Single-molecule imaging of transcription at damaged chromatin

Alexandra C. Vítor, Sreerama C. Sridhara, João C. Sabino, Ana I. Afonso, Ana R. Grosso†, Robert M. Martin‡, Sérgio F. de Almeida†

How DNA double-strand breaks (DSBs) affect ongoing transcription remains elusive due to the lack of single-molecule resolution tools directly measuring transcription dynamics upon DNA damage. Here, we established new reporter systems that allow the visualization of individual nascent RNAs with high temporal and spatial resolution upon the controlled induction of a single DSB at two distinct chromatin locations: a promoter-proximal (PROP) region downstream the transcription start site and a region within an internal exon (EX2). Induction of a DSB resulted in a rapid suppression of preexisting transcription initiation regardless of the genomic location. However, while transcription was irreversibly suppressed upon a PROP DSB, damage at the EX2 region drove the formation of promoter-like nucleosome-depleted regions and transcription recovery. Two-color labeling of transcripts at sequences flanking the EX2 lesion revealed bidirectional break-induced transcription initiation. Transcrip tome analysis further showed pervasive bidirectional transcription at endogenous intragenic DSBs. Our data provide a novel framework for interpreting the reciprocal interactions between transcription and DNA damage at distinct chromatin regions.

RESULTS

Engineering reporter genes to image single-molecule transcription upon controlled induction of a DSB

We used human embryonic kidney (HEK) 293 cells containing a single Flp recombinase target (FRT) genomic locus for Flp-in homologous recombination of a single copy of a modified immunoglobulin M (IgM) gene driven by a doxycycline-inducible cytomegalovirus (CMV) promoter (Fig. 1A). Two independent reporter genes were generated upon insertion at distinct regions of a unique ScI restriction site, which is absent from the human genome: insertion 61 base pairs (bp) downstream the transcription start site at a promoter-proximal region (PROP reporter; Fig. 1B and fig. S1A) and insertion inside exon 2 (EX2 reporter; Fig. 1B and fig. S1B). To image transcription in real time using the PROP reporter, we inserted 24 repeats of MS2 sequences immediately downstream the I-ScI site. In the EX2 reporter, we flanked the I-ScI site with 24 repeats of MS2 and PP7 phage sequences. Transcription is visualized through the recruitment of fluorescently tagged MS2- and PP7-binding proteins to single nascent transcripts (Fig. 1, C and F) (8). The two-color labeling strategy designed for the EX2 reporter allows for interrogation of transcription at both regions flanking the DSB. In intact (i.e., without DSB induction) EX2 reporter cells, analysis of transcription revealed a strong correlation between the MS2 and PP7 fluorescent signals throughout the entire imaging period (Fig. 1F). The two-color labeling had a sufficient resolution to capture the temporal delay between transcription of the MS2 sequences and transcription of the PP7 sequences located downstream (Fig. 1F and fig. S2A). Moreover, we could detect the increase in fluorescence intensity corresponding to the sequential binding of the MS2-RFP (red fluorescent protein) and PP7-GFP (green fluorescent protein) proteins to their target sequences on single RNA transcripts as they emerge from RNAPII (fig. S2, B and C). By computing the MS2 or PP7 sequence length with the time intervals of different metrics between initial detection of fluorescence values and/or maximum values, we determined the average transcription rates ranging from 3.8 to 4.9 kb/min at the MS2 and PP7 sequences, respectively (fig. S2D). These values, which are in agreement with the approximate rates of 4 kb/min estimated in previous studies (9, 10), reveal that the reporter genes are transcribed normally.

INTRODUCTION

DNA is the template for vital cellular processes such as replication, transcription, and DNA repair. Thus, it is expected that sometimes, two of these processes simultaneously use the exact same sequence and collide. Such conflicts may have catastrophic consequences and are an important source of aberrant gene expression programs and genome instability (1, 2). Intriguingly, recent studies revealed that close encounters between transcription and the DNA repair machineries may have important physiological outcomes. For instance, complexes involved in scheduled DNA damage have been shown to coactivate...
Fig. 1. New reporter genes allow visualization of transcription upon DNA damage with single-molecule resolution. (A) Illustration of a HEK293 cell containing a single FRT genomic locus for single copy integration of the reporter gene. (B) Schematic of the two reporter genes with the I–SceI site and MS2/PP7 sequences highlighted. (C and D) Representative images of PROP (C) and EX2 (D) reporter cells. Images were acquired at the indicated time points after doxycycline treatment. Examples of pre-mRNAs diffusing throughout the nucleoplasm are numbered 1 to 4. The transcription site (TS) of the reporter gene is indicated in each cell. Corresponding images depicting the diffraction limited objects in the highest intensity plane are shown below. (E and F) Time-lapse series of imaged PROP (E) and EX2 (F) reporter cells upon doxycycline treatment. The number (n) of transcripts was plotted over time in line graphs. Enlarged line plots depict a complete cycle of fluorescence gain and loss for each of the time-lapse series.
**Fig. 2. PROP DSBs silence transcription.** (A) PROP reporter cells were transfected with the I–SceI–GR–RFP fusion protein and imaged for 8 min after TA treatment. Images acquired at the indicated time points illustrating the translocation of the enzyme into the nucleus are shown. (B) Dynamics of transcription revealed by MS2-RFP fluorescence intensity at the transcription site and of MDC1-GFP recruitment to the transcription site in PROP cells transfected with I–SceI–GR fusion proteins and live-imaged after TA treatment. Representative images acquired at the indicated time points are shown. The dual-line plot on the right-hand side represents the average labeling intensity of MDC1-GFP and MS2-RFP quantified for a region of interest (ROI) defined around the transcription site over time. One representative experiment from more than 30 individual cells recorded is shown. a.u., arbitrary units. (C and D) Time-lapse, multiplane, spinning disk confocal microscopy was used to monitor live cell transcription dynamics of PROP reporter cells expressing MS2-GFP in the absence of (C) or upon (D) TA treatment. Representative line plots of number (n) of transcripts over time are shown. Bar graphs on the right-hand side represent the frequency of cells that exhibit a pattern of continuous (Contin.) or suppressed/terminated (Term.) transcription revealed by the lack of MS2-GFP fluorescence. The absolute number (n) of imaged cells is shown above each bar.
Regulated induction of a single DSB at distinct genic regions

To control the timing of DSB induction, we transfected the reporter cells with an I–SceI enzyme fused to the glucocorticoid receptor (I–SceI–GR). This fusion was fully able to digest the reporter genes in vitro (fig. S3A). Within 3 min after cell treatment with triamcinolone acetonide (TA), cytoplasmic I–SceI–GR translocated into the nucleus (Fig. 2A). In the nucleus, I–SceI–GR has access to its restriction site on each reporter gene to induce a single genomic DSB. To confirm this, we measured the dynamics of mediator of DNA damage checkpoint 1 (MDC1), a factor that plays an early role in the DNA damage response (DDR) (11, 12), through live cell microscopy imaging of PROP (Fig. 2B) and EX2 (fig. S3B) reporter cells. Approximately 12 min after the addition of TA, we observed MDC1 recruitment to the transcription site, confirming that I–SceI translocation into the nucleus cut the reporter genes. In parallel to MDC1 recruitment, we also observed a strong reduction of the transcription levels of both reporter genes (Fig. 2B and fig. S3B), suggesting that a DSB leads to rapid suppression of preexisting transcription initiation.

Fig. 3. Break-induced transcription initiation. Time-lapse, multiplane, spinning disk confocal microscopy recordings of EX2 reporter transcription dynamics in live cells revealed by the dual labeling of nascent transcripts with MS2-RFP and PP7-GFP in the absence ([A] no TA) or upon ([B] and [C] +TA) DNA damage. The dual-line plots represent the number (n) of transcripts detected at the transcription site over time. The bar graph in (A) shows the frequency of cells that exhibit a pattern of continuous (Contin.) and suppressed/terminated (Term.) transcription. The bar graph in (B) shows the frequency of cells, where preexisting transcription was suppressed in response to TA treatment and either resumes (Resume) or remains silenced (Term.) during the entire imaging period. A representative pattern of transcription resuming is depicted on the images and dual-line plot in (B). (C) Representative images and dual-line plot of a recorded cell where only the PP7-GFP fluorescence is observed after the damage-induced termination and before dual-labeled transcripts are detected. The bar graph depicts the frequency of cells that exhibit a similar uncoupling of the MS2-RFP and PP7-GFP fluorescent signal (Uncoupl.) and those where transcription resumed as illustrated in (B) (Coupl.). (D) Western blot analysis of DNA ligase IV in control (ctrl) or DNA ligase IV small interfering RNA (siRNA)–depleted cells. α-Tubulin and total histone H3 serve as loading controls. The bar graph depicts the frequency of cells that either resume or irreversibly terminate transcription after induction of the DSB in DNA ligase IV–depleted cells. kd, knockdown; RNAi, RNA interference.

The impact of a DSB on transcription depends on its location within the gene

Recent studies reported break-induced transcription at the DNA ends of a DSB (3–6). While suppression of the canonical preexisting transcription initiation was the immediate response to a DSB regardless of its location within the reporter gene, we sought evidence of break-induced transcription initiation upon a DSB. In our reporters, the PP7 and/or MS2 sequences were inserted in close proximity to the I–SceI site. This implies that any transcription event at the DSB should be readily detected by the recruitment of fluorescent PP7- and/or MS2-binding proteins. To address this hypothesis, we first imaged nondamaged PROP reporter cells for 50 min. Under these conditions, we continuously detected transcription of the reporter gene, yielding a number of nascent RNA transcripts that oscillated between 1 and 18 (see Materials and Methods and figs. S4 and S5 for details on the estimation of the transcript number) (Fig. 2C). As expected, induction of the PROP DSB resulted in a rapid suppression of full-length transcription of the reporter gene within approximately 10 min after the
Fig. 4. Break-induced chromatin modifications. (A) Nucleosome occupancy measured by total histone H3 ChIP at two regions (marked A and B in the schematics of the reporter gene) flanking the I–SceI site of the EX2 reporter. (B) Levels of the indicated histone modifications at the A and B regions of the EX2 reporter gene measured by ChIP before and after (30 min and 1 hour) the addition of TA to induce a DSB. (C) Nucleosome occupancy and (D) levels of histone modifications at a region (marked C) immediately downstream the I–SceI site of the PROP reporter gene. All data were normalized against the nucleosome occupancy on non-TA-treated cells and represent mean and SEM from at least three independent experiments. *P < 0.05 and **P < 0.01, obtained using two-tailed Student’s t test.
addition of TA in 87% of imaged cells (Fig. 2D). Transcription of the PROP reporter gene remained silenced throughout the remaining imaging period, as revealed by the lack of any MS2-GFP fluorescence recovery (Fig. 2D). We conclude that the PROP DSB suppresses full-length transcription of the reporter gene and that this is irreversible on the time scale of the experiment.

In EX2 reporter cells, canonical transcription was continuously observed in the absence of DNA damage, as revealed by the overlap between MS2-RFP and PP7-GFP fluorescence signals (Fig. 3A). Upon TA treatment, induction of the DSB led to a complete suppression of preexisting transcription initiation detected by the loss of both MS2-RFP and PP7-GFP signals at the transcription site (Fig. 3B). Increasing the temporal resolution to 10-s imaging intervals further showed that, after a DSB, RNAPII elongates until the break site, as revealed by the detection of MS2-RFP fluorescence alone, before a complete transcription shutdown is observed (Fig. S3C). However, in contrast to the PROP reporter, full-length transcription of the damaged EX2 reporter gene recovered in the majority (81%) of imaged cells (Fig. 3B). Strikingly, the pattern of transcription recovery was not unique. In 54% of the cells, we detected a late transcription recovery, approximately 30 to 40 min after DSB induction, whereas in the remaining cells, transcription recovered immediately after suppression of the preexisting transcription (Fig. 3C). Moreover, the late transcription recovery was significantly ($P < 0.05$) suppressed in DNA ligase IV–depleted cells (Fig. 3D), which are incompetent for nonhomologous end joining (13). This suggests that DNA repair is necessary before RNAPII resumes elongation across the damage site. The early transcription recovery was detected by the accumulation of PP7-GFP but not of MS2-RFP proteins (Fig. 3C), indicating that transcription was initiated between the MS2 and PP7 sequences, consistently with break-induced transcription initiation from the DSB.

**Locus-specific changes to the chromatin landscape surrounding DSBs**

The finding that recovery of transcription is greatly determined by the lesion site highlights chromatin structure as an important determinant of this outcome. Previous studies revealed that changes in chromatin structure, such as local nucleosome depletion, are sufficient to drive transcription initiation from noncanonical promoter sequences (14, 15). With this in mind, we sought evidence of break-induced changes in

---

**Fig. 5. Intragenic DNA damage drives bidirectional break-induced transcription initiation.** (A) Schematics of the EX2-AS reporter constructed to directly inspect break-induced antisense transcription initiation. (B) Representative images of a total of eight imaged cells and dual-line plot of a recorded cell where break-induced antisense transcription initiation was detected. The bar graph depicts the frequency of cells where antisense transcription was either not observed (No AS transcr., 50%) or observed (AS transcr., 50%).
nucleosome occupancy at the vicinity of the DSB. Histone H3 chromatin immunoprecipitation (ChIP) assays before and after induction of the DSB revealed a significant loss of nucleosomes at the chromatin region near the I–Scel site of the EX2 reporter (Fig. 4A). Nucleosome depletion was further confirmed by micrococcal nuclease (MNase) digestion assays (fig. S6). In agreement with the finding that transcription was not suppressed following the intragenic DSB, we did not observe a reduction in the relative levels of histone modifications that mark active chromatin, H3K9Ac and H3K36me3 (Fig. 4B). We observed slightly increased levels of H3K9Ac, a histone modification that decorates active promoter regions, at the DSB site (Fig. 4B), which may suggest break-induced tailoring of a promoter-like chromatin landscape and is compatible with the observed break-induced transcription initiation.

Induction of a DSB at the I–Scel site of the PROP reporter gene did not cause nucleosome depletion (Fig. 4C). Moreover, the levels of H3K36me3 decreased significantly within 30 min after the DSB, and no changes were detected in the levels of H3K9Ac (Fig. 4D). These data are consistent with the irreversible suppression of transcription observed upon induction of a PROP DSB.

Intragenic DSBs induce bidirectional transcription initiation
Nucleosomes play an important role in restricting antisense transcription (16, 17). The finding that intragenic DSBs reduce nucleosome occupancy raises the hypothesis that transcription may initiate from the broken DNA ends in the sense and antisense direction. To further test the possibility of bidirectional transcription in response to a DSB, we imaged transcription of a new reporter gene obtained by inverting the 24 MS2 repeat tract of the EX2 reporter (Fig. 5A and figs. S1C and S3D). In the resulting EX2-antisense (EX2-AS) reporter, recruitment of MS2-RFP proteins to the nascent RNA transcript occurs only if the MS2 sequences are transcribed in the anti-sense direction. Only PP7-GFP proteins were detected at the site of transcription of the EX2-AS reporter gene in the absence of DNA damage (Fig. 5B). Upon TA treatment, the I–Scel–induced DSB suppressed preexisting sense transcription (revealed by the loss of PP7-GFP) and drove antisense transcription initiation in 50% of cells, as shown by the accumulation of MS2-RFP fluorescence at the site of transcription (Fig. 5B).

Pervasive break-induced transcription initiation at endogenous DSBs
To gain further insight into bidirectional break-induced transcription initiation, we analyzed DSBCapture and transcriptome data obtained from normal human epidermal keratinocytes (NHEKs) (18). Analysis of DSBCapture data provided a genome-wide profile of endogenous DSBs at a single-nucleotide resolution. From these, we selected those that map within the body of active genes and interrogated transcriptome data for antisense transcription within 1 kb from the DSB (Fig. 6A). This analysis revealed a marked peak immediately upstream the DSB sites (Fig. 6A). Permutation analysis further revealed statistically significant levels of antisense transcription at the intragenic DSB loci when compared to random genic regions (Fig. 6B). In agreement with the data...
obtained with our reporter gene, DSB and antisense transcription peaks overlapped with nucleosome-depleted regions [deoxirribonuclease sequencing (DNase-seq)], indicating an association between regulatory chromatin and antisense break-induced transcription (Fig. 6A). Together, these data suggest that bidirectional transcription initiation is a widespread feature of intragenic DSBs that is related with the formation of break-induced nucleosome-depleted regions.

**DISCUSSION**

Here, we imaged single nascent RNAs at the site of transcription to gather quantitative information about the dynamics of transcription at damaged chromatin. We provide accurate kinetic parameters of the timing of transcription suppression upon a DSB, complementing previous reports that correlated DNA damage with repression of transcription (19–25). Moreover, we demonstrate that transcription recovery after break-induced suppression depends on the location of the DSB. While a lesion in the beginning of the transcription unit at a promoter proximal region prevented recovery during the entire imaging period, DNA damage within the gene body permitted transcription recovery with seemingly distinct kinetic patterns: late (i.e., 30 to 40 min after induction of a DSB) and early (immediately after suppression of the preexisting transcription initiation). The two-color labeling of the nascent RNAs revealed that RNAPII elongated through the I–SceI site during the late recovery, suggesting that the broken DNA ends had already been resealed. In agreement, preventing DNA repair through depletion of DNA ligase IV significantly decreased the number of cells where transcription recovery was observed. In contrast, the early transcription recovery is not consistent with the kinetics of DNA repair, which requires protein effectors that are recruited to the DSB several minutes after the lesion (26). Instead, our data suggest that broken DNA ends are competent to drive bidirectional transcription initiation both in our reporter genes and widespread across the human genome. This is supported by the finding that an intragenic DSB nucleates a promoter-like chromatin landscape characterized by nucleosome depletion. Our data further suggest that each DSB promotes transcription initiation in one direction only, with no apparent preference for either sense or antisense. This is illustrated by the failure to simultaneously observe PP7- and MS2-containing transcripts at the Ex2-A5 reporter. This notion is consistent with the detection of break-induced transcription initiation at the Ex2 reporter (which does not allow visualization of antisense transcription) in only approximately half of the cells. These data raise the hypothesis that, at each DSB, break-induced transcription initiation occurs exclusively in a single direction, and suggest the existence of a mechanism preventing RNAPII firing in the opposite direction. The observed 50% frequency of antisense transcription suggests that the direction of break-induced transcription is decided in a stochastic manner, perhaps determined by the kinetics of RNAPII binding to one of the two broken DNA ends. One possibility is that this binding is sufficient to directly hinder the binding of polymerases to the opposite broken DNA. Alternatively, additional factors (perhaps those involved in DNA repair) may be required to suppress simultaneous bidirectional transcription initiation at DSBs. Further research is needed to elucidate the role, if any, of break-induced RNAs in the DDR, a notion that was previously suggested (7, 27). Regardless of the function of these transcripts, our data are consistent with a model that places bidirectional transcription initiation as a widespread outcome of DSBs within the body of active genes, disclosing a determinant role for chromatin in the functional interactions between DNA damage and transcription.

**MATERIALS AND METHODS**

**Plasmids and genetic constructs**

The plasmid pcDNA5-FRT-TO-IgM-1.7k-PY-CFPPTS-24MS2SL-I–SceI–24PP7SL was constructed with the restriction site for I–SceI inserted within the exon II of a mouse immunoglobulin μ (IgM) reporter gene, referred to as EX2 reporter. The 24 PP7 stem-loop sequence was excised from pCR4-24xPP7SL [Addgene plasmid no. 31864 (28)] by Bam HI and Bgl II digest and inserted into the Bam HI site of pCMV5-24xMS2SL, generating pCMV5-24xMS2SL–24xPP7SL. To insert the I–SceI restriction site in pCMV5-24xMS2SL–24xPP7SL, two oligo sequences encoding the I–SceI and two Bam HI sites were hybridized (sequences are shown in table S1) and amplified by polymerase chain reaction (PCR), purified (NZYGelpure kit, NZYTech), and Bam HI digested and ligated into the same site to generate pCMV5-24xMS2SL–I–SceI–24xPP7SL. The Ecl136I–I–SceI–24xMS2SL–I–SceI–24xPP7SL fragment was cloned into the blunted Xho I site of the pcDNA5-FRT-TO-IgM-1.7k-PY-CFPPTS reporter (10) using the Paperclip protocol (29). The plasmid pcDNA5-FRT-TO-IgM-1.7k-PY-CFPPTS was designed with the I–SceI site in the promoter proximal region of the IgM reporter gene, referred to as the PROP reporter. An array of two tet operator (tetO) sequences followed by an I–SceI site, 24 MS2 stem loops from two nonidentical stem-loop sequences [as described for Addgene plasmid no. 31865 (30)], including five additional nonidentical spacer sequences, to further decrease redundancy was de novo designed (synthesis by GeneArt, Invitrogen). The 2tetO–(I–SceI)–24xMS2SL fragment was cloned by Ecl136I and Hind II digest into the same sites in pcDNA5-FRT-TO-IgM-1.7k-PY-CFPPTS. The third reporter gene referred to as EX2–AS was designed on the basis of the human ubiquitin B gene (HGNC:12463; ENST00000302182.7) with an I–SceI in exon II followed by an open reading frame of two head-to-tail ubiquitin units (synthesis by GeneArt, Invitrogen). The 24 MS2 stem-loop sequence with five nonidentical spacers [lacking the 2tetO–(I–SceI) part] was ligated in antisense orientation into exon I of the reporter gene and the 24 PP7 stem-loop sequence in sense orientation into the exon II 3’ untranslated region. The complete reporter unit was ligated into the pcDNA5-FRT-TO (Invitrogen, Thermo Fisher Scientific) backbone. The plasmids encoding tandem dimers of MS2 and PP7 coat proteins (named MCP and PCP thereafter, respectively) fused to fluorescent proteins were constructed the following way: First, pEGFP-C1 was Vsp I and Bsp 1407I digested, blunted, and religated to remove the CMV promoter and enhanced GFP (EGFP) sequences. The Eco RI–Not I fragment (ubiquitin C promoter) from phage-ubc-nls-ha-tdPCP-gfp [Addgene plasmid no. 40650 (31)] was ligated into the Eco RI–Bsp 120I sites generating a PUBC vector. Upstream restriction sites were removed by religating blunted EcoRI–BglII sites and EGFP including multicloning site were reintegrated by ligating a Bsh TI–XbaI EGFP fragment from pEGFP-C1 into Cfr9I–XbaI sites in PUBC to create pUBC-GFP-C1. The pUBC-GFP-nls-tdMCP-GFP plasmid (referred to as MS2-GFP) was generated by ligating a Bsp120I–BglII fragment from phage-ubc-nls-ha-tdMCP-gfp [Addgene plasmid no. 40649 (31)] into Bsp120I–BglII sites of PUBC-GFP-C1. pUBC-GFP-nls-tdPCP-GFP plasmid (referred to as PP7-GFP) was generated by ligating a Bsp120I–BglII fragment from phage-ubc-nls-ha-tdPCP-gfp [Addgene plasmid no. 40650 (31)] into Bsp120BamHI sites of PUBC-GFP-C1. The pUBC-TagRFPT-nls-tdMCP-TagRFPT (referred to as MS2-RFPT) was cloned by PCR amplification of the TagRFPT sequence from phage-ubc-nls-ha-2X MCP-TagRFPT [Addgene plasmid no. 64541 (32)] using primers with AleI and BglII sites (sequences are shown in table S1) to ligate...
it into the same sites of pUBC-GFP-C1 to generate pUBC-TagRFPt. The Bsp 120I-Bam HI fragment containing nls-tdMCP-TagRFPt from Addgene plasmid no. 64541 was ligated into the same sites in pUBC-TagRFPt. The pl–SceI–GR–iRFP (near-infrared fluorescent protein) construct was made by removing RFP from pl–SceI–GR–RFP [Addgene plasmid no. 17654 (33)] by religating Ecl136II-blunted AarI sites, creating an in-frame deletion. Next, the BamHI+BglII fragment from pIRES-H2BIRFP was ligated into the same sites of the i–SceI–GR containing plasmid to generate pl–SceI–GR–iRFP.

**Stable cell lines**
The Flp-In T-REx 293 cell line (Thermo Fischer Scientific) was used to allow for single genomic integration of the reporter genes by Flp recombinase–mediated integration. Cells were grown as monolayers in high-glucose (4.5 g/liter) Dulbecco’s modified Eagle medium (DMEM) supplemented with 10% (v/v) fetal bovine serum and 1% (v/v) l-glutamine (Thermo Fisher Scientific). To create the PROP, EX2, and EX2-AS reporter cells, the pOG44 Flp recombinase expression vector was cotransfected with the respective reporter construct. Transfected cells were maintained under selection with hygromycin B (200 μg/ml; Roche) and blasticidin (15 μg/ml; InvivoGen) at 37°C in a humidified atmosphere with 5% CO2.

**Live cell experiments and transfections**
Where indicated, the reporter cell lines were seeded in DMEM with Hepes, without phenol red (Gibco) supplemented with 10% (v/v) charcoal-stripped fetal bovine serum (Sigma-Aldrich). For the live cell imaging experiments, cells were plated on 35-mm petri dishes with 10-mm glass-bottom microwell (coverglass thickness no. 1.5) (MatTek Corporation). Cells were transiently transfected with one or more plasmids simultaneously 16 to 24 hours before experiments using Lipofectamine 3000 reagent (Thermo Fisher Scientific), according to the manufacturer’s protocol. Transcription of the reporter genes was induced with doxycycline (0.5 μg/ml; D9891, Sigma-Aldrich), where indicated, and nuclear translocation of pl–SceI–GR–iRFP or pl–SceI–GR–RFP was induced by adding TA (T6501, Sigma-Aldrich), prepared in dimethyl sulfoxide (DMSO), and, at the time of treatment, diluted in the abovementioned culture medium (final concentration, 10−7 M).

**In vitro I–SceI cleavage efficiency assays**
To determine the in vitro cleavage efficiency by the restriction enzyme I–SceI, we first linearized each reporter gene plasmid by a restriction enzyme digest, followed by enzyme inactivation, and extracted the DNA using a column purification protocol (NZYTech): The EX2 reporter construct was Xho I digested, and the PROP and EX2-AS reporter constructs were each Hind III digested (all enzymes are from Thermo Fisher Scientific). Incubation of the predigested EX2 plasmid with the I–SceI restriction enzyme (Thermo Fisher Scientific) generated 1.5– and 1.2–kilo-bp (kbp) fragments. Aliquots were removed at different time points (2.5, 5, 10, and 15 min after incubation with I–SceI). Predigested PROP and EX2-AS plasmids incubated for 15 min with I–SceI generated one additional fragment of 1.5 and 2.0 kbp, respectively, and all digests were analyzed on 1% agarose gels, together with non–I–SceI–digested controls. DNA was visualized by GelRed fluorescence (Biotium), and agarose gel images were acquired using the Bio-Rad gel imager. The bands were quantified using ImageJ software. The in vitro I–SceI cleavage efficiency was calculated by measuring the background-corrected integrated band intensity of the uncleaved 2.7 kbp and the total of the cleaved 1.5- + 1.2-kbp fragments at the indicated times of I–SceI incubation.

**Live cell spinning disk confocal imaging and image analysis**
Live cell imaging was performed on a 3i Marianas SDC spinning disk confocal imaging system (Intelligent Imaging Innovations Inc.) using a similar microscopy setup previously described (34). The system is based on an Axio Observer Z1 inverted microscope (Carl Zeiss MicroImaging Inc., Germany) equipped with a Yokogawa CSU-X1 spinning disk confocal head (Yokogawa Electric, Tokyo, Japan) and 100-mW solid-state lasers (Coherent Inc., Santa Clara, CA) coupled to an acousto-optic tunable filter. The axial position of the sample was controlled with a piezo-driven stage (Applied Scientific Instrumentation, Eugene, OR). Each MatTek dish was placed in an incubation chamber (PeCon P-Set 2000, PeCon GmbH, Erbach, Germany) mounted on the microscope stage and connected to CO2 (CO2 module S, PeCon) and humidity (Heating Device Humidity 2000, PeCon) controllers. The whole microscope body excluding lasers, camera, and spinning disk head was maintained inside a large plexiglass environmental chamber (PeCon, Erbach, Germany). The temperature in both the microscope and top stage incubation chambers was controlled by a common unit and set to 37°C. The environment inside the top stage incubation chamber was further set to 5% CO2 and 100% humidity. Samples were illuminated with λ = 488 nm for GFP, λ = 561 nm for TagRFPt, and λ = 640 nm for iRFP. Images were acquired using a 100× (Plan Apo, 1.4 numerical aperture) oil immersion objectives (Carl Zeiss Microimaging Inc.) under the control of SlideBook 6.0 software (Intelligent Imaging Innovations, Denver, CO). Three-dimensional (3D) time-lapse image stacks of 12 to 16 optical slices separated by 0.4 μm were collected every 30 s for 50 to 60 min, with exposure acquisition times between 40 and 50 ms. Digital images (16-bit) were acquired using a back thinned air-cooled electron-multiplying charge-coupled device camera (Evolve 512, Photometrics, Tucson, AZ). 3D time-lapse sequences were analyzed with the spot tracking and quantification software tool, STaQTool, as described previously (34, 35). Briefly, this tool was used to track single transcription sites in the cell nucleus over time and calculate the total fluorescence intensity (TFI) for each diffusion-limited spot in a time-lapse sequence by performing a 2D Gaussian fit on the volume of interest at the Z plane corresponding to the highest intensity value (35). Image was used to quantify labeling intensities of a selected ROI defined around the transcription site to determine MDC1 accumulation. Briefly, time-lapse frames in both green (GFP) and red (TagRFPt) were first corrected for fluorescence loss due to photobleaching, as described previously (36). A circular ROI was used to define the transcription site in the RFP channel. The same ROI was then duplicated in the GFP channel and used to determine the average intensity over time of MDC1 in the transcription site for the whole time-lapse sequence.

**Microscopy calibration for single transcript detection**
Calibration experiments to relate TFI measurements to single molecule and single transcript numbers were carried out as follows: HEK293 cells were transfected with pEGFP-C1 (Clontech) and harvested on ice the next day. Cytoplasmic extract was prepared by swelling and lysing cells in RSB (resuspension buffer) (10 mM tris, 10 mM NaCl, and 3 mM MgCl2). Cytoplasmic extract was then diluted in 1× phosphate-buffered saline (PBS) at 1:50 to 1:1000, and 1 to 2 μl were placed between two 30-mm no. 1.5 coverslips (Thermo Fisher Scientific) and placed in a PeCon LOCmini-2 cell chamber as described (34). The chamber was placed in the same stage incubator as described above, and GFP...
molecules were imaged with the same laser intensity as in the live cell experiments but with exposure times of 1000 ms over 100 time points in one plane. Image series were analyzed using the StaQTool to detect fluorescent objects and measure TFI values. Single GFP molecules were identified by detecting events irreversible and rapid loss of fluorescence bleaching, where the corresponding difference of the TFI value was fitted and recorded as TFI of a single GFP molecule (fig. S4A). Live cell experiments to determine the labeling rate of single reporter gene transcripts were performed using PROP and EX2 reporter cell lines transiently transfected with MS2-GFP or PP7-GFP, respectively. One day after transfection, the reporter gene transcription was induced with doxycycline (0.5 μg/ml) for 1 hour, and single imaging plane time series of 100 time points with 500-ms intervals were recorded. Particles with high mobility in the cell nucleus represent single-labeled mRNA transcripts in contrast to stationary larger objects representing sites of reporter gene transcription with multiple-labeled RNAs present. Analysis of the brightness of single-labeled reporter gene mRNAs was performed using the STAqTool by determining the TFI of single-labeled mRNA molecules in the image frame (fig. S4B). The mean TFI value of single GFP molecules was related to the mean TFI of MS2-GFP– or PP7-GFP–labeled transcripts to determine the labeling ratio (fig. S4C). Additional live and fixed cell calibration measurements to detect single mRNA transcripts labeled by MS2-GFP, MS2-RFP, and PP7-GFP were performed accordingly to determine the single transcript labeling mean TFI values for the respective labeling protein and imaging settings used in the live cell I–SceI cutting assays (fig. S5). Particles with high mobility in the cell nucleus represent single-labeled mRNA transcripts in contrast to stationary larger objects representing sites of reporter gene transcription with multiple-labeled RNAs present (fig. S5, A to D). To improve the TFI measurements of double-labeled transcripts from the EX2 reporter gene, cells were double transfected (MS2-RFP and PP7-GFP) and, after transcription induction for 1 hour, fixed for 10 min in 3.7% formaldehyde, washed in PBS, and embedded in VECTASHIELD mounting medium (Vector Laboratories). Image stacks of 30 to 40 optical slices at 0.27-μm intervals were taken using the same imaging conditions as in live cell experiments. Analysis was performed using the STAqTool by determining the TFI of single-labeled mRNA molecules in the image frame (live cells) or plane (fixed cells) of highest intensity (fig. S5, E and F) in cells, where the site of reporter gene transcription was identified as a much larger and brighter spot within the nucleus and excluded from analysis. Our results show that around 13 (MS2) to 15 (PP7) of 24 RNA stem loops were occupied, on average, and confirm a less than 100% labeling ratio detected before.

**Chromatin immunoprecipitation**

ChIP was performed as described (14, 37). Briefly, EX2 and PROP cells were transfected with I–SceI–GR-RFP and, 24 hours later, treated with TA or DMSO (nontreated) as mentioned above. Then, cells were cross-linked in 1% (v/v) formaldehyde for 10 min and subsequently quenched with 250 mM glycine for 5 min. Precleared chromatin was incubated with 3 μg of a specific histone H3 antibody (ab1791, Abcam), H3K36me3 (ab9050, Abcam), or H3K9Ac (ab4441, Abcam) overnight at 4°C. One-tenth of the sample was taken as input control. Immunoprecipitated (IP) DNA were subsequently purified and quantified. The relative occupancy of the IP protein at the specified DNA sites was estimated by real-time quantitative PCR (RT-qPCR) as follows: $2^{(C_t - C_t \_input)}$. Where $C_t$ is the mean threshold cycles of RT-qPCR done in duplicate on DNA samples from input and specific immunoprecipitations, respectively. Data were normalized against the ChIP values obtained in control cells (i.e., without TA). Primer sequences are shown in table S1.

**MNase assay**

This assay was performed as described (14), with some modifications. Briefly, EX2 cells were prepared as described for the ChIP protocol and, after TA incubation, collected by trypsinization. Cells were resuspended in ice-cold RSBG40 buffer [10 mM tris (pH 7.4), 10 mM NaCl, 3 mM MgCl₂, 10% glycerol, and 0.5% NP-40] and centrifuged (3 min, 7000 rpm at 4°C) to isolate nuclei. Nuclei were washed with digestion buffer [10 mM tris (pH 7.4), 15 mM NaCl, 60 mM KCl, and 1 mM CaCl₂], centrifuged (3 min, 7000 rpm at 4°C), and resuspended in digestion buffer. One-half was digested with 10 U of MNase (Fermentas) for 30 min at 37°C before incubation with stop buffer [100 mM EDTA and proteasine K (20 μg/ml)] for 30 min at 55°C, and another half was treated with buffer alone (nondigested). Mononucleosome-sized DNA obtained after 30-min digestion and nondigested DNA were used after standard phenol-chloroform purification as templates for RT-qPCR. The amount of MNase-resistant DNA at each gene specific region was estimated as follows: $2^{(C_t \_nondigested - C_t \_MNase digested)}$. Where $C_t$ is the mean threshold cycles of RT-qPCR done in duplicate on DNA samples from nondigested ($t_0$) and 30-min MNase-digested ($t_{30}$) samples, respectively. Data were normalized against the MNase values were obtained in control cells (i.e., without TA). Primer sequences are shown in table S1.

**RNA interference**

Cells were seeded into p35 dishes and simultaneously transfected with 5 μl of 10 μM duplex siRNAs (sequences given in table S1) (TriFecta Kit DsiRNA Duplex, Integrated DNA Technologies) using Opti-MEM (Invitrogen) and Lipofectamine RNAiMAX (Invitrogen), according to the manufacturer’s instructions. The following day, cells were split and seeded into two glass-bottom microscopy dishes in the medium indicated above for live cell imaging experiments and transfected again with siRNA using the same protocol. After 8 hours, around 30% of the growth medium was exchanged, and the cells were transfected with the plasmids encoding fluorescent proteins as indicated above.

**Western blot**

Whole-cell protein extracts were prepared by cell lysis with SDS-polyacrylamide gel electrophoresis (PAGE) buffer [80 mM tris-HCl (pH 6.8), 16% glycerol, 4.5% SDS, 450 mM dithiothreitol, and 0.01% bromophenol blue], with benzonase (200 U/ml; Sigma-Aldrich) and 50 μM MgCl₂ and boiling for 5 min. Equal amounts of protein extracts were resolved by SDS-PAGE and transferred to a nitrocellulose membrane. Immunoblotting was performed with the following antibodies: rabbit anti-DNA ligase IV (ab193353, Abcam), mouse anti–α-tubulin (T5168, Sigma-Aldrich), rabbit anti–total histone H3 (ab1791, Abcam) and using the secondary antibodies goat anti-mouse horseradish peroxidase (HRP) (170-6516, Bio-Rrad) and goat anti-rabbit HRP (170-6515, Bio-Rad).

**Statistical analysis**

Where indicated, data were subjected to a two-tailed Student’s t test to resolve statistical significance. The measured TFI values of GFP molecules and single-labeled mRNA transcripts were plotted into histograms and fitted with a Gaussian function according to the formula $y = y_0 + \frac{A}{a \sqrt{\pi/2}} e^{-\frac{(x-x_0)^2}{a^2}}$ to determine the mean value $x_0$ of the intensity
distribution in OriginPro 8. The mean TFI value for single GFP molecules was related to the mean single transcript intensities via conversion of the exposure time at identical imaging settings. The comparison of DNA ligase IV knockdown and control live cell experiment datasets was performed using Fisher’s exact test for $2 \times 2$ contingency tables and a $\chi^2$ test for contingency tables, both at a level of significance of $P = 0.05$.

Bioinformatics analyses
Endogenous DSBs on NHEK cells were obtained from a previous study (18), considering only the high-confidence 84,946 DSBs common to DSBCapture replicates. Transcription was assessed using polyA-depleted and strand-specific RNA sequencing (RNA-seq) data from nuclear fractions of NHEK cells (GSM2072453). DNA accessibility profiles (DNase-seq) for NHEK cells were previously identified by the National Institutes of Health Roadmap Epigenomics Mapping Consortium. DSB and DNA accessibility coordinates were converted from hg19 into hg38 human genome versions using liftOver (38). Gene annotations were obtained from GENCODE (v26 version), merged into a single transcript model per gene and removed overlapping genes using BEDTools (39). Only DSBs located in the gene body region [500 bp downstream TSS (transcription start-site) to TTS (transcription termination-site)] of transcriptionally active genes were considered in the downstream analyses. Transcriptionally active genes were defined as those with expression levels [TPMs (transcripts per million)] from Kallisto (40) higher than the 50th percentile (median). DSBs with antisense transcription were determined with a cutoff of 0.01 RPKMs (reads per kilobase per million mapped reads) in the 500-bp flanking regions, according to the strand-specific information. For the metaprofiles, DSBs were aligned by the median region, and the read density for the flanking 1 kb was averaged in a 20-bp window. All profiles and heatmaps were plotted on normalized RPKMs considering 20-bp windows. A set of in-house scripts for data processing and graphical visualization were written in bash and in the R environmental language (www.R-project.org). SAMtools (41) and BEDTools were used for alignment manipulation, filtering steps, file format conversion, and comparison of genomic features. Fold enrichment over random and for alignment manipulation, filtering steps, file format conversion, and application was performed using Fisher’s exact test for $2 \times 2$ contingency tables and a $\chi^2$ test for contingency tables, both at a level of significance of $P = 0.05$.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/1/eaau1249/DC1
Fig. S1. Schematics of the reporter genes.
Fig. S2. Detection of single RNAs and measurement of transcription rates.
Fig. S3. I-Scel cleaves the reporter genes driving MDC1 recruitment to damaged loci.
Fig. S4. Single-molecule calibration measurements.
Fig. S5. Single transcript calibration measurements.
Fig. S6. Break-induced nucleosome depletion.
Table S1. Sequence of primers used in this study.

REFERENCES AND NOTES


Acknowledgments: We thank R.H. Singer, J.A. Chao, T. Misteli, M. Carmo-Fonseca, E. Gomes, J. Lukas, and C. M Azzalin for comments and gifts of plasmids and reagents. We are also indebted to J. Rino, A. Temudo, and A. Nascimento from the iMM Bioimaging Facility and S. Carvalho for excellent technical support. Funding: This work was funded by PTDC/BIM-ONC/0016/2014, PTDC/MED-OUT/32271/2017, and PTDC/BIA-MOL/30438/2017 from Fundação para a Ciência e a Tecnologia (FCT), Portugal and by LISBOA-01-0145-FEDER-007391, project cofunded by FEDER through POR Lisboa, Portugal 2020–Programa Operacional Regional de Lisboa, and FCT. A.R.G. is the recipient of a FCT Investigator grant (IF/00510/2014). A.C.V. and J.C.S. are recipients of FCT PhD fellowships SFRH/BD/52322/2013 and PD/BD/128292/2017.

Author contributions: A.C.V. performed most cell culture, microscopy, and biochemical experiments and acquired, analyzed, and interpreted data. S.C.S. and J.C.S. performed biochemical assays and analyzed data. A.I.A. performed cloning experiments. A.R.G. analyzed and interpreted RNA-seq, DSBCapture, and DNase-seq data. R.M.M. generated the reporter genes, performed microscopy experiments, and acquired, analyzed, and interpreted data.

S.F.d.A. conceived the study, analyzed and interpreted data, and wrote the manuscript with the assistance of all authors. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.
Single-molecule imaging of transcription at damaged chromatin
Alexandra C. Vitó, Sreerama C. Sridhara, João C. Sabino, Ana I. Alfonso, Ana R. Grosso, Robert M. Martin and Sérgio F. de Almeida

DOI: 10.1126/sciadv.aau1249