mukhopadhyay, A. Kapanidis, V. Mekler, E. Kortkhonjia, Y. Ebright, R. Ebright, *Cell* **106**, 453-463 (2001).

S2. J. Mukhopadhyay, V. Mekler, E. Kortkhonjia, A. Kapanidis, Y. Ebright, R. Ebright, *Meths. Enzymol.*, 144-159 (2003).

S3. A. Kapanidis, N. Lee, T. Laurence, S. Doose, E. Margeat, S. Weiss, *Proc. Natl. Acad. Sci. USA* **101**, 8936-8941 (2004).

S4. T. Gaal, M. Bartlett, W. Ross, C. Turnbough, R. Gourse, *Science* **278**, 2092-2097 (1997).

S5. A. Revyakin, R. Ebright, T. Strick, *Proc. Natl. Acad. Sci. USA* **101**, 4776-4780 (2004).

S6. A. Kapanidis, E. Margeat, T. Laurence, S. Doose, S. Ho, J. Mukhopadhyay, E. Kortkhonjia, V. Mekler, R. Ebright, S. Weiss, *Mol. Cell* **20**, 347-356 (2005).

S7. E. Margeat, A. Kapanidis, P. Tinnefeld, Y. Wang, J. Mukhopadhyay, R. H. Ebright, S. Weiss, *Biophys. J.* **90**, 1419-1431 (2006).

S8. N. Lee, A. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. Ebright, S. Weiss, *Biophys. J.* **88**, 2939-2953 (2005).

S9. E. Nir, X. Michalet, K. Hamadani, T. Laurence, D. Neuhauser, Y. Kovchegov, S. Weiss, *J. Phys. Chem.*, in press.

S10. E. Margeat, A. Kapanidis, P. Tinnefeld, Y. Wang, J. Mukhopadhyay, R. H. Ebright, S. Weiss, *Biophys. J.* **90**, 1419-1431 (2006).

S11. A. Kapanidis, Y. Ebright, R. Ludescher, S. Chan, R. Ebright, *J. Mol. Biol.* **312**, 453-468 (2001).

S12. C. Lawson, D. Swigon, K. Murakami, S. Darst, H. Berman, R. Ebright, *Curr. Opin. Structl. Biol.* **14**, 10-20 (2004).

S13. E. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, S. Darst, *Cell* **104**, 901-912 (2001).

Supplement: Supplemental Figure Legends

Fig. S1. DNA fragments.
Black boxes, transcription start site (with arrow), promoter -10 element, and promoter -35 element. Blue legend, figure numbers for figures presenting results with the DNA fragment.
(A) DNA fragments used in experiments in main-text figures.
(B) DNA fragments used in experiments in supplemental figures.

Fig. S2. RNAP leading-edge moves relative to downstream DNA in initial transcription: additional measurements.
Experiments documenting movement of the RNAP leading edge relative to downstream DNA [tetramethylrhodamine as donor at σ^{70} residue 366 or residue 396 (both located in σR2, the σ^{70} domain responsible for recognition of the promoter -10 element); Cy5 as acceptor at DNA position +25, +20 or +15]. *Left:* Structural model of RP_o (*S12*) showing positions of donor (green circle) and acceptor (red square). RNAP core is in gray; σ^{70} is in yellow; the DNA template and nontemplate strands are in red and pink, respectively. *Right:* E* histograms for RP_o and $RP_{itc,\leq 7}$. The vertical line and vertical dashed line mark mean E* values for RP_o and $RP_{itc,\leq 7}$, respectively. *Bottom:* Predictions of the three models.

Fig. S3. RNAP leading-edge moves relative to downstream DNA in initial transcription: control experiment.
Experiment documenting that addition of rifampicin (an inhibitor that prevents synthesis of RNA products >2 nt in length; *S13*) prevents movement of RNAP leading-edge relative to downstream DNA [tetramethylrhodamine as donor at σ^{70} residue 366 located in σR2, the σ^{70} domain responsible for recognition of the promoter -10 element); Cy5 as acceptor at DNA position +20]. E* histograms as in Fig S2.

Fig. S4. RNAP trailing-edge does not move relative to upstream DNA in initial transcription: additional measurements.
Experiment documenting absence of movement of the RNAP trailing edge relative to downstream DNA [tetramethylrhodamine as donor at σ^{70} residue 596 (located in σR4, the σ^{70} domain responsible for recognition of the promoter -35 element); Cy5 as acceptor at DNA position -39]. Subpanels as in Fig. S2.

Fig. S5. RNAP trailing-edge does not move relative to upstream DNA in initial transcription: control measurements.
Experiment documenting that our approach is able to detect a 5-bp increase in distance between the RNAP trailing edge and upstream DNA. *Left:* E* histograms for RP_o containing tetramethylrhodamine as donor at σ^{70} residue 569 (located in σR4, the σ^{70} domain responsible for recognition of the promoter -35 element) and Cy5 as acceptor at DNA position -39 (top) or position -44 (bottom). *Right:* E* histograms for RP_o containing tetramethylrhodamine as donor at σ^{70} residue 596 (located in σR4, the σ^{70} domain responsible for recognition of the promoter -35 element) and Cy5 as acceptor at DNA position -39 (top) or position -44 (bottom).

Fig. S6. Initial transcription does not involve inchworming: additional measurements.
(A) Experiment documenting absence of movement of the RNAP leading edge relative to -10/-35 spacer DNA [tetramethylrhodamine as donor at σ^{70} residue 396 (located in σR2, the σ^{70} domain responsible for recognition of the promoter -10 element); Alexa647 as acceptor at DNA position -20]. Subpanels as in Fig. S2.
(B) Experiment documenting absence of movement of the RNAP trailing edge relative to -10/-35 spacer DNA [tetramethylrhodamine as donor at σ^{70} residue 596 (located in σR4, the σ^{70} domain responsible for

recognition of the promoter -35 element); Alexa647 as acceptor at DNA position -20]. Subpanels as in Fig. S2.

Fig. S7. Initial transcription involves scrunching: additional measurement.
Experiment documenting contraction of DNA between positions -15 and +20 [Cy3B as donor at DNA position -15; Alexa647 as acceptor at DNA position +20]. Subpanels as in Fig. S2. The additional, low-E*, donor-acceptor species is free DNA [arising from dissociation of non-specific complexes upon heparin challenge during preparation of RP_o (see Materials and Methods: Transcription complexes); detected because DNA contains both donor and acceptor in this experiment].

Fig. S8. Initial transcription involves scrunching: control experiment.
Experiment documenting that addition of rifampicin--an inhibitor that prevents synthesis of RNA products >2 nt in length (*S13*)--prevents contraction of DNA between positions -15 and +15 [Cy3B as donor at DNA position -15; Alexa647 as acceptor at DNA position +15]. E* histograms as in Fig. S2. The additional, low-E*, donor-acceptor species is free DNA [arising from dissociation of non-specific complexes during heparin challenge during preparation of RP_o (see Materials and Methods: Transcription complexes); detected because DNA contains both donor and acceptor in this experiment].

Supplement: Supplemental Table S1
Summary of observed energy-transfer efficiencies and corresponding distances[1]

donor location	acceptor location	figure	R_o (Å)	RP_o			$RP_{itc,\leq 7}$			ΔR (Å)
				E^* [mean(±SEM)]	E	R (Å)	E^* [mean(±SEM)]	E	R (Å)	
σ^{366}	DNA^{+20}	2A	61	0.248(±0.002)	**0.29**	**71**	0.321(±0.003)	**0.43**	**64**	**–7**
σ^{569}	DNA^{-39}	2B	65	0.663(±0.003)	**0.83**	**50**	0.671(±0.003)	**0.83**	**50**	**0**
σ^{366}	DNA^{-20}	3A	61	0.445(±0.006)	**0.45**	**63**	0.451(±0.006)	**0.45**	**63**	**0**
σ^{569}	DNA^{-20}	3B	65	0.571(±0.005)	**0.64**	**59**	0.570(±0.005)	**0.64**	**59**	**0**
DNA^{-15}	DNA^{+15}	4A	68	0.223(±0.001)	**0.34**	**76**	0.256(±0.001)	**0.42**	**72**	**–4**
σ^{366}	DNA^{+25}	S2	61	0.239(±0.002)	**0.13**	**83**	0.276(±0.003)	**0.24**	**75**	**–8**
σ^{396}	DNA^{+20}	S2	58	0.185(±0.003)	**0.06**	**91**	0.237(±0.002)	**0.17**	**75**	**–16**
σ^{396}	DNA^{+15}	S2	58	0.494(±0.006)	**0.56**	**56**	0.602(±0.007)	**0.68**	**51**	**–5**
σ^{596}	DNA^{-39}	S4	64	0.935(±0.004)	**0.97**	**35**	0.935(±0.004)	**0.97**	**35**	**0**
σ^{569}	DNA^{-39}	S5	65	0.663(±0.003)	**0.83**	**50**	0.671(±0.003)	**0.83**	**50**	**0**
σ^{569}	DNA^{-44}	S5	65	0.380(±0.004)	**0.54**	**64**				
σ^{596}	DNA^{-39}	S5	64	0.935(±0.004)	**0.97**	**35**	0.935(±0.004)	**0.97**	**35**	**0**
σ^{596}	DNA^{-44}	S5	64	0.757(±0.004)	**0.88**	**46**				
σ^{396}	DNA^{-20}	S6	58	0.347(±0.006)	**0.23**	**71**	0.336(±0.007)	**0.22**	**72**	**1**
σ^{596}	DNA^{-20}	S6	64	0.579(±0.007)	**0.63**	**59**	0.570(±0.006)	**0.62**	**59**	**0**
DNA^{-15}	DNA^{+20}	S7	68	0.165(±0.002)	**0.20**	**86**	0.188(±0.002)	**0.27**	**80**	**–6**

[1] Values of R_o have precision of ±1 Å. Values of E have precision of ±0.01. Values of R have precision of ~±1 Å for $0.15 \leq E \leq 0.97$ and ~±1.5 Å for $0.05 < E < 0.15$. ΔR is defined as the difference between R in $RP_{itc,\leq 7}$ and R in RP_o.

Figure S1

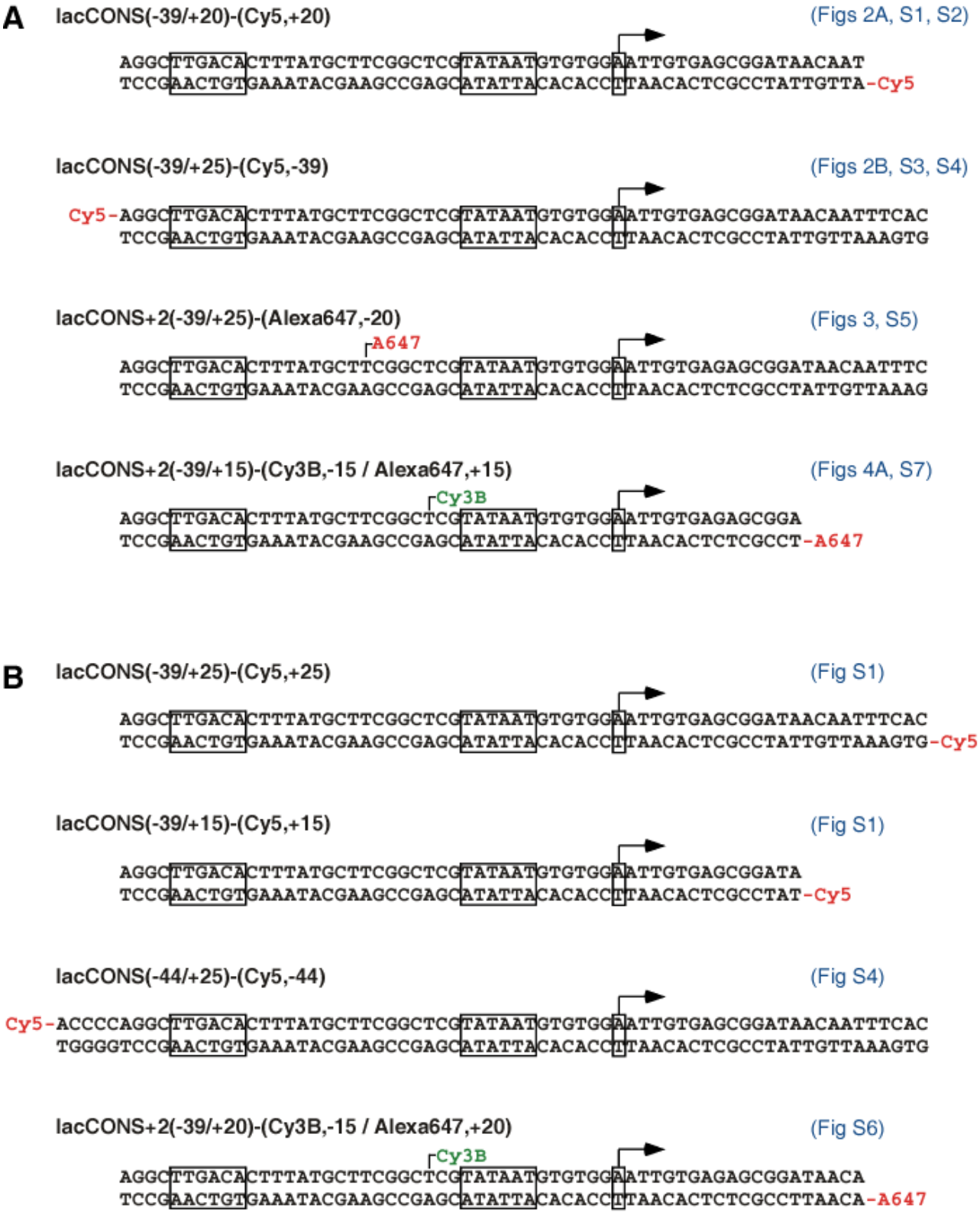

A
lacCONS(-39/+20)-(Cy5,+20)
(Figs 2A, S1, S2)
AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAAT
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCGCCTATTGTTA-Cy5
lacCONS(-39/+25)-(Cy5,-39)
(Figs 2B, S3, S4)
Cy5-AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCAC
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCGCCTATTGTTAAAGTG
lacCONS+2(-39/+25)-(Alexa647,-20)
(Figs 3, S5)
A647
AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGAGCGGATAACAATTTC
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCTCGCCTATTGTTAAAG
lacCONS+2(-39/+15)-(Cy3B,-15 / Alexa647,+15)
(Figs 4A, S7)
Cy3B
AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGAGCGGA
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCTCGCCT-A647
B
lacCONS(-39/+25)-(Cy5,+25)
(Fig S1)
AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCAC
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCGCCTATTGTTAAAGTG-Cy5
lacCONS(-39/+15)-(Cy5,+15)
(Fig S1)
AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGCGGATA
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCGCCTAT-Cy5
lacCONS(-44/+25)-(Cy5,-44)
(Fig S4)
Cy5-ACCCCAGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCAC
TGGGGTCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCGCCTATTGTTAAAGTG
lacCONS+2(-39/+20)-(Cy3B,-15 / Alexa647,+20)
(Fig S6)
Cy3B
AGGCTTGACACTTTATGCTTCGGCTCGTATAATGTGTGGAATTGTGAGAGCGGATAACA
TCCGAACTGTGAAATACGAAGCCGAGCATATTACACACCTTAACACTCTCGCCTTAACA-A647

Figure S2

distance between RNAP leading-edge and downstream DNA: additional measurements

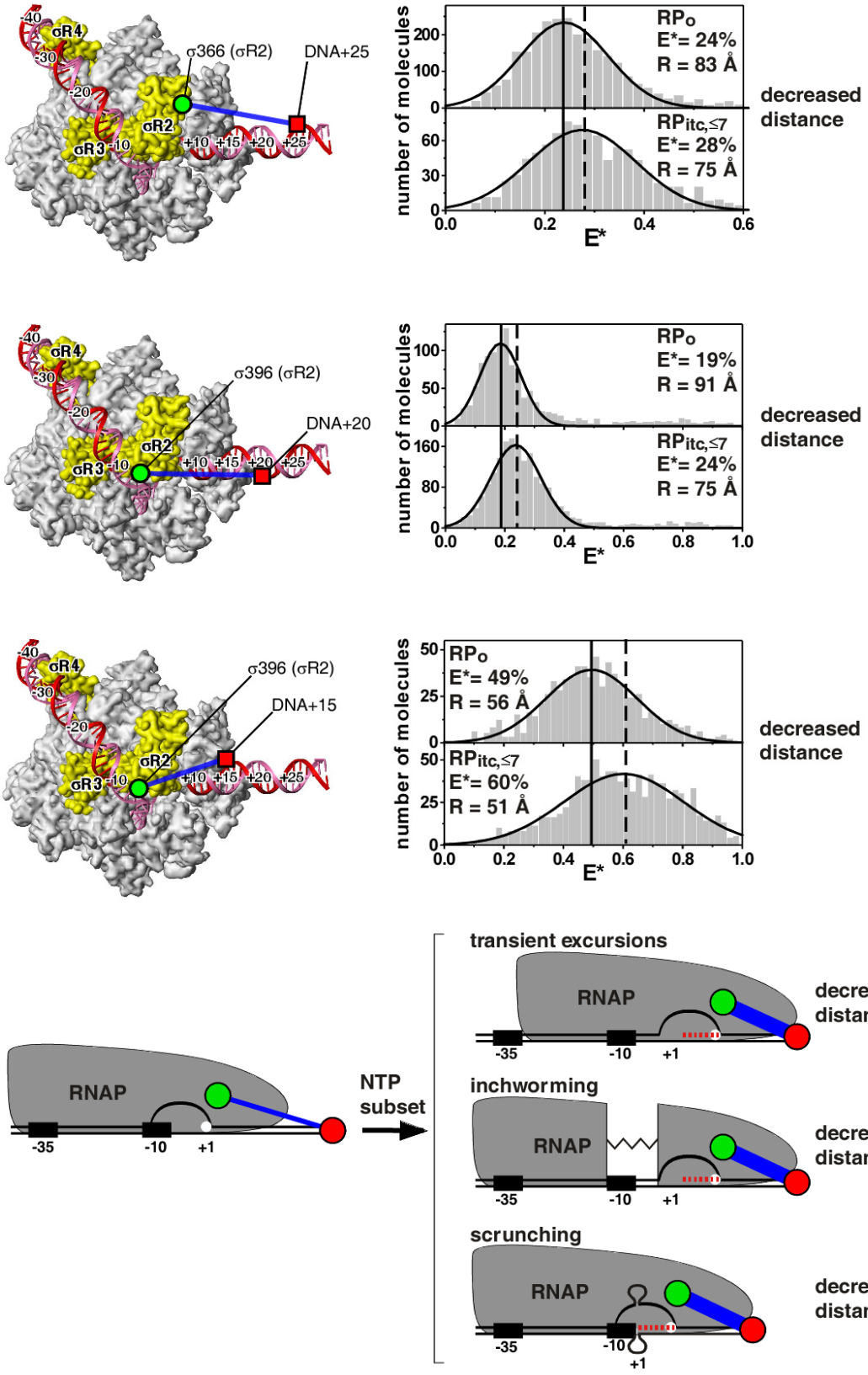


Figure S3

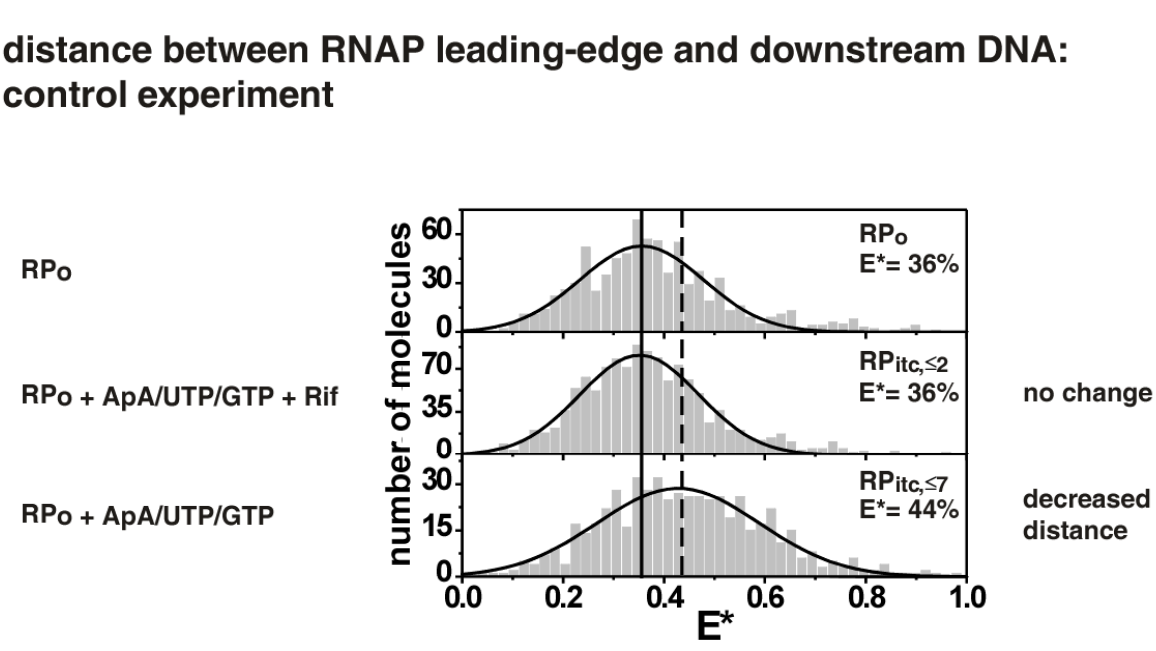

distance between RNAP leading-edge and downstream DNA: control experiment
RPo
RPo + ApA/UTP/GTP + Rif
RPo + ApA/UTP/GTP
number of molecules
60
30
0
70
35
0
30
15
0
0.0
0.2
0.4
0.6
0.8
1.0
E*
RPo
E*= 36%
RPitc,≤2
E*= 36%
RPitc,≤7
E*= 44%
no change
decreased distance

Figure S4

distance between RNAP trailing-edge and upstream DNA: additional measurement

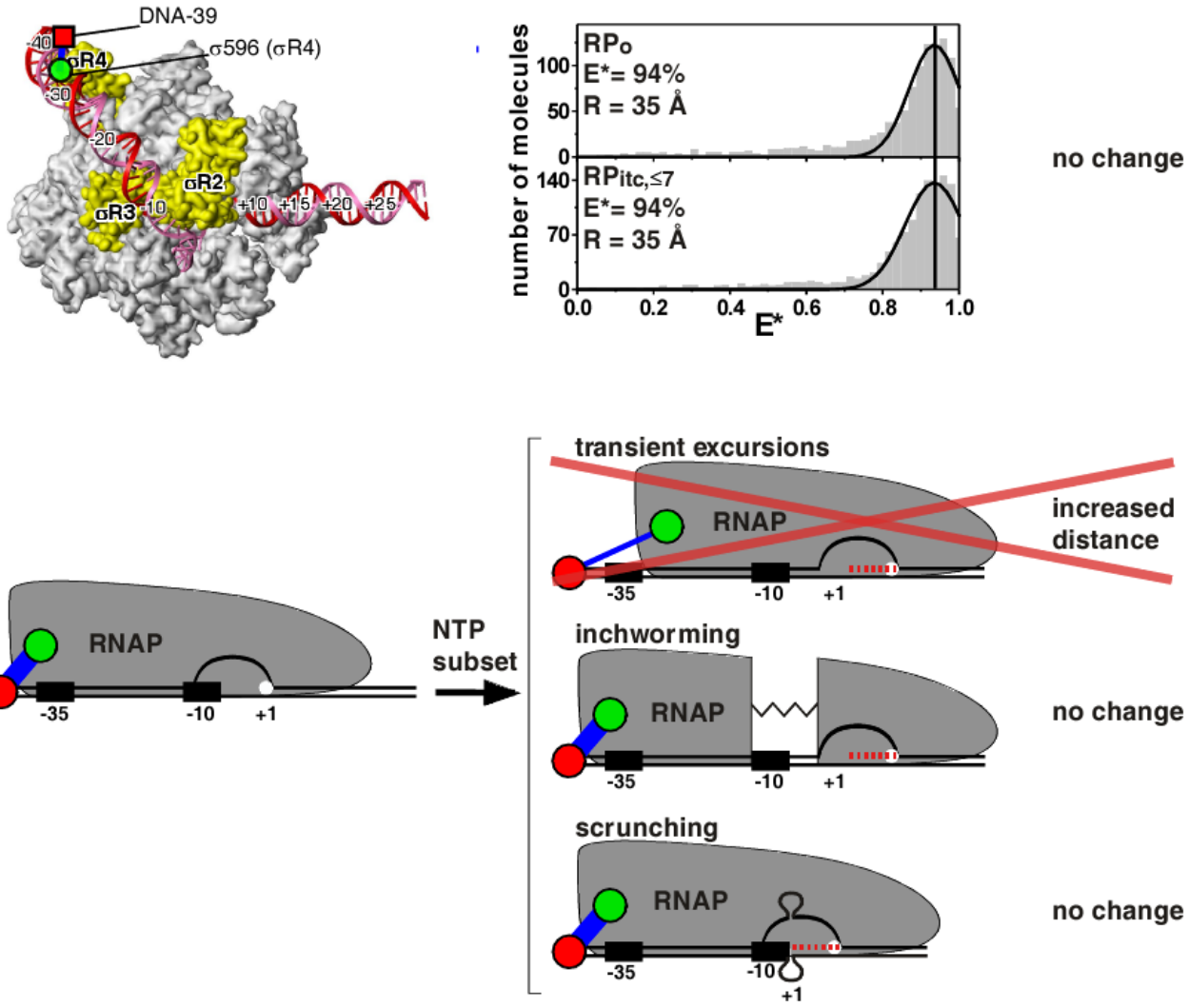


Figure S5

distance between RNAP trailing-edge and upstream DNA: control experiments

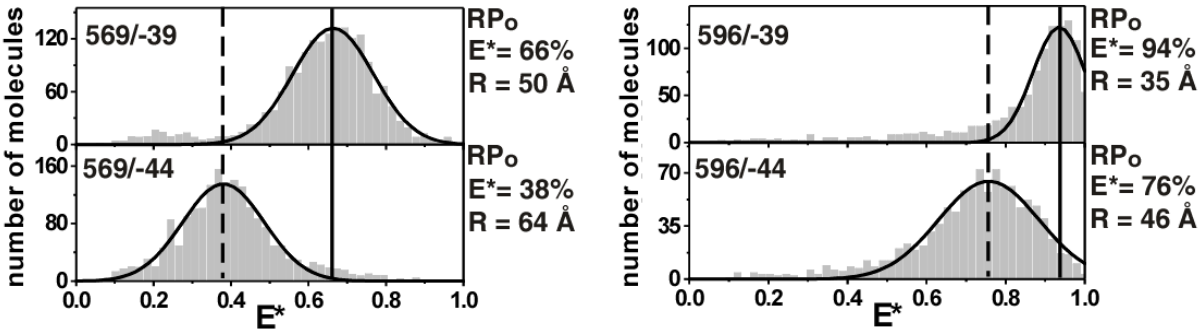


Figure S6

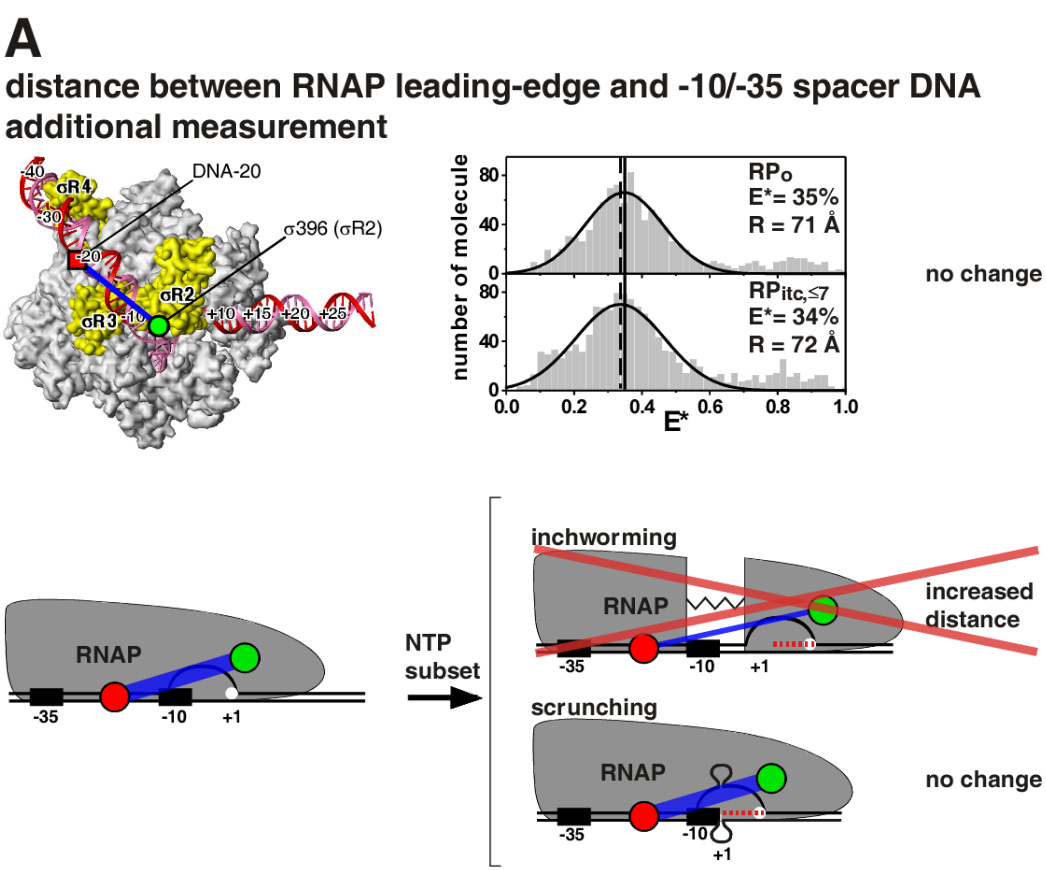

A
distance between RNAP leading-edge and -10/-35 spacer DNA
additional measurement
DNA-20
σ396 (σR2)
σR4
σR2
σR3
number of molecule
RPo
E*= 35%
R = 71 Å
RPitc,≤7
E*= 34%
R = 72 Å
E*
no change
RNAP
-35
-10
+1
NTP
subset
inchworming
increased
distance
scrunching
no change
B
distance between RNAP trailing-edge and -10/-35 spacer DNA
additional measurement
σ596 (σR4)
DNA-20
number of molecules
RPo
E*= 58%
R = 59 Å
RPitc,≤7
E*= 57%
R = 59 Å
no change
inchworming
scrunching

Figure S7

distance between -10/-35 spacer DNA and downstream DNA: additional measurement

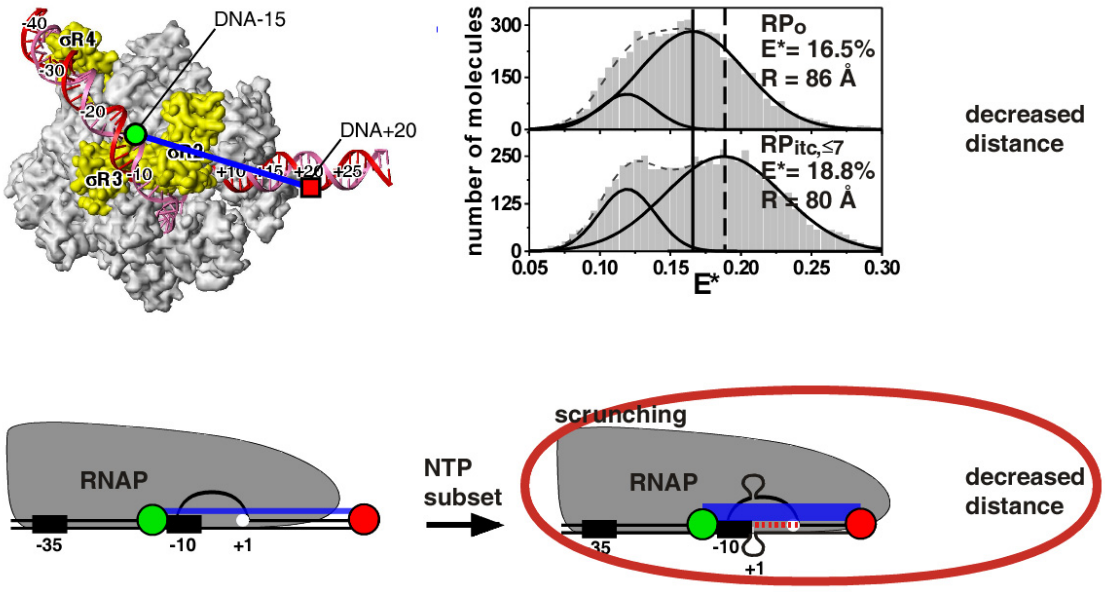
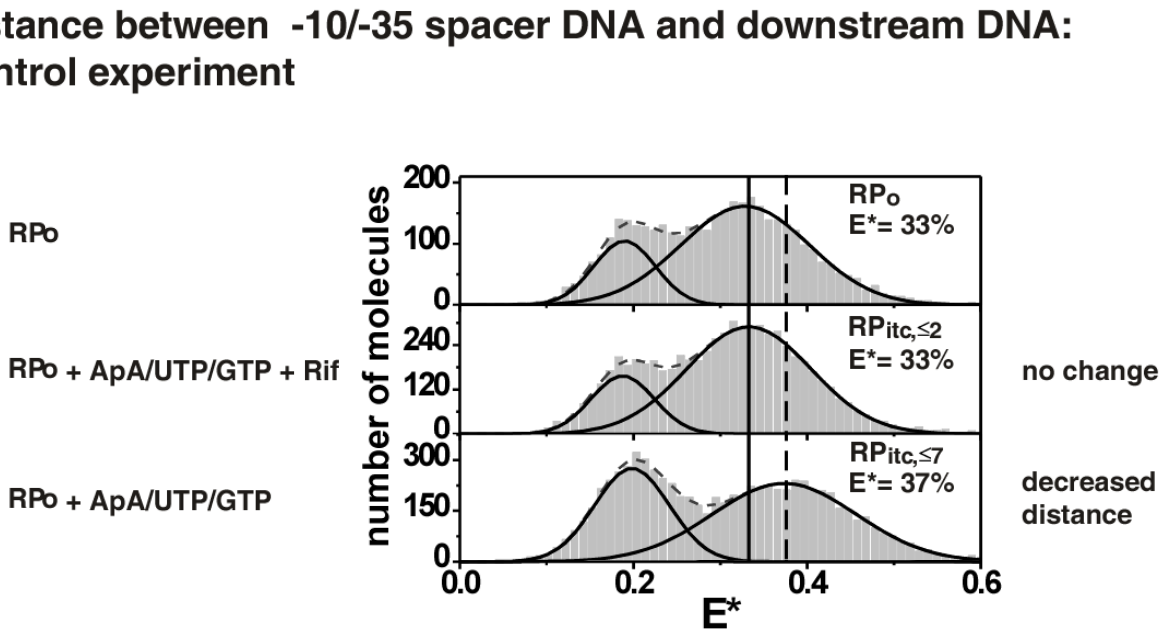


Figure S8


distance between -10/-35 spacer DNA and downstream DNA:
control experiment
RPo
RPo + ApA/UTP/GTP + Rif
RPo + ApA/UTP/GTP
number of molecules
200
100
0
240
120
0
300
150
0
RPo
E*= 33%
RPitc,≤2
E*= 33%
RPitc,≤7
E*= 37%
no change
decreased
distance
0.0
0.2
0.4
0.6
E*