Supporting Information

Steurer et al. 10.1073/pnas.1717920115

SI Materials and Methods

Generation of GFP-RPB1 KI Cells. A sgRNA sequence targeting *POLR2A* (*RPB1*) was selected using the ZitFit Targeter Program (1). The sgRNA with overhangs for BbsI restriction was inserted into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene plasmid ID: 42230) expressing a human codon-optimized Cas9. The homologous repair template, pBs-Rpb1-LA-Atg-Flag-Ggp-RPB1-RA, was generated using GIBSON assembly of the linearized pBs (pBluescript) vector and three PCR products encoding the left homology arm (LA), the GFP-FLAG linker fragment and the right homology arm (RA), respectively. LA is identical to the 750-bp upstream of the RPB1 start codon. RA corresponds to the 750-bp downstream of the RPB1 start codon but contains three (in bold) silent point mutations (ATG CAC GGG GGT GGC → ATG CA**T** GG**A** GG**G** GGC) to prevent Cas9 from recutting once RPB1 has been targeted. The homology arms were amplified from sequenced subclones of genomic DNA isolated from VH10 cells. 0.9 μg of pX330-U6-Chimeric_BB-CBh-hSpCas9_sgRNA and 3.6 μg of pBs-RPB1-LA-Atg-Flag-Gfp-Rpb1-RA were transiently transfected in WT MRC-5 cells with the JetPei transfection reagent (Polyplus-transfection) according to the manufacturer's instructions. GFP^{+} cells were FACS-sorted and seeded in a limiting dilution to 96-well plates to expand single clones. Primers used to genotype single clones are listed in Dataset S2.

Immunofluorescence and EU Incorporation. Cells were grown to 80% confluency on glass coverslips, treated as indicated, and fixed with 2% PFA in PBS for 15 min at room temperature. After permeabilization with 0.1% Triton in PBS for 10 min and blocking in 1.5% BSA and 0.15% glycine in PBS for 15 min, cells were incubated with primary antibodies as listed in Dataset S2. After washing in 0.1% triton in PBS for 10 min, cells were incubated with respective secondary antibodies coupled to the indicated Alexa fluorophores for 2 h at room temperature. Transcription levels were measured by pulse labeling with EU (Jena Bioscience) for 30 min with 1 μM EU in Ham's F-10 medium supplemented with 10% dialyzed FCS (Gibco). To measure EU in corporation after transcription inhibition, cells were treated with inhibitors for 1 h before the EU pulse and inhibitors were present during the EU pulse. Subsequently, cells were washed with PBS, and fixed with 2% PFA in PBS for 15 min. After permeabilization with 0.1% Triton in PBS for 10 min, click chemistry-based azide coupling was performed by incubation for 30 min with 60 μM Atto594 Azide (Attotec) in 50 mM Tris buffer (pH 8) with 4 mM $CuSO_4$ (Sigma Aldrich) and 10 mM freshly prepared ascorbic acid. Coverslips were washed with PBS and mounted with Vectashield containing DAPI (Brunschwieg Chemie). Cells were imaged with a Zeiss LSM 700 Axio Imager Z2 upright microscope equipped with a 63× Plan-Apochromat oil immersion lens (NA 1.40).

FRAP. For FRAP at pixel size 24.6 × 24.6 μM, a strip of 512 × 32 pixels spanning the nucleus was imaged every 400 ms with 400 Hz. Twenty-five frames were recorded before the bleach pulse. The average, background-corrected fluorescence intensity of frames 10–20 of these prebleach measurements were used to calculate the prebleach fluorescence intensity. GFP fluorescence in the strip was bleached for one frame with 100% laser power. The recovery of fluorescence was monitored for 4 min (600 frames) within and outside the strip. For half-nucleus bleach at pixel size 123 × 123 μM, an image of 1,024 × 1,024 pixels was recorded every minute for 120 min at 200 Hz. Regions of interest were selected over half nuclei and were photobleached with 100% laser intensity for one frame. FRAP and FLIP curves were corrected for background fluorescence outside the nucleus and normalized to prebleach fluorescence in the region of interest.

Cellular Fractionation. Cells were lysed in lysis buffer [30 mM Hepes pH 7.6, 1 mM $MgCl_2$, 130 mM NaCl, 0.5% Triton, 0.5 mM DTT, 50 μM Mg132, EDTA-free protease inhibitor (Roche) and phosphatase inhibitor mixture 2 (Sigma Aldrich)] on ice for 15 min, and then scraped in lysis buffer. Next, 500 U of benzonase (Millipore) was added for 1 h on ice. For fractionation, whole-cell lysates were centrifuged for 15 min at 16,100 × *g* and 4 °C. Supernatant containing nucleoplasmic RPB1 was collected. The pellet containing chromatin-bound RPB1 was washed twice with lysis buffer and resuspended in lysis buffer. Supernatant and pellet were diluted with SDS Page loading buffer (4% SDS, 0.2% bromophenol blue, 20% glycerol, 200 mM β-mercaptoethanol) and separated on a 6% SDS Page gel. Transfer was done overnight at 4 °C and 40 V in 2× transfer buffer (25 mM Tris, 190 mM Glycine) without methanol. Blots were blocked with 1.5% BSA in PBS and stained with the same primary antibodies as used for immunofluorescence. Secondary antibodies were coupled to IRDyes (LiCor) and imaged with an Odyssey CLx infrared scanner (LiCor). Quantification of Western blots was done using Odyssey CLx application software version 3.0.21 with the following background correction setting: average border with = 2, segments = all.

Determination of Number of Nuclear Pol II. Lyophilized eGFP was purchased from Biovison and dissolved in PBS. To determine the concentration of fluorescently active molecules, we determined the absorbance at 489 nm (ε489 nm = 55,000 M^{-1} Cm^{-1}). A 0.4 μM GFP stock solution in culture medium was serially diluted to 0.3, 0.2, and 0.1 μM. *Z*-stacks of GFP-RPB1 cells were recorded in medium containing free GFP at a pixel size of 82 × 82 μM at 200 Hz. Fluorescence intensities of free extracellular GFP was quantified alongside nuclear GFP-RPB1 fluorescence using the LAS AF Lite software. The nuclear volume of MRC-5 cells was determined from *z*-stacks of MRC-5 cells after segmenting the nucleus using a Gaussian blur with a radius of 2 pixels and a fixed threshold. To get the number of nuclear Pol II for a diploid cell, the number of measured nuclear Pol II was corrected for the hypertetra-ploidy of MRC-5 cells, which showed an average chromosome count of 75 after chromosome spread.

Chromatin Immunoprecipitation. Chromatin was cross-linked using 1% formaldehyde in culture medium for 10 min. Cross-linking was quenched by adding glycine to a final concentration of 0.125 M for 5 min. Cells were washed twice with PBS and scraped in PBS before pelleting by centrifugation. Pellets were lysed for 10 min in sonication buffer (0.1% SDS, 10 mM Tris·HCl pH 8, 1 mM EDTA, 0.5 mM EGTA) including complete protease inhibitor mixture without EDTA (Roche), phosphatase inhibitor mixture 2 and 3 (Sigma). Samples were sonicated for 15 min using Bioruptor (Diagenode, 15-s on, 15-s off, amplitude High) to fragments of 200–500 bp and centrifuged for 10 min at 16,100 × *g* to collect chromatin. Chromatin was diluted using dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris·HCl pH 8, 167 mM NaCl) and precleared for 30 min using Pierce Protein G Agarose beads (Thermo Scientific). RPB1 was

immunoprecipitated overnight at 4 °C using 5 μg GFP antibody (5 mg/mL, ab290; Abcam) and for an additional hour with Pierce Protein G Magnetic beads (Thermo Fisher). Precipitates were washed for 3 min with low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris·HCl pH 8, 150 mM NaCl), high-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris·HCl pH 8, 500 mM NaCl), lithium chloride buffer (0.25 M LiCl, 1% Nonidet P-40, 1% Sodium Deoxychelate, 1 mM EDTA, 10 mM Tris·HCl pH 8), and twice with Tris-EDTA buffer (10 mM Tris·HCl pH 8, 1 mM EDTA). DNA was eluted from beads twice using elution buffer (1% SDS, 0.1 M $NaHCO_3$) and decross-linked for 4 h with NaCl at 65 °C and 950 rpm in a thermoshaker. Proteins were removed by incubating 1 h at 45 °C and 700 rpm with Proteinase K, Tris·HCl pH 6.5 and EDTA before cleaning the DNA using a DNA ChIP Clean & Concentrator kit (Zymo Research).

siRNA Transfection and RT-PCR. Cells were transfected with siRNA using the Lipofectamine RNAiMAX reagent according to the manufacturer's instructions. Experiments were performed 48 h after transfection. Knockdown efficiency was checked by qPCR. siRNA and primer sequences are listed in Dataset S2. mRNA expression levels were normalized to B2M (β-2-Microglobulin). qPCR was performed using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad) and PowerUp SYBR green master mix (Thermo Fisher). Primers are listed in Dataset S2 and the used program was: 50 °C for 2 min, 95 °C for 2 min, 45 cycles of 15 s at 95 °C and 1 min at 58 °C followed by a dissociation curve: 95 °C for 10 s, 65 °C for 5 s and heating from 5 °C to 95 °C.

MC-Based Computational Model of Pol II Kinetics. To simulate photobleaching, the 3D intensity profile of the focused laser beam was experimentally determined by bleaching fixed cells that expressed GFP, as well as cells expressing H2B-GFP, using a stationary laser beam. Care was taken to avoid saturation (in this experiment reflected by complete photobleaching in the focus). The obtained fluorescence intensities before and after the bleach were used to determine the probability per unit time for a GFP molecule to get bleached considering its 3D position within the focused laser. The size of the ellipsoid model nucleus (x, y, and z diameters) was based on averages from experimental data. The simulation of the FRAP curve was run using discrete time steps of 20 ms. Diffusion was simulated at each new time step $t + \Delta t$ by deriving new locations ($x_{t+\Delta t}$, $y_{t+\Delta t}$, $z_{t+\Delta t}$) of mobile molecules from their current positions (x_t, y_t, z_t) by $x_{t+\Delta t} = x_t + G(r_1)$, $y_{t+\Delta t} = y_t + G(r_2)$, and $z_{t+\Delta t} = z_t + G(r_3)$, where r_i is a random number ($0 \le r_i \le 1$) chosen from a uniform distribution, and $G(r_i)$ is an inversed cumulative Gaussian distribution with $\sigma^2 = 2D\Delta t$, where D is the diffusion coefficient. A fixed diffusion coefficient of 2.0 $\mu m^2/s$ was used based on the estimated molecular weight the Pol II complex and based on previous measurements of free GFP and chains of GFP and a nonfluorescent GFP variant with increasing molecular weight (2). Immobilization was simulated using simple binding kinetics described by: $k_{on}/k_{off} = F_{imm}/(1 - F_{imm})$, where F_{imm} is the relative number of immobile molecules. The probability for each particle to become immobilized (representing chromatin-binding) per unit time is $P_{immobilize} = k_{on} = k_{off} \times F_{imm}/(1 - F_{imm})$, where k_{on} and k_{off} are effective on- and off-rates with dimension s^{-1}, $k_{off} = 1/T_{imm}$, and T_{imm} is the average time spent in the immobile state. The probability per unit time to be released is $P_{mobilize} = k_{off} = 1/T_{imm}$. In simulations of two or three immobile fractions with different kinetics, three immobilization/mobilization probabilities were evaluated each unit time step. Simulations of the FRAP curve were performed at every unit time step by counting the number of unbleached molecules in the bleached region after simulations of diffusion and binding during that time step. In addition, three pairs of on- and off-rates were used representing short, medium, and long residence times in immobile state and corresponding fractions (see above). The fraction with long residence time (long fraction) was then determined by taking the average of the 10 best-fitting simulated curves. Simulations were then run again with fixed long fraction size, and subsequently, the residence time corresponding to this fraction was determined by taking the average of the 10 best-fitting curves. Subsequently, new simulations were run with fixed on- and off-rates corresponding to the determined fraction and long residence time. This procedure was repeated for the fractions with medium and short residence times.

RNA-Seq Data Analysis. Nascent RNA sequencing data from human WT foreskin fibroblasts (HF1) (Gene Expression Omnibus: GSM1612077, run SRR1806546) (3) was trimmed using TrimmomaticSE v0.32 (4) with the steps: "ILLUMINACLIP: <illumina_adapter_list_file>:2:40:15 LEADING:3 TRAILING:3 MINLEN:15". The data were then aligned to the hg19 reference genome using bowtie2 v2.2.6 (5) with the local option. Stranded reads (ignoring duplicates) with a minimum 2 base overlap for any transcript variant of each gene were counted. Individual genes with multiple nonoverlapping transcripts were excluded. RPKM was computed using the total number of mapped reads counted. The per gene read distribution was computed as a fraction of all counted reads and the individual gene reads per kilobase were computed based on the length of the combined gene-transcript variants.

RNA Isolation and Sequencing. Total RNA was isolated in duplicate from a mixed population of MRC-5 WT cells, a single clone of MRC-5 WT cells, and two single clones of MRC-5 GFP-Rpb1 KI cells using the miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Total RNA concentration and quality was determined with an Agilent 2100 Bioanalyzer using a RNA 6000 Nano LabChip. Sequencing libraries were prepared using the Illumina TruSeq Total RNA-seq library preparation kit according to the manufacturer's protocol. The 150-bp paired-end sequencing files from the Illumina system were quality checked by FastQC (v0.11.3; available online at: www.bioinformatics.babraham.ac.uk/projects/fastqc). Low-quality reads and sequencing adapters were removed by Trimmomatic (v0.35) (4) and the resulting FASTQ files were aligned to the human reference genome (hg19) using TopHat (v2.0.9) (6). The aligned reads were analyzed by HTseq (v0.6.0) (7) to quantify gene expression. Genes with fewer than five reads were removed. Differentially expressed genes were identified with Deseq (available online at: https://bioconductor.org/packages/3.7/bioc/vignettes/DESeq/inst/doc/DESeq.pdf) with a cut-off of Log2FC (fold-change) >1.5 and $P < 0.05$. The statistics plots were created by Rstudio integrated development for R [v0.99.486; Team R (2015), available online at www.rstudio.com/].

1. Sander JD, et al. (2010) ZiFiT (Zinc Finger Targeter): An updated zinc finger engineering tool. *Nucleic Acids Res* 38:W462–W468.
2. van Royen ME, et al. (2009) Fluorescence recovery after photobleaching (FRAP) to study nuclear protein dynamics in living cells. *Methods Mol Biol* 464:363–385.
3. Andrade-Lima LC, Veloso A, Paulsen MT, Menck CF, Ljungman M (2015) DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes. *Nucleic Acids Res* 43:2744–2756.
4. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
5. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
6. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-seq. *Bioinformatics* 25:1105–1111.
7. Anders S, Pyl PT, Huber W (2015) HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.

Fig. S1. (*A*) Schematic representation of GFP-RPB1 KI strategy: Cas9 is targeted to the POLR2A locus by a specifically designed sgRNA. Cas9 induces a DSB directly downstream of the POLR2A start codon (ATG). The DSB can be repaired by homologous recombination (HR) using the repair template containing the GFP-FLAG sequence for homology. The RA of the repair template contains three point mutations to prevent Cas9 from recutting after HR. The LA is identical to the 750 bp upstream of the sgRNA target site. (*B*) Schematic procedure of Cas9-mediated GFP-labeling of RPB1 in MRC-5 cells. A plasmid expressing the sgRNA and Cas9 was cotransfected with a repair template into MRC-5 cells. GFP^+ cells were isolated using FACS and seeded as single cells into a 96-well plate by limiting dilution. Single cells were expanded and subsequently genotyped. (*C*) GFP-RPB1 KI was checked by genotyping genomic DNA of MRC-5 WT and KI cells by PCR with the indicated primers (*Upper*). PCR products were analyzed by DNA gel-electrophoresis (*Lower*). (*D*) Representative images of cells that were pulse labeled with EU for 1 h followed by click-chemistry–based EU coupling to ATTO594 of NT MRC-5 WT or GFP-RPB1 KI cells. DNA was stained with DAPI; RPB1 was visualized by αGFP immunostaining. (*E*) Pearson correlation coefficients (*Left*) and principal component analysis (*Right*) of RNA expression levels of a mixed population of MRC-5 WT cells, a single-cell clone of MRC-5 WT cells (WT_CL), and two single-cell clones of MRC-5 cells expressing GFP-RPB1 (KI_CL). To replicates of each sample were analyzed.

Fig. S2. (*A*) Representative images of cells that were pulse-labeled with EU for 1 h followed by click-chemistry–based EU coupling to ATTO594 of NT KI cells and after treatment with the indicated transcription inhibitors. DNA was stained with DAPI; RPB1 was visualized by αGFP immunostaining. (*B*) Quantification of EU incorporation shown (*Left*) and quantification of GFP fluorescence (*Right*). FI, Fluorescence intensity, mean ± SEM, $n > 55$ cells. (*C*) FRAP of the bleached and FLIP of the nonbleached half of nuclei of NT cells and cells treated with 100 μM Cordycepin for 90 min are plotted against time. RFI, relative fluorescence intensity, mean of $n = 10$ cells. (*D*) Relative TFIIS mRNA levels measured by qPCR of the TFIIS paralogues TCEA1 and TCEA2 after siRNA-mediated knockdown. Values are normalized to TCEA1 and TCEA2 expression of cells transfected with a nontargeting siRNA. (*E*) Propidium-Iodide content-based cell cycle distribution of NT MRC-5 cells or cells treated with 1 μM of THZ1 for 1.5 h.

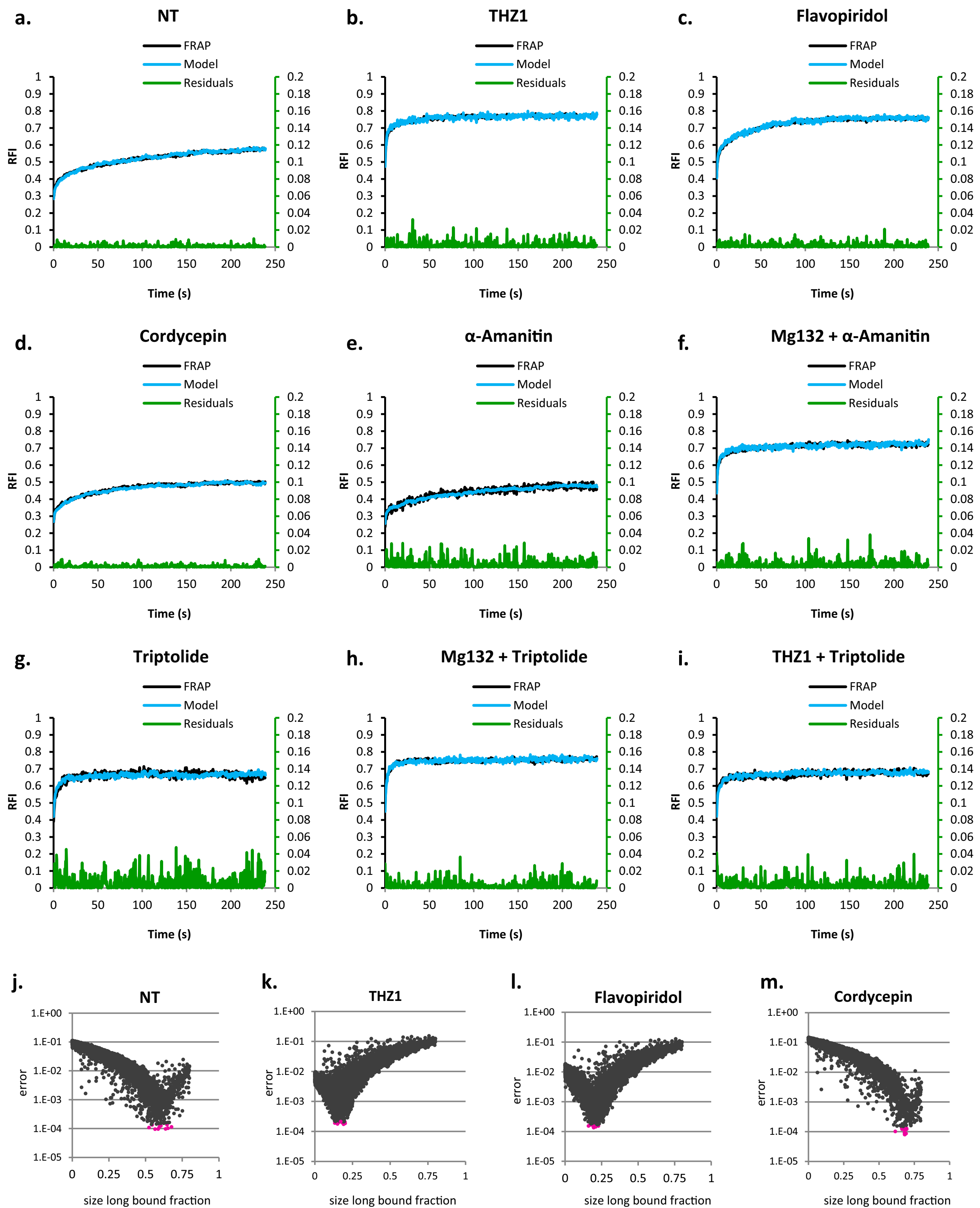

Fig. S3. Fitting of FRAP curves described in Figs. 2*C* and 4 *A* and *B* to the best-fitting curve generated by MC-based modeling. MRC-5 GFP-RPB1 KI cells were NT (*A*) or treated for 1 h with 1 μM THZ1 (*B*), 1 μM Flavopiridol (*C*), or 100 μM Cordycepin (*D*). Cells were treated for 2 h with 100 μg/mL α-Amanitin with (*F*) or without (*E*) pretreatment with 50 μM Mg132 for 1 h. Cells were treated for 1 h with 0.5 μM Tritptolide with (*H*) or without (*G*) pretreatment with 50 μM Mg132 for 1 h, or after pretreatment with 1 μM THZ1 for 1 h (*I*). Residuals represent the squared differences between measured and simulated FRAP curves. Errors of all MC-modeled fraction sizes of the long bound Pol II fraction in NT or inhibitor-treated cells, as described above. The 10 fraction sizes with the smallest error, (i.e., the best fit) that were averaged for Fig. 3*B*, are marked in magenta (*J*–*M*).

Fig. S4. (*A*) Residence times and fraction sizes of Pol II fractions obtained from MC-based modeling of 20 cells of two experiments (NT original, same as NT Fig. 2*C*), the average of three individually modeled FRAP curves (NT avg. NT1, -2, and -3), and three individually modeled FRAP experiments (NT1, -2, and -3). NT1, -2, and -3 and NT original show the mean ± SD of the 10 best fits. NT avg NT1, -2, and -3 represents the mean ± SD of the individually modeled parameters. (*B*) Residence times (*Center*) and fraction sizes (*Right*) of Pol II fractions obtained from MC-based modeling of one individual FRAP curve without or with addition of Gaussian noise ($\sigma = 2$). (*C*, *Lower*) Western blot after cell fractionation of NT MRC-5 KI cells or cells treated with the indicated inhibitors for 90 min. (*C*, *Upper*) Fraction sizes of chromatin-bound (Pel) and nucleoplasmic (Sup) Pol II after fractionation and Western blot analysis compared with fraction sizes predicted by MC-based modeling (Fig. 3*B*). The long-bound fraction is presented as chromatin-bound, whereas the sum of the free, short-, and medium-bound fractions is presented as nucleoplasmic. Mean ± SD of three Western Blots. One representative Western blot is shown.

Fig. S5. (*A*) Representative images of cells that were pulse-labeled with EU for 1 h followed by click-chemistry–based EU coupling to ATTO594 of NT KI cells and after treatment with the indicated transcription inhibitors. DNA was stained with DAPI; RPB1 was visualized by αGFP immunostaining. Quantification of ATTO594 and GFP fluorescence intensities are plotted in the chart (*Right*). Mean ± SEM, $n > 50$ cells. (*B*) FRAP of GFP-RPB1 in NT cells and cells treated with the proteasome inhibitor Mg132 (*Left*). Column chart shows mean ± SD of the prebleach FI of 10 cells analyzed by FRAP as a measure of Pol II protein levels (*Right*). (*C*) Western blot of GFP-RPB1 KI whole-cell lysates after indicated treatments. All samples originate from the same blot. Mg132, proteasome inhibitor. (*D* and *E*) Residence times of short-bound, medium-bound, and long-bound Pol II fractions obtained from MC-based modeling of GFP-RPB1 FRAP data (Fig. 3 *A* and *C*) in NT KI cells or after treatment with Triptolide (*D*) or α-Amanitin (*E*). Mean ± SEM of the 10 best fitting simulations.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)