DNA DAMAGE-INDUCED TRANSCRIPTION STRESS
A focus on RNA polymerase II

Barbara Steurer

DNA DAMAGE-INDUCED TRANSCRIPTION STRESS
A focus on RNA polymerase II

TRANSCRIPTIE STRESS VEROORZAAKT DOOR DNA SCHADE
Een focus op RNA polymerase II

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board

The public defense shall be held on

Tuesday, December 10, 2019 at 11:30

by

Barbara Steurer

born in Bregenz, Austria

Erasmus University Rotterdam

PROMOTOR

W. Vermeulen

OTHER MEMBERS

J. H. J. Hoeijmakers
R.A. Poot
P.J. Verschure

COPROMOTOR

J.A.F. Marteijn

TABLE OF CONTENTS

SCOPE OF THIS THESIS

Accurate expression of our genes is facilitated by the tight regulation of RNA Polymerase II (Pol II) mediated transcription and by maintaining the integrity of the DNA template. DNA lesions that are located in the transcribed strand of genes may impede or completely block Pol II transcription. Such transcription-blocking DNA lesions (TBLs) cause DNA-damaged-induced transcription stress, which may result in reduced cell function or cell death. To counteract these serious effects, most TBLs are resolved by transcription-coupled nucleotide excision repair (TC-NER). TC-NER is a highly conserved multi-step DNA-repair pathway that is initiated when elongating RNA polymerase II stalls on a TBL. While most key proteins involved in this versatile pathway have been extensively studied since the discovery of the pathway three decades ago, several aspects regarding the spatio-temporal coordination of Pol II regulation and the adaptation of transcription to stressed conditions are not yet fully understood. These aspects are thoroughly reviewed and discussed in **chapter 1**. For example, while the physical block of elongating Pol II on DNA lesions (*in cis*) has long been acknowledged as the main cause for reduced Pol II transcription rates upon UV irradiation, recently accumulating evidence suggests that the signal-transduced regulation of Pol II throughout the nucleus (*in trans)* is an additional means to remotely control transcription in response to genotoxic stress.

To study putative effects of transcription stress on Pol II *in trans,* it is crucial that TBLs are limited to a relatively small number of genes. Therefore we developed a highly sensitive method that allows monitoring Pol II behavior in living cells after the induction of very low damage loads. In **chapter 2** we describe the CRISPR/Cas9-mediated generation of GFP-RPB1 (RPB1 is the largest subunit of Pol II) knock-in cells and their application as a live-cell imaging tool to determine the in vivo kinetics of endogenous Pol II. In contrast to the methods usually applied to study Pol II behavior, such as ChIP- or run-on-sequencing, this approach allows direct, real time assessment of Pol II dynamics without the use of transcription inhibitors or fixatives. Combined with computational modeling this approach allowed to kinetically dissect promoter-paused Pol II from initiating and elongating Pol II and showed that initiation and promoter proximal pausing are surprisingly dynamic events due to premature termination of Pol II promoter-binding. Our study provides new insights into Pol II dynamics and suggests that the iterative release and re-initiation of promoter-bound Pol II is an important component of transcriptional regulation.

Investigating the - possibly subtle - role of newly identified factors in their contribution to DNA repair efficiency is challenging, as the currently available assays to quantify UV-induced DNA damage and repair are confined to endpoint measurements, are often limited by antibody specificity, and cannot be performed in living cells or at the single cell level. To circumvent these challenges, in **chapter 5**, we describe the generation of fluorescently-tagged UV lesion-specific photolyases (PLs) as a highly sensitive new tool to monitor DNA repair kinetics in living cells. CPD-PLs and 6-4PP-PLs efficiently recognize and specifically bind to the respective photo lesion, allowing direct, real-time quantitation of UV-induced DNA damage *in vivo*. Furthermore, we showed that, using the 405 nm laser during live cell imaging experiments, PLs can be enzymatically activated to specifically photo-reactivate CPDs or 64PP lesions. This not only facilitates studying the behavior of repair factors upon instantaneous DNA repair in living cells, but also enables investigating whether 6-4PP and CPD lesions might trigger distinct cellular responses.

In **Chapter 6** the main findings of this thesis and their implications on our current understanding of the cellular responses triggered by UV-induced DNA damage are highlighted and discussed. Finally, we propose several future research directions that may help to expand our insights into the DNA damage-induced transcription stress response.

Chapter 1

INTRODUCTION

0Barbara Steurer[1], Jurgen A. Marteijn[1]*

[1] Department of Molecular Genetics, Erasmus MC, Wytemaweg 80,
Rotterdam 3015 CN, The Netherlands
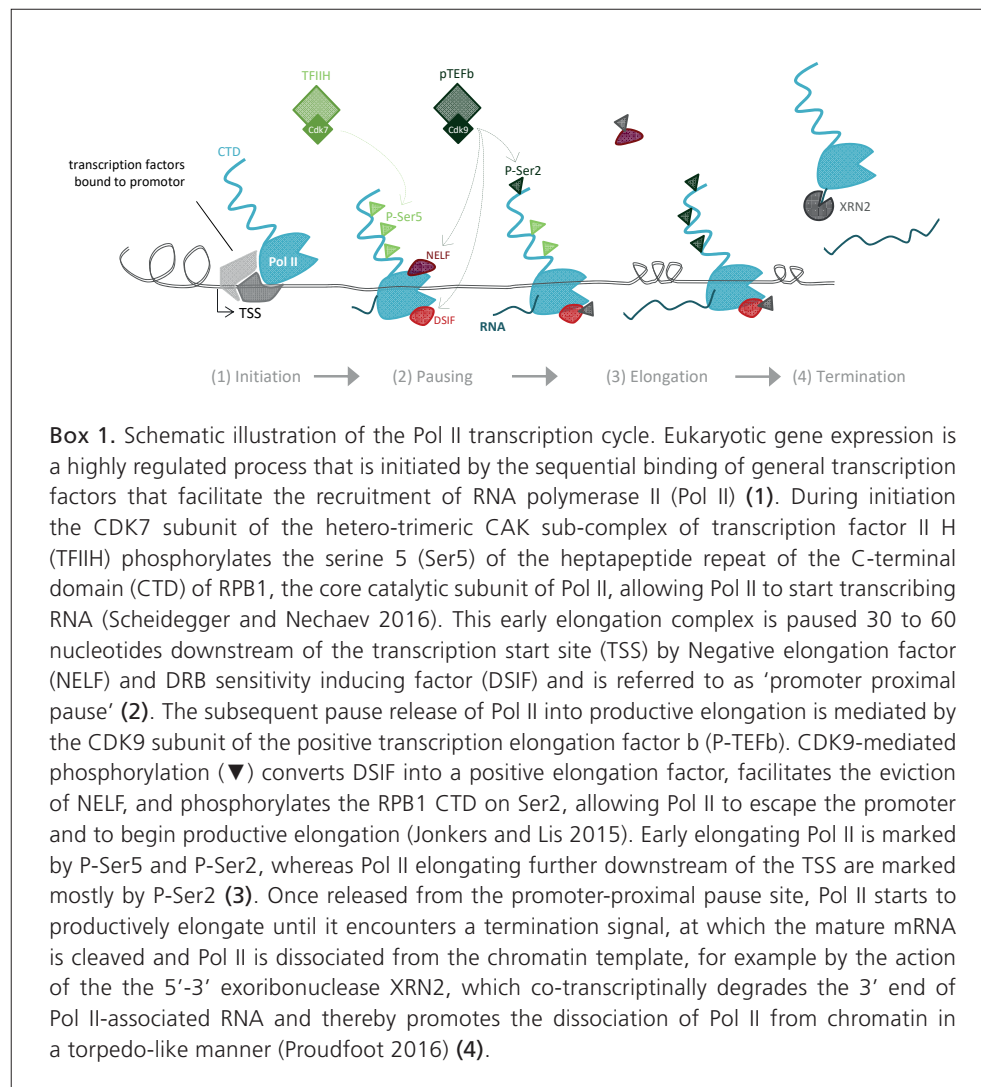* Corresponding author: Jurgen A. Marteijn

ABSTRACT

The faithful transcription of eukaryotic genes by RNA polymerase II (Pol II) is crucial for proper cell function and tissue homeostasis. However, transcription-blocking DNA lesions of both endogenous and environmental origin continuously challenge the progression of elongating Pol II. The stalling of Pol II on a transcription-blocking lesion triggers a series of highly regulated events, including Pol II processing to make the lesion accessible for DNA repair, R-loop-mediated DNA damage signaling, and the initiation of transcription-coupled DNA repair. The correct execution and coordination of these processes is vital for resuming transcription following the successful repair of transcription-blocking lesions. Here we outline recent insights into the molecular consequences of Pol II stalling on transcription-blocking DNA lesions and how these lesions are resolved to restore mRNA synthesis.

INTRODUCTION

The accurate transcription of genes by RNA polymerase II (Pol II) is crucial for proper cell function and is therefore tightly regulated at each step of the Pol II transcription cycle (Box1) [1]. However, DNA damage continuously compromises the efficiency and fidelity of DNA transcription and threatens cell viability and genome integrity. Many different DNA damaging agents, of both endogenous and environmental origin, can cause DNA injuries that block or strongly hinder RNA polymerase II (Pol II) transcription elongation. Furthermore, in cycling cells, advancing replication forks can collide with stalled Pol II complexes [2]. The arrest of Pol II on transcription-blocking lesions (TBLs) leads to a lack of newly synthesized RNA molecules or may result in mutant mRNA. Not only these effects on RNA expression but also the prolonged arrest of Pol II itself are both highly cytotoxic. The stalling of Pol II on lesions for extended periods of time can arrest cell cycle progression and lead to apoptosis [3, 4]. If TBLs remain unrepaired, this blocked transcription can cause severe cellular dysfunction, eventually resulting in DNA-damage-induced aging [5-7]. The structural complexity of lesion-stalled Pol II requires that an intricate protein network needs to be activated to ensure removal of genomic roadblocks and to overcome blocked transcription. The stalling of elongating Pol II on DNA lesions initiates transcription-coupled DNA repair (TC-NER), which is a multistep pathway that efficiently removes DNA lesions specifically from the transcribed strand of active genes. TC-NER is a sub-pathway of the multistep DNA repair pathway nucleotide excision repair (NER). NER can also be initiated via global genome NER (GG-NER), which recognizes helix-destabilizing DNA lesions throughout the genome (Box 2) [8]. Only upon completion of TC-NER stalled transcription will restart [9]. The biological relevance of this DNA repair pathway is best demonstrated by the severe phenotypes of human disorders that are related to defective TC-NER [7, 9, 10]. However, even though the concept of TC-NER was discovered three decades ago [9, 11], many questions remain unanswered about how cells coordinate transcription arrest and TBL repair, and subsequently restart mRNA synthesis. Here, we discuss the multifaceted cellular response that is triggered following the stalling of Pol II on TBLs.

Fates of lesion-stalled Pol II

To repair TBLs, TC-NER faces a significant steric problem: Pol II may be trapped near to or right on top of a TBL, severely obstructing the access of repair factors to the lesion [7, 9] (Fig. 1a). Different types of TBLs differentially inhibit the forward translocation of the transcription machinery [9, 12]. For example, UV-induced cyclobutane-pyrimidine dimers cause the arrest of Pol II on top of the TBL. The 35-nucleotide footprint of the stalled Pol II is asymmetrically located around the lesion, covering 10 nucleotides downstream and 25 nucleotides upstream of the UV-lesion [13-15]. By contrast, cisplatin-induced inter-strand crosslinks stall Pol II before the lesion can enter the polymerase's active site [16]. Oxidative DNA lesions such as 8-oxo-7,8-dihydroguanine (8-oxo-G), which

Box 1. Schematic illustration of the Pol II transcription cycle. Eukaryotic gene expression is a highly regulated process that is initiated by the sequential binding of general transcription factors that facilitate the recruitment of RNA polymerase II (Pol II) **(1)**. During initiation the CDK7 subunit of the hetero-trimeric CAK sub-complex of transcription factor II H (TFIIH) phosphorylates the serine 5 (Ser5) of the heptapeptide repeat of the C-terminal domain (CTD) of RPB1, the core catalytic subunit of Pol II, allowing Pol II to start transcribing RNA (Scheidegger and Nechaev 2016). This early elongation complex is paused 30 to 60 nucleotides downstream of the transcription start site (TSS) by Negative elongation factor (NELF) and DRB sensitivity inducing factor (DSIF) and is referred to as 'promoter proximal pause' **(2)**. The subsequent pause release of Pol II into productive elongation is mediated by the CDK9 subunit of the positive transcription elongation factor b (P-TEFb). CDK9-mediated phosphorylation (▼) converts DSIF into a positive elongation factor, facilitates the eviction of NELF, and phosphorylates the RPB1 CTD on Ser2, allowing Pol II to escape the promoter and to begin productive elongation (Jonkers and Lis 2015). Early elongating Pol II is marked by P-Ser5 and P-Ser2, whereas Pol II elongating further downstream of the TSS are marked mostly by P-Ser2 **(3)**. Once released from the promoter-proximal pause site, Pol II starts to productively elongate until it encounters a termination signal, at which the mature mRNA is cleaved and Pol II is dissociated from the chromatin template, for example by the action of the the 5'-3' exoribonuclease XRN2, which co-transcriptinally degrades the 3' end of Pol II-associated RNA and thereby promotes the dissociation of Pol II from chromatin in a torpedo-like manner (Proudfoot 2016) **(4)**.

are induced by endogenous reactive oxygen species, also interfere with transcription. However, the damage-induced transcription stalling of Pol II does not appear to be caused by 8-oxo-G itself, but rather indirectly by base excision repair intermediates [17-19].

To overcome persistent Pol II stalling and to facilitate access of the DNA repair machinery, cells have evolved three different mechanisms to displace lesion-stalled Pol II: reverse translocation, degradation, and lesion bypass (Fig. 1b).

Reverse translocation, or backtracking, of Pol II not only occurs in the presence of TBLs, but also when Pol II encounters DNA sequences that are difficult to transcribe [20]. In bacteria, the DNA helicase UvrD travels along with elongating Pol II and moves the complex backwards upon encountering a DNA lesion [21]; and a similar backtracking mechanism has also been suggested for higher eukaryotes, though this
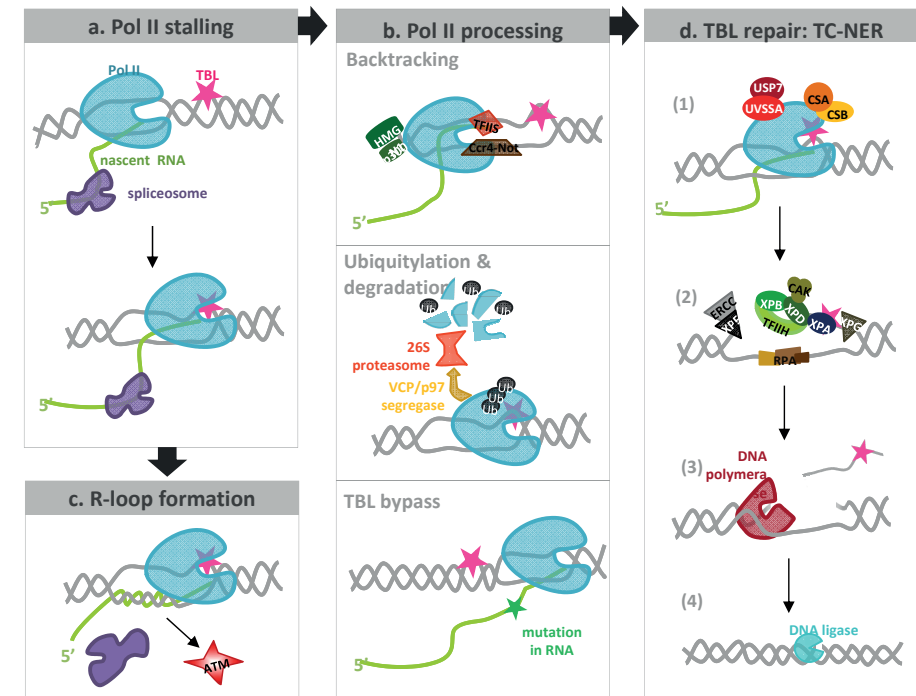

Figure 1. The arrest of elongating RNA polymerase II (POL II) on a transcription blocking DNA lesion (TBL) triggers a series of cellular events. **(a)** Elongating POL II runs into a TBL and stalls **(b)** R-loops can be formed by hybridisation of pre-mRNA with template ssDNA adjacent to the transcription bubble. TBL induced R-loop formation activates non-canonical ATM signaling, whch in turn results in eviction of co-transcriptional spliceosomes. **(c)** To allow access of the repair machinery to TBLs, the damage-stalled POL II has to be removed from the lesion. POL II processing might occur via TFIIS and Ccr4-Not mediated backtracking (top panel). Alternatively RPB1, the largest subunit of the POL II complex, might be ubiquitylated and proteasomally degraded. Segregase activity is needed to extract RPB1 from chromatin (middle panel). Lesion bypass of POL II might also make the lesion accessible, however this might result in mutant RNA (bottom panel). **(d)** Transcription Coupled DNA repair (TC-NER) is initiated when POL II stalls at a TBL during transcript elongation. Whether TC-NER stimulates backtracking or backtracking is needed to initiate TC-NER is unknown. During transcript elongation UV-stimulated scaffold protein A (UVSSA), ubiquitin-specific-processing protease 7 (USP7) and Cockayne syndrome protein B (CSB) transiently interact with POL II. Upon stalling at a TBL, the affinity of CSB for RNA Pol II increases, which recruits the WD40 protein CSA. CSA and CSB complex are required for the subsequent steps of the NER reaction (step1). After damage recognition, the TFIIH (transcription initiation factor IIH) complex is recruited to the lesion and the structure- specific endonuclease XPG binds to the pre-incision NER complex. The helicase activity of TFIIH further opens the double helix around the lesion. The TFIIH helicase subunit XPD unwinds the DNA 5'–3' and verifies the existence of lesions with the help of the ATPase activity of the TFIIH subunit XPB subunit and XPA. XPA and RPA then recruit the endonuclease XPF–ERCC1, which creates an incision 5' to the TBL. This results in the activation of XPG, which cuts the damaged strand 3' to the lesion. This excises the lesion within a 22–30 nucleotide-long strand (step3). Gap filling synthesis by DNA Pol δ, DNA Pol κ or DNA Pol ε can begin immediately after the 5' incision is made. The NER reaction is completed through sealing the final nick by DNA ligase 1 or DNA ligase 3 (step4).
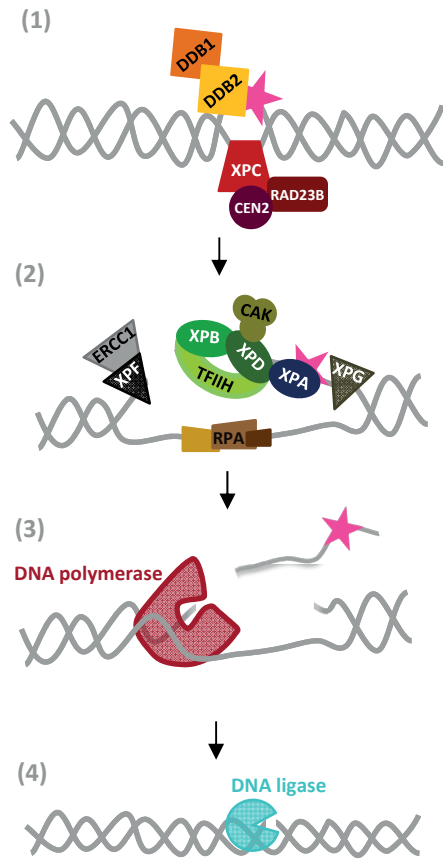
Box 2. Global Genome NER (GG-NER). The main damage sensor in GG-NER is XPC. In complex with RAD23B and centrin 2 (Cen2) (Hoogstraten, Bergink et al. 2008), XPC constantly probes the DNA for helix-distorting lesions (Nishi, Okuda et al. 2005). Mildly helix distorting lesions such as CPDs, which are poor substrates for XPC, require the action of the UV-DDB2 complex, consisting of DDB1 and DDB2. The latter binds damaged base pairs, kinks the DNA backbone to extrude the lesion into its binding pocket and thereby facilitates XPC binding opposite to the DNA lesion (Sugasawa, Okamoto et al. 2001) **(1)**. Lesion-bound XPC subsequently recruits transcription factor II H to sites of DNA damage (TFIIH) (Riedl, Hanaoka et al. 2003). After damage recognition, GG-NER and TC-NER converge into a mutual pathway, collectively referred to as NER, which is described in Figure 1.

has not yet been confirmed [8, 20-22]. To resume transcription after backtracking, the protruding nascent RNA needs to be cleaved to reposition the 3' end of the RNA in the active site of the polymerase [23]. In eukaryotes, this reaction is mediated by transcription factor IIS (TFIIS), which stimulates the intrinsic 3'–5' exonuclease activity of Pol II [13, 24-26]. A recent study showed that TFIIS recruitment to elongating Pol II is increased by the Ccr4-Not complex, and consequently the authors suggested that TFIIS

and Ccr4-Not work together to reactivate arrested Pol II [27]. In addition, Ccr4-Not may promote the resumption of elongation by binding to the emerging transcript protruding from the polymerase [28].

Pol II backtracking upon collision with a TBL would provide the space needed for the TC-NER machinery to repair the TBL. This principle was elegantly demonstrated by researchers in the Hanawalt laboratory [13, 29], who showed that photolyases, which specifically bind UV-induced DNA lesions, could only recognize TBLs following the TFIIS-mediated backtracking of arrested Pol II [29]. Furthermore, TFIIS was shown to be involved in the efficient recovery of transcription following UV irradiation, emphasizing its role in TC-NER [30]. While little is still known about factors that mediate Pol II backtracking, the process may be facilitated by sliding of the upstream nucleosomes by the histone acetyltransferase p300 and the nucleosome binding protein HMGN1, both of which interact with stalled Pol II [9, 31]. In addition, the key TC-NER protein Cockayne syndrome B (CSB) may be involved in the displacement of stalled Pol II, as it contains a SWItch/sucrose non-fermentable (SWI/SNF2) ATPase domain and has chromatin remodeling activity that is stimulated by the histone chaperone NAP1 [32-34].

If backtracking fails, arrested Pol II may be degraded instead, most likely to prevent genomic roadblocks being caused by its persistent stalling. Ubiquitylation and degradation of RPB1, the largest and core catalytic subunit of Pol II, also occurs during basal transcription elongation [35, 36]; however, it is greatly increased following genotoxic stress [37, 38]. After a decade of discovering the individual factors that are involved in RPB1 degradation [39-42], Harremann and colleagues clarified the pathway in yeast by ordering the actions of distinct and sequentially acting ubiquitin ligases and de-ubiquitylating enzymes (DUBs) [43].

In yeast, the HECT ubiquitin ligase Rsp5 binds to the C-terminal domain of RPB1 [39] and modifies the subunit with a K63-linked polyubiquitin chain, which by itself does not trigger proteolysis. This K63-polyubiquitin chain is then trimmed by the DUB Ubp2 [43]. The residual monoubiquitin on RPB1 can either be hydrolyzed by Ubp3, rescuing RPB1 from degradation [44], or extended to K48-linked-polyubiqutin by the Elc1/Cul3 ligase complex, marking RPB1 for proteasomal degradation [43]. Finally, the ring-like AAA+ ATPase CDC48/p97 is required to segregate the K48-polyubiquitylated yeast Rpb1 from chromatin, and to facilitate RPB1 degradation by the 26s proteasome [45]. Remarkably, RPB1 is the only subunit of the 12-subunit Pol II complex that is degraded following UV exposure [46, 47].

It is currently unknown whether the DNA damage-induced degradation of Pol II by the successive action of different ubiquitin ligases is conserved in mammals. The mammalian RSP5 homolog Nedd4 was found to ubiquitylate RPB1 in human cells, resulting in its degradation upon genotoxic stress [48]. However, Nedd4-depleted cells are not sensitive to UV light [48], indicating that it is not the only factor required to modify RPB1 upon UV exposure. It has also been shown that the von Hippel-Lindau tumor suppressor protein (pVHL) can bind RPB1 in a proline-hydroxylation-dependent manner and functions as an

E3 ligase that targets elongating RPB1 for ubiquitylation and degradation in response to UV light. pVHL negative cells were shown to accumulate elongating RPB1 and undergo apoptosis in response to UV, whereas cells expressing pVHL were less apoptotic [49]. These results clearly indicate that pVHL plays a role in eukaryotic RPB1 degradation. pVHL is a crucial component of the VHL-E3 ubiquitin ligase complex, which consists of Elongin BC, Cullin2, and Rbx1. In this complex, pVHL serves as substrate recognition unit and the Cullin/Rbx module functions as a ubiquitin-activating enzyme [50]. The mammalian ElonginA-ElonginBC-Cul5/Rbx2 complex can also efficiently ubiquitylate RPB1 in vitro [51]. However, rather than pVHL induced degradation of elongating RPB1 (Ser2-phosphorylated), Elongin A and Cul5 interact with initiating RPB1 (Ser5-phosphorylated) upon exposure to UV light. Furthermore, Ser5-phosphorylated RPB1 has also been shown to be a substrate for the BRCA1/BARD1 ligase complex [52]. Interestingly, BRCA1 also ubiquitylates the Pol II subunit RPB8 in response to UV irradiation, but this ubiquitylation does not result in RPB8 degradation [53].

The observation that initiating Ser5-phosphorylated RPB1 is targeted by specific E3 ligases raises the exciting possibility that the collision of Pol II with TBLs in the gene body may also have consequences for transcription-initiating Pol II complexes at the promoter. Since the regulation of transcription is far more complex in eukaryotes than in yeast, including, for example, promoter-proximal pausing [54-56], it is tempting to speculate that Pol II stalling initiates a much more sophisticated cellular response in mammalian cells compared with yeast. In support of this speculation, it was recently shown that UV irradiation results in the loss of Pol II at the promoters of many transcribed genes [57], suggesting a genome-wide mechanism that regulates transcription initiation in response to TBLs.

It is unclear whether the valosin-containing protein VCP/p97, which is the human homologue of the yeast ATPase CDC48/p97, is required for chromatin extraction of mammalian RPB1. Even though several key players in the ubiquitylation of mammalian RPB1 have been identified, our understanding of RPB1 degradation in mammals is incomplete. The identification of many distinct ligases that are involved in the degradation of mammalian RPB1 highlights the importance of Pol II regulation by ubiquitin, but further research is needed to fully understand the precise interplay of all of the factors involved.

Finally, DNA lesions that are encountered by Pol II may be bypassed, although this occurs infrequently [9, 58]. If the helix distortion of a TBL is minimal, such as at a-basic sites or single-strand breaks, it might translocate into the Pol II active site. The subsequent translocation is disfavored, but not totally blocked [58]. Lesion bypass can be stimulated by various transcription factors, such as CSB [59] or TFIIF [60], but this is at the cost of transcriptional mutagenesis [61]. Nucleotide mis-incorporation due to lesion bypass can have serious consequences for the cell if the faulty nucleotide leads to changes in the amino acid coding and the expression of mutant proteins.

Which type of Pol II processing ultimately occurs upon Pol II stalling at a TBL (backtracking, degradation, or bypass) and how these options are regulated in the cell remains largely unknown. The pathway choice is probably influenced by the nature of the TBL and the chromatin environment, but may also be affected by cell type, cell cycle stage, or gene-specific regulation of transcription.

TBL arrest of Pol II induces R-loops, spliceosome eviction, and non-canonical ATM signaling

The association of multi-megadalton spliceosomes with nascent RNA may pose another steric challenge to the repair of TBLs. It was recently reported that late-stage spliceosomes, composed of U2, U5, and U6 small nuclear ribonucleoproteins (snRNPs) are rapidly excluded from DNA damage sites in response to UV-induced TBLs [62]. This displacement of co-transcriptional spliceosomes from arrested Pol II most likely results in an increase in R-loop formation through the hybridization of pre-mRNA with template ssDNA adjacent to the transcription bubble [62, 63] (Fig. 1b). Persistent R-loops are genotoxic, as they can interfere with transcription and replication, increase the probability of replication fork collapse following collisions with stalled transcription complexes, and promote unscheduled replication by transcription-associated recombination. Furthermore, the ssDNA in the R-loop poses a further threat to genome fidelity, as it is sensitive to mutagens, can undergo spontaneous hydrolysis, and is prone to the formation of secondary structures such as G-quadruplexes. To counteract R-loop toxicity, cells are equipped with specialized RNA hydrolases (RNaseH1 and H2) or helicases (e.g., Pif1, DHX9, and senataxin) that unwind the RNA:DNA hybrid [64-67]. In the context of TBLs, R-loop formation leads to non-canonical activation of the ataxia-telangiectasia mutated (ATM) protein kinase, which signals the further mobilization of spliceosomes from elongating polymerases, as well as those that are located distal to Pol II-blocking DNA lesions. The exact molecular mechanism by which TBLs activate ATM remains unclear [62, 63]. Interestingly, ATM, via R-loop formation, relays the local (*cis*) event of Pol II arrest to the genome-wide (*trans*) modulation of alternative splicing, adapting global gene expression, and shaping the proteome in response to TBLs [62, 68].

## Initiation of transcription-coupled repair (TC-NER)

To counteract the fatal implications of lesion-stalled Pol II, TC-NER has evolved to specifically remove a wide range of helix-distorting lesions that impede the elongation of Pol II from actively transcribed genes. TC-NER is initiated by the recruitment of Cockayne syndrome A and B (CSA, CSB) [8, 9], and the UV-stimulated scaffold protein A (UVSSA) [69-71] to lesion-stalled Pol II (Fig. 1d). CSB has ATPase-dependent chromatin remodeling activity and may locally modify the DNA conformation [32, 33, 72]. CSB recruits CSA, which is part of a Cullin-RING ubiquitin E3 ligase complex, that was described to target CSB for ubiquitylation and degradation [73]. CSB degradation is counteracted by UVSSA, which recruits the de-ubiquitylating enzyme USP7 and thereby stabilizes CSB at the site of damage [69, 71]. Although CSA is dispensable for the attraction of the excision repair

machinery, in combination with CSB it is essential for the recruitment of xeroderma pigmentosum group A (XPA)-binding protein 2 (XAB2), a pre-mRNA splicing factor that is involved in TC-NER [73-75]. Following damage detection, the transcription factor II H (TFIIH) complex unwinds a stretch of approximately 30 nucleotides surrounding the damage site. XPA then stimulate the damage-verification activity of TFIIH. XPA and replication protein A (RPA) orient the XPF/excision repair cross-complementing 1 (ERCC1) and XPG endonucleases, which subsequently excise the damaged DNA. The resulting gap is filled by DNA synthesis and sealed by DNA ligases [76, 77]. In addition to NER, another mechanism to remove UV-induced DNA lesions exists, namely the light-activated catalytic reversion of CPD and 6-4PP lesions by photolyases (Box 3). However, this mechanism was lost in eukaryotes during evolution.

Restarting transcription upon repair of TBLs

Although the successful repair of a TBL is necessary, this in itself is not sufficient for transcription restart following genotoxic stress, which is essential for cell survival. Several factors that have explicit roles in TC-NER-associated transcription restart, but not in repair itself, have been identified over the past few years. For instance, the eleven-nineteen lysine rich leukemia (ELL) protein, which interacts with TFIIH via the Cdk7 subunit of the CDK-activating kinase (CAK) complex, was found to be essential for transcription resumption following the removal of TBLs, and yet was not involved in the repair of TBLs [78]. Moreover, downregulation of ELL increased Pol II chromatin retention in a UV-dependent manner. Together, these findings suggest that ELL serves as a docking site for proteins involved in the regulation of Pol II-mediated transcription restart once repair has been completed [78].

The chromatin environment (i.e., histone chaperones, histone variants, and post-translational modifications of histones) also plays an important role during the restart of transcription following DNA damage. For example, knockdown of the histone chaperone HIRA impairs the recovery of RNA synthesis following UV damage to an extent that is comparable to that seen in TC-NER-deficient cells, but does not affect the recruitment of repair factors. HIRA accumulates at sites of DNA damage, where it deposits the histone variant H3.3, which is crucial for facilitating transcription recovery upon the repair of TBLs [79]. In addition, H2A/H2B dimer exchange has also been found to increase at sites of UV-induced DNA damage [80]. This damage-induced histone exchange is mediated by the histone chaperone facilitates chromatin transcription (FACT). FACT is a heterodimer consisting of the SPT16 and SSRP1 subunits, and is a known H2A/H2B chaperone [81]. Although both FACT subunits are recruited to sites of UV damage, only SPT16 depletion results in a loss of damage-induced H2A/H2B exchange. Spt16 is required for the efficient restart of RNA synthesis following UV damage. This suggests that the FACT subunit SPT16 plays a specific role in damage-induced chromatin dynamics and transcription recovery [80]. In addition, knockdown of the methyltransferase disruptor of telomeric silencing 1-like (DOT1L) results in UV-sensitivity, whereas DNA damage is removed
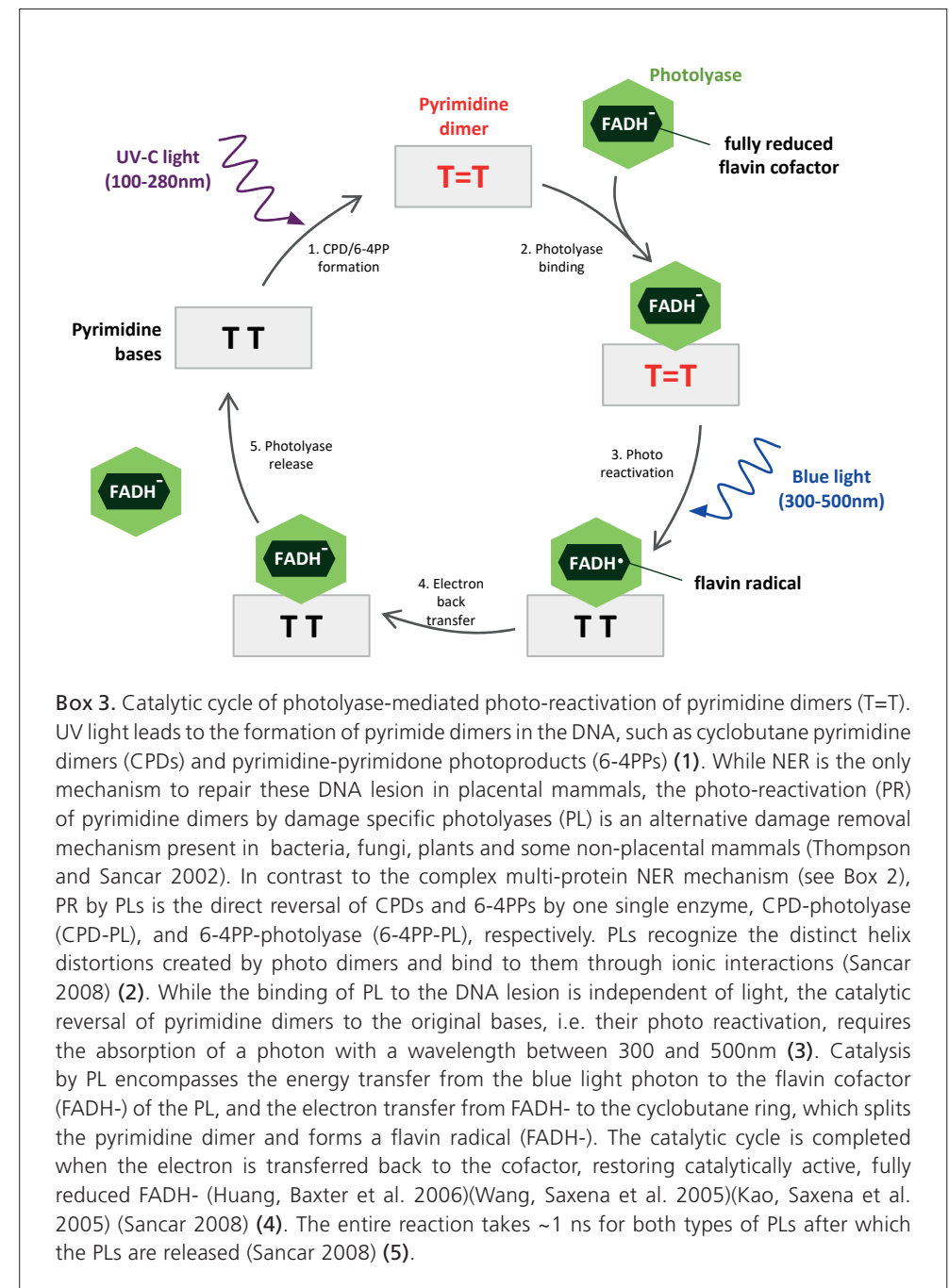
Box 3. Catalytic cycle of photolyase-mediated photo-reactivation of pyrimidine dimers (T=T). UV light leads to the formation of pyrimide dimers in the DNA, such as cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone photoproducts (6-4PPs) **(1)**. While NER is the only mechanism to repair these DNA lesion in placental mammals, the photo-reactivation (PR) of pyrimidine dimers by damage specific photolyases (PL) is an alternative damage removal mechanism present in bacteria, fungi, plants and some non-placental mammals (Thompson and Sancar 2002). In contrast to the complex multi-protein NER mechanism (see Box 2), PR by PLs is the direct reversal of CPDs and 6-4PPs by one single enzyme, CPD-photolyase (CPD-PL), and 6-4PP-photolyase (6-4PP-PL), respectively. PLs recognize the distinct helix distortions created by photo dimers and bind to them through ionic interactions (Sancar 2008) **(2)**. While the binding of PL to the DNA lesion is independent of light, the catalytic reversal of pyrimidine dimers to the original bases, i.e. their photo reactivation, requires the absorption of a photon with a wavelength between 300 and 500nm **(3)**. Catalysis by PL encompasses the energy transfer from the blue light photon to the flavin cofactor (FADH-) of the PL, and the electron transfer from FADH- to the cyclobutane ring, which splits the pyrimidine dimer and forms a flavin radical (FADH-). The catalytic cycle is completed when the electron is transferred back to the cofactor, restoring catalytically active, fully reduced FADH- (Huang, Baxter et al. 2006)(Wang, Saxena et al. 2005)(Kao, Saxena et al. 2005) (Sancar 2008) **(4)**. The entire reaction takes ~1 ns for both types of PLs after which the PLs are released (Sancar 2008) **(5)**.

normally [82]. Thus, the activities of HIRA, FACT, and DOT1L are thought to generate the proper chromatin environment or provide the correct chromatin plasticity needed for efficient transcription recovery following the removal of TBLs [83]. Interestingly, transcription restart following treatment with the transcription inhibitor 5,6-dichloro-

1-beta-D-ribofuranosylbenzimidazole (DRB) occurs independently of DOT1L and HIRA [79, 82]. This indicates that transcriptional restart following DNA damage removal and basal transcription initiation are distinctly regulated. Together, these findings highlight that repair of the transcribed strand alone is not sufficient for the cell to restore mRNA expression. Transcription restart requires the synergy of many factors, including not only the discussed chromatin remodellers, but most likely also additional transcriptional regulators [83, 84].

Activating transcription factor 3 (ATF3) is one example of a transcriptional regulator that is involved in transcription restart upon DNA damage but not repair [85]. ATF3 expression is dramatically upregulated by various stress signals, including UV damage. The binding of ATF3 to its target genes usually silences them [86]. However, although the transcription of ATF3 target genes recovers 12–24 h following UV damage in TC-NER-proficient cells, the ATF3 target gene repression is prolonged in CSB-deficient cells, likely due to ATF3 impeding Pol II access to the promoter. Supporting this, silencing ATF3 rescues the transcription restart defect in CSB-deficient cells. These findings allocate a new role to CSB besides its key function in sensing lesion-arrest of Pol II: CSB may also be involved in overcoming the silencing of ATF3-dependent genes. Furthermore, these results imply that there is a direct link between the stalling of Pol II in the gene body and the inhibition of transcription that is regulated via the promoter [85].

This raises an interesting question: does the restart of Pol II transcription upon TBL only occur locally at the site of damage, or does it also occur genome-wide at non-arrested polymerases? (See Fig. 2.) Although the suggested TFIIS and Ccr4-Not mediated backtracking of Pol II would allow to resume elongation of the same transcript, there is currently no experimental proof for this mechanism. Interestingly, recent genome-wide analyses of nascent RNA-Seq data suggest that transcription recovers in a wave from the 5'-end of genes upon TBL induction by either UV irradiation or treatment with the topoisomerase 1 inhibitor camptothecin [87-89]. A wave-like recovery of transcription following genotoxic stress would implicate two interesting new concepts: (1) a significant part of transcription restarts at the beginning of genes, rather than at the sites where Pol II initially stalled; and (2) transcription does not restart stochastically upon repair of individual genes, but rather simultaneously, in a regulated manner in most genes [12].

Perspective

Over the past few decades, we have acquired an impressive body of knowledge about the cellular response to transcription-blocking DNA damage. However, to further improve our understanding of the post-repair transcription restart process, several questions remain to be answered.

Pol II processing upon DNA damage has been thoroughly studied, with Pol II displacement by backtracking, degradation, or lesion bypass now being widely accepted mechanisms. However, what guides this choice of pathways remains largely known. The degradation of arrested Pol II is assumed to be a last resort mechanism that occurs
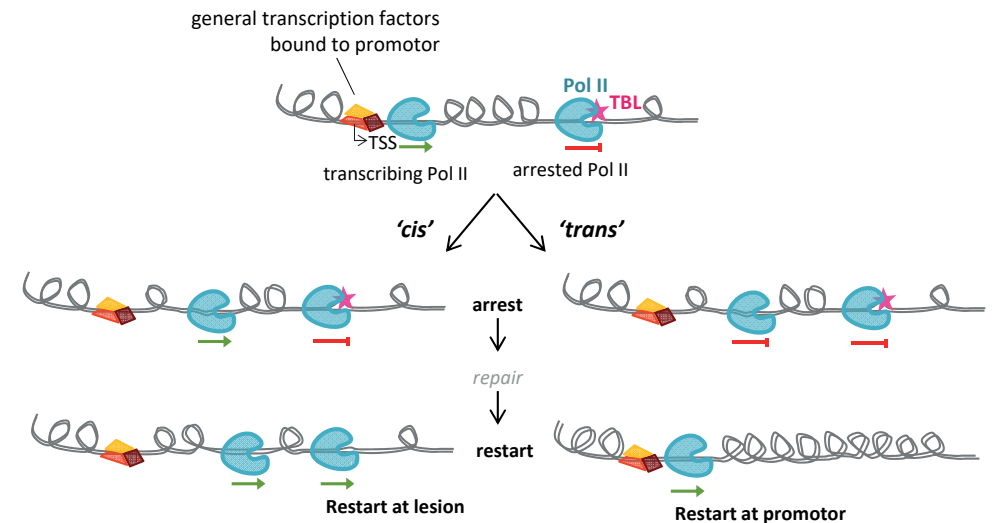
Figure 2. Potential mechanism of transcription arrest and restart in 'cis' or 'trans' upon stalling of RNA Polymerase 2 (POL II) on a transcription blocking lesion (TBL). A regulation in 'cis' implies that only those POL II that hit a TBL will stall and cause transcription inhibition (indicated by ⊣). Other POL II transcribing the same or other genes are not affected and keep transcribing (indicated by →). Upon repair of TBLs arrested POL II might resume transcription at the site of stalling (right panel). A regulation in 'trans' would arrest also other polymerases on the damaged gene, maybe including both initiating and elongating POL II. 'Trans' regulation might even include arrest of POL II on other undamaged genes (not shown). Restart of transcription upon TBL repair may occur at the site of arrest (not shown), but also by re-initiation of POL II at the promoter. If the latter scenario happens in a regulated manner at many promoters, transcription would recovery as a wave from the 5' start of genes (left panel).

only when lesion-stalled Pol II cannot be resolved, as occurs, for example, in the absence of TC-NER proteins [90]. However, although preserving Pol II and its transcript from degradation by means of Pol II backtracking intuitively seems to be the most favorable scenario, experimental evidence to support such regulation is scarce. It is possible that Pol II degradation is favored over backtracking above a certain threshold of damage. Alternatively, the pathway choice to process stalled Pol II may be guided by the complexity of the lesion or the chromatin environment, or may even be gene specific.

A better insight into the fate of lesion-stalled Pol II may also improve our understanding of TC-NER-associated phenotypes. TC-NER defects in humans cause Cockayne syndrome (CS) or UV sensitivity syndrome (UVsS). CS and UVsS cells are equally deficient in TC-NER in vitro, and yet the patients exhibit strikingly distinct clinical symptoms: CS patients display severe developmental, neurological, and premature aging features, whereas UVsS individuals present with a much milder phenotype that is mostly restricted to UV hypersensitivity [9, 10, 91]. How molecular defects within the same pathway can lead to such strikingly diverse phenotypes remains unresolved, but may be associated with the specific functions of the CS proteins outside TC-NER [92, 93], such as transcription

1

initiation [94], the maintenance of mitochondrial DNA stability [95, 96], or the regulation of specific transcriptional programs [97]. Another hypothesis suggests that aberrant processing of lesion-stalled Pol II may explain the differences between the UVsS and CS phenotypes. Here, it is proposed that in CS cells, which lack functional CSB, Pol II cannot be degraded or displaced [9], and so the lack of TC-NER combined with the persistent arrest of Pol II leads to apoptosis and senescence, causing the severe CS phenotype. By contrast, in UVsS cells, which lack functional UVSSA, stalled Pol II may still be ubiquitylated or displaced in a CSA/CSB-dependent manner, making the lesion accessible for alternative DNA repair mechanisms, including global genome NER or base excision repair, and thus resulting in the milder UVsS phenotype [8, 98].

To date, no study has investigated whether damage-induced R-loop formation and non-canonical ATM activation contribute to the phenotypes of these TC-NER syndromes. It has been reported that CSB is required to resolve R-loops, whereas XPC (the protein that initiates global genome NER) is not. However, CSB-mediated R-loop removal results in DNA breaks [66]. CSB may not only promote R-loop removal by excision [66], but also by resolving lesion-stalled Pol II, which is an important source of R-loops [66, 99]. Furthermore, R-loop-induced spliceosome displacement may promote TBL repair by facilitating Pol II backtracking or removal [62, 68]. Alternatively, the loss of the co-transcriptional splicing machinery may be linked to a regulated inhibition of transcription via non-canonical ATM signaling. However, the role of R-loop-induced ATM activation as a new mechanism of transcription-stress signaling requires further investigation.

TBLs may have strikingly different outcomes in different organs and cell types [8, 100-103]. A clear example of this is the extreme damage sensitivity of photoreceptor cells in the retina of TC-NER-deficient mice and the neurodegeneration in CS patients [100, 104, 105]. This suggests that DNA damage induction, recognition, repair, and signaling also differ between tissues and cell types, which would result in respective changes in the level of damage-induced mutagenesis, senescence or cell death. Several factors may influence the differential cellular consequences to TBL exposure, including transcription levels, chromatin states, or differential activity of the DNA-repair pathways [106]. Finally, differences in replication rates may also lead to strong differences in the cellular consequences of TBLs. In contrast to post-mitotic differentiated cells, in replicating cells advancing replication forks may collide with TBL-stalled Pol II complexes, which can have severe cellular outcomes [64, 107]. However, comprehensive studies on tissue-specific regulation of lesion-stalled Pol II and its underlying mechanisms are currently lacking.

ACKNOWLEDGEMENTS

1

REFERENCES

1. Zhou, Q., T. Li, and D.H. Price, *RNA polymerase II elongation control.* Annu Rev Biochem, 2012. **81**: p. 119-43.
2. Stirling, P.C. and P. Hieter, *Canonical DNA Repair Pathways Influence R-Loop-Driven Genome Instability.* J Mol Biol, 2016.
3. Harper, J.W. and S.J. Elledge, *The DNA damage response: ten years after.* Mol Cell, 2007. **28**(5): p. 739-45.
4. Jiang, G. and A. Sancar, *Recruitment of DNA damage checkpoint proteins to damage in transcribed and nontranscribed sequences.* Mol Cell Biol, 2006. **26**(1): p. 39-49.
5. Ljungman, M. and F. Zhang, *Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis.* Oncogene, 1996. **13**(4): p. 823-31.
6. Ljungman, M., *Activation of DNA damage signaling.* Mutat Res, 2005. **577**(1-2): p. 203-16.
7. Vermeulen, W. and M. Fousteri, *Mammalian transcription-coupled excision repair.* Cold Spring Harb Perspect Biol, 2013. **5**(8): p. a012625.
8. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.
9. Hanawalt, P.C. and G. Spivak, *Transcription-coupled DNA repair: two decades of progress and surprises.* Nat Rev Mol Cell Biol, 2008. **9**(12): p. 958-70.
10. Nance, M.A. and S.A. Berry, *Cockayne syndrome: review of 140 cases.* Am J Med Genet, 1992. **42**(1): p. 68-84.
11. Mellon, I., G. Spivak, and P.C. Hanawalt, *Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene.* Cell, 1987. **51**(2): p. 241-9.
12. Andrade-Lima, L.C., A. Veloso, and M. Ljungman, *Transcription Blockage Leads to New Beginnings.* Biomolecules, 2015. **5**(3): p. 1600-17.
13. Tornaletti, S., D. Reines, and P.C. Hanawalt, *Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA.* J Biol Chem, 1999. **274**(34): p. 24124-30.
14. Tornaletti, S., *Transcription arrest at DNA damage sites.* Mutat Res, 2005. **577**(1-2): p. 131-45.
15. Brueckner, F., et al., *CPD damage recognition by transcribing RNA polymerase II.* Science, 2007. **315**(5813): p. 859-62.
16. Damsma, G.E., et al., *Mechanism of transcriptional stalling at cisplatin-damaged DNA.* Nat Struct Mol Biol, 2007. **14**(12): p. 1127-33.
17. Kitsera, N., et al., *8-Oxo-7,8-dihydroguanine in DNA does not constitute a barrier to transcription, but is converted into transcription-blocking damage by OGG1.* Nucleic Acids Res, 2011. **39**(14): p. 5926-34.
18. Yanamadala, S. and M. Ljungman, *Potential role of MLH1 in the induction of p53 and apoptosis by blocking transcription on damaged DNA templates.* Mol Cancer Res, 2003. **1**(10): p. 747-54.
19. Menoni, H., et al., *The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage.* Nucleic Acids Res, 2018. **46**(15): p. 7747-7756.
20. Nudler, E., *RNA polymerase backtracking in gene regulation and genome instability.* Cell, 2012. **149**(7): p. 1438-45.
21. Epshtein, V., et al., *UvrD facilitates DNA repair by pulling RNA polymerase backwards.* Nature, 2014. **505**(7483): p. 372-7.
22. Mullenders, L., *DNA damage mediated transcription arrest: Step back to go forward.* DNA Repair (Amst), 2015. **36**: p. 28-35.
23. Izban, M.G. and D.S. Luse, *The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII.* Genes Dev, 1992. **6**(7): p. 1342-56.
24. Reines, D., *Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II.* J Biol Chem, 1992. **267**(6): p. 3795-800.

25. Kettenberger, H., K.J. Armache, and P. Cramer, *Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage.* Cell, 2003. **114**(3): p. 347-57.

26. Sigurdsson, S., A.B. Dirac-Svejstrup, and J.Q. Svejstrup, *Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability.* Mol Cell, 2010. **38**(2): p. 202-10.

27. Dutta, A., et al., *Ccr4-Not and TFIIS Function Cooperatively To Rescue Arrested RNA Polymerase II.* Mol Cell Biol, 2015. **35**(11): p. 1915-25.

28. Kruk, J.A., et al., *The multifunctional Ccr4-Not complex directly promotes transcription elongation.* Genes Dev, 2011. **25**(6): p. 581-93.

29. Donahue, B.A., et al., *Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template.* Proc Natl Acad Sci U S A, 1994. **91**(18): p. 8502-6.

30. Jensen, A. and L.H. Mullenders, *Transcription factor IIS impacts UV-inhibited transcription.* DNA Repair (Amst), 2010. **9**(11): p. 1142-50.

31. Lans, H., J.A. Marteijn, and W. Vermeulen, *ATP-dependent chromatin remodeling in the DNA-damage response.* Epigenetics Chromatin, 2012. **5**: p. 4.

32. Citterio, E., et al., *ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor.* Mol Cell Biol, 2000. **20**(20): p. 7643-53.

33. Cho, I., et al., *ATP-dependent chromatin remodeling by Cockayne syndrome protein B and NAP1-like histone chaperones is required for efficient transcription-coupled DNA repair.* PLoS Genet, 2013. **9**(4): p. e1003407.

34. Selzer, R.R., et al., *Differential requirement for the ATPase domain of the Cockayne syndrome group B gene in the processing of UV-induced DNA damage and 8-oxoguanine lesions in human cells.* Nucleic Acids Res, 2002. **30**(3): p. 782-93.

35. Somesh, B.P., et al., *Multiple mechanisms confining RNA polymerase II ubiquitylation to polymerases undergoing transcriptional arrest.* Cell, 2005. **121**(6): p. 913-23.

36. Somesh, B.P., et al., *Communication between distant sites in RNA polymerase II through ubiquitylation factors and the polymerase CTD.* Cell, 2007. **129**(1): p. 57-68.

37. Svejstrup, J.Q., *Rescue of arrested RNA polymerase II complexes.* J Cell Sci, 2003. **116**(Pt 3): p. 447-51.

38. Ratner, J.N., et al., *Ultraviolet radiation-induced ubiquitination and proteasomal degradation of the large subunit of RNA polymerase II. Implications for transcription-coupled DNA repair.* J Biol Chem, 1998. **273**(9): p. 5184-9.

39. Huibregtse, J.M., J.C. Yang, and S.L. Beaudenon, *The large subunit of RNA polymerase II is a substrate of the Rsp5 ubiquitin-protein ligase.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3656-61.

40. Beaudenon, S.L., et al., *Rsp5 ubiquitin-protein ligase mediates DNA damage-induced degradation of the large subunit of RNA polymerase II in Saccharomyces cerevisiae.* Mol Cell Biol, 1999. **19**(10): p. 6972-9.

41. Ribar, B., L. Prakash, and S. Prakash, *Requirement of ELC1 for RNA polymerase II polyubiquitylation and degradation in response to DNA damage in Saccharomyces cerevisiae.* Mol Cell Biol, 2006. **26**(11): p. 3999-4005.

42. Ribar, B., L. Prakash, and S. Prakash, *ELA1 and CUL3 are required along with ELC1 for RNA polymerase II polyubiquitylation and degradation in DNA-damaged yeast cells.* Mol Cell Biol, 2007. **27**(8): p. 3211-6.

43. Harreman, M., et al., *Distinct ubiquitin ligases act sequentially for RNA polymerase II polyubiquitylation.* Proc Natl Acad Sci U S A, 2009. **106**(49): p. 20705-10.

44. Kvint, K., et al., *Reversal of RNA polymerase II ubiquitylation by the ubiquitin protease Ubp3.* Mol Cell, 2008. **30**(4): p. 498-506.

45. Verma, R., et al., *Cdc48/p97 mediates UV-dependent turnover of RNA Pol II.* Mol Cell, 2011. **41**(1): p. 82-92.

46. Wild, T. and P. Cramer, *Biogenesis of multisubunit RNA polymerases.* Trends Biochem Sci, 2012. **37**(3): p. 99-105.

47. Chen, X., C. Ruggiero, and S. Li, *Yeast Rpb9 plays an important role in ubiquitylation and degradation of Rpb1 in response to UV-induced DNA damage.* Mol Cell Biol, 2007. **27**(13): p. 4617-25.

48. Anindya, R., O. Aygun, and J.Q. Svejstrup, *Damage-induced ubiquitylation of human RNA polymerase II by the ubiquitin ligase Nedd4, but not Cockayne syndrome proteins or BRCA1.* Mol Cell, 2007. **28**(3): p. 386-97.

49. Kuznetsova, A.V., et al., *von Hippel-Lindau protein binds hyperphosphorylated large subunit of RNA polymerase II through a proline hydroxylation motif and targets it for ubiquitination.* Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2706-11.

50. Stebbins, C.E., W.G. Kaelin, Jr., and N.P. Pavletich, *Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function.* Science, 1999. **284**(5413): p. 455-61.

51. Yasukawa, T., et al., *Mammalian Elongin A complex mediates DNA-damage-induced ubiquitylation and degradation of Rpb1.* EMBO J, 2008. **27**(24): p. 3256-66.

52. Starita, L.M., et al., *BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II.* J Biol Chem, 2005. **280**(26): p. 24498-505.

53. Wu, W., et al., *BRCA1 ubiquitinates RPB8 in response to DNA damage.* Cancer Res, 2007. **67**(3): p. 951-8.

54. Keaveney, M. and K. Struhl, *Activator-mediated recruitment of the RNA polymerase II machinery is the predominant mechanism for transcriptional activation in yeast.* Mol Cell, 1998. **1**(6): p. 917-24.

55. Alexander, R.D., et al., *Splicing-dependent RNA polymerase pausing in yeast.* Mol Cell, 2010. **40**(4): p. 582-93.

56. Scheidegger, A. and S. Nechaev, *RNA polymerase II pausing as a context-dependent reader of the genome.* Biochem Cell Biol, 2016. **94**(1): p. 82-92.

57. Gyenis, A., et al., *UVB induces a genome-wide acting negative regulatory mechanism that operates at the level of transcription initiation in human cells.* PLoS Genet, 2014. **10**(7): p. e1004483.

58. Walmacq, C., et al., *Mechanism of translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage.* Mol Cell, 2012. **46**(1): p. 18-29.

59. Selby, C.P. and A. Sancar, *Cockayne syndrome group B protein enhances elongation by RNA polymerase II.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11205-9.

60. Charlet-Berguerand, N., et al., *RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors.* EMBO J, 2006. **25**(23): p. 5481-91.

61. Saxowsky, T.T. and P.W. Doetsch, *RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis?* Chem Rev, 2006. **106**(2): p. 474-88.

62. Tresini, M., et al., *The core spliceosome as target and effector of non-canonical ATM signalling.* Nature, 2015. **523**(7558): p. 53-8.

63. Sordet, O., et al., *Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks.* EMBO Rep, 2009. **10**(8): p. 887-93.

64. Aguilera, A. and T. Garcia-Muse, *R loops: from transcription byproducts to threats to genome stability.* Mol Cell, 2012. **46**(2): p. 115-24.

65. Hamperl, S. and K.A. Cimprich, *The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability.* DNA Repair (Amst), 2014. **19**: p. 84-94.

66. Sollier, J. and K.A. Cimprich, *Breaking bad: R-loops and genome integrity.* Trends Cell Biol, 2015. **25**(9): p. 514-22.

67. Skourti-Stathaki, K. and N.J. Proudfoot, *A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression.* Genes Dev, 2014. **28**(13): p. 1384-96.

68. Tresini, M., J.A. Marteijn, and W. Vermeulen, *Bidirectional coupling of splicing and ATM signaling in response to transcription-blocking DNA damage.* RNA Biol, 2016. **13**(3): p. 272-8.

69. Nakazawa, Y., et al., *Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair.* Nat Genet, 2012. **44**(5): p. 586-92.

70. Zhang, X., et al., *Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair.* Nat Genet, 2012. **44**(5): p. 593-7.

71. Schwertman, P., et al., *UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair.* Nat Genet, 2012. **44**(5): p. 598-602.

72. Beerens, N., et al., *The CSB protein actively wraps DNA.* J Biol Chem, 2005. **280**(6): p. 4722-9.

73. Fousteri, M., et al., *Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo.* Mol Cell, 2006. **23**(4): p. 471-82.

74. Nakatsu, Y., et al., *XAB2, a novel tetratricopeptide repeat protein involved in transcription-coupled DNA repair and transcription.* J Biol Chem, 2000. **275**(45): p. 34931-7.

75. Kuraoka, I., et al., *Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair.* J Biol Chem, 2008. **283**(2): p. 940-50.

76. Riedl, T., F. Hanaoka, and J.M. Egly, *The comings and goings of nucleotide excision repair factors on damaged DNA.* EMBO J, 2003. **22**(19): p. 5293-303.

77. Volker, M., et al., *Sequential assembly of the nucleotide excision repair factors in vivo.* Mol Cell, 2001. **8**(1): p. 213-24.

78. Mourgues, S., et al., *ELL, a novel TFIIH partner, is involved in transcription restart after DNA repair.* Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17927-32.

79. Adam, S., S.E. Polo, and G. Almouzni, *Transcription recovery after DNA damage requires chromatin priming by the H3.3 histone chaperone HIRA.* Cell, 2013. **155**(1): p. 94-106.

80. Dinant, C., et al., *Enhanced chromatin dynamics by FACT promotes transcriptional restart after UV-induced DNA damage.* Mol Cell, 2013. **51**(4): p. 469-79.

81. Belotserkovskaya, R., et al., *FACT facilitates transcription-dependent nucleosome alteration.* Science, 2003. **301**(5636): p. 1090-3.

82. Oksenych, V., et al., *Histone methyltransferase DOT1L drives recovery of gene expression after a genotoxic attack.* PLoS Genet, 2013. **9**(7): p. e1003611.

83. Mandemaker, I.K., W. Vermeulen, and J.A. Marteijn, *Gearing up chromatin: A role for chromatin remodeling during the transcriptional restart upon DNA damage.* Nucleus, 2014. **5**(3): p. 203-10.

84. Polo, S.E. and G. Almouzni, *Chromatin dynamics after DNA damage: The legacy of the access-repair-restore model.* DNA Repair (Amst), 2015.

85. Kristensen, U., et al., *Regulatory interplay of Cockayne syndrome B ATPase and stress-response gene ATF3 following genotoxic stress.* Proc Natl Acad Sci U S A, 2013. **110**(25): p. E2261-70.

86. Hai, T., et al., *ATF3 and stress responses.* Gene Expr, 1999. **7**(4-6): p. 321-35.

87. Andrade-Lima, L.C., et al., *DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes.* Nucleic Acids Res, 2015. **43**(5): p. 2744-56.

88. Veloso, A., et al., *Genome-wide transcriptional effects of the anti-cancer agent camptothecin.* PLoS One, 2013. **8**(10): p. e78190.

89. Ljungman, M. and P.C. Hanawalt, *The anti-cancer drug camptothecin inhibits elongation but stimulates initiation of RNA polymerase II transcription.* Carcinogenesis, 1996. **17**(1): p. 31-5.

90. Wilson, M.D., M. Harreman, and J.Q. Svejstrup, *Ubiquitylation and degradation of elongating RNA polymerase II: the last resort.* Biochim Biophys Acta, 2013. **1829**(1): p. 151-7.

91. Spivak, G., *UV-sensitive syndrome.* Mutat Res, 2005. **577**(1-2): p. 162-9.

92. Sarasin, A., *UVSSA and USP7: new players regulating transcription-coupled nucleotide excision repair in human cells.* Genome Med, 2012. **4**(5): p. 44.

93. Cleaver, J.E., *Photosensitivity syndrome brings to light a new transcription-coupled DNA repair cofactor.* Nat Genet, 2012. **44**(5): p. 477-8.

94. Proietti-De-Santis, L., P. Drane, and J.M. Egly, *Cockayne syndrome B protein regulates the transcriptional program after UV irradiation.* EMBO J, 2006. **25**(9): p. 1915-23.

95. Aamann, M.D., et al., *Cockayne syndrome group B protein promotes mitochondrial DNA stability by supporting the DNA repair association with the mitochondrial membrane.* FASEB J, 2010. **24**(7): p. 2334-46.

96. Kamenisch, Y., et al., *Proteins of nucleotide and base excision repair pathways interact in mitochondria to protect from loss of subcutaneous fat, a hallmark of aging.* J Exp Med, 2010. **207**(2): p. 379-90.

97. Wang, Y., et al., *Dysregulation of gene expression as a cause of Cockayne syndrome neurological disease.* Proc Natl Acad Sci U S A, 2014. **111**(40): p. 14454-9.

98. Schwertman, P., W. Vermeulen, and J.A. Marteijn, *UVSSA and USP7, a new couple in transcription-coupled DNA repair.* Chromosoma, 2013. **122**(4): p. 275-84.

99. Santos-Pereira, J.M. and A. Aguilera, *R loops: new modulators of genome dynamics and function.* Nat Rev Genet, 2015. **16**(10): p. 583-97.

100. Gorgels, T.G., et al., *Retinal degeneration and ionizing radiation hypersensitivity in a mouse model for Cockayne syndrome.* Mol Cell Biol, 2007. **27**(4): p. 1433-41.

101. Garinis, G.A., et al., *DNA damage and ageing: new-age ideas for an age-old problem.* Nat Cell Biol, 2008. **10**(11): p. 1241-7.

102. Lans, H. and W. Vermeulen, *Tissue specific response to DNA damage: C. elegans as role model.* DNA Repair (Amst), 2015. **32**: p. 141-8.

103. Nouspikel, T. and P.C. Hanawalt, *DNA repair in terminally differentiated cells.* DNA Repair (Amst), 2002. **1**(1): p. 59-75.

104. Laugel, V., *Cockayne syndrome: the expanding clinical and mutational spectrum.* Mech Ageing Dev, 2013. **134**(5-6): p. 161-70.

105. Jaarsma, D., et al., *Cockayne syndrome pathogenesis: lessons from mouse models.* Mech Ageing Dev, 2013. **134**(5-6): p. 180-95.

106. de Waard, H., et al., *Cell-type-specific consequences of nucleotide excision repair deficiencies: Embryonic stem cells versus fibroblasts.* DNA Repair (Amst), 2008. **7**(10): p. 1659-69.

107. Gan, W., et al., *R-loop-mediated genomic instability is caused by impairment of replication fork progression.* Genes Dev, 2011. **25**(19): p. 2041-56.

Chapter 2

LIVE-CELL ANALYSIS OF ENDOGENOUS GFP-RPB1 UNCOVERS RAPID TURNOVER OF INITIATING AND PROMOTER-PAUSED RNA POLYMERASE II

Barbara Steurer[1,2], Roel C. Janssens[1,2],
Bart Geverts[3], Marit E. Geijer[1,2], Franziska Wienholz[1],
Arjan F. Theil[1], Jiang Chang[1], Shannon Dealy[1],
Joris Pothof[1], Wiggert A. van Cappellen[3],
Adriaan B. Houtsmuller[3], Jurgen A. Marteijn[1,2,] *

[1] Department of Molecular Genetics, Erasmus MC,
Rotterdam, The Netherlands
[2] Oncode Institute, Erasmus MC, Rotterdam, The Netherlands
[3] Department of Pathology, Optical Imaging Centre, Erasmus MC,
Rotterdam, The Netherlands
*Corresponding author: Jurgen A. Marteijn

2

ABSTRACT

Initiation and promoter-proximal pausing are key regulatory steps of RNA Polymerase II (Pol II) transcription. To study the *in vivo* dynamics of endogenous Pol II during these step we generated fully functional GFP-RPB1 knock-in cells. GFP-RPB1 photo-bleaching combined with computational modeling revealed 4 kinetically distinct Pol II fractions and showed that on average 7% of Pol II are freely diffusing, while 10% are chromatin-bound for 2.4 s during initiation, and 23% are promoter-paused for only 42 s. This unexpectedly high turnover of Pol II at promoters is most likely caused by premature termination of initiating and promoter-paused Pol II, and is in sharp contrast to the 23 min that elongating Pol II resides on chromatin. Our live-cell imaging approach provides new insights into Pol II dynamics and suggests that the continuous release and re-initiation of promoter-bound Pol II is an important component of transcriptional regulation.

SIGNIFICANCE STATEMENT

Transcription by RNA Polymerase II (Pol II) is a highly dynamic process that is tightly regulated at each step of the transcription cycle. We generated GFP-RPB1 knock-in cells and developed photo-bleaching of endogenous Pol II combined with computational modeling to study the *in vivo* dynamics of Pol II in real time. This approach allowed to dissect promoter paused Pol II from initiating and elongating Pol II and showed that initiation and promoter proximal pausing are surprisingly dynamic events due to premature termination of Pol II. Our study provides new insights into Pol II dynamics and suggests that the iterative release and re-initiation of promoter-bound Pol II is an important component of transcriptional regulation.

INTRODUCTION

Eukaryotic gene expression is a highly regulated process initiated by the sequential binding of transcription factors that facilitate the recruitment of RNA polymerase II (Pol II) and the assembly of the preinitiation complex (PIC) [1]. During initiation the CDK7 subunit of TFIIH phosphorylates serine (Ser) 5 of the C-terminal domain (CTD) of RPB1, the core catalytic subunit of Pol II, allowing Pol II to engage the DNA template and to start transcribing a short stretch of RNA [2, 3]. This early elongation complex is paused 30 to 60 nucleotides downstream of the transcription start site by Negative elongation factor (NELF) and DRB sensitivity inducing factor (DSIF) [4, 5]. The subsequent pause release into productive elongation is mediated by the positive transcription elongation factor b (P-TEFb), whose Cdk9 kinase converts DSIF into a positive elongation factor, facilitates the eviction of NELF, and phosphorylates the RPB1 CTD on Ser2 [2].

Traditionally it was thought that transcription is primarily regulated by Pol II recruitment and initiation, but owing to the advances in genome-wide sequencing technologies we know today that mRNA output is also controlled by the tight coordination of post-initiation steps [2, 5]. For example, promoter proximal pausing is a key rate-limiting step of RNA synthesis that serves as a checkpoint for 5' capping of the nascent RNA and maintains an open chromatin structure near promoters [6]. Originally discovered as a regulatory switch of stimulus responsive genes in Drosophila [7], recent genome-wide studies have revealed that promoter pausing is a widespread phenomenon occurring on most metazoan genes [5, 8]. Yet, despite this prevalence the dynamics of promoter-paused Pol II remain under debate. The currently prevailing model suggests that Pol II pauses at promoters with a half-life of 5-15 minutes [8-12], serving as an integrative hub to control pause release into productive elongation, while promoter proximal termination is infrequent. However, conflicting studies have reported that promoter-paused Pol II is less stable due to repeated premature termination and chromatin release proximal to the promoter, which is accompanied by the release of short transcription start site-associated RNAs RNAs [13-16].

Thus far, genome-wide dynamics of promoter-paused Pol II have been studied by Gro-seq [8], ChIP-Seq [10, 11] or methyltransferase footprinting [15] after inhibiting Pol II initiation. While these techniques provide gene-specific snapshots of Pol II transcription, relative abundance, or position at a given time, they do not allow to measure steady state Pol II kinetics, i.e. chromatin binding times, in real time. Though these studies have gained insights into the turnover of paused Pol II, most experiments have been performed after inhibiting transcription initiation by Triptolide[8, 10-12]. This covalent XPB inhibitor severely affects Pol II levels [17, 18] and has been recently shown to have a slow mode of action [16], which makes it less suitable to study a potentially rapid cellular process. To overcome these limitations we developed photo-bleaching of endogenously expressed GFP-RPB1 followed by computational modeling to quantitatively assess the kinetics of Pol II in unperturbed living cells.

2

Here we show that GFP-RPB1 knock in (KI) cells generated by CRISPR/Cas9-mediated gene targeting are fully functional and provide a promising new tool to study the steady-state kinetics of endogenous Pol II. By photo-bleaching of GFP-RPB1 we identified three kinetically distinct fractions of chromatin-bound Pol II. Using Monte Carlo-based-based modeling of Pol II kinetics we assessed the quantitative framework of the Pol II transcription cycle and elucidated its timeframe and quantitative set-up. Our findings are highly supportive of a model in which Pol II initiation and promoter pausing are highly dynamic events of iterative cycles of Pol II chromatin binding and release.

RESULTS

Generation and characterization of GFP-RPB1 cells

To study the *in vivo* kinetics of endogenous Pol II we generated a GFP-RPB1 (POLR2A) knock-in (KI) cell line (MRC-5 sv40) fluorescently labeling the largest subunit of Pol II. We transiently expressed a single guide RNA (sgRNA) to induce a CRISPR-associated protein 9 (Cas9)-mediated double strand break (DSB) downstream of the RPB1 transcriptional start site. A repair template containing GFP cDNA flanked by homology arms comprised of the genomic RPB1 sequence [19] (Fig S1a) was co-transfected to allow repair of the DSB by homologous recombination. GFP positive cells were isolated by FACS and single-cell clones of homozygous knock-in (KI) cells were selected (Fig S1b) in which wild type (WT) RPB1 was replaced by GFP-RPB1, as shown by western blot (Fig 1a) and genotyping (Fig S1c). Importantly, the GFP-RPB1 expression level and phosphorylation status in KI cells were similar to WT cells (Fig 1a), the GFP tag did not compromise the basal transcription rate (Fig 1b, S1d), and did not alter gene expression profiles (Fig S1e). Live-cell confocal imaging of KI cells revealed a non-homogenous, nuclear distribution of GFP-RPB1 with multiple bright foci and exclusion from nucleoli (Fig 1c). Localization of GFP-RPB1 was similar to endogenous RPB1, shown by immunofluorescent stainings of the RPB1 CTD and NTD (Fig 1d). Together these data illustrate that GFP-RPB1 is fully functional.

As GFP-RPB1 is expressed from its endogenous gene loci and RPB1 only translocates to the nucleus as part of the fully assembled Pol II complex [20], nuclear GFP fluorescence can be used as a direct readout for endogenous Pol II localization and concentration in living cells. To estimate the number of Pol II complexes, we compared the nuclear GFP intensity of KI cells to the extracellular fluorescence of known, increasing concentrations of recombinant fluorescent GFP added to the culture medium (Fig 1e). This direct comparison of fluorescence on the same microscope slide revealed a nuclear Pol II concentration of 0.18 µM. Based on the determined average nuclear volume of the KI cells of 734µm³ this equals ~50,000 Pol II per diploid genome (Fig 1e), which is within the range of previously determined Pol II quantities [21].
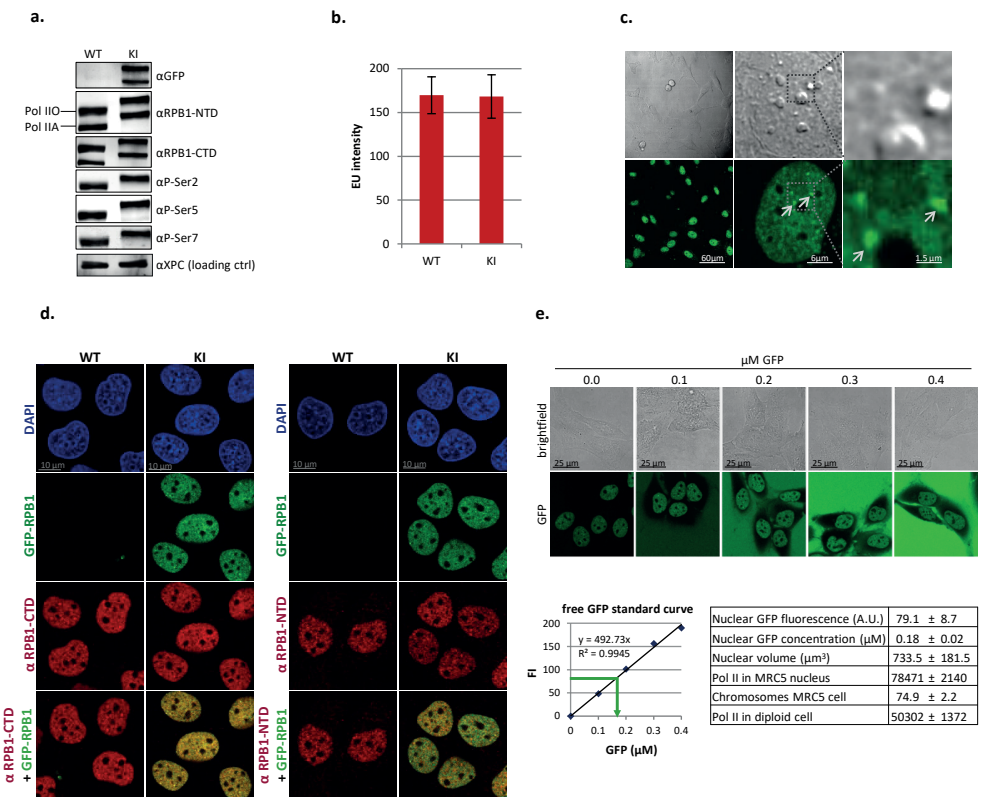

Figure 1. Characterization of GFP-RPB1 knock-in cells. Western blot of MRC-5 wild type (WT) and GFP-RPB1 knock in (KI) cells. NTD= N-terminal domain, CTD= C-terminal domain, Pol IIA = hypo-phosphorylated Pol II, Pol IIO= hyper-phosphorylated Pol II, P-Ser = phosphoserine. EU incorporation levels of WT and KI cells. Mean total nuclear fluorescence intensity (FI) ± SD, n= 60 cells, 2 independent experiments. Live-cell images of KI cells. Arrows indicate foci of locally enriched Pol II. Immunofluorescent stainings of RPB1-CTD (left) and RPB1-NTD (right) in WT and KI cells. DNA is stained with DAPI, GFP fluorescence from endogenously expressed GFP-RPB1. Representative images of KI cells with increasing concentrations of free GFP in the culture medium. Extracellular GFP fluorescence was plotted against GFP concentration to make a standard curve (graph, bottom left) and nuclear Pol II concentration was calculated based on the standard curve (indicated with the green arrow). Numbers used to calculate the amount of Pol II in a diploid nucleus are summarized in the table (bottom right). Table shows mean ± SD, n= 40 cells of 2 independent experiments.

Initiation, promoter pausing, and productive elongation are characterized by distinct Pol II kinetics

This multitude of Pol II complexes is divided over the different stages of the transcription cycle, including initiation, promoter pausing, and productive elongation. Previous studies have indicated that Pol II mobility is linked to its state of engagement during transcription [22-24], however, thus far promoter-proximally paused Pol II could not be discriminated and quantified in living cells. To test whether this fraction could be identified

in GFP-RPB1 KI cells expressing Pol II at endogenous levels, we determined the kinetic framework of the Pol II transcription cycle by fluorescence recovery after photo-bleaching (FRAP) [25, 26].

First, we measured the redistribution of Pol II after photo-bleaching in half the nucleus (Fig 2a). Half-nucleus FRAP revealed an initial, fast redistribution of Pol II followed by a slow recovery of apparently less mobile Pol II. In line with previous studies [22-24] we assumed that the latter fraction represents long-term chromatin-bound, elongating Pol II that are released from chromatin over time when transcription terminates. Pol II fluorescence was fully recovered 80 min after the bleach pulse, likely reflecting the engagement of a small fraction of Pol II with very long or slowly transcribed genes. (Fig 2a)

For a more detailed analysis of Pol II kinetics we performed FRAP in a narrow strip spanning the nucleus (Strip-FRAP) [26], allowing fluorescence measurements every 0.4 seconds. In line with half-nucleus FRAP, Strip-FRAP of GFP-RPB1 in non-treated cells (NT) showed a long-term immobilization of a large fraction of Pol II (Fig 2b), hereafter referred to as long-bound fraction (Fig 2b, for clarification approximated with green dotted line). To investigate the nature of this long-bound fraction we measured Pol II mobility after treatment for 1hr with 100µM Cordycepin, a 3'-deoxy nucleoside that inhibits transcript elongation when incorporated into RNA [27] (Fig 2d). Indeed, Cordycepin reduced EU incorporation (Fig S2a, S2b) without affecting Pol II protein levels, determined by quantifying GFP fluorescence either prior to FRAP (Fig 2c right panel), by immunofluorescence (Fig S2b bottom panel), or western blot (Fig 2e). The chromatin release of specifically the long-bound fraction was markedly decreased after Cordycepin (Fig 2c), indicating that the long-bound fraction mainly represents elongating Pol II. Cordycepin-mediated inhibition of elongation also markedly delayed the redistribution of Pol II after half-nucleus FRAP (Fig S2c), corroborating that the long-bound fraction mainly reflects the behaviour of productively elongating Pol II. This allocation was further validated by the depletion of the transcription factor IIS (TFIIS) (Fig 2g, S2d). TFIIS stimulates the intrinsic cleavage activity of Pol II needed to reactivate complexes that have been arrested during productive elongation [28]. Similar as after Cordycepin (Fig 2c), simultaneous siRNA-mediated knockdown of the TFIIS paralogues TCEA1 and TCEA2 (Fig2f, S2d) decreased the chromatin release of long-bound Pol II (Fig2g).

In addition to the large fraction of elongating Pol II with long-term chromatin binding, Strip-FRAP revealed a fraction of short-term chromatin-bound Pol II and a fraction with intermediate chromatin binding, hereafter referred to as the short-bound and the medium-bound fractions (Fig 2b, approximated with grey and purple dotted lines, respectively). To investigate the nature of these fractions we performed Strip-FRAP using inhibitors that specifically arrest Pol II at defined, consecutive stages of the transcription cycle (Fig 2d).

THZ1 inhibits CDK7 [29], the TFIIH subunit that phosphorylates Ser5 of the RPB1 CTD and thereby prevents promoter pausing, mRNA capping and subsequently Pol II transition into productive elongation [30]. As expected, treatment with 1 µM of THZ1 for 90' resulted in a loss of hyper-phosphorylated Pol II (Pol IIo), P-Ser2 (Fig 2e) and loss of EU
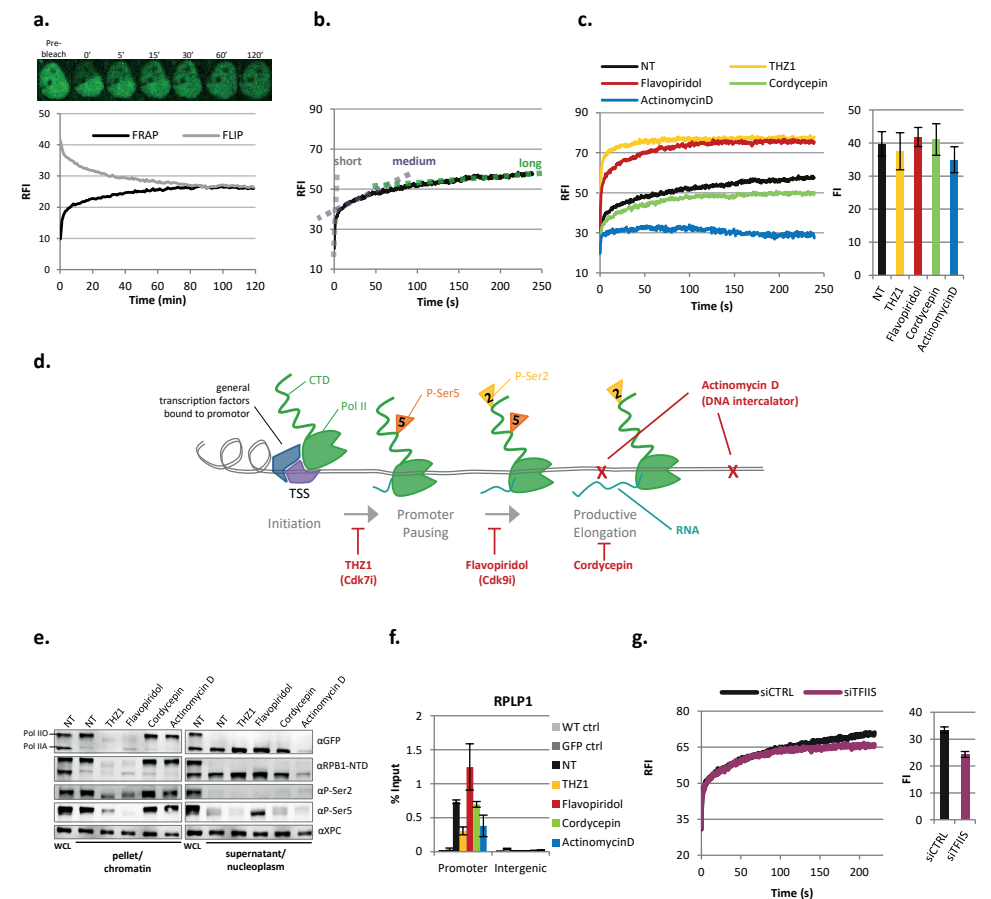

Figure 2. Real-time measurements of Pol II kinetics at different transcription cycle stages. Fluorescence recovery after photo-bleaching (FRAP) of the bleached and fluorescence loss in photo-bleaching (FLIP) of the non-bleached half of nuclei are plotted in time. Images are snapshots of a representative cell at indicated time points. RFI=Relative fluorescence intensity, Mean of n=20 cells of 2 independent experiments. Strip-FRAP analysis of GFP-RPB1 in non-treated (NT) cells. GFP-RPB1 was bleached in a narrow strip spanning the nucleus. Fluorescence recovery was measured every 0.4s for 4min, background corrected and normalized to pre-bleach fluorescence intensity. The dotted lines indicate kinetically distinct Pol II fractions, which mainly represent Pol II complexes that are short-term (grey), medium-term (purple), or long-term (green). Mean ± SD, n=20 cells of 2 independent experiments.Strip-FRAP of GFP-RPB1 in NT cells and after 1h of treatment with the indicated transcription inhibitors. Column chart on the right shows average GFP-RPB1 pre-bleach fluorescence intensities (FI) of cells analysed by FRAP as a measure for Pol II protein levels. n>16 cells per condition measured in 2 independent experiments. FI chart shows mean ± SD. Schematic representation of Pol II transcription cycle and points of action of the used transcription inhibitors. THZ1 inhibits the phosphorylation of Ser5 by Cdk7. Flavopiridol inhibits the phosphorylation of Ser2 by Cdk9. Cordycepin is a 3'deoxy adenosine analogue that stalls chain elongation when incorporated into the mRNA. Actinomycin D is a DNA intercalator. Pol II= RNA Polymerase II, CTD= C-terminal domain, P-Ser= phosphoserine. c Western blot of GFP-RPB1 KI whole cell lysates (WCL) and chromatin and nucleoplasm fractions of non-treated (NT) cells and cells treated for 90min with ▶

‣ the indicated transcription inhibitors. NTD= N-terminal domain of RPB1. Pol II binding to the RPLP1 promoter measured by chromatin-IP. Wild type (WT) and cells expressing free GFP (GFP) were analysed as controls (ctrl). Mean ±SEM of n=3 independent experiments. FRAP of GFP-RPB1 after transfection with a non-targeting control siRNA (siCTRL) or siRNAs targeting the TFIIS paralogues TCEA1 and TCEA2 (siTFIIS). Mean of n=20 cells of 2 independent experiments

incorporation (Fig S2a, S2b top panel), indicating a run-off of most engaged, elongating Pol II, without affecting Pol II protein levels (Fig 2c right panel, S2a, S2b bottom panel) or the cell cycle (Fig S2e). In addition to the loss of promoter-paused Pol II [30], THZ1-treated cells showed a loss of phosphorylated Ser5 (P-Ser5) (Fig 2e) and reduced Pol II promoter-binding (Fig 2f). Accordingly, FRAP after THZ1 showed a decreased slope of the curve compared to NT cells at time points >100 seconds (Fig 2c, yellow line), i.e. a loss of elongating Pol II. In addition, THZ1 extensively enlarged the short-bound fraction of Pol II, indicating that Pol II is not stably chromatin-bound. This was confirmed by the increased detection of non-phosphorylated Pol II (Pol IIo) in the nucleoplasm after cellular fractionation (Fig 2e), concomitant with a loss of Pol II in the pellet fraction. As Pol II can still be incorporated into the PIC in the absence of CDK7 activity [30], this indicated that the short-bound fraction mainly represents free and initiating Pol II. Remarkably, the very fast replacement of bleached Pol II by new and fluorescent Pol II in THZ1-treated cells suggests that initiation is a highly dynamic process of continuous cycles of Pol II release and re-binding of new Pol II. In addition to the increase of short-bound Pol II, THZ1 treatment resulted in the loss of both the long and the medium-bound Pol II fractions (Fig 2c). As THZ1 prevents both Pol II promoter pausing and productive elongation (Fig 2d,f and S2a,b) [30], and we assigned the long-bound fraction to productively elongating Pol II, we hypothesized that the medium-bound fraction might represent promoter-paused Pol II.

To test this we inhibited the transition of Pol II from pausing to productive elongation with the CDK9 inhibitor Flavopiridol. The CDK9 activity of the PTEF-b complex licenses Pol II pause release by phosphorylating DSIF, NELF, and Ser2 of the RPB1 CTD [2, 8, 31, 32]. As expected, inhibition of Pol II pause release by incubation with 1µM Flavopiridol for 90 min increased Pol II promoter binding (Fig2f), and inhibited productive elongation (Fig 2e, Fig S2a, S2b top panel) [8, 31, 32]. Similar to THZ1, Flavopiridol markedly enlarged the fraction of short-bound Pol II (Fig 2c) and resulted in a loss of productively elongating Pol II, as shown by the horizontal slope of the slow fraction at time points >100 seconds. Most importantly and in striking contrast to THZ1, the medium-bound fraction of Pol II remained after Flavopiridol (Fig 2c yellow versus red curve), strongly suggesting that Pol II with intermediate kinetics represent mostly promoter-paused Pol II. Similar results were obtained with another CDK9 inhibitor, DRB. Intriguingly, the fluorescence recovery after photo-bleaching in Flavopiridol-treated cells was complete after ~100sec (Fig 2c), which implied that within this short time frame all Pol II that were paused during the bleach pulse were already replaced by new, fluorescent Pol II. This FP-induced accumulation

of a fraction of Pol II that is only transiently chromatin-bound was corroborated by an accumulation of P-Ser5 positive Pol II in the nucleoplasm after cellular fractionation (Fig 2e). Together these results suggested that Pol II promoter pausing is surprisingly dynamic and presumably consists of iterative rounds of pause entry and promoter-proximal termination, i.e. release from chromatin and not release into productive elongation. Importantly, the expression of non-coding genomic loci such as enhancer RNAs and upstream antisense RNAs, which would results in short Pol II chromatin binding times, has been shown to be dependent on CDK9 activity [33-35] and hence we conclude that their contribution to the medium-bound fraction, which is not reduced but rather increased after Flavopiridol (Fig 2c,f), is limited.

Treatment with 1 µg/ml of Actinomycin D, a DNA intercalator [36], for 90 min , completely inhibited transcription (Fig S2a, and S2b top panel). In line with the loss of nucleoplasmic Pol II after Actinomycin D (Fig 2e) both the short and the medium-bound fractions were completely lost (Fig 2c). Most Pol II remained chromatin-bound after Actinomycin D resembling the Pol II FRAP curve in PFA-fixed cells, suggesting that after Actinomycin D most Pol II are stably trapped on DNA.

Computational Model of steady state Pol II kinetics

By inhibiting specific transitions in the Pol II transcription cycle we demonstrated that Pol II engaged in initiating, promoter pausing, and productive elongation are characterized by significantly distinct kinetics. While the use of transcription inhibitors allowed to assign kinetically distinct Pol II fractions to specific stages of the transcription cycle, we cannot exclude possible additional effects of these inhibitors on Pol II dynamics and stability. For example, it has recently been shown that Cdk9 inhibition by Flavopiridol also affects Pol II initiation [11]. It is therefore important to extract information about Pol II kinetics from unperturbed cells. To do so, we used Monte Carlo (MC) modeling to computationally simulate Pol II kinetics and fitted the simulated curves to the Pol II Strip-FRAP curves. MC modeling is based on the assumption that (molecular) events have a certain probability to occur within a specific short period of time (in a living cell) [37]. As every transcription cycle starts with Pol II binding to the promoter we defined one single binding step upon initiation (Fig 3a). Supported by our analysis with specific transcription inhibitors (Fig 2c) we proposed, in addition to freely diffusing Pol II, three interconnected Pol II fractions with three possible exit points: Pol II release after initiation, after promoter pausing, and after transcription termination [13-15, 24, 38] (Fig 3a). The best simulations obtained from this model fit the experimental FRAP curves very well (Fig S3a-i). From the average of the ten best fitted MC simulations (Fig S3j-m) we extracted quantitative information about the relative size of the kinetically distinct Pol II fractions, their respective residence times on chromatin, and their binding constants (k_{on} and k_{off}). This approach favored a model in which 60% of all nuclear Pol II are productively elongating with a residence time on chromatin for ~23 min on average (Fig 3b, 3c right panel). These results were in line with the large fraction of long-term chromatin-bound Pol II observed during FRAP experiments
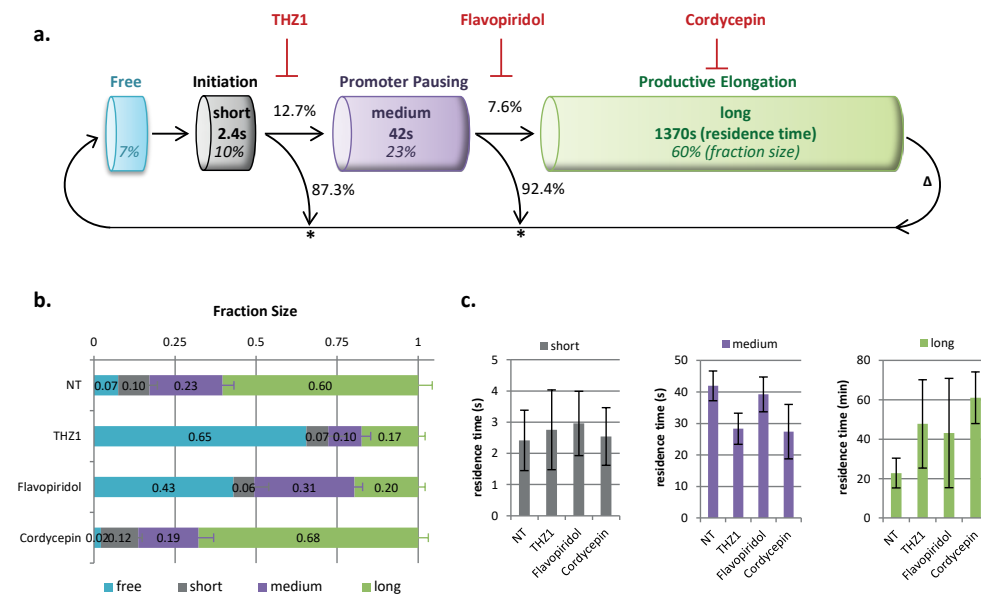
a.

b.

c.

Figure 3. Monte Carlo Modeling of live cell Pol II kinetics. Schematic representation of possible binding and release steps of Pol II to DNA. Modeled fraction sizes are depicted as percentages and written in *italics*. Modeled residence times of Pol II in different stages are written in bold. Chances for Pol II to pass through or exit a stage were calculated from k_{on} and k_{off} rate constants (see methods) and are given as percentages. * marks points of abortive Pol II release, Δ marks Pol II release after transcription termination. Fraction sizes and Residence times of short, medium, and long-bound Pol II fractions obtained from Monte Carlo-based modeling of GFP-RPB1 FRAP data (Fig 2b) of non-treated (NT) GFP-RPB1 knock in cells or after 1hr of treatment with 1µM THZ1, or 1µM Flavopiridol, or 100µMCordycepin. Mean ±SD of the 10 best fitting simulations.

(Fig 2a,b). When combined with the number and length of expressed genes, these results allowed to deduce the average Pol II elongation speed in unperturbed conditions. Therefore we determined the number and average length of all actively transcribed genes in immortalized human fibroblasts using existing nascent RNA-Seq data [39] (Dataset S1). We included all genes with >0.2 reads per kb of transcript per million mapped reads (RPKM) and a gene length >300 bp in the analysis, which resulted in ~13.600 actively transcribed genes with an average, expression frequency-corrected gene length of 46 kb (see Dataset S1, sheet 2). Together, these numbers resulted in an average elongation speed of 2 kb/min in living cells (Dataset S1, sheet3), which is in line with previously reported Pol II elongation speeds determined by run-on sequencing [8, 40].

Promoter-bound Pol II are rapidly turned over

MC modeling of RPB1 kinetics further allowed to extract relative fraction sizes and revealed that ~10% of Pol II are involved in initiation (short-bound fraction) and ~23% are promoter-bound during promoter pausing (medium-bound fraction). This implied

that most Pol II (93%) are chromatin-bound and that only 7% are freely diffusing in the nucleus (Fig 3b, 3c). In line with the fast recovery of photo-bleached Pol II after THZ1 (<10 seconds) and Flavopiridol (<100 seconds), MC modeling of Pol II kinetics in unperturbed cells revealed that initiating Pol II remains chromatin-bound for only 2.4 seconds (Fig 3c left panel) and promoter-paused Pol II for merely 42 seconds. Interestingly, these big differences in the residence time of initiating (2.4 sec), promoter pausing (42 sec) and elongating (23min) Pol II implied that only a small fraction of initiating and pausing Pol II can proceed to the subsequent step of the transcription cycle, and hence that most initiating and pausing Pol II must be released from chromatin at the promoter (Fig 3a-c). In line, analysis of the rate constants of the distinct Pol II fractions showed that only 1 out 8 Pol II that attempt initiation will proceed to promoter pausing, and that only 1 out of 13 promoter-paused Pol II will continue to productive elongation (Fig 3a). When combined, this implied that only 1 out of 100 transcription initiation events finally results in mRNA production.

Of note, the average of three independently modeled FRAP experiments resulted in highly similar estimates of fraction sizes and residence times and illustrates the variation between biological experiments (Fig S4a). Furthermore, the modeled parameters were hardly affected by adding artificial noise to FRAP curves (Fig S4b), further validating our MC-based modeling approach.

Modeling of Pol II kinetics after transcription inhibition

One of the current models of Pol II promoter pausing suggests that the release of promoter-paused Pol II occurs mainly by transition to productive elongation [5]. Our results however, indicated a very rapid turnover of promoter-paused Pol II, which most likely cannot not be explained solely by release of promoter-paused Pol II into productive elongation. Therefore we tested whether Cdk9-mediated pause release influences the half-life of promoter-paused Pol II by determining the residence time of paused Pol II after blocking pause release with FP. Interestingly, MC-based modeling indicated that Flavopiridol did not increase the residence-time of promoter-paused Pol II, (Fig 3c middle panel), illustrating that Pol II turnover at the promoter is hardly affected by Pol II release into productive elongation, and suggesting that the half-life of paused Pol II is mainly determined by promoter-proximal termination. Although Flavopiridol is known to accumulate paused Pol II at promoters by inhibiting the transition to productive elongation [8, 16, 41], it only resulted in a 35% increase of the fraction size of paused Pol II (Fig 3b), which is comparable to observed increases in Pol II promoter occupancy [8] (Fig S2e). The increase of paused Pol II was concomitant with a decrease of initiating Pol II and is in line with the recent finding that paused Pol II inhibits Pol II initiation [11].

Modeling of Pol II kinetics after THZ1 or Flavopiridol revealed a substantial increase in freely diffusing Pol II, whereas the fraction of initiating Pol II was slightly reduced (Fig 3b), suggesting that Pol II initiation frequency does not directly depend on the number

of freely diffusing Pol II. The very small fraction of elongating Pol II that remained after THZ1 and Flavopiridol resided longer on the gene body (Fig 3c, right panel) and might represent a sub-fraction of Pol II that did not run off within the 90 min of inhibitor treatment, likely because they transcribe very long genes or genes with a low elongation speed. Interestingly, Cordycepin increased the fraction size of elongating Pol II (Fig2f), and reduced the number of free Pol II (Fig2f). In addition to the shift of fraction sizes, Cordycepin increased the residence time of elongating Pol II three-fold (Fig 3f right panel), arguing for a strongly decreased elongation speed after treatment with this elongation inhibitor, without affecting transcription initiation. MC-modeled fraction sizes of Pol II were well in line with Pol II quantities detected in the chromatin-bound or nucleoplasmic fraction following cell fractionation, in both non-treated conditions or after treatment with different transcription inhibitors (Fig S4c).

Kinetics of Pol II after treatment with Triptolide or α-Amanitin

We demonstrated, that GFP-RPB1 KI cells are a suitable tool to quantitatively assess real time Pol II transcription kinetics and Pol II concentration. Next we applied this tool to study the effects of Triptolide and α-Amanitin, 2 widely used compounds that inhibit Pol II transcription and induce Pol II degradation. Triptolide inhibits the ATPase activity of the TFIIH subunit XPB, thereby preventing the opening of the transcription bubble, which consequently blocks transcription and triggers Pol II degradation [18, 42, 43] (Fig 4a right panel, S5a). Cells that were treated with 0.5 µM Triptolide for 90min showed a sharp increase of the short-bound Pol II fraction and a loss of both the medium and long-bound Pol II fractions (Fig 4a, 4b), similar to what we observed after THZ1 (Fig 2c). The residence time of short-bound Pol II after Triptolide or THZ1 was similar to NT conditions (Fig S5d), indicating that the turnover of initiating Pol II binding is barely affected by its transition to promoter pausing. In contrast to THZ1, 90 min of Triptolide resulted in a loss of ~50% of Pol II compared to non-treated cells as determined by live-cell imaging (Fig 4a right panel). Triptolide-induced Pol II degradation could be rescued by proteasome inhibition (Mg132, 50µM for 1hr before Triptolide), which further increased the short-bound Pol II fraction (Fig 4a), most likely due to accumulation of free Pol II (Fig 4b). This indicates that the fraction of Pol II that is targeted for degradation after Triptolide is highly mobile. Of note, Mg132 alone did not affect Pol II mobility or protein levels (Fig S5b). Interestingly, rescue of Triptolide-induced Pol II degradation by Mg132 accumulated P-Ser5 Pol II, suggesting that CDK7 can phosphorylate CTD Ser5 independently of transcription bubble opening and that this Ser5-phosphorylation might be a prerequisite for Pol II degradation after Triptolide. To test this hypothesis, we inhibited the Triptolide-induced Ser5-phosphorylation with THZ1 (Fig S5c). The addition of THZ1 completely rescued the Triptolide-induced Pol II degradation (Fig 3a right panel), in line with a model in which in the absence of an open transcription bubble due to Triptolide, P-Ser5 Pol II fails to escape the PIC and is subsequently degraded.
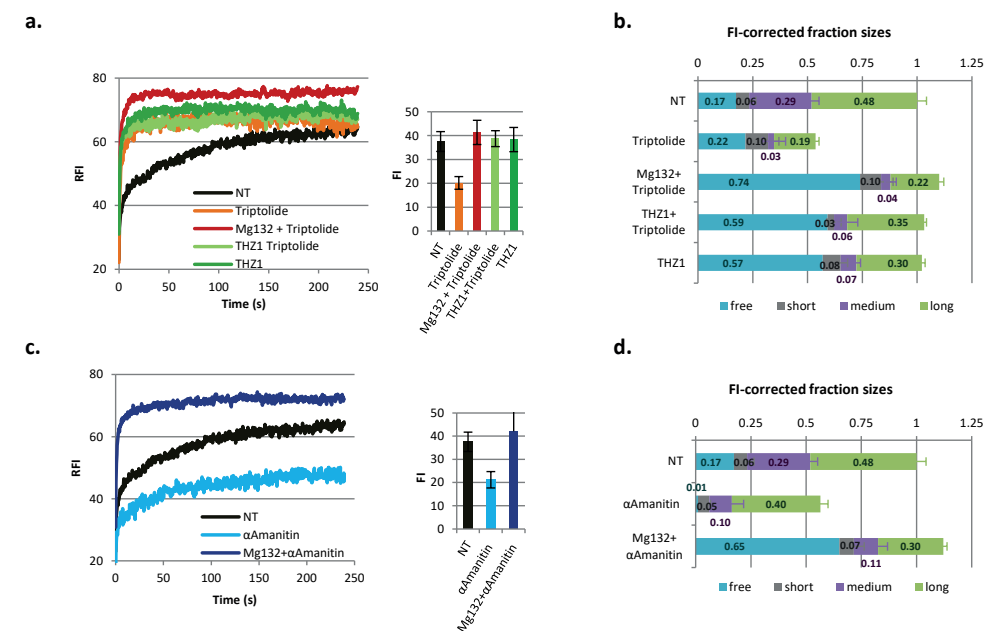

Figure 4. Live-cell Pol II kinetics after Triptolide and α-Amanitin. FRAP of GFP-RPB1 in non-treated (NT) cells and cells treated with 0.5µM Triptolide for 1h **(a)** or 100µg/ml α-Amanitin for 2h **(c)**. Mg132=proteasome inhibitor. Mean of n=20 cells of 2 independent experiments, FI chart shows mean ± SD. Modeled, FI-corrected Pol II fraction sizes in NT KI cells or after treatment with Triptolide **(b)** or α-Amanitin **(d)**. Mean ±SD of the 10 best fitting simulations.

Another widely used transcription inhibitor is α-Amanitin. α-Amanitin traps Pol II in a conformation that prevents translocation of the transcript and thereby inhibits nucleotide incorporation [42]. Thus far the effect of α-Amanitin on Pol II in living cells could not be extensively studied as most studies using fluorescently tagged Pol II were performed using an exogenously expressed α-Amanitin-resistant Pol II mutant [24, 44, 45]. FRAP of GFP-RPB1 after treatment with 100 µg/ml α-Amanitin for 90 min showed a loss of free, short-bound and medium-bound Pol II, whereas the long-bound, elongating fraction did not change in size (Fig 4d), but increased in residence time (Fig S5e), illustrating the reduced productive elongation rate after α-Amanitin [46]. In addition, α-Amanitin led to degradation of Pol II (Fig 4c right panel, Fig S5a). Interestingly, rescue of α-Amanitin-induced Pol II degradation by Mg132 tremendously increased the free fraction of Pol II (Fig 4d), as shown by an increase of the short-bound fraction during FRAP (Fig 4c), suggesting that stalled and presumably ubiquitylated Pol II is evicted from chromatin before proteasomal degradation.

DISCUSSION

Detailed knowledge of the dynamics of Pol II transcription is crucial to understand the complex regulation of Pol II transcription. To study the in vivo dynamics of endogenously expressed Pol II we generated GFP-RPB1 KI cells and used them to monitor real time Pol II kinetics. Photo-bleaching experiments revealed three kinetically distinct fractions of Pol II that remained chromatin-bound for either 2.4, 42, or 1370 seconds (Fig 3c). Using dedicated inhibitors that block transcription at specific stages we allocated these fractions to Pol II engaged in initiation, promoter pausing, or productive elongation (Fig 2b). Allocation of these fractions allowed to model Pol II kinetics and to construct a kinetic framework of steady-state Pol II transcription in unperturbed conditions.

FRAP allows the discrimination of promoter-paused Pol II based on its distinct kinetics

Interestingly, a previous study on the in vivo kinetics of YFP-tagged α-Amanitin-resistant Pol II on a lacO array integrated into U2OS cells also identified three different Pol II populations with chromatin binding times of 6, 54, and 517 sec, respectively [24]. The authors linked the fastest Pol II fraction to Pol II transiently interacting with promoters, while the slowest fraction was allocated to elongating Pol II. However, the fraction with intermediate kinetics was linked to initiating Pol II. Although the residence time of their intermediate fraction was strikingly similar to ours (54 s vs. 42 s), which would suggest that this might represent the same population of Pol II, their fraction was insensitive to P-TEFb inhibition, while we observed an increase in the fraction sizes of initiating and promoter-paused Pol II (Fig 3b). Hence, for the first time, our study discerned the kinetic fraction of Pol II that corresponds to promoter-paused complexes. While THZ1 and Triptolide abrogate promoter pausing and productive elongation, Flavopiridol still allows Pol II initiation and pausing but inhibits the release of Pol II from pausing to productive elongation. In line, we observed that in Flavopiridol-treated cells, the short and the medium-bound Pol II fractions remained, while the long-bound fraction was lost, strongly suggesting that the medium-bound fraction mainly represents promoter-paused Pol II. Of note, Flavopiridol accumulated transcription-competent but promoter-paused Pol II, as tested by performing FRAP after washing out Flavopiridol after 60 minutes of treatment, which resulted in Pol II kinetics similar to non-treated cells within minutes.

Pol II promoter pausing is a highly dynamic event

The average of the ten best fitting, simulated chromatin binding times of promoter-paused Pol II was 42 seconds in unperturbed conditions. Although we cannot exclude that Pol II is stably paused with residence times longer than 100 sec on some specific loci, as for example shown by tracking photoactivatable Pol II at the uninduced Hsp70 transgene of Drosophila polytene chromosomes [9], we found that on a genome wide average 23% of nuclear Pol II is paused for merely 42 seconds in human cells(Fig 2b,

Fig 2c middle panel). While this result differs strikingly from previous studies that found Pol II to be stably paused at promoters with half-lives of 5-15 minutes [8-12], it is in line with the very fast (<100sec) fluorescence recovery of Pol II in Flavopiridol-treated cells (Fig 2c). To our best knowledge the most likely explanation for this discrepancy may be that these previous studies were performed in the presence of Triptolide [8, 10-12], which disturbs transcriptional regulation and Pol II stability [17, 18],(Fig S5a). Furthermore, Triptolide may not be fully functional instantly upon administration and it is therefore difficult to exactly time the Pol2 half-life at promoters [15, 16]. Stable pausing of Pol II on promoters also has been challenged recently by DNA foot printing experiments in Drosophila cells, that have revealed that most promoter-paused Pol II is lost within 2.5 min (the earliest time point assessed by the authors) after Triptolide [15]. Rapid promoter-proximal termination was also found to be the most plausible explanation for the drastic increase of promoter-paused Pol II within a minute after H2O2 administration to U2OS cells [16]. In agreement with their conclusions, our findings strongly support a non-processive model of transcriptional idling of promoter-paused Pol II, defined by iterative rounds of Pol II pause entry and promoter-proximal termination [13-16]. As blocking the transition to productive elongation with Flavopiridol did not increase the residence-time of promoter-paused Pol II (Fig 3c middle panel), we further suggest that the Pol II turnover at the promoter is independent of its release into productive elongation but is mainly a result of continuous premature termination [15]. Importantly, our findings additionally elucidate the timeframe and the quantitative set up of these processes. In light of the increasing evidence supporting the dynamic turnover of promoter-bound Pol II, it is very tempting to speculate that the release of Pol II into productive elongation might be, at least partially, regulated by inhibiting promoter proximal Pol II termination.

Not only promoter-paused, but also initiating Pol II complexes are rapidly turned over

In addition to termination of Pol II after promoter pausing, our results further suggested that Pol II is frequently released from chromatin after initiation (Fig 3a). This abortive release after 2.4 seconds occurred even when the transition from initiation to pausing was blocked by THZ1 (Fig 3c) or Triptolide (Fig S5d), implying that the turnover of initiating Pol II does not depend on its transition to promoter pausing. The very fast recovery of fluorescence after photo-bleaching of initiating Pol II after THZ1 and Triptolide (Fig 2c) additionally implied that the Pol II that is abortively released after initiation (which is bleached during FRAP) is not the same complex that re-initiates (which is fluorescent and causes the fast fluorescence recovery). Similar observations have been made for the TATA-box binding protein [47, 48], further corroborating a model in which Pol II PICs are rapidly assembled and disassembled, rendering initiation highly dynamic. However, it has also been reported that a subset of the transcription machinery remains at the promoter after transcription initiation, forming a reinitiation scaffold that facilitates high levels of transcription inititation [49].

2

Implications of dynamic Pol II turnover at promoters

As a consequence of the frequent abortive release after initiation and pausing, according to our model, only one mRNA transcript will be synthesized per 100 Pol II initiation events on average. Although this seeming inefficiency of transcriptional initiation, which is consistent with Pol II measurements on a lacO array [24], is intriguing, it highlights the importance of initiation and promoter pausing as key regulatory control events [2]. However, the exact purpose of Pol II idling during initiation and pausing remains unknown. The continuous rounds of initiation and termination may contribute to keep promoters free of nucleosomes at active genes [6, 50], or alternatively, idling may be important to maintain promoter enhancer contacts [51]. Pol II idling at promoters would also facilitate transcriptional bursting, which refers to the episodic release of multiple Pol II during a short period of time [52, 53].

Quantitative framework of the Pol II transcription cycle

In addition to the assessment of Pol II dynamics, endogenous expression of GFP-RPB1 allowed to estimate the number of nuclear Pol II complexes. This analysis resulted in ~50,000 in a diploid cell (Fig 1e). With the assumption that productive elongation represents a single kinetic step the fraction size of elongating Pol II was modeled to 60%, approximating to ~30,000 elongating Pol II. Combined with the expression frequency-corrected average gene length of 46 kb (see Dataset S1, sheet 2) this suggests that on average ~1.1 Pol II are elongating per gene body (Dataset S1, sheet 3). Of note, while some less frequently expressed genes are probably not continuously transcribed in every cell and therefore do not have an engaged Pol II constantly, other highly expressed genes are likely continuously transcribed by multiple Pol II complexes [54]. Modeling further indicated that 10% of nuclear Pol II are initiating, and 23% are paused proximal to the promoter, which equals to ~16.500 Pol II located at promoter regions. Compared to 30,000 Pol II dispersed over the average gene body of 46 kb, these 16.500 Pol II are bound to the relatively short promoter region, coinciding with the high density of Pol II at the transcription start site observed in Chip-Seq experiments [11].

In summary, we show that GFP-RPB1 KI cells provide a powerful tool to study endogenous Pol II in living cells and uncover valuable new insights into the mechanism and the dynamics of Pol II initiation and promoter pausing. We demonstrate that photo-bleaching combined with modeling is a valuable extension of the existing tools to study the dynamics of Pol II transcription.

FUNDING

MATERIALS AND METHODS

Generation of GFP-RPB1 knock in (KI) cells and cell culture

KI cells were generated from Sv40-immortalized MRC-5 fibroblasts as described here [19]. Detailed protocol in SI Materials and Methods. Cells were cultured in a 1:1 mixture of Ham's F10 and DMEM (Gibco) supplemented with antibiotics and 10% fetal calf serum, at 37°C; 20% O_2, and 5% CO_2 in a humidified incubator.

Immunofluorescence and EU incorporation

Click it-based EU incorporation was performed as described previously [55]. Detailed protocol in SI Materials and Methods. Immunofluorescence was performed as described in SI Materials and Methods. Transcription inihibitors, antibodies and respective concentrationsare listed in Dataset S2.

Cellular Fractionation, Western Blot Analysis

The fractionation procedure was adapted from [56], and is described in detail in SI Materials and Methods. Fractions were separated on a 6% SDS Page gel, blotted overnight at 40V. Blots were blocked with 1.5% BSA in PBS and stained with primary antibodies listed in Dataset S2. Secondary antibodies were coupled to IRDyes (LiCor) and imaged with an Odyssey CLx infrared scanner (LiCor).

Live cell confocal microscopy and quantification of nuclear Pol II

Live cell imaging was performed on a Leica SP5 confocal laser scanning microscope with a HCX PL APO CS 63x, 1.40-NA oil immersion lens. Images were recorded with a 488nm Argon laser and a 500–600nm bandpass filter. For detailed RPB1 FRAP procedure please see SI Materials and Methods. For Pol II concentration measurements eGFP (Biovision) was dissolved in PBS, serially diluted in culture medium and the fluorescence intensity was measured alongside nuclear GFP-RPB1 fluorescence. For detailed procedure see SI Materials and Methods.

Monte-Carlo Modeling

Experimental FRAP curves were fit by least squares, to a large set of computer simulation-generated FRAP curves that were computed based on a model that simulates diffusion of molecules (here Pol II), and binding to and releasing from immobile elements, (here chromatin), in an ellipsoidal volume (here the nucleus). For details see SI Materials and Methods.

Detailed protocols for **RNA Seq data analysis, Chromatin immunoprecipitation, siRNA transfection and RT-PCR** can be found in SI Materials and Methods

ACKNOWLEDGEMENTS

2

2

REFERENCES

1. Sainsbury, S., C. Bernecky, and P. Cramer, *Structural basis of transcription initiation by RNA polymerase II.* Nat Rev Mol Cell Biol, 2015. **16**(3): p. 129-43.
2. Jonkers, I. and J.T. Lis, *Getting up to speed with transcription elongation by RNA polymerase II.* Nat Rev Mol Cell Biol, 2015. **16**(3): p. 167-77.
3. Heidemann, M., et al., *Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription.* Biochim Biophys Acta, 2013. **1829**(1): p. 55-62.
4. Scheidegger, A. and S. Nechaev, *RNA polymerase II pausing as a context-dependent reader of the genome.* Biochem Cell Biol, 2016. **94**(1): p. 82-92.
5. Adelman, K. and J.T. Lis, *Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans.* Nat Rev Genet, 2012. **13**(10): p. 720-31.
6. Gilchrist, D.A., et al., *Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation.* Cell, 2010. **143**(4): p. 540-51.
7. Rougvie, A.E. and J.T. Lis, *The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged.* Cell, 1988. **54**(6): p. 795-804.
8. Jonkers, I., H. Kwak, and J.T. Lis, *Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons.* Elife, 2014. **3**: p. e02407.
9. Buckley, M.S., et al., *Kinetics of promoter Pol II on Hsp70 reveal stable pausing and key insights into its regulation.* Genes Dev, 2014. **28**(1): p. 14-9.
10. Henriques, T., et al., *Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals.* Mol Cell, 2013. **52**(4): p. 517-28.
11. Shao, W. and J. Zeitlinger, *Paused RNA polymerase II inhibits new transcriptional initiation.* Nat Genet, 2017.
12. Chen, F., X. Gao, and A. Shilatifard, *Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide.* Genes Dev, 2015. **29**(1): p. 39-47.
13. Brannan, K., et al., *mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription.* Mol Cell, 2012. **46**(3): p. 311-24.
14. Wagschal, A., et al., *Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII.* Cell, 2012. **150**(6): p. 1147-57.
15. Krebs, A.R., et al., *Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters.* Mol Cell, 2017.
16. Nilson, K.A., et al., *Oxidative stress rapidly stabilizes promoter-proximal paused Pol II across the human genome.* Nucleic Acids Res, 2017. **45**(19): p. 11088-11105.
17. Wang, Y., et al., *Triptolide (TPL) inhibits global transcription by inducing proteasome-dependent degradation of RNA polymerase II (Pol II).* PLoS One, 2011. **6**(9): p. e23993.
18. Manzo, S.G., et al., *Natural product triptolide mediates cancer cell death by triggering CDK7-dependent degradation of RNA polymerase II.* Cancer Res, 2012. **72**(20): p. 5363-73.
19. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-308.
20. Boulon, S., et al., *HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II.* Mol Cell, 2010. **39**(6): p. 912-24.
21. Jackson, D.A., A. Pombo, and F. Iborra, *The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells.* FASEB J, 2000. **14**(2): p. 242-54.
22. Hieda, M., et al., *Different populations of RNA polymerase II in living mammalian cells.* Chromosome Res, 2005. **13**(2): p. 135-44.
23. Kimura, H., K. Sugaya, and P.R. Cook, *The transcription cycle of RNA polymerase II in living cells.* J Cell Biol, 2002. **159**(5): p. 777-82.
24. Darzacq, X., et al., *In vivo dynamics of RNA polymerase II transcription.* Nat Struct Mol Biol, 2007. **14**(9): p. 796-806.

25. Houtsmuller, A.B. and W. Vermeulen, *Macromolecular dynamics in living cell nuclei revealed by fluorescence redistribution after photobleaching.* Histochem Cell Biol, 2001. **115**(1): p. 13-21.

26. van Royen, M.E., et al., *Fluorescence recovery after photobleaching (FRAP) to study nuclear protein dynamics in living cells.* Methods Mol Biol, 2009. **464**: p. 363-85.

27. Muller, W.E., et al., *Effect of cordycepin on nucleic acid metabolism in L5178Y cells and on nucleic acid-synthesizing enzyme systems.* Cancer Res, 1977. **37**(10): p. 3824-33.

28. Fish, R.N. and C.M. Kane, *Promoting elongation with transcript cleavage stimulatory factors.* Biochim Biophys Acta, 2002. **1577**(2): p. 287-307.

29. Kwiatkowski, N., et al., *Targeting transcription regulation in cancer with a covalent CDK7 inhibitor.* Nature, 2014. **511**(7511): p. 616-20.

30. Nilson, K.A., et al., *THZ1 Reveals Roles for Cdk7 in Co-transcriptional Capping and Pausing.* Mol Cell, 2015. **59**(4): p. 576-87.

31. Chao, S.H., et al., *Flavopiridol inhibits P-TEFb and blocks HIV-1 replication.* J Biol Chem, 2000. **275**(37): p. 28345-8.

32. Peterlin, B.M. and D.H. Price, *Controlling the elongation phase of transcription with P-TEFb.* Mol Cell, 2006. **23**(3): p. 297-305.

33. Sigova, A.A., et al., *Transcription factor trapping by RNA in gene regulatory elements.* Science, 2015. **350**(6263): p. 978-81.

34. Flynn, R.A., et al., *Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome.* Proc Natl Acad Sci U S A, 2011. **108**(26): p. 10460-5.

35. Hah, N., et al., *Enhancer transcripts mark active estrogen receptor binding sites.* Genome Res, 2013. **23**(8): p. 1210-23.

36. Sobell, H.M., *Actinomycin and DNA transcription.* Proc Natl Acad Sci U S A, 1985. **82**(16): p. 5328-31.

37. Geverts, B., M.E. van Royen, and A.B. Houtsmuller, *Analysis of biomolecular dynamics by FRAP and computer simulation.* Methods Mol Biol, 2015. **1251**: p. 109-33.

38. Saunders, A., L.J. Core, and J.T. Lis, *Breaking barriers to transcription elongation.* Nat Rev Mol Cell Biol, 2006. **7**(8): p. 557-67.

39. Andrade-Lima, L.C., et al., *DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes.* Nucleic Acids Res, 2015. **43**(5): p. 2744-56.

40. Veloso, A., et al., *Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications.* Genome Res, 2014. **24**(6): p. 896-905.

41. Yu, M., et al., *RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II.* Science, 2015. **350**(6266): p. 1383-6.

42. Bensaude, O., *Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity?* Transcription, 2011. **2**(3): p. 103-108.

43. Alekseev, S., et al., *Transcription without XPB Establishes a Unified Helicase-Independent Mechanism of Promoter Opening in Eukaryotic Gene Expression.* Mol Cell, 2017. **65**(3): p. 504-514 e4.

44. Becker, M., et al., *Dynamic behavior of transcription factors on a natural promoter in living cells.* EMBO Rep, 2002. **3**(12): p. 1188-94.

45. Lux, C., et al., *Transition from initiation to promoter proximal pausing requires the CTD of RNA polymerase II.* Nucleic Acids Res, 2005. **33**(16): p. 5139-44.

46. Rudd, M.D. and D.S. Luse, *Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes.* J Biol Chem, 1996. **271**(35): p. 21549-58.

47. Sprouse, R.O., et al., *Regulation of TATA-binding protein dynamics in living yeast cells.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13304-8.

48. de Graaf, P., et al., *Chromatin interaction of TATA-binding protein is dynamically regulated in human cells.* J Cell Sci, 2010. **123**(Pt 15): p. 2663-71.

49. Yudkovsky, N., J.A. Ranish, and S. Hahn, *A transcription reinitiation intermediate that is stabilized by activator.* Nature, 2000. **408**(6809): p. 225-9.

50. Gilchrist, D.A., et al., *NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly.* Genes Dev, 2008. **22**(14): p. 1921-33.

51. Ghavi-Helm, Y., et al., *Enhancer loops appear stable during development and are associated with paused polymerase.* Nature, 2014. **512**(7512): p. 96-100.

52. Fukaya, T., B. Lim, and M. Levine, *Enhancer Control of Transcriptional Bursting.* Cell, 2016. **166**(2): p. 358-368.

53. Hnisz, D., et al., *A Phase Separation Model for Transcriptional Control.* Cell, 2017. **169**(1): p. 13-23.

54. Tantale, K., et al., *A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting.* Nat Commun, 2016. **7**: p. 12248.

55. Nakazawa, Y., et al., *A semi-automated non-radioactive system for measuring recovery of RNA synthesis and unscheduled DNA synthesis using ethynyluracil derivatives.* DNA Repair (Amst), 2010. **9**(5): p. 506-16.

56. Aygun, O., J. Svejstrup, and Y. Liu, *A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin.* Proc Natl Acad Sci U S A, 2008. **105**(25): p. 8580-4.

57. Sander, J.D., et al., *ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W462-8.

58. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.

59. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.

60. Trapnell, C., L. Pachter, and S.L. Salzberg, *TopHat: discovering splice junctions with RNA-Seq.* Bioinformatics, 2009. **25**(9): p. 1105-11.

61. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.

2

2

SUPPLEMENTAL MATERIALS AND METHODS

Generation of GFP-RPB1 knock in (KI) cells

A single guide RNA (sgRNA) sequence targeting *POLR2A* (*RPB1*) was selected using the ZitFit Targeter Program [57]. 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 3 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 3 silent point mutations (ATG CAC GGG GGT GGC à ATG CA**T** GG**A** GG**G** GGC) to prevent Cas9 from re-cutting once RPB1 has been targeted. The homology arms were amplified from sequenced sub clones of genomic DNA isolated from VH10 cells. 0.9µg of pX330-U6-Chimeric_BB-CBh-hSpCas9_sgRNA and 3.6µg of pBs-RPB1-LA-Atg-Flag-Gfp-Rpb1-RA were transiently transfected in wild type MRC-5 cells with the JetPei transfection reagent (Polyplus-transfection) according to manufacturer's instructions. GFP-positive 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 Table S2.

Immunofluorescence and EU incorporation

Cells were grown to 80% confluency on glass coverslips, treated as indicated, fixed with 2% PFA in PBS for 15min at room temperature. After permeabilisation with 0.1% Triton in PBS for 10min and blocking in 1.5% BSA and 0.15% glycine in PBS for 15 min cells were incubated with primary antibodies as listed in TableS2. After washing in 0.1% triton in PBS for 10min cells were incubated with respective secondary antibodies coupled to the indicated Alexa fluorophores for 2h at RT. Transcription levels were measured by pulse labeling with 5ethynyl uridine (EU) (Jena Bioscience). for 30' with 1µM EU in Ham's F10 medium supplemented with 10% dialyzed fetal calf serum (Gibco). To measure EU corporation after transcription inhibition cells were treated with inhibitors for 1h before the EU pulse and inhibitors were present during the EU pulse. Subsequently, cells were washed with PBS, fixed with 2% PFA in PBS for 15min. After permeabilisation with 0.1% triton in PBS for 10min click chemistry-based azide coupling was performed by incubation for 30min with 60µM Atto594 Azide (Attotec, Germany) in 50mM Tris buffer (pH8) with 4mM CuSO$_4$ (Sigma Aldrich) and 10mM 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 63x Plan-Apochromat oil immersion lens (NA 1.40).

FRAP

For FRAP at pixel size 24.6x24.6 µM a strip of 512x32 pixels spanning the nucleus was imaged every 400ms with 400Hz. 25 frames were recorded before the bleach pulse. The average, background-corrected fluorescence intensity of frame 10 to 20 of these pre-bleach measurements were used to calculate the pre-blech fluorescence intensity. GFP fluorescence in the strip was bleached for 1 frame with 100% laser power. The recovery of fluorescence was monitored for 4min (600frames) within and outside the strip. For half nucleus bleach at pixel size 123x123µM an image of 1024x1024 pixels was recorded every minute for 120min at 200 Hz. Regions of interest were selected over half nuclei and were photo-bleached with 100% laser intensity for one frame. FRAP and FLIP curves were corrected for background fluorescence outside the nucleus and normalized to pre-bleach fluorescence in the region of interest.

Cellular Fractionation

Cells were lysed in lysis buffer (30mM Hepes pH7.6, 1mM MgCl$_2$, 130mM NaCl, 0.5% Triton, 0.5mM DTT, 50µM Mg132, EDTA free protease inhibitor (Roche) and phosphatase inhibitor cocktail 2 (Sigma Aldrich)) on ice for 15 min and then scraped in lysis buffer. 500U of benzonase (Millipore) was added for 1h on ice. For fractionation whole cell lysates were centrifuged for 15 min at 13000 rpm and 4°C. Supernatant containing nucleoplasmic RPB1 was collected. The pellet containing chromatin-bound RPB1 was washed twice with lysis buffer and re-suspended in lysis buffer. Supernatant and pellet were diluted with SDS Page loading buffer (4% SDS, 0.2% bromophenol blue, 20% glycerol, 200mM β-mercaptoethanol) and separated on separated on a 6% SDS Page gel. Transfer was done overnight at 4°C and 40V in 2x transfer buffer (25mM TRIS, 190mM 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 by the absorbance at 489 nm ($\varepsilon489$ 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 82x82 µM at 200Hz. 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 hyper-tetra-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 10min. Crosslinking was quenched by adding glycine to a final concentration of 0.125M 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, 10mM Tris-HCL pH8, 1mM EDTA, 0.5mM EGTA) including complete protease inhibitor cocktail without EDTA (Roche), phosphatase inhibitor cocktail 2 and 3 (Sigma). Samples were sonicated for 15 min using Bioruptor (Diagenode, 15 sec on, 15 sec off, amplitude High) to fragments of 200-500 base pairs and centrifuged for 10 min at 14000 rpm to collect chromatin. Chromatin was diluted using dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCL pH8, 167mM NaCl) and pre-cleared for 30min using Pierce Protein G Agarose beads (Thermo Scientific). RPB1 was immuno-precipitated overnight at 4°C using 5 μg GFP antibody (5mg/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, 2mM EDTA, 20mM Tris-HCL pH8, 150mM NaCl), high salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCL pH8, 500mM NaCl), lithium chloride buffer (0.25M LiCl, 1% NP-40, 1% Sodium Deoxychelate, 1mM EDTA, 10mM Tris-HCL pH8) and twice with Tris-EDTA buffer (10mM Tris-HCL pH8, 1mM EDTA). DNA was eluted from beads twice using elution buffer (1% SDS, 0.1M NaHCO$_3$) and decrosslinked for 4h with NaCl at 65°C and 950rpm in a thermoshaker. Proteins were removed by incubating 1h at 45°C and 700rpm 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 manufacturer's instruction. Experiments were performed 48h after transfection. Knockdown efficiency was checked by qPCR. siRNA and primer sequences are listed in Table S2. mRNA expression levels were normalized to B2M (Beta-2-Microglobulin). qPCR was performed using the CFX96 Touch™ Real-Time PCR Detection System (Biorad) and PowerUp SYBR green master mix (Thermo Fisher). Primers are listed in Table S2 and the used program was: 50°C for 2 min, 95°C for 2 min, 45 cycles of 15 sec at 95°C and 1 min at 58°C followed by a dissociation curve: 95°C for 10 sec, 65°C for 5 sec and heating from 5°C to 95°C.

Monte-Carlo-based computational Model of Pol II kinetics

To simulate photo-bleaching the three-dimensional 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 photo-bleaching 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 3-D 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 ($0 \leq r_i \leq 1$) chosen from a uniform distribution, and $G(r_i)$ is an inversed cumulative Gaussian distribution with $s^2 = 2D\Delta t$, where D is the diffusion coefficient. A fixed diffusion coefficient of 2.0 μm^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 non-fluorescent GFP variant with increasing molecular weight [26]. 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 $P_{immobilise} = k_{on} = k_{off} \cdot F_{imm} / (1 - F_{imm})$, 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 $P_{mobilise} = k_{off} = 1 / T_{imm}$. 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 ten 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 ten 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 wild type foreskin fibroblasts (HF1) (GEO: GSM1612077, run SRR1806546) [39] was trimmed using TrimmomaticSE v0.32 [58] with

the steps: "ILLUMINACLIP:<illumina_adapter_list_file>:2:40:15 LEADING:3 TRAILING:3 MINLEN:15". The data was then aligned to the hg19 reference genome using bowtie2 v2.2.6 [59] 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 non-overlapping 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 MRC5 wild type cells, a single clone of MRC5 WT cells, and 2 single clones of MRC5 GFP-Rpb1 knock in cells using the miRNeasy Mini Kit (Qiagen) according to 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 pair-ended sequencing files from the Illumina system were quality checked by FastQC (v0.11.3) (available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Low quality reads and sequencing adapters were removed by Trimmomatic (v0.35)[58] and the resulting FASTQ files were aligned to the human reference genome (hg19) using Tophat (v2.0.9)[60]. The aligned reads were analysed by HTseq (v.0.6.0) [61] to quantify gene expression. Genes with less than 5 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-value <0.05. The statistics plots were created by Rstudio (v0.99.486) (Team, R. (2015). RStudio: integrated development for R. Available online at *http://www. rstudio. com*.

SUPPLEMENTAL REFERENCES

1. Sander JD*, et al.* (2010) ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res* 38(Web Server issue):W462-468.
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(5):2744-2756.
4. Bolger AM, Lohse M, & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114-2120.
5. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357-359.
6. Trapnell C, Pachter L, & Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105-1111.
7. Anders S, Pyl PT, & Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2):166-169.

2

56

LIVE-CELL ANALYSIS OF ENDOGENOUS GFP-RPB1
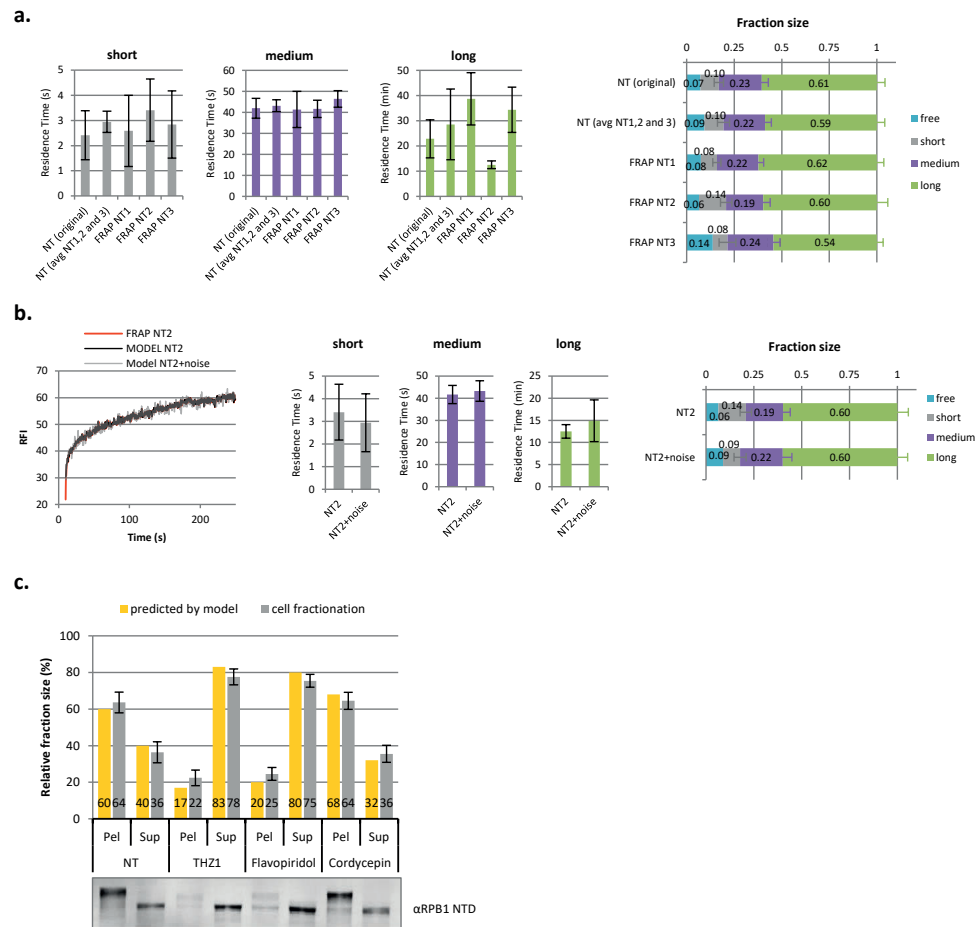
SUPPLEMENTAL FIGURES

2
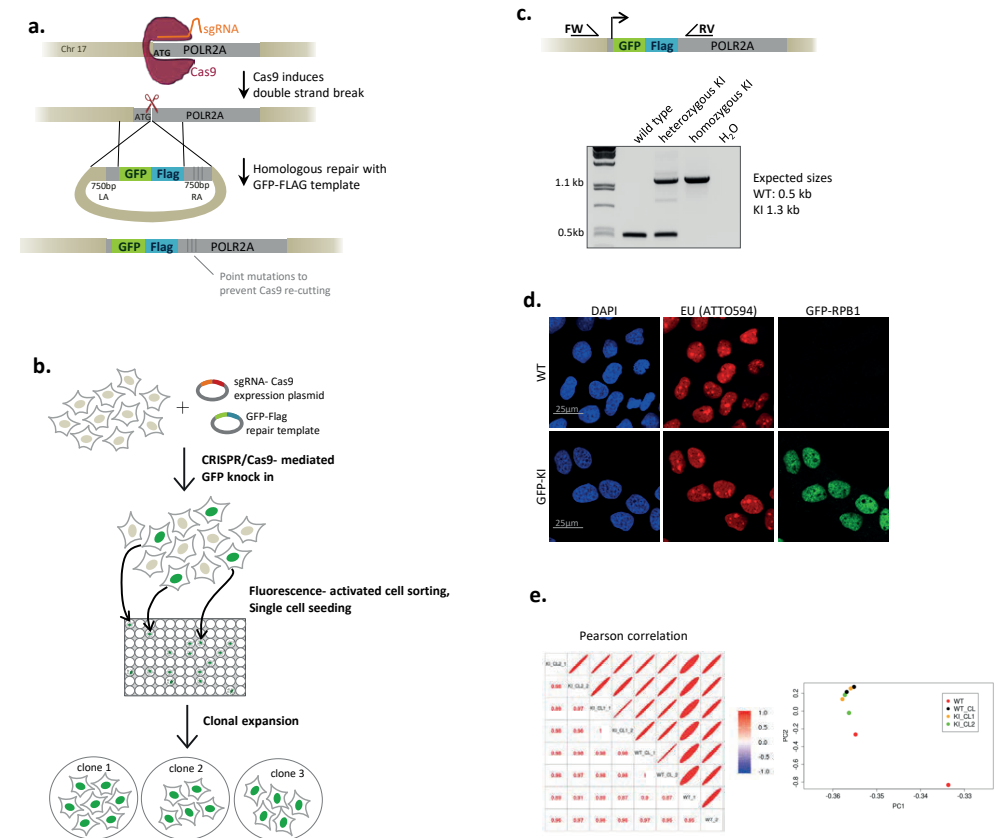
Supplemental Figure 1. Schematic representation of GFP-RPB1 KI strategy: CRISPR associated protein 9 (Cas9) is targeted to the POLR2A locus by a specifically designed single guide RNA (sgRNA). Cas9 induces a double strand break (DSB) directly downstream of the POLR2A start codon (ATG). The DSB can be repaired by homologous recombination (HR) using the repair template containing the GFP-FLAG sequence for homology. The right homology arm (RA) of the repair template contains three point mutations to prevent Cas9 from re-cutting after HR. The left homology arm (LA) is identical to the 750 bp upstream of the sgRNA target site. Schematic procedure of Cas9-mediated GFP-labelling of RPB1 in MRC-5 cells. A plasmid expressing the sgRNA and Cas9 was co-transfected with a repair template into MRC-5 cells. GFP positive cells were isolated using fluorescence-activated cell sorting and seeded as single cells into a 96-well plate by limiting dilution. Single cells were expanded and subsequently genotyped. GFP-RPB1 knock-in (KI) was checked by genotyping genomic DNA of MRC-5 wild type (WT) and KI cells by PCR with the indicated primers (top panel). PCR products were analysed by DNA gel-electrophoresis (bottom panel). Representative images of cells that were pulse labeled with EU for 1h followed by click-chemistry-based EU coupling to ATTO594 of non-treated MRC5 wild type (WT) or GFP-RPB1 knock in (KI) cells. DNA was stained with DAPI, RPB1 was visualized by αGFP immunostaining (left) and principal component analysis (right) of RNA expression levels of a mixed population of MRC5 wild type (WT) cells, a single cell clone of MRC5 WT cells (WT_CL), and 2 single cell clones of MRC5 cells expressing GFP-RPB1 (KI_CL). To replicates of each sample were analysed.
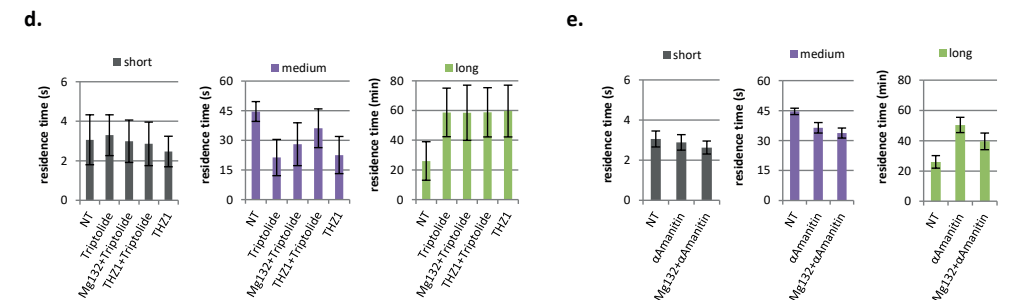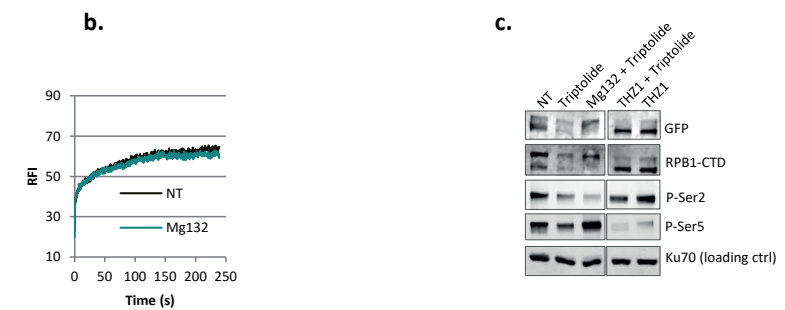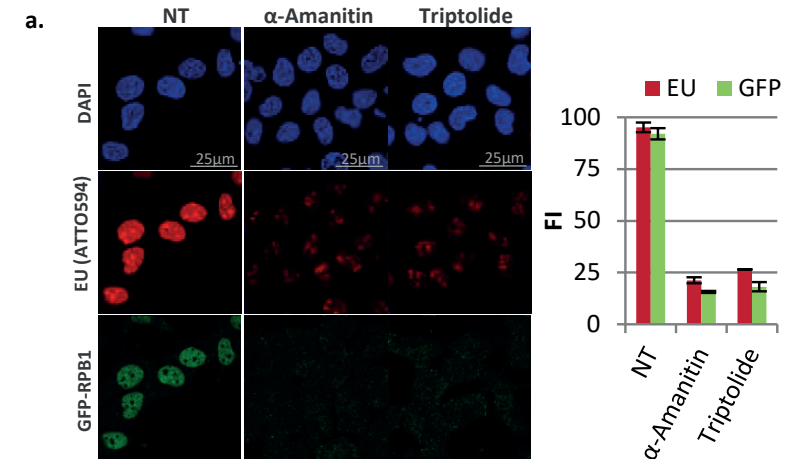
57

2
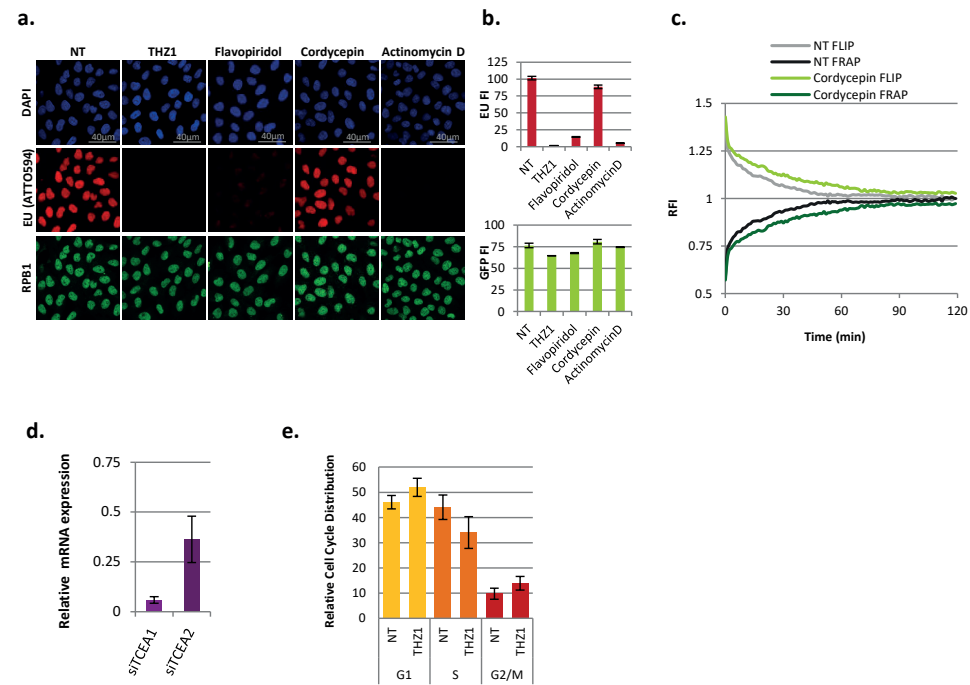
Supplemental Figure 2. Representative images of cells that were pulse labeled with EU for 1h followed by click-chemistry based EU coupling to ATTO594 of non-treated (NT) KI cells and after treatment with the indicated transcription inhibitors. DNA was stained with DAPI, RPB1 was visualized by αGFP immunostaining. Quantification of EU incorporation shown in S2b (left panel) and quantification of GFP fluorescence (right panel). FI= Fluorescence intensity, mean ±SEM, n>55 cells. Fluorescence recovery after photo-bleaching (FRAP) of the bleached and fluorescence loss in photo-bleaching (FLIP) of the non-bleached half of nuclei of non-treated (NT) cells and cells treated with 100μM Cordycepin for 90 min are plotted against time. RFI=Relative fluorescence intensity, mean of n=10 cells . 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 non-targeting siRNA. Propidium-Iodide content-based cell cycle distribution of non-treated (NT) MRC5 cells or cells treated with 1μM of THZ1 for 1.5 hr.
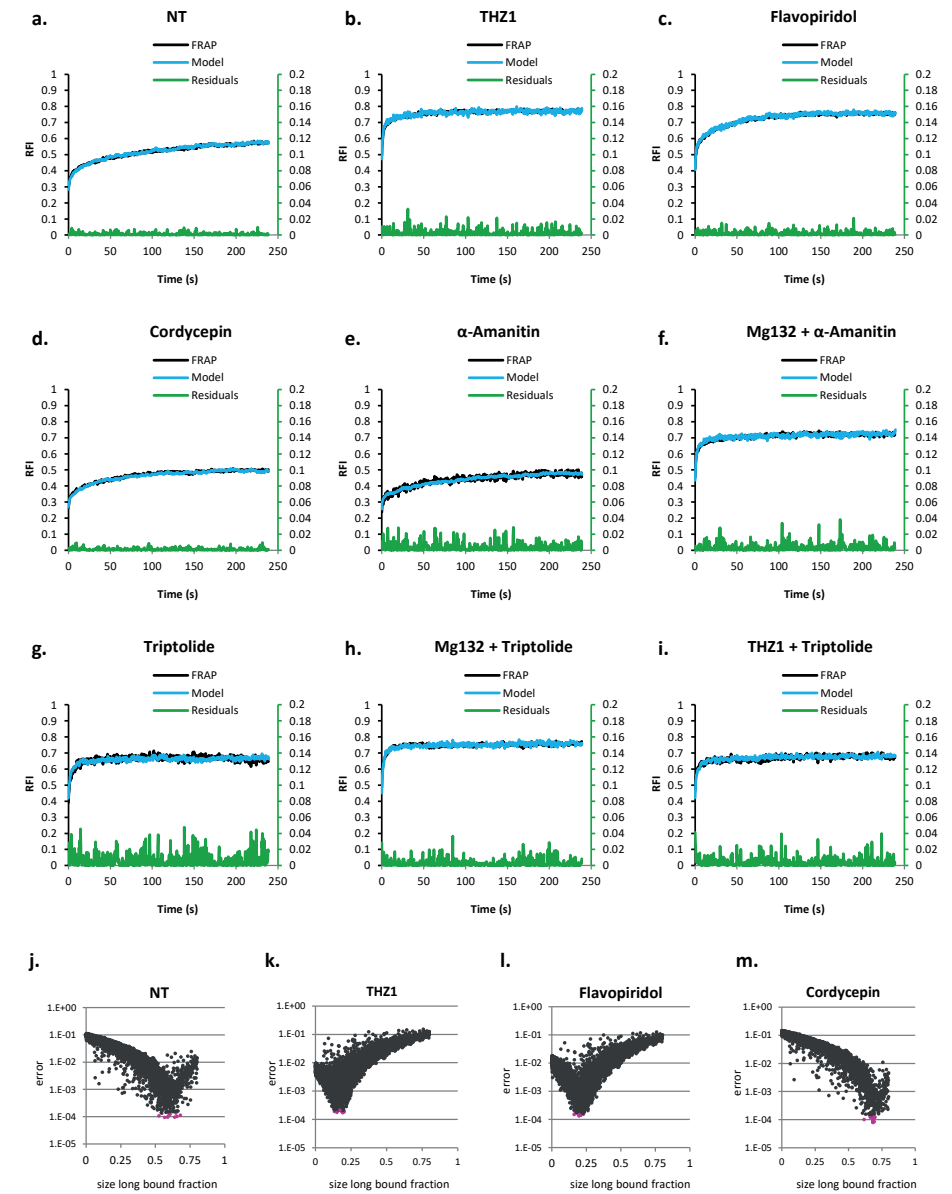
Supplemental Figure 3. Fitting of FRAP curves described in Figures 2c, 4a, and 4b to the best fitting curve generated by MC-based modeling. MRC5 GFP-RPB1 knock in cells were non-treated (NT) (a) or treated for 1hr with 1μM THZ1(b), 1μM Flavopiridol (c), or 100μM Cordycepin (d). Cells were treated for 2hr with 100μg/ml α-Amanitin with (f) or without (e) pre-treatment with 50μM Mg132 for 1 hr. Cells were treated for 1hr with 0.5μM Tritptolide with (h) or without (g) pre-treatment with 50μM Mg132 for 1 hr, or after pre-treatment with 1μM THZ1 for 1hr (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 3b, are marked in magenta (j-m).

Supplemental Figure 4. Residence times and fraction sizes of Pol II fractions obtained from MC-based modelling of 20 cells of 2 experiments (NT original, same as NT Fig 2c), the average of three individually modeled FRAP curves (NT avg. NT1,2, and3), and 3 individually modeled FRAP experiments (NT1, NT2, and NT3). NT1,2,3 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. Residence times (middle panel) and fraction sizes (right panel) of Pol II fractions obtained from MC-based modelling of one individual FRAP curve without or with addition of gaussian noise (σ=2). Bottom panel: Western blot after cell fractionation of non-treated (NT) MRC5 KI cells or cells treated with the indicated inhibitors for 90 min. Top panel: Fraction sizes of chromatin-bound (Pel) and nucleoplasmic (Sup) Pol II after fractionation and western blot analysis compared to fraction sizes predicted by MC-based modeling (Fig 3b). The long bound fraction is presented as chromatin-bound, whereas the sum of the free, short, and medium bound fractions is presented as nucleoplasmic. NTD= N-terminal domain. Mean ±SD of 3 Western Blots. One representative western blot is shown.

Supplemental Figure 5. Representative images of cells that were pulse labeled with EU for 1h followed by click-chemistry-based EU coupling to ATTO594 of non-treated (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 in the right. mean ±SEM, n>50 cells FRAP of GFP-RPB1 in non-treated (NT) cells and cells treated with the proteasome inhibitor Mg132 (left panel). Column chart shows mean ±SD of the pre-bleach fluorescence intensities (FI) of 10 cells analysed by FRAP as a measure of Pol II protein levels (right panel). Western blot of GFP-RPB1 KI whole cell lysates after indicated treatments. All samples originate from the same blot. NT= non-treated. Mg132 =proteasome inhibitor. and e. Residence times of short bound, medium bound, and long bound Pol II fractions obtained from Monte Carlo-based modeling of GFP-RPB1 FRAP data (Fig 3a,c, and e) in NT KI cells or after treatment with Triptolide (d) or α-Amanitin (e). Mean ±SEM of the 10 best fitting simulations.

Chapter 5

FLUORESCENTLY-LABELLED CPD AND 6-4PP PHOTOLYASES: NEW TOOLS FOR LIVE-CELL DNA DAMAGE QUANTIFICATION AND LASER-ASSISTED REPAIR

Barbara Steurer[1,#], Yasemin Turkyilmaz[1,#], Marvin van Toorn[1], Wessel van Leeuwen[1], Paula Escudero-Ferruz[1], Jurgen A. Marteijn[1,*]

[1] Erasmus MC, University Medical Center Rotterdam, Department of Molecular Genetics, Oncode Institute, Wytemaweg 80, 3015 CN, Rotterdam, The Netherlands
#contributed equally
* Corresponding author: Jurgen A. Marteijn

ABSTRACT

UV light induces cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone (6-4) photoproducts (6-4PPs), which can result in carcinogenesis and aging, if not properly repaired by nucleotide excision repair (NER). Assays to determine DNA damage load and repair rates are invaluable tools for fundamental and clinical NER research. However, most current assays to quantify DNA damage and repair cannot be performed in real time. To overcome this limitation, we made use of the damage recognition characteristics of CPD and 6-4PP photolyases (PLs). Fluorescently-tagged PLs efficiently recognize UV-induced DNA damage without blocking NER activity, and therefore can be used as sensitive live-cell damage sensors. Importantly, FRAP-based assays showed that PLs bind to damaged DNA in a highly sensitive and dose-dependent manner, and can be used to quantify DNA damage load and to determine repair kinetics in real time. Additionally, PLs can instantly reverse DNA damage by 405 nm laser-assisted photo-reactivation during live-cell imaging, opening new possibilities to study lesion-specific NER dynamics and cellular responses to damage removal. Our results show that fluorescently-tagged PLs can be used as a versatile tool to sense, quantify and repair DNA damage, to study NER kinetics and UV-induced DNA damage response in living cells.

INTRODUCTION

Our genome is continuously exposed to various types of DNA damage. If not repaired correctly, DNA lesions may result in mutations, cellular senescence or cell death, which can eventually lead to various pathological conditions including carcinogenesis and aging [1]. To counteract these deleterious effects of DNA damage, cells have evolved a variety of mechanisms, including several DNA repair pathways [2]. Nucleotide excision repair (NER) is one of the most versatile DNA repair pathways, as it removes a wide variety of DNA helix-destabilizing lesions. Prominent examples of NER substrates are the UV-induced cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone (6-4) photoproducts (6-4PPs). The biological importance of NER is illustrated by the severe clinical symptoms of human disorders caused by inherited NER defects, including the cancer-prone xeroderma pigmentosum (XP) syndrome or the premature aging disorder Cockayne's syndrome (CS) [3].

NER is initiated by two sub-pathways that differ in their mode of damage recognition. Global genome NER (GG-NER) detects lesions in the entire genome, by the main DNA damage binding protein XPC [4]. XPC recognizes DNA-helix distortions such as induced by 6-4PP lesions, but needs the activity of the UV-DDB complex, composed of DDB1 and DDB2, to detect mildly helix-destabilizing CPD lesions [5, 6]. Transcription-coupled NER (TC-NER) is initiated when DNA damage located in the actively transcribed strand blocks elongating RNA polymerase II, which results in the recruitment of the TC-NER factors CSA, CSB, and UVSSA [7, 8]. Once the DNA lesion is recognized, general transcription factor II H (TFIIH) is recruited [9, 10] to unwind the DNA surrounding the damage [11] and to verify the lesion together with XPA [12, 13]. The endonucleases XPG and ERCC1/XPF subsequently remove a ~30 nucleotide long fragment of DNA around the lesion [14]. Finally, the DNA is restored back to its original state by DNA synthesis and ligation steps [15, 16].

Recent studies have shown that NER is a tightly regulated, multistep pathway that requires many proteins and post-translational modifications for the efficient and accurate transition between the successive reaction steps [3, 17-19]. Additionally, as NER takes place in the complex chromatin and nuclear environment, many factors involved in chromatin remodeling [3, 20, 21], transcription [22], or replication [23] influence NER activity, and most likely many other involved factors are awaiting their discovery. Therefore, assays to quantify DNA damage and repair rates are invaluable tools to investigate the roles of such factors and to obtain new fundamental insights into the molecular mechanism of NER. Moreover, assays to detect impairments or deficiencies in NER activity have been crucial for the diagnosis of NER-deficient patients and can be used as indicators for predispositions to mutations, the onset of cancer, or DNA damage-induced aging [24-27].

Over the years, several assays were developed to quantitatively monitor UV-induced DNA damage and NER-mediated repair. Traditionally NER activity is measured by determining the rate of UV-induced DNA repair synthesis, the last step of the NER

reaction [28-30], or by determining the levels of CPDs in the DNA in time using T4 endonuclease V [31]. Over the years several other assays have been developed to monitor upstream NER activity, including UV-damage removal [32], NER-induced incisions [33] or quantification of excision products [34]. TC-NER is often determined indirectly by quantifying the recovery of RNA synthesis (RRS) [35, 36], or by using host cell reactivation assays [37]. Alternatively, TC-NER can be measured in a direct manner by strand-specific repair assays [38], or by more recently developed single-cell assays, such as the modified COMET-FISH procedure [39], or the TC-NER specific UDS assay [40]. Direct detection and quantification of UV-induced DNA damage and its removal in time can be accomplished using antibodies specifically recognizing CPD or 6-4PP lesions in combination with immunofluorescence or ELISA procedures [32]. Although proven to be useful in studying UV-induced DNA repair, these assays depend highly on the quality of the antibodies and have specific limitations. For instance, antibody-based detection of CPD or 6-4PP lesions requires DNA denaturation, to allow DNA damage recognition by these antibodies. For example in immunofluorescence experiments, this denaturation may interfere with co-staining of other proteins of interest. Importantly, most of these assays require cell fixation, which makes them incompatible with live-cell applications, and therefore can only provide endpoint measurements. To overcome these issues, measurements of the DNA damage binding kinetics of fluorescently-tagged NER factors can be used to evaluate repair activity in living cells [41-43]. However, these binding kinetics do not provide a direct measurement of DNA damage quantities, as the DNA-damage-induced binding of the NER proteins is not influenced exclusively by the DNA damage load, but can also be regulated by post-translational modifications or chromatin remodelers [42, 44-49].

While NER is the only mechanism to repair UV-induced DNA damage in placental mammals, an alternative damage removal mechanism known as photo-reactivation (PR) remained preserved through evolution in other branches of life, ranging from bacteria to non-placental mammals, [50, 51]. In contrast to NER-mediated repair, which is a complex mechanism that requires the activity of at least 30 proteins [3], PR is the direct reversal of CPD or 6-4PP lesions by one single damage specific photolyase (PL). PLs recognize the helix distortions created by CPD and 6-4PPs and bind to them through moderately strong ionic interactions. These interactions further destabilize the distorted DNA helix and lead to a flipping out of the DNA lesion into the active site of the PL, forming a highly stable complex [52, 53]. In contrast to the binding of PLs to DNA lesions, which is independent of light, the catalytic reversal of pyrimidine dimers to the original bases requires the absorption of a photon. Catalysis by PLs is achieved by light-initiated cycloconversion of the cyclobutane ring joining the two pyrimidines, which encompasses first the adsorption of a 333-500 nm photon by the chromophore MTHF, second the energy transfer from the blue light photon to the Flavin cofactor (FADH-), and third the electron transfer from FADH- to the cyclobutane ring, which splits the pyrimidine dimer and forms

a flavin radical (FADH·). The catalytic cycle is completed when the electron is transferred back to the cofactor, restoring catalytically active, fully reduced FADH- [53-55]. The entire reaction takes ~1 ns for both types of PLs [53]. The repair-independent binding of PLs to CPDs or 6-4PPs and their very fast damage removal makes PLs an attractive tool to study UV-induced damage and its repair. However thus far, PLs have mainly been used to test the specific cellular responses to either CPD or 6-4PP after removing the other type of lesion by PR [56-59].

Here we show, that fluorescently labelled PLs provide a versatile and sensitive tool to locate, quantify and repair UV-induced DNA damage in real time in living cells. Fluorescence recovery after photobleaching (FRAP)-based mobility studies of PLs allow to quantitatively determine DNA damage load, as well as repair kinetics. Furthermore, we show that PLs can be activated by the 405 nm laser light during live cell imaging experiments to photo reactivate DNA damage, which facilitates studying the behavior of NER factors and the DNA damage response upon DNA repair in living cells.

MATERIAL AND METHODS
Cell Lines and Constructs
VH10 hTERT immortalized human fibroblasts, XP4PA SV40 immortalized XP-C fibroblasts and HCT116 human colon cancer cells were cultured in DMEM/F10 and RPMI/F10 media, respectively, containing 10% FCS and 1% penicillin-streptomycin in a humidified incubator at 37°C and 5% CO_2. To generate a lentiviral 6-4PP-PL-mCherry-3xNLS-HA expression vector, *Arabidopsis thaliana* 6-4PP-PL cDNA [60], missing the first 57 nucleotides corresponding to a mitochondrial localization signal, was first cloned into pENTR/D-TOPO vector (Invitrogen). mCherry-3xNLS-HA was ordered as gBlocks Gene Fragment (Integrated DNA technologies) and ligated to the C-terminal of 6-4PP-PL cDNA in the pENTR/D-TOPO vector, using AscI and EcoRI. Then 6-4PP-PL-mCherry-3xNLS-HA was cloned into pLenti CMV Puro DEST using Gateway cloning (Invitrogen). CPD-PL-mCherry [48], or 6-4PP-PL-mCherry-3xNLS-HA expressing lentiviral vectors were used to make the corresponding lentiviruses using the third generation system [61]. GFP-DDB2 expressing VH10 cells [44], or GFP-XPC expressing HCT116 cells were transduced with the generated lentiviruses and cells stably expressing CPD-PL-mCherry or 6-4PP-PL-mCherry-3xNLS-HA were selected with puromycin.

GFP-XPC expressing HCT116 cells were generated by a CRISPR/Cas9 mediated knock-in strategy where HCT116 cells were co-transfected with a lentiCRISPR v2 vector containing an XPC guide RNA (5'-GCTCGGAAACGCGCGGCCGG-3') targeting right after the XPC start codon and, a linearized homology-directed repair (HDR) template. GFP-XPC DDB2-/- HCT116 cells were generated by transfection of GFP-XPC HCT116 cells with a lentiCRISPR v2 vector containing a DDB2 guide RNA (5'-TATTACGCCCCAGGAACAAG-3'). The HDR template to generate a GFP-XPC knock-in was generated in a single PCR step using 200 bp primers. The primers were designed in

a way that 30 bp of each primer anneals to the FLAG-GFP construct and the remaining 170 bp anneals to the human genomic XPC sequence. Furthermore, the PAM sequence was mutated by 5 silent mutations which were introduced in the region targeted by the XPC gRNA to prevent Cas9 cutting the integrated HDR template. The following primers were used to generate the HDR template: forward primer (5'CCGCAGTTTTTTAGTG GCCACGGGTATGGGGTGGAGCTTCCTTTAGGGGCGTGACTAGGCCTCCAACGA AGGGGCGTGGCCAAGCGCACCGCCTCGGGGCGGGGCCGGCGTTCTAGCG CATCGCGGCCGGGTGCGTCACTCGCGAAGTGGAATTTGCCCAGACAAGCA ACATGGACTACAAGGACGACGATGACAAGGTG-3'), reverse primer (5'-GCCTCTGGGCC TCCTCCGCCCACCGGCGGCGTCTCCCGCGAAGCCCGCTGGGCCTCGCTCTCACCCTCC TCCTCCTCCTCACGCCGGGCCTTGCTCTTGGCCTTGGATTTCTGGCTGCGCAGTTCG CGTCCCCGCGGCTCCCCGCCTGCGGCTCTCTTCCGAGCGAGATGCTTGTACAGCT CGTCCATGCCGAGAGTGAT-3'). The PCR generated template was cloned into pCR-Blunt II-TOPO vector and then the vector was digested with EcoRI to generate the linearized GFP-XPC construct.

Transfected HCT116 cells were selected by puromycin for 2 days and stable GFP-XPC expressing cells were FACS sorted. Then single cell clones were picked and clones were selected using genotyping, and western blotting to check for expression of the full-length GFP-XPC protein and the concomitant loss of wild type XPC expression.

RNA interference

Cells were transfected with the indicated siRNAs (150 pmol) using RNAiMax (Thermo Fisher Scientific) 48-72h prior to the experiment, according to manufacturer's protocol. The siRNAs were purchased from Dharmacon: control (siGENOME Non_Targeting siRNA#5, D-001210-05) and XPF (siGENOME ERCC4 siRNA, M-019946-00).

Inflication of UV-induced DNA damage

Cells were washed with PBS, and after PBS removal the cells were exposed to UV-C light from a 254 nm germicidal lamp (Philips). Local UV-C damage was inflicted through an isopore membrane filter (Millipore) with a pore size of 5 µm [62].

Photo-reactivation

After PBS wash, cells were covered with a thin layer of HBSS (Thermofisher) and then placed at a distance of 10 cm under white-light tubes (General Electric Lightning Polylux LX F36W/840) for 10 minutes at 37°C. Mock-treated samples were covered with aluminum foil during photo-reactivation (PR).

Western blotting

Cells were lysed in 2x sample buffer and boiled for 10 minutes at 95°C. The proteins were subsequently separated by SDS-PAGE and transferred to PVDF membranes (0.45 µm).

Membranes were blocked with 5% BSA in PBS-T (PBS containing 0.05% Tween 20) for 1 hour at room temperature (RT) and blotted with the following primary antibodies: CPD-PL and 6-4PP-PL (rabbit polyclonal, 1:500) [58, 59], RFP mCherry (rat monoclonal, 1:1000, 5F8, Chromotek), DDB2 (rabbit monoclonal, 1:1000, EPR981, abcam), Ku70 (goat polyclonal, 1:1000, M-19, sc-1487, Santa Cruz), tubulin (mouse monoclonal, 1:3000, B-5-1-2, sc-23948, Santa Cruz), XPC (rabbit polyclonal, 1:1000, A301-122A, Bethyl) or XPF (mouse monoclonal, 1:500, 3F2/3, sc-136153, Santa Cruz). After five times washing with PBS-T, the membranes were blotted with the following corresponding secondary antibodies from Sigma Aldrich: CF™ 680 Goat anti-Rabbit IgG (1:5000) and CF™ 770 Goat anti-Mouse IgG (1:5000).. The blots were imaged with the Odyssey CLx Infrared Imaging System (LI-COR Biosciences).

Immunofluorescence

Cells were grown on 24 mm coverslips and fixed in 2% paraformaldehyde containing PBS Triton X-100 (0.1%). After five times washing with PBS Triton X-100, the coverslips were blocked in PBS+ (PBS containing 0.15% glycine and 0.5% BSA). A denaturation step of 5 minutes using freshly diluted NaOH (0.07M) in PBS was performed to make DNA lesions accessible for the CPD (mouse monoclonal, 1:1000, TDM-2, Cosmo Bio) or 6-4PP (mouse monoclonal, 1:300, 64M2, Cosmo Bio) primary antibodies. Following an incubation of 1-2h at RT with primary antibodies diluted in PBS+, the coverslips were washed with PBS Triton X-100 five times and PBS+ once. Then the coverslips were incubated with 488, 555 or 639 Alexa Fluor secondary antibody conjugates (Invitrogen) diluted in PBS+ for 1h at RT. After the coverslips were washed again as described above, they were embedded in Vectashield Mounting Medium with DAPI (Vector Laboratories). The coverslips were imaged using a LSM 700 microscope equipped with a Plan-Apochromat 40x/1.3 NA oil immersion lens (Carl Zeiss MicroImaging Inc.). The ImageJ software [63] was used to quantify the CPD and 6-4PP signals in the generated images. The DAPI signal was used to determine the nuclei and the mean fluorescence intensities measured in the nuclei were used to plot the graphs. For CPD and 6-4PP removal assay cells were globally UV irradiated with 10J/m^2 and 16 10J/m^2 respectively, and fixed after the indicated time points. CPD and 6-4PP staining was performed as described above. Fluorescence levels were quantified in at least 70 cells per sample by measuring the background-corrected overall nuclear fluorescence, which was set at 100% for 0h after UV irradiation.**Colony survival assay**

Cells were seeded in triplicate in 6-well plates (300 cells/well) and next day treated with the indicated UV-C doses. After a week, the colonies were fixed and stained with 0.1% Brilliant Blue R (Sigma), and counted using GelCount (Oxford Optronix Ltd.).

Live cell confocal laser-scanning microscopy

All live cell imaging experiments were performed at 37°C and 5% CO$_2$ using a Leica SP5 laser-scanning confocal microscope with a 63x/1.4 NA HCX PL APO CS oil immersion

objective. Fluorescence recovery after photobleaching (FRAP) experiments were performed as described previously [42], in short; a narrow strip (512 x 16 pixels at zoom 9) along the nucleus was bleached 94 ms with 100% power of 488 nm laser for the GFP and 42 ms with 100% power of 561 nm laser for the mCherry signal. The signal in this strip was measured pre-bleach for 3.6 s and post-bleach for 20 s every 400 ms with 0.2% power of the 488 nm laser for GFP-XPC. FRAP of the mCherry-tagged PLs was performed by measuring pre-bleach for 2.5 s and post-bleach for 20 s every 100 ms with 3% power of the 561 nm laser for the mCherry signal of the PLs. To analyze fluorescence recovery, measured fluorescence intensities were first background corrected, then normalized to the average pre-bleach fluorescence signal which was set at 1. Immobile fractions were calculated using the following formula: Immobile fraction (%) = 1 – ((average fluorescence intensity of UV-C irradiated cells – the first post-bleach data point) / (average fluorescence intensity of mock-treated cells–the first post-bleach data point)). The average fluorescence intensities are calculated over the measurements of the last 10 s. For local repair during live cell imaging experiments, the fluorescence intensity of PL-mCherry was monitored every 2.585 s, both inside and outside the local damage within the nucleus (at zoom 10). PR was performed by exposure of the DNA damage to 5 frames of 5% 405 nm laser light. The power output of the 405 nm laser was measured to be 0.063 mW at 10% laser power. Data were corrected for background fluorescence signal outside the cell and normalized to average fluorescence signal at the local damage before PR, which was set at 1.

RESULTS

Generation and characterization of CPD-PL and 6-4PP-PL-expressing cells

To develop a method to quantify UV-induced DNA damage and its repair kinetics directly in living cells, we first tested whether the ability of PLs (PLs) to specifically detect UV-induced lesions could be exploited to generate live cell damage markers by fluorescently labeling them. For this purpose, we generated lentiviral vectors [64] expressing *Potorous tridactylis* CPD-PL [58] or *Arabidopsis thaliana* 6-4PP-PL [65] tagged with mCherry fluorescent protein at their C-terminus. In addition, 3 NLS sequences were added after the mCherry-tag of the 6-4PP PL to ensure nuclear expression. These lentiviruses were used to transduce GFP-DDB2 expressing VH10 (hTERT immortalized human fibroblast) cells [44] to stably express either CPD PL-mCherry or 6-4PP PL-mCherry (referred as CPD-PL or 6-4PP-PL, respectively). Western blot analysis showed that the generated VH10 cell lines express full-length PL-mCherry fusion proteins (Figure 1A). To be able to use these PL-mCherry proteins as UV damage markers, it is important that PL expression does not interfere with NER-mediated repair of UV-induced lesions. As shown by UV colony survival experiments, both CPD-PL and 6-4PP-PL-expressing cells showed a similar UV sensitivity as wild type (WT) VH10 cells (Figure 1B), indicating that the expression of these fusion proteins does not affect endogenous DNA repair activity. To corroborate this, we
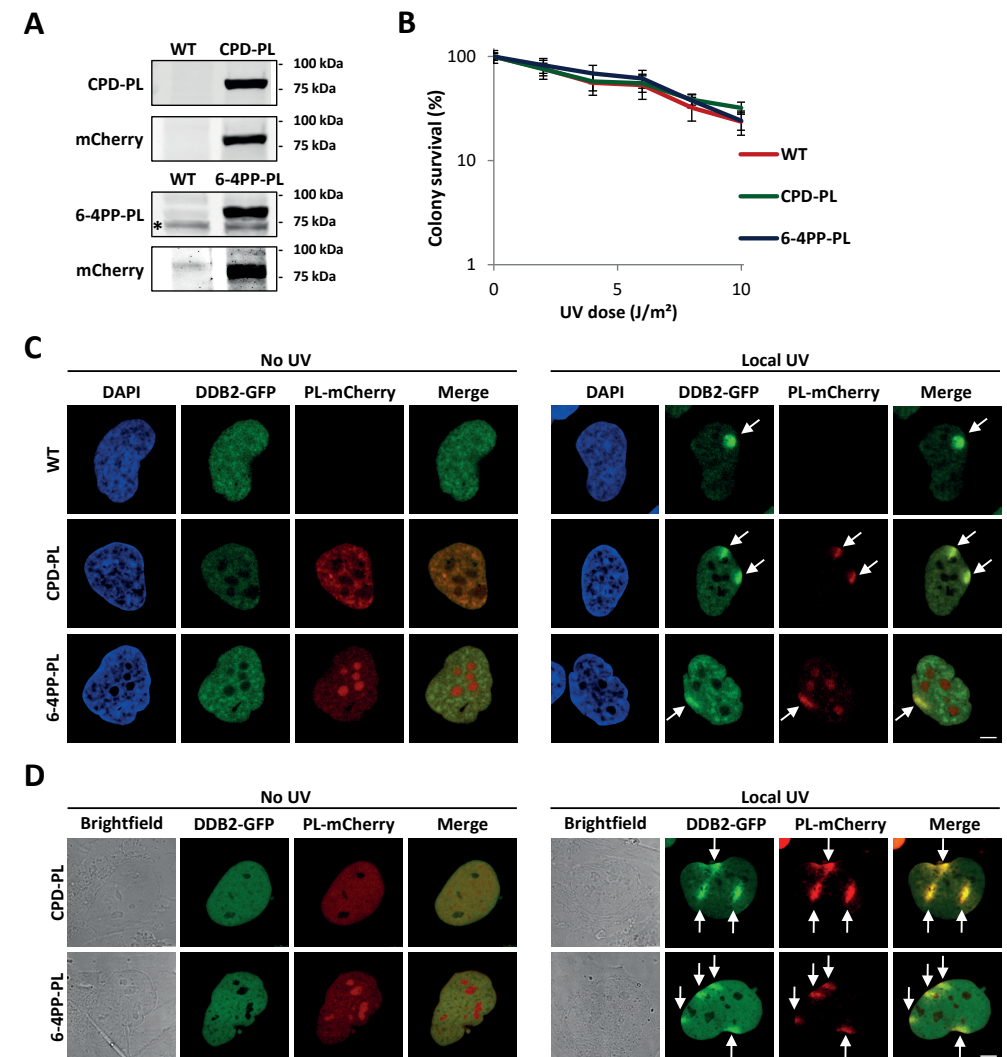

Figure 1. Characterization of mCherry-tagged photolyase-expressing cells. (A) Western blot of lysates of VH10 cells stably expressing GFP-DDB2 (WT), or co-expressing GFP-DDB2 and either CPD-PL-mCherry (CPD-PL, upper panel, expected size 75kDa) or 6-4PP-PL-mCherry (6-4PP-PL, lower panel, expected size 85kDa). Blots were stained with the indicated antibodies. Relevant marker sizes are indicated and * indicates an unspecific band. (B) UV-C sensitivity of WT or PL-expressing VH10 cells, determined by colony-forming ability (mean ± SEM). Percentage of surviving cells is plotted against the applied UV-C dose, colony number at 0 J/m^2 is set at 100 %. (C and D) Representative images of WT and PL-expressing VH10 cells. Cells were either non-irradiated (no UV, left panel) or locally irradiated with 60 J/m^2 UV-C (Local UV, right panel). Cells were either fixed directly after DNA damage induction (C) or monitored directly by live cell imaging (D). Arrows indicate local UV damage. Scale bar: 5 μm.

compared the kinetics of 6-4PP and CPD removal in PL-expressing cells and WT VH10 cells (Supplementary Figure S1A and B). This revealed that the endogenous repair of CPDs and 6-4PPs by NER was similar in PL-expressing cells and in wild type cells, but was strongly diminished in NER deficient XP-C cells.

The fusion to a mCherry tag allowed direct visualization of PLs and showed that both CPD-PL and 6-4PP-PL are expressed mainly in the nucleus (left panels of Figure 1C and D). While the CPD-PL is excluded from the nucleoli, the 6-4PP-PL was enriched in the nucleoli (bottom panels of Figure 1C and 1D), however, the mechanism behind this different nucleolar localization is unknown. As PLs bind CPDs and 6-4PPs light independently, but need white light to initiate catalysis, we subsequently tested whether the PLs were capable of binding to UV-induced DNA damage, while cells were kept in the dark. Both CPD-PL and 6-4PP-PL accumulated at local UV damage induced through micropore filters [62], as shown by a co-localization with the DNA damage recognizing protein DDB2 (Figure 1C, right panel). Of note, the exogenous expression of the PLs did not block DDB2 recruitment to sites of DNA damage. Furthermore, both CPD-PL and 6-4PP-PL co-localized with the respective lesion-specific antibodies (Supplementary Figure S1C). Importantly, CPD-PL and 6-4PP-PL were efficiently recruited to locally induced DNA damage in living cells (Figure 1D, right panel), demonstrating that PLs can be used to directly detect UV-induced CPD and 6-4PP lesions in living cells, which is not possible with photo lesion-specific antibodies.

CPD-PL and 6-4PP-PL as quantitative, real-time, damage and repair markers in living cells

Both CPD-PL and 6-4PP-PL were able to precisely detect the UV-induced DNA damage without interfering with NER activity (Figure 1B-D, Supplementary Figures S1A and B). Binding of repair proteins to DNA damage generally immobilizes them on chromatin, which can be quantified by fluorescence recovery after photo-bleaching (FRAP) [43, 66]. Therefore, we performed FRAP experiments to quantitatively assess differences in the chromatin-bound fraction of PLs in response to different UV doses. FRAP of PLs showed that both CPD-PL and 6-4PP-PL are highly mobile in unperturbed cells (no UV) (Figure 2A and B), indicating that PLs are not stably bound to chromatin in the absence of DNA damage. Interestingly, both CPD-PL and 6-4PP-PL were immobilized in a dose-dependent manner after UV irradiation (Figure 2A and B). From these FRAP curves, we determined the immobile fractions of the PLs (Supplementary Figure S2A and B), which revealed a linear increase for both CPD-PL and 6-4PP-PL with increasing UV doses up to 10 J/m². To assess whether the PL immobilization correlates with the actual quantity of CPDs and 6-4PPs, we quantified the relative amount of CPDs and 6-4PPs induced at these UV doses by immunofluorescence using photo lesion-specific antibodies [32] (Supplementary Figure S2C-E). Importantly, this revealed that the PL immobilization determined by PL FRAP (Figure 2C and D, primary Y-axis) at the indicated UV doses correlates very well with antibody-detected CPD and 6-4PP damage loads (Figure 2C and D, secondary Y-axis). This
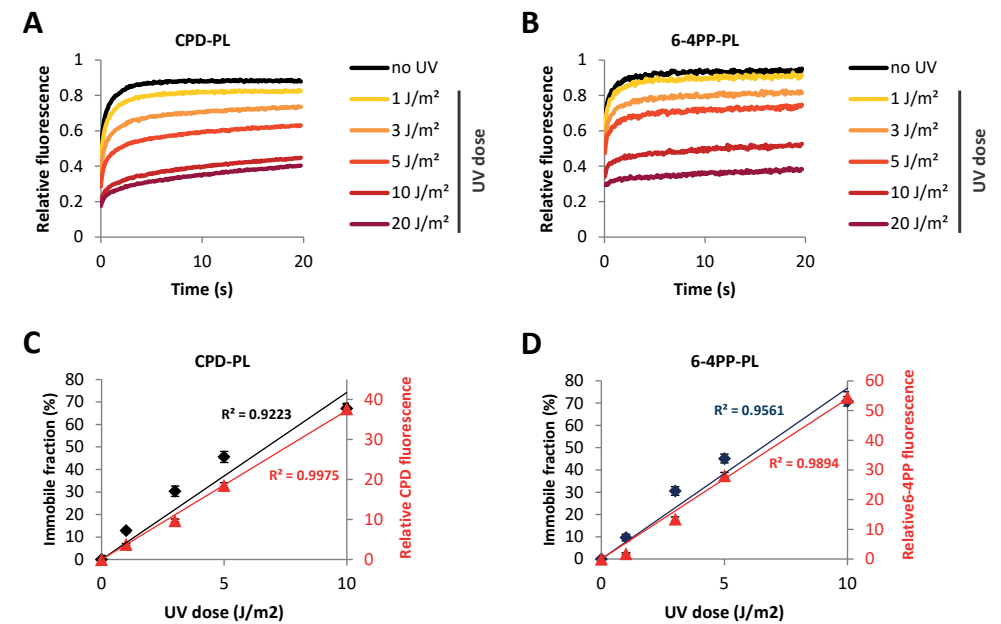

Figure 2. mCherry-tagged PLs as quantitative and real-time damage markers in living cells. (A and B) UV dose-dependent immobilization of CPD-PL (**A**) and 6-4PP-PL (**B**). PL-expressing VH10 cells were non-irradiated or global UV-irradiated with the indicated UV doses and were directly analyzed by FRAP. The plotted values were normalized over the average pre-bleach signal (n=20 cells from 2 independent experiments). (C and D) Immobile fractions of CPD-PL (**C**) and 6-4PP-PL (**D**) at the indicated UV-C doses were plotted together with the relative quantity of CPD and 6-4PP lesions at the same UV-C doses as determined by the mean fluorescence intensities in immunofluorescence assays using lesion-specific antibodies.

shows that FRAP of PLs allows a direct, relative quantification of UV-induced DNA damage in living cells. Interestingly, above 10 J/m², hardly any increase in PL immobilization was observed. This is most likely caused by limiting amounts of non-bound PLs at higher UV doses, in line with the almost complete immobilization of PLs at 10J/m² (Figure 2A and B). This may indicate that PL expression levels influence UV-induced PL immobilization. To test this, we compared the UV-induced PL immobilization in cells with low and high PL expression levels. This revealed that PL expression levels determine the dynamic range of PL mobility (Supplementary Figure S2F and G). Cells with low PL expression levels showed an increased immobilization at lower UV doses (e.g. 1 and 3 J/m²). However, this dose-dependent increase in immobilization levelled off around 5J/m² (Supplementary Figure S2F and G, left panels). In contrast, cells with high PL expression showed a reduced immobilization at low UV doses, but PL immobilization continued to increase at high damage loads (e.g. 10-20 J/m²) (Supplementary Figure S2F and G, right panels). These experiments demonstrate the importance of using cells with similar PL expression levels to avoid variation due to differences in the dynamic range of PL immobilization.

To test whether FRAP of PLs can also be used to study live-cell repair kinetics of CPDs and 6-4PPs, we UV irradiated cells and determined PL immobilization in time (Figure 3, Supplementary Figure S3A and B). In line with the previous experiments, both CPD-PL and 6-4PP-PL were strongly immobilized immediately upon UV exposure (10 J/m²). As expected, this immobilization decreased over time, reflecting the repair of CPD and 6-4PP lesions. While 6-4PP-PL was quickly mobilized, with a 50% reduction at 7h post UV and an almost complete mobilization at 24h post UV, the reduction in binding to damaged DNA by CPD-PL was much slower, in line with previously shown differences in repair rates of CPD and 6-4PP lesions [42, 44, 67]. The mobilization of PLs over time was almost completely blocked by siRNA-mediated depletion of the NER factor XPF (Figure 3 lower panel, Supplementary Figure S3A and B, lower panel and Supplementary Figure S3C), indicating that the mobilization of PLs in time represented repair of CPD and 6-4PP by NER. These results demonstrate that FRAP of PLs enables the real-time monitoring of DNA damage load in living cells and thus provides a sensitive method to detect perturbations of the NER reaction in living cells.

Lesion-specific repair of UV-induced DNA damage in living cells

In addition to their use as quantitative live-cell damage markers, PLs can be used to specifically remove either CPD or 6-4PP lesions by direct reversal of the DNA damage using energy from near UV light (300–500 nm) [68, 69]. First, we determined optimal PR times for the repair of CPD and 6-4PP lesions (Figure 4A and B). PL-expressing cells were UV-irradiated and DNA damage was photo-reactivated with white light for the indicated times. While 5 min of PR was not enough for complete removal of DNA damages, 10 min PR resulted in a PL mobility similar to that of non-irradiated cells, indicative of an almost complete removal of DNA lesions (Figure 4A and B, Supplementary Figure S4A and B). Of note, the PL mobility was not affected in cells that were shielded from the white light during PR (UV + 10min mock).

Having determined the optimal PR conditions, we assessed the previously described PR specificity of each PL. For this purpose, UV-induced DNA damage was photo-reactivated and the CPD or 6-4PP lesions were detected using immunofluorescence with specific antibodies. As expected, we observed an almost complete loss of CPDs following PR in CPD-PL-expressing cells, while the quantity of 6-4PP lesions was not affected (Supplementary Figure S4C). In 6-4PP-PL-expressing cells, removal of only 6-4PP lesions, but not of CPD, was observed upon PR (Supplementary Figure S4C).

After confirming that PL-expressing cells can specifically repair CPD or 6-4PP lesions, we made use of this feature to study live-cell DNA binding kinetics of XPC, the main damage sensor in GG-NER [4]. For this purpose, we co-expressed GFP-XPC and CPD-PL or 6-4PP-PL in HCT116 cells (Supplementary Figure S4D, left panel), and performed FRAP experiments to simultaneously asses the mobility of mCherry-tagged PLs and GFP-tagged XPC (Figure 4C and D). As shown in the FRAP curves (Figure 4C and D) and
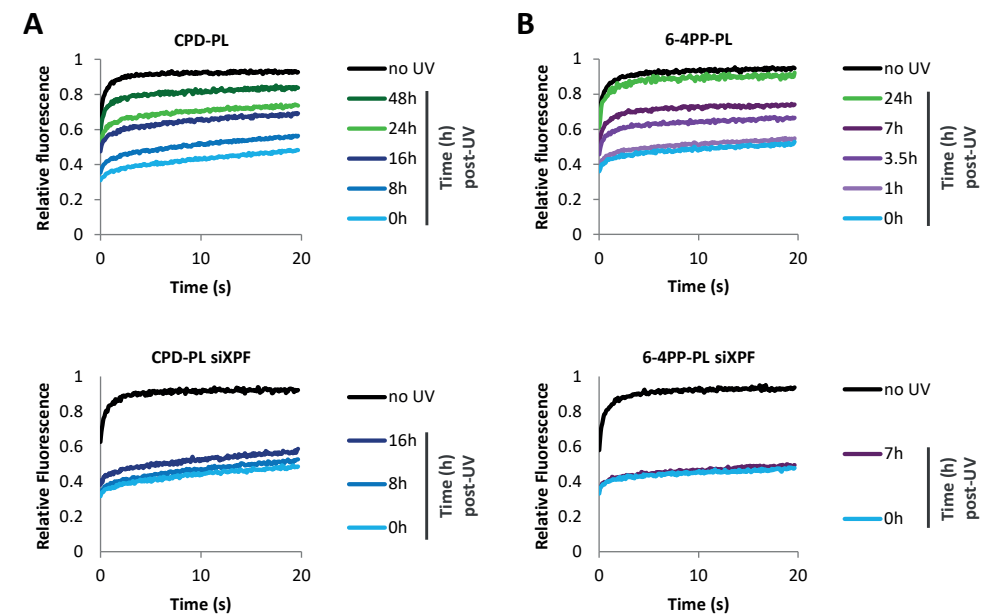

Figure 3. mCherry-tagged PLs to monitor DNA repair in living cells. (**A**) CPD-PL and (**B**) 6-4PP-PL expressing VH10 cells were transfected with the control (upper panel) or XPF siRNAs (siXPF) (lower panel). The plotted PL mobilities in non-irradiated or globally UV-C irradiated (10 J/m²) cells were determined by FRAP at the indicated time points post UV irradiation. (n ≥ 25 cells from 2 independent experiments for control siRNA experiment, n ≥ 15 for siXPF).

the respective immobile fractions (Supplementary Figure S4E and F), UV irradiation led to the binding of GFP-XPC to damaged DNA resulting in its immobilization. Upon PR of each type of photo lesion, GFP-XPC immobilization was reduced, however not to the same extent as in non-irradiated cells. This is most likely due to the fact that XPC has affinity for both CPD and 6-4PP lesions [4, 70, 71]. PR of 6-4PPs and CPDs was successful as shown by the mobilization of both PLs upon PR (Supplementary Figure S4G and H). We observed more increase in the GFP-XPC mobilization upon PR of 6-4PPs compared to CPDs. Although XPC is able to directly recognize 6-4PP lesions, DDB2 facilitates this recognition and is crucial for XPC to detect CPD lesions [5, 6, 72]. This suggests that the residual damage binding of XPC, following 6-4PP removal, represents DDB2-mediated binding to CPDs. To test this, we performed the same FRAP experiments in DDB2-deficient cells (Supplementary Figure S4D, right panel). In line with a stimulatory effect of DDB2 on XPC damage recognition, the UV-induced XPC immobilization was reduced by approximately 50% in DDB2 deficient cells (Figure 4E and F, Supplementary Figure S4E and F). Furthermore, in the absence of DDB2, GFP-XPC immobilization fully recovered after PR of 6-4PPs (Figure 4F and Supplementary Figure S4J), indicating that the UV-induced immobilization of XPC in DDB2-deficient cells is caused solely by 6-4PPs. In contrast, PR of CPDs (Supplementary Figure S4I) did not affect the XPC immobilization
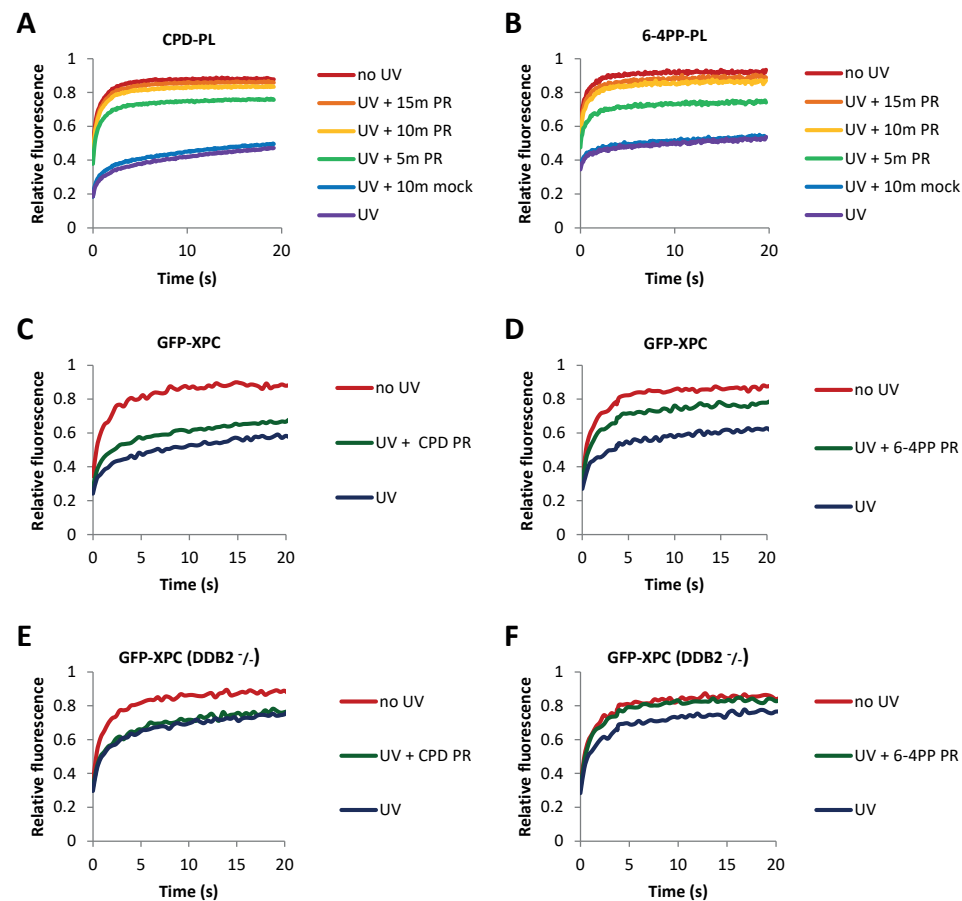
Figure 4. Lesion-specific repair of UV-induced DNA damage in living cells. (A and B) Mobility of (A) CPD-PL and (B) 6-4PP-PL as determined by FRAP analysis. Cells were non-irradiated (no UV), globally UV-C irradiated with 10 J/m² (UV), or globally UV-irradiated with 10 J/m² and photo-reactivated (UV + PR) for the indicated times by exposure to white light at 37°C. Cells were directly analyzed by FRAP after each treatment. ''UV + 10min mock'' cells were UV-irradiated and mock photo-reactivated by shielding from white light during PR. (n ≥ 20 cells from 2 independent experiments) (C and D) GFP-XPC mobility was determined by FRAP in non-irradiated, globally UV-C irradiated (10 J/m²), or globally UV-C irradiated (10 J/m²) and photo-reactivated (10 min PR) in CPD-PL (C) or 6-4PP-PL (D) expressing cells. (E and F) Similarly, GFP-XPC mobility analysis was performed in DDB2 deficient cells (GFP-XPC DDB2-/-) (n ≥ 20 cells from 2 independent experiments).

(Figure 4E), which confirms that XPC does not bind CPDs in DDB2-deficient cells. These experiments illustrate firstly, that the PL-mediated removal of specific UV-induced lesions can provide important quantitative insights into the behavior of NER factors on specific types of DNA lesions. Secondly, the direct comparison of PL and XPC mobility by FRAP in the same cell following the same UV exposure illustrated that fluorescently-labelled PLs can quantify DNA damage with a bigger dynamic range than XPC, as shown by

the bigger immobile fraction of PLs (Supplementary Figure S4K). Thirdly, the mobility of PLs was not affected by the presence or absence of DDB2, which like PLs directly binds to DNA lesions [3] (compare Supplementary Figure S4G and I, and Supplementary Figure S4H and J). This shows that the dynamic range of PLs as live cell damage markers is not influenced by competitive substrate binding of PLs and DDB2.

Local repair of UV-induced DNA damage in living cells

A limitation of the PR-based DNA damage removal described above is that full PR takes ~10 minutes and needs to be performed before live-cell imaging. This interferes with the real-time measurement of the effects of DNA damage removal on proteins of interest in the cells. To improve our system, we set out to perform PR during live-cell imaging. As *Potorous tridactylis* CPD-PL and *Arabidopsis thaliana* 6-4PP-PL have absorption spectra that peak between 360 nm and 450 nm [65, 73], we tested whether it was possible to remove UV-induced damage by activating the PLs using a 405 nm laser during live cell imaging. CPD-PL-expressing cells were locally UV-irradiated resulting in accumulation of CPD-PL at sites of DNA damage (Figure 5A, upper panel). These locally accumulated CPD-PLs were subsequently exposed to different intensities of the 405 nm laser, which almost instantaneously released the damage-accumulated PLs already at 0.5 % 405 nm laser power (Figure 5A and B), reaching complete PR at 1 % laser power. To exclude that the loss of fluorescence at the damage site was caused by photo bleaching of mCherry, the 405 nm laser was also activated at a region outside the damage within the nucleus, which did not result in any reduction in signal intensity (Supplementary Figure S5A). To further confirm that PL exposure to the 405nm laser induced CPD removal by PR, we first photo-reactivated damaged DNA in a specific region (marked with the cross) in the nucleus and then stained the cells with a CPD specific antibody (Figure 5C). CPD lesions within the marked area were completely removed. In line with this, the mCherry signal of CPD-PL was reduced in the 405 nm laser-exposed region. This can be explained by its release and its subsequent binding to the areas in the nucleus where the damage is not removed. Additionally, 6-4PP lesions could also be removed upon PR by 6-4PP-PL, however, this required slightly higher 405 nm laser intensities (>5%) (Figure 5D and 4E, and Supplementary Figure S5B).

Importantly, this live-cell PR is compatible with GFP imaging, as the PR-based repair is hardly triggered by the 488 nm laser at intensities that are commonly used for imaging GFP-tagged factors (Supplementary Figure S5C-H). Altogether, these results show that PLs can be used to photo reactivate UV-induced DNA damage in real-time in living cells, using the 405 nm laser. In conclusion, while the induction of DNA damage in living cells has been an available tool for many years [74] and resulted in many important mechanistic insights in the repair reaction, in this study, we introduce the repair of specific UV-induced DNA damage in living as a unique tool to study the dissociation of DNA repair factors and behavior of other cellular processes upon damage removal.
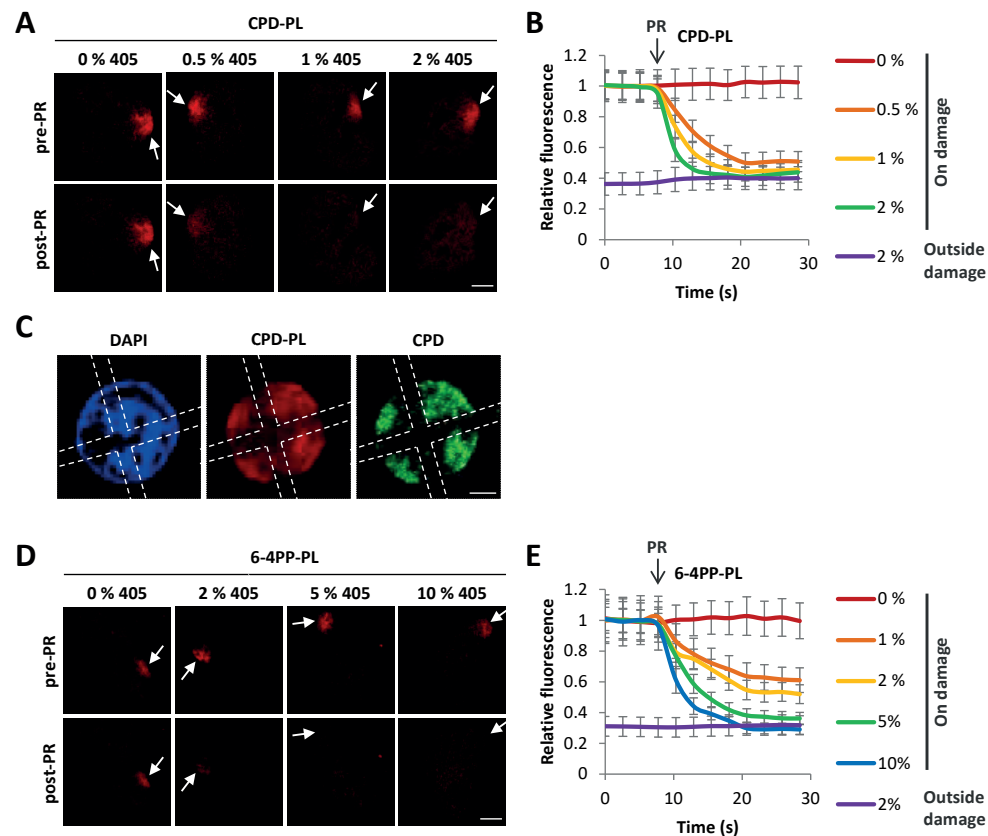
Figure 5. Local repair of UV-induced DNA damage in living cells. (A) Representative images of CPD-PL-expressing VH10 cells before and 13 s after PR (pre-PR and post-PR, respectively) using 405 nm laser at the indicated intensities. Arrows indicate local UV damage. Cells were locally UV-irradiated (60 J/m²), the local DNA damage spot and a region of the exact same size outside the damage within the nucleus were exposed after 7.8 s (indicated by arrow and PR) to the indicated intensities of 405 nm laser for 13 s. Scale bar: 5 μm. (B) Relative mCherry fluorescence signal of CPD-PL was quantified inside and outside the DNA damage within the nucleus and normalized to pre-PR intensities at the local damage. (n = 8 cells, mean ± SEM) (C) Representative immunofluorescence images of CPD-PL-expressing VH10 cells after PR. Global UV-irradiation (10 J/m²) of the cells was followed by photo-reactivation of the damaged DNA by 5% 405 nm laser for 13 s in a specific region (region marked with dotted line) in the nucleus and the cells were subsequently fixed and stained with CPD antibody using immunofluorescence. (D) Representative images of locally UV-irradiated (60 J/m²) 6-4PP-PL-expressing VH10 cells before and 13 s after PR (pre-PR and post-PR, respectively) using 405 nm laser at the indicated intensities as described above in Figure 4A. Arrows are indicating the local UV-C damage spots. Scale bar: 5 μm. (E) Relative mCherry fluorescence signal of 6-4PP-PL was quantified inside and outside the DNA damage within the nucleus and normalized to pre-PR intensities at the local damage. (n = 8 cells, mean ± SEM)

DISCUSSION

The currently available assays to investigate UV-induced DNA damage and repair have proven to be invaluable tools to study NER factors in both fundamental and clinical research. However, these assays cannot be performed in living cells, and are therefore confined to endpoint measurements instead of monitoring the DNA damage quantities in real time. Therefore, in this study, we developed a novel method using fluorescently-tagged PLs to directly recognize and quantify UV-induced DNA damage in a highly sensitive manner in living cells.

For this purpose, we made use of the high affinity of PLs for UV-induced DNA damage, which was confirmed by their accumulation at locally induced UV damage (Figure 1C and D), and their immobilization on damaged DNA during FRAP. Both CPD-PL and 6-4PP-PL showed a strong and reproducible UV dose-dependent immobilization. This approach allowed to quantitatively monitor the relative DNA damage loads (Figure 2A-D) and NER-mediated repair kinetics in a highly sensitive manner (Figure 3). FRAP of PLs proved to be highly sensitive and enabled the detection of physiological relevant damage loads as low as 1 J/m², which are difficult to quantify with other techniques. Furthermore, we observed a linear and relatively large dynamic range of PL immobilization between 0-10 J/m² UV, enabling precise quantification of the DNA damage loads. Of note, at higher UV doses (20 J/m²) the dose-dependent immobilization was not linear anymore, which might be caused by limiting amounts of non-chromatin bound PLs, in line with the almost complete immobilization of PLs at 10J/m² (Figure 2A and B).

In line with this, in cells with higher PL expression levels this levelling off of PL immobilization at higher UV doses was reduced, indicating that cells with higher PL levels are more suitable to quantify high damage loads (>5J/m2) (Supplementary Figure S2F and G, right graphs). On the other hand, our experiments show that cells with low PL expression levels allow a more sensitive detection of low damage loads (<5J/m2) (Supplementary Figure S2F and G, left graphs). Together these data show that the dynamic range of FRAP-based UV damage detection using fluorescently-tagged PLs can be adjusted to experimental needs by choosing cells with distinct PL expression levels. PL expression levels can easily be fine-tuned in the used lentiviral transduction system by the choice of promoter [64]. These data furthermore show the importance of using cells with similar PL expression levels when studying PL kinetics in different conditions. To achieve a very homogenous expression of fluorescently-tagged PLs, thereby potentially even increasing the precision of PL-mediated damage quantification, CRISPR/Cas9-mediated genomic targeting of PL expression cassettes to safe harbor loci like ROSA26 or AAVS1 [75] could be used.

Furthermore, mutated PLs that are still capable of binding the UV-induced DNA damage, but are incapable of PR, might be developed, as these catalytically dead PLs will be insensitive to unintentional day light exposure during experimental handling. Of note, the FRAP-based PL assay is already very sensitive, as shown by the direct comparison

of the PL immobilization to that of XPC (Supplementary Figure S4K), the main damage sensor of GG-NER [4]. XPC is one of NER factors that shows the highest immobilization on UV-damaged DNA [42, 76, 77], however our FRAP data show that the fluorescently labeled PLs detect damage with even higher sensitivity than XPC.

The precise correlation between PL immobilization and DNA damage load is most likely explained by the fact that exogenously expressed PLs, which function as single proteins, are most likely not regulated by the activity of other proteins, post-translational modifications or other forms of regulation in mammalian cells. Especially, these types of regulation have been shown to influence the direct correlation of the NER factor immobilization with the DNA damage quantity [17-19, 45, 47-49, 78]. The direct recognition of DNA damage with high affinity makes PLs ideally suited to visualize DNA damage in both living and fixed cells. Fluorescently-tagged PLs can therefore be used as sensitive, lesion-specific quantitative damage markers while studying the accumulation of other proteins at sites of local UV-induced damage [62, 74].

Of note, although the PLs have high affinity for DNA damage and are highly immobilized on damaged DNA, they did not interfere with NER. No differences were observed in the UV survival of parental and PL-expressing cells (Figure 1B). In line with this, PL expression did not block the accumulation of DDB2 on local UV damages (Figure 1C) or inhibit the NER-mediated repair, as shown by PL immobilization in time (Figure 3). The absence of interference with the NER reaction might be explained by a transient binding of the PLs to DNA damage, thereby allowing NER factors to access DNA lesions. In line with this, even though a large fraction of the PLs were immobilized following UV exposure, these PLs were not long-term immobilized on chromatin but were rather continuously released and rebound as evidenced by the continuous increase in the FRAP curves over time (Figure 2A and B). Overall, our data showed that the use of fluorescently-labeled PLs is a robust and sensitive new method for the direct detection and quantification of UV-induced DNA damage in living cells.

In addition to their use as sensitive DNA damage markers, the fluorescently-tagged PLs can also be used to revert the UV-induced damage by PR. In this case, the fluorescent tag could be used to monitor directly the DNA damage reversion by assessing PL immobilization. This feature allowed us to determine the minimally required PR times for CPD-PL and 6-4PP-PL (Figure 4A and B). Our results showed that 10-15 minutes of PR by exposure to white light was enough for both PLs to repair almost all of the lesions (Figure 4A and B). This minimal PR duration is much shorter than the PR times of 1-4 h used in most previous studies performed in mammalian cells [6, 79-83]. The lesion-specific repair by CPD-PL and 6-4PP-PLs can be used to investigate the lesion-specific behavior of NER factors [6, 84]. These fluorescently-tagged PLs can be also used in combination with NER factors with different fluorescent tags, to simultaneously study the dynamic behavior of NER factor of interest while confirming the successful PR of CPD or 6-4PP lesions by PLs in the same FRAP experiment. As a proof of principle, we compared the GFP-XPC kinetics

before and after PR, by simultaneously monitoring PR-mediated repair by PL mobility (Figure 4E and F). This approach could be applied to investigate the behavior of any fluorescently-tagged repair factor while monitoring the progress of either PR-based or endogenous repair over time in living cells.

Despite the short required PR times, white light mediated-PR is technically incompatible with live cell imaging. To study cellular processes directly upon, or even during PR we developed PR using a 405 nm laser as a new method to repair DNA lesions during live-cell imaging. As the CPD-PL and 6-4PP-PL absorption spectra peak between 360 nm and 450 nm, we photo-reactivates PLs using the 405 nm laser, which is commonly available in live cell imaging setups. Using laser pulses as short as 12.5 s, with relatively low laser power, allowed us to efficiently photo-reactivate DNA damage in live cell imaging experiments (Figure 5). Of note, while the 405 nm laser light can also be used to generate DNA damage, [85, 86], the laser intensity used for efficient PR is more than 10-fold lower than the intensities required to induce DNA damage [74]. Our results indicated that PR is rather specific for 405 nm laser, as PLs were not efficiently activated by the 488 nm laser at settings normally used for imaging GFP-tagged factors (Supplementary Figure S5C-H).

Thus far, mainly association kinetics (K_{on}) of NER factors were studied using local DNA damage infliction during live cell imaging [74]. This has been a powerful tool to study the accumulation kinetics and recruitment order of fluorescently-tagged NER factors and has revealed crucial information about molecular mechanism and interdependencies of NER factors [41, 87]. Thus far the dissociation kinetics (K_{off}) of NER factors following DNA repair were more difficult to address, as the endogenous NER-mediated repair is expected to happen in a stochastic manner over time. Interestingly, our 405 nm laser-assisted live-cell repair method enables to almost instantaneously remove DNA damage. This approach can be used to gain important insights into the release of NER factors, repair times and stability of NER intermediates. In addition, as the PLs are lesion-specific, these parameters could be specifically attributed to CPD or 6-4PP lesions. PR activity of the CPD and 6-4PP-PLs can also be used simultaneously to repair the vast majority of UV-induced lesions directly. This allows testing whether specific cellular effects are caused by the DNA damage itself, or by other types of damages generated by UV exposure, such as membrane, protein, or RNA damage [88-90]. Additionally, 405 nm laser-mediated PR can be used to instantly repair sub-nuclear regions, which could be used to determine the contribution of DNA damage (*in cis*) or signaling pathways (*in trans*) to transcription inhibition, replication stress, or other cellular effects following UV damage [91-93].

In conclusion, here we describe how fluorescently-labeled PLs can be used as highly sensitive UV-induced DNA damage markers to quantitatively determine damage load and repair in real-time, in living cells. Moreover, the instant repair of DNA damage by activating PLs during live cell imaging opens new possibilities to assess the cellular effects following damage removal. In addition, lentiviral expression is highly efficient to stably express PLs in a broad range of cell lines. Overall, the methods described here are a valuable

5

5

extension of the current toolbox to study factors involved in the UV-induced DNA damage response, and will contribute to a better understanding of the molecular mechanism of NER in living cells.

ACKNOWLEDGEMENTS

FUNDING

CONFLICT OF INTEREST

The authors declare they have no conflict of interest.

REFERENCES

1. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.
2. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.
3. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.
4. Sugasawa, K., et al., *Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair.* Mol Cell, 1998. **2**(2): p. 223-32.
5. Wakasugi, M., et al., *DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair.* J Biol Chem, 2002. **277**(3): p. 1637-40.
6. Fitch, M.E., et al., *In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product.* J Biol Chem, 2003. **278**(47): p. 46906-10.
7. Vermeulen, W. and M. Fousteri, *Mammalian transcription-coupled excision repair.* Cold Spring Harb Perspect Biol, 2013. **5**(8): p. a012625.
8. Hanawalt, P.C. and G. Spivak, *Transcription-coupled DNA repair: two decades of progress and surprises.* Nat Rev Mol Cell Biol, 2008. **9**(12): p. 958-70.
9. Yokoi, M., et al., *The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA.* J Biol Chem, 2000. **275**(13): p. 9870-5.
10. Volker, M., et al., *Sequential assembly of the nucleotide excision repair factors in vivo.* Mol Cell, 2001. **8**(1): p. 213-24.
11. Compe, E. and J.M. Egly, *TFIIH: when transcription met DNA repair.* Nat Rev Mol Cell Biol, 2012. **13**(6): p. 343-54.
12. Sugasawa, K., et al., *Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning.* Mol Cell, 2009. **36**(4): p. 642-53.
13. Camenisch, U., et al., *Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair.* Nat Struct Mol Biol, 2006. **13**(3): p. 278-84.
14. Staresincic, L., et al., *Coordination of dual incision and repair synthesis in human nucleotide excision repair.* EMBO J, 2009. **28**(8): p. 1111-20.
15. Ogi, T., et al., *Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells.* Mol Cell, 2010. **37**(5): p. 714-27.
16. Moser, J., et al., *Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner.* Mol Cell, 2007. **27**(2): p. 311-23.
17. van Cuijk, L., W. Vermeulen, and J.A. Marteijn, *Ubiquitin at work: the ubiquitous regulation of the damage recognition step of NER.* Exp Cell Res, 2014. **329**(1): p. 101-9.
18. Dijk, M., et al., *Insight in the multilevel regulation of NER.* Exp Cell Res, 2014. **329**(1): p. 116-23.
19. Dantuma, N.P. and H. van Attikum, *Spatiotemporal regulation of posttranslational modifications in the DNA damage response.* EMBO J, 2016. **35**(1): p. 6-23.
20. Lans, H., J.A. Marteijn, and W. Vermeulen, *ATP-dependent chromatin remodeling in the DNA-damage response.* Epigenetics Chromatin, 2012. **5**: p. 4.
21. Mandemaker, I.K., W. Vermeulen, and J.A. Marteijn, *Gearing up chromatin: A role for chromatin remodeling during the transcriptional restart upon DNA damage.* Nucleus, 2014. **5**(3): p. 203-10.
22. Steurer, B. and J.A. Marteijn, *Traveling Rocky Roads: The Consequences of Transcription-Blocking DNA Lesions on RNA Polymerase II.* J Mol Biol, 2017. **429**(21): p. 3146-3155.
23. Novarina, D., et al., *Mind the gap: keeping UV lesions in check.* DNA Repair (Amst), 2011. **10**(7): p. 751-9.
24. de Boer, J. and J.H. Hoeijmakers, *Nucleotide excision repair and human syndromes.* Carcinogenesis, 2000. **21**(3): p. 453-60.

5

25. Lehmann, A.R., *DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy.* Biochimie, 2003. **85**(11): p. 1101-11.

26. DiGiovanna, J.J. and K.H. Kraemer, *Shining a light on xeroderma pigmentosum.* J Invest Dermatol, 2012. **132**(3 Pt 2): p. 785-96.

27. Decordier, I., K.V. Loock, and M. Kirsch-Volders, *Phenotyping for DNA repair capacity.* Mutat Res, 2010. **705**(2): p. 107-29.

28. Latimer, J.J. and C.M. Kelly, *Unscheduled DNA synthesis: the clinical and functional assay for global genomic DNA nucleotide excision repair.* Methods Mol Biol, 2014. **1105**: p. 511-32.

29. Cleaver, J.E., *Defective repair replication of DNA in xeroderma pigmentosum.* Nature, 1968. **218**(5142): p. 652-6.

30. Friedberg, E.C., *The discovery that xeroderma pigmentosum (XP) results from defective nucleotide excision repair.* DNA Repair (Amst), 2004. **3**(2): p. 183, 195.

31. Mellon, I., et al., *Preferential DNA repair of an active gene in human cells.* Proc Natl Acad Sci U S A, 1986. **83**(23): p. 8878-82.

32. Kobayashi, N., et al., *Quantitation and visualization of ultraviolet-induced DNA damage using specific antibodies: application to pigment cell biology.* Pigment Cell Res, 2001. **14**(2): p. 94-102.

33. Langie, S.A., et al., *Development and validation of a modified comet assay to phenotypically assess nucleotide excision repair.* Mutagenesis, 2006. **21**(2): p. 153-8.

34. Hu, J., et al., *Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution.* Genes Dev, 2015. **29**(9): p. 948-60.

35. Mayne, L.V. and A.R. Lehmann, *Failure of RNA synthesis to recover after UV irradiation: an early defect in cells from individuals with Cockayne's syndrome and xeroderma pigmentosum.* Cancer Res, 1982. **42**(4): p. 1473-8.

36. Nakazawa, Y., et al., *A semi-automated non-radioactive system for measuring recovery of RNA synthesis and unscheduled DNA synthesis using ethynyluracil derivatives.* DNA Repair (Amst), 2010. **9**(5): p. 506-16.

37. Johnson, J.M. and J.J. Latimer, *Analysis of DNA repair using transfection-based host cell reactivation.* Methods Mol Biol, 2005. **291**: p. 321-35.

38. Bohr, V.A., et al., *DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall.* Cell, 1985. **40**(2): p. 359-69.

39. Guo, J., P.C. Hanawalt, and G. Spivak, *Comet-FISH with strand-specific probes reveals transcription-coupled repair of 8-oxoGuanine in human cells.* Nucleic Acids Res, 2013. **41**(16): p. 7700-12.

40. Wienholz, F., W. Vermeulen, and J.A. Marteijn, *Amplification of unscheduled DNA synthesis signal enables fluorescence-based single cell quantification of transcription-coupled nucleotide excision repair.* Nucleic Acids Res, 2017. **45**(9): p. e68.

41. Marteijn, J.A., et al., *Nucleotide excision repair-induced H2A ubiquitination is dependent on MDC1 and RNF8 and reveals a universal DNA damage response.* J Cell Biol, 2009. **186**(6): p. 835-47.

42. van Cuijk, L., et al., *SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair.* Nat Commun, 2015. **6**: p. 7499.

43. Vermeulen, W., *Dynamics of mammalian NER proteins.* DNA Repair (Amst), 2011. **10**(7): p. 760-71.

44. Pines, A., et al., *PARP1 promotes nucleotide excision repair through DDB2 stabilization and recruitment of ALC1.* J Cell Biol, 2012. **199**(2): p. 235-49.

45. Robu, M., et al., *Poly(ADP-ribose) polymerase 1 escorts XPC to UV-induced DNA lesions during nucleotide excision repair.* Proc Natl Acad Sci U S A, 2017. **114**(33): p. E6847-E6856.

46. Niida, H., et al., *Phosphorylated HBO1 at UV irradiated sites is essential for nucleotide excision repair.* Nat Commun, 2017. **8**: p. 16102.

47. Zhang, L., et al., *The chromatin remodeling factor BRG1 stimulates nucleotide excision repair by facilitating recruitment of XPC to sites of DNA damage.* Cell Cycle, 2009. **8**(23): p. 3953-9.

48. Aydin, O.Z., et al., *Human ISWI complexes are targeted by SMARCA5 ATPase and SLIDE domains to help resolve lesion-stalled transcription.* Nucleic Acids Res, 2014. **42**(13): p. 8473-85.

49. Jiang, Y., et al., *INO80 chromatin remodeling complex promotes the removal of UV lesions by the nucleotide excision repair pathway.* Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17274-9.

50. Dulbecco, R., *Reactivation of ultra-violet-inactivated bacteriophage by visible light.* Nature, 1949. **163**(4155): p. 949.

51. Thompson, C.L. and A. Sancar, *Photolyase/cryptochrome blue-light photoreceptors use photon energy to repair DNA and reset the circadian clock.* Oncogene, 2002. **21**(58): p. 9043-56.

52. Husain, I., et al., *Mechanism of damage recognition by Escherichia coli DNA photolyase.* J Biol Chem, 1987. **262**(27): p. 13188-97.

53. Sancar, A., *Structure and function of photolyase and in vivo enzymology: 50th anniversary.* J Biol Chem, 2008. **283**(47): p. 32153-7.

54. Huang, Y., et al., *Crystal structure of cryptochrome 3 from Arabidopsis thaliana and its implications for photolyase activity.* Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17701-6.

55. Wang, H., et al., *Femtosecond dynamics of flavin cofactor in DNA photolyase: radical reduction, local solvation, and charge recombination.* J Phys Chem B, 2005. **109**(4): p. 1329-33.

56. Sancar, G.B., *Enzymatic photoreactivation: 50 years and counting.* Mutat Res, 2000. **451**(1-2): p. 25-37.

57. Eker, A.P., et al., *DNA repair in mammalian cells: Direct DNA damage reversal: elegant solutions for nasty problems.* Cell Mol Life Sci, 2009. **66**(6): p. 968-80.

58. Schul, W., et al., *Enhanced repair of cyclobutane pyrimidine dimers and improved UV resistance in photolyase transgenic mice.* EMBO J, 2002. **21**(17): p. 4719-29.

59. Jans, J., et al., *Powerful skin cancer protection by a CPD-photolyase transgene.* Curr Biol, 2005. **15**(2): p. 105-15.

60. Jans, J., et al., *Differential role of basal keratinocytes in UV-induced immunosuppression and skin cancer.* Mol Cell Biol, 2006. **26**(22): p. 8515-26.

61. Dull, T., et al., *A third-generation lentivirus vector with a conditional packaging system.* J Virol, 1998. **72**(11): p. 8463-71.

62. Mone, M.J., et al., *Local UV-induced DNA damage in cell nuclei results in local transcription inhibition.* EMBO Rep, 2001. **2**(11): p. 1013-7.

63. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.

64. Campeau, E., et al., *A versatile viral system for expression and depletion of proteins in mammalian cells.* PLoS One, 2009. **4**(8): p. e6529.

65. Nakajima, S., et al., *Cloning and characterization of a gene (UVR3) required for photorepair of 6-4 photoproducts in Arabidopsis thaliana.* Nucleic Acids Res, 1998. **26**(2): p. 638-44.

66. Houtsmuller, A.B. and W. Vermeulen, *Macromolecular dynamics in living cell nuclei revealed by fluorescence redistribution after photobleaching.* Histochem Cell Biol, 2001. **115**(1): p. 13-21.

67. Mitchell, D.L., *The relative cytotoxicity of (6-4) photoproducts and cyclobutane dimers in mammalian cells.* Photochem Photobiol, 1988. **48**(1): p. 51-7.

68. Sancar, G.B., *DNA photolyases: physical properties, action mechanism, and roles in dark repair.* Mutat Res, 1990. **236**(2-3): p. 147-60.

69. Zhao, X., et al., *Reaction mechanism of (6-4) photolyase.* J Biol Chem, 1997. **272**(51): p. 32580-90.

70. Batty, D., et al., *Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites.* J Mol Biol, 2000. **300**(2): p. 275-90.

71. Hey, T., et al., *The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay.* Biochemistry, 2002. **41**(21): p. 6583-7.

72. Moser, J., et al., *The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions.* DNA Repair (Amst), 2005. **4**(5): p. 571-82.

73. Yasui, A., et al., *A new class of DNA photolyases present in various organisms including aplacental mammals.* EMBO J, 1994. **13**(24): p. 6143-51.

74. Dinant, C., et al., *Activation of multiple DNA repair pathways by sub-nuclear damage induction methods.* J Cell Sci, 2007. **120**(Pt 15): p. 2731-40.

5

5

75. Sadelain, M., E.P. Papapetrou, and F.D. Bushman, *Safe harbours for the integration of new DNA in the human genome.* Nat Rev Cancer, 2011. **12**(1): p. 51-8.

76. Nishi, R., et al., *UV-DDB-dependent regulation of nucleotide excision repair kinetics in living cells.* DNA Repair (Amst), 2009. **8**(6): p. 767-76.

77. Hoogstraten, D., et al., *Versatile DNA damage detection by the global genome nucleotide excision repair protein XPC.* J Cell Sci, 2008. **121**(Pt 17): p. 2850-9.

78. Puumalainen, M.R., et al., *Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity.* Nat Commun, 2014. **5**: p. 3695.

79. Chigancas, V., A. Sarasin, and C.F. Menck, *CPD-photolyase adenovirus-mediated gene transfer in normal and DNA-repair-deficient human cells.* J Cell Sci, 2004. **117**(Pt 16): p. 3579-92.

80. Asahina, H., et al., *Expression of a mammalian DNA photolyase confers light-dependent repair activity and reduces mutations of UV-irradiated shuttle vectors in xeroderma pigmentosum cells.* Mutat Res, 1999. **435**(3): p. 255-62.

81. de Lima-Bessa, K.M., et al., *CPDs and 6-4PPs play different roles in UV-induced cell death in normal and NER-deficient human cells.* DNA Repair (Amst), 2008. **7**(2): p. 303-12.

82. You, Y.H., et al., *Cyclobutane pyrimidine dimers are responsible for the vast majority of mutations induced by UVB irradiation in mammalian cells.* J Biol Chem, 2001. **276**(48): p. 44688-94.

83. Chigancas, V., et al., *Photorepair prevents ultraviolet-induced apoptosis in human cells expressing the marsupial photolyase gene.* Cancer Res, 2000. **60**(9): p. 2458-63.

84. Chigancas, V., et al., *Defective transcription/repair factor IIH recruitment to specific UV lesions in trichothiodystrophy syndrome.* Cancer Res, 2008. **68**(15): p. 6074-83.

85. Martin, R.F., et al., *Comparative studies of UV-induced DNA cleavage by analogues of iodoHoechst 33258.* Int J Radiat Biol, 1994. **66**(5): p. 517-21.

86. Martin, R.F., et al., *Radiation sensitization by an iodine-labelled DNA ligand.* Int J Radiat Biol, 1990. **57**(5): p. 939-46.

87. Luijsterburg, M.S., et al., *Heterochromatin protein 1 is recruited to various types of DNA damage.* J Cell Biol, 2009. **185**(4): p. 577-86.

88. Schwarz, T., *UV light affects cell membrane and cytoplasmic targets.* J Photochem Photobiol B, 1998. **44**(2): p. 91-6.

89. Pattison, D.I. and M.J. Davies, *Actions of ultraviolet light on cellular structures.* EXS, 2006(96): p. 131-57.

90. Wurtmann, E.J. and S.L. Wolin, *RNA under attack: cellular handling of RNA damage.* Crit Rev Biochem Mol Biol, 2009. **44**(1): p. 34-49.

91. Ciccia, A. and S.J. Elledge, *The DNA damage response: making it safe to play with knives.* Mol Cell, 2010. **40**(2): p. 179-204.

92. Sirbu, B.M. and D. Cortez, *DNA damage response: three levels of DNA repair regulation.* Cold Spring Harb Perspect Biol, 2013. **5**(8): p. a012724.

93. Tresini, M., J.A. Marteijn, and W. Vermeulen, *Bidirectional coupling of splicing and ATM signaling in response to transcription-blocking DNA damage.* RNA Biol, 2016. **13**(3): p. 272-8.
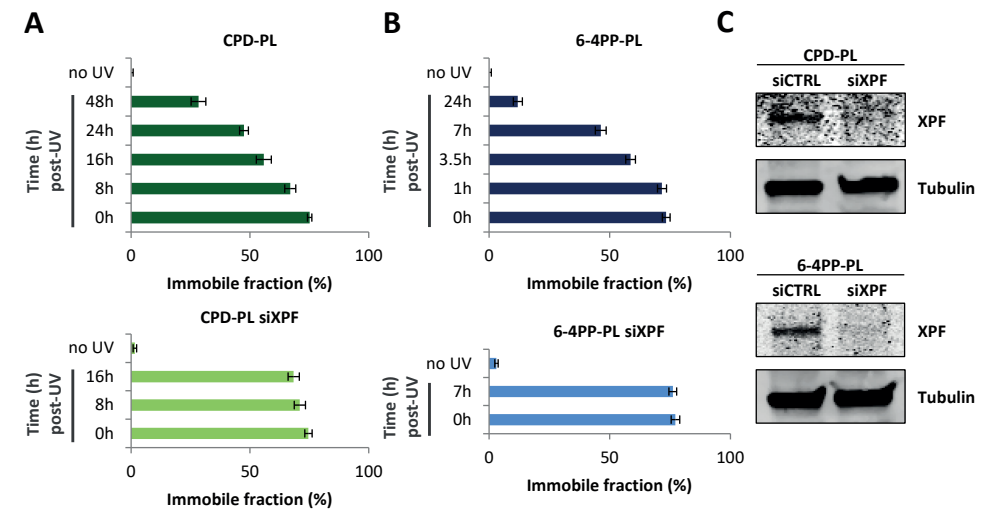
SUPPLEMENTARY FIGURES
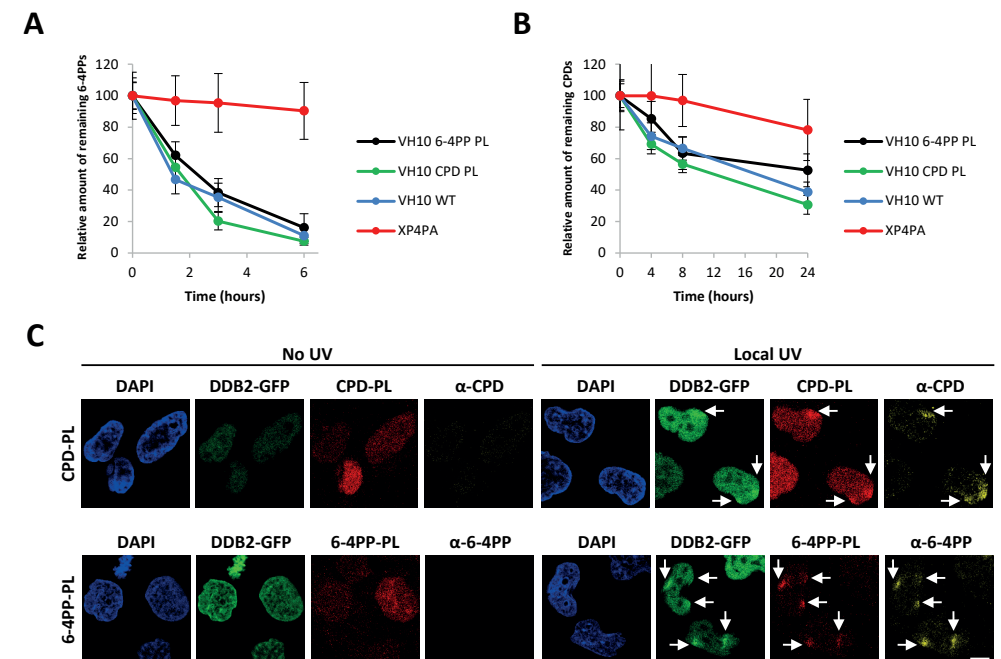
Supplementary Figure 1. (A and B) The kinetics of endogenous DNA damage removal by NER was determined by quantifying the levels of 6-4PPs in time after 16 J/m^2 (A) and CPDs in time after 10 J/m^2 (B) by immunofluorescence using 6-4PP and CPD specific antibodies. VH10 wild type cells (WT), VH10 cells expressing CPD-PL or 6-4PP-PL, and NER-compromised (XP-C) XP4PA cells were UV-irradiated and allowed to repair for the indicated time points. Relative fluorescence directly after UV exposure was set at 100% and average fluorescence intensities were plotted in time (n>150 cells of 2 independent experiments +/- SEM). (C) Representative immunofluorescence images of GFP-DDB2 expressing VH10 cells that were transduced with either CPD-PL (upper panel) or 6-4PP-PL (lower panel). Cells were non-irradiated or locally UV-C irradiated (60 J/m^2), directly fixed and stained with CPD or 6-4PP antibodies as indicated. Arrows indicate local UV damages. Scale bar: 7.5 µm.
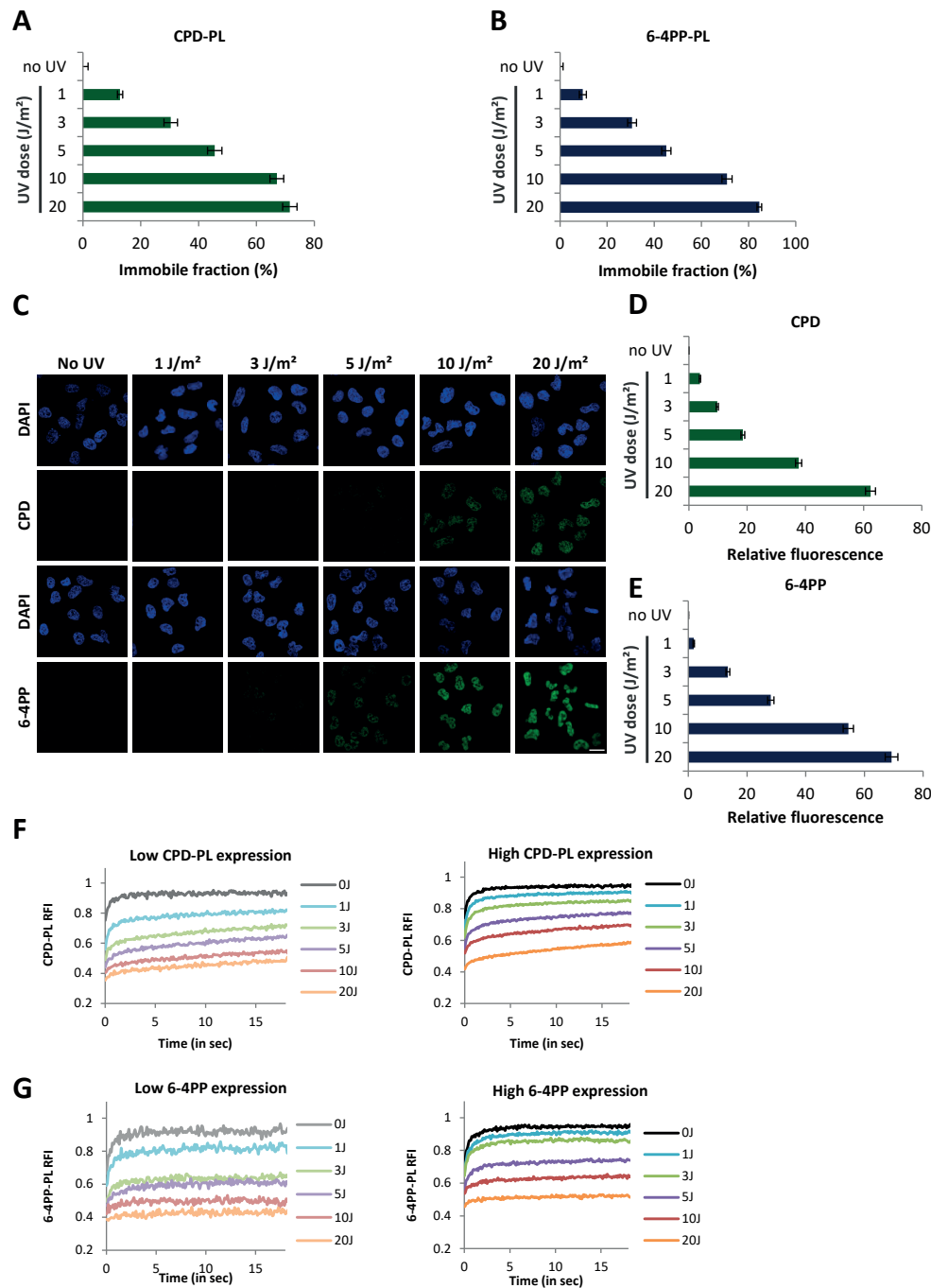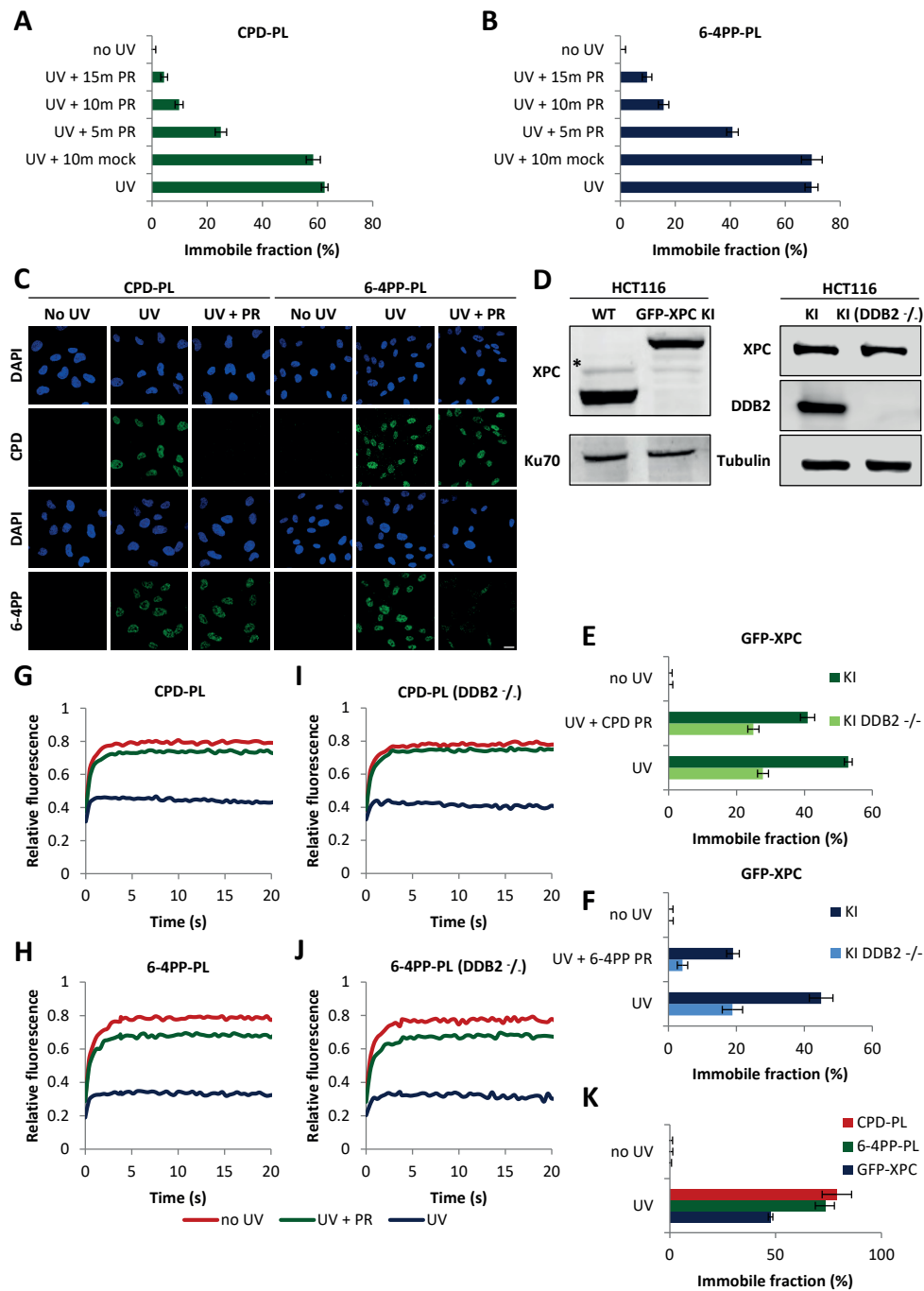
Supplementary Figure 2. (**A** and **B**) Immobile fractions of CPD-PL (**A**) and 6-4PP-PL (**B**) expressing VH10 cells in non-irradiated or globally UV-C irradiated at the indicated UV doses as determined by FRAP analyses shown in Figure 2A and 2B. Immobile fractions are calculated using the following formula: Immobile fraction (%) = 1 − ((average fluorescence intensity of UV-irradiated cells − the first post-bleach data point) / (average fluorescence intensity of non-irradiated cells−the first post-bleach data point)). The average fluorescence intensities are calculated over the measurements of the last 10 s. (n = 20 cells from 2 independent experiments, mean ± SEM). (**C**) Representative immunofluorescence images of non-irradiated (no UV) or globally UV-irradiated VH10 cells with the indicated UV doses, directly fixed and stained with CPD or 6-4PP antibodies as indicated. Scale bar: 25 µm. (**D**) CPD or (**E**) 6-4PP lesions (Supplementary Figure S3C) were quantified by determining the mean relative fluorescence intensities in immunofluorescence assays using lesion-specific antibodies. (n ≥ 50 cells, mean ± SEM). UV-treated conditions were background corrected by subtracting the mean fluorescence intensity of the non-irradiated condition. (**F** and **G**) UV dose-dependent immobilization of CPD-PL (**F**) and 6-4PP-PL (**G**) expressing VH10 cells with low (left panel) or high (right panel) PL expression levels. Non-irradiated or globally UV-irradiated cells were analyzed directly after irradiation with the indicated UV doses. Relative fluorescence intensity (RFI) values were normalized to the average pre-bleach signal (n=20 cells from 2 independent experiments).

Supplementary Figure 3. (**A** and **B**) Immobile fractions of CPD-PL (**A**) and 6-4PP-PL (**B**) in VH10 cells, which were transfected with control (upper panel) or XPF (lower panel) siRNAs, were determined by FRAP analyses shown in Figure 3E and 3F. (n ≥ 15 cells from 2 independent experiments, mean ± SEM). (**C**) siRNA-mediated XPF knockdown was assessed by immunoblotting VH10 lysates with XPF antibody, tubulin staining was used as loading control.

◀ **Supplementary Figure 4.** (**A** and **B**) Immobile fractions of non-irradiated, globally UV-irradiated (10 J/m², or globally UV-irradiated (10 J/m²) and photo-reactivated CPD-PL (**A**) and 6-4PP-PL (**B**) as determined by the FRAP analyses depicted in Figure 3A and 3B. (**C**) Representative immunofluorescence images of CPD-PL or 6-4PP-PL-expressing VH10 cells using 6-4PP or CPD lesion specific antibodies as indicated. Cells were non-irradiated, globally UV-irradiated (10 J/m2), or globally UV-irradiated (10 J/m2) and photo-reactivated (10 min PR), directly fixed and stained using immunofluorescence. Scale bar: 25 μm. (**D**) Upper panel; Expression of the full-length GFP-XPC protein and the concomitant loss of wild type (WT) XPC expression was confirmed by western blotting the lysates from WT and GFP-XPC knock-in HCT116 cell lines with an XPC antibody. Ku70 staining was used as loading control. * indicates an unspecific band. Lower Panel; CRISPR/Cas9-mediated DDB2 knock-out in GFP-XPC HCT116 cells was confirmed by western blotting with a DDB2 antibody. (**E** and **F**) Immobile fractions of non-irradiated, globally UV-irradiated (10 J/m²), or globally UV-irradiated (10 J/m²) and photo-reactivated (10 min) CPD-PL (**E**) and 6-4PP-PL (**F**) expressing GFP-XPC or GFP-XPC DDB2⁻/⁻ (DDB2⁻/⁻) HCT116 cells determined in the FRAP analyses depicted in Figure 3C-F. (**G** and **H**) FRAP analyses of PL-mCherry in non-irradiated, globally UV-irradiated (10 J/m²), or globally UV-irradiated (10 J/m²) and photo-reactivated (10 min) CPD-PL (**G**) and 6-4PP-PL (**H**) expressing GFP-XPC HCT116 cells. (**I** and **J**) FRAP analyses of PL-mCherry in non-irradiated, globally UV-irradiated (10 J/m²), or globally UV-irradiated (10 J/m²) and photo-reactivated (10 min) CPD-PL (**I**) and 6-4PP-PL (**J**) expressing GFP-XPC DDB2⁻/⁻ HCT116 cells (n ≥ 20 cells from 2 independent experiments, mean ± SEM). (**K**) Direct comparison of immobile fractions of CPD-PL, 6-4PP-PL and GFP-XPC in non-irradiated or globally UV-irradiated (XX J/m²) cells (n ≥ 20 cells from 2 independent experiments, mean ± SEM).
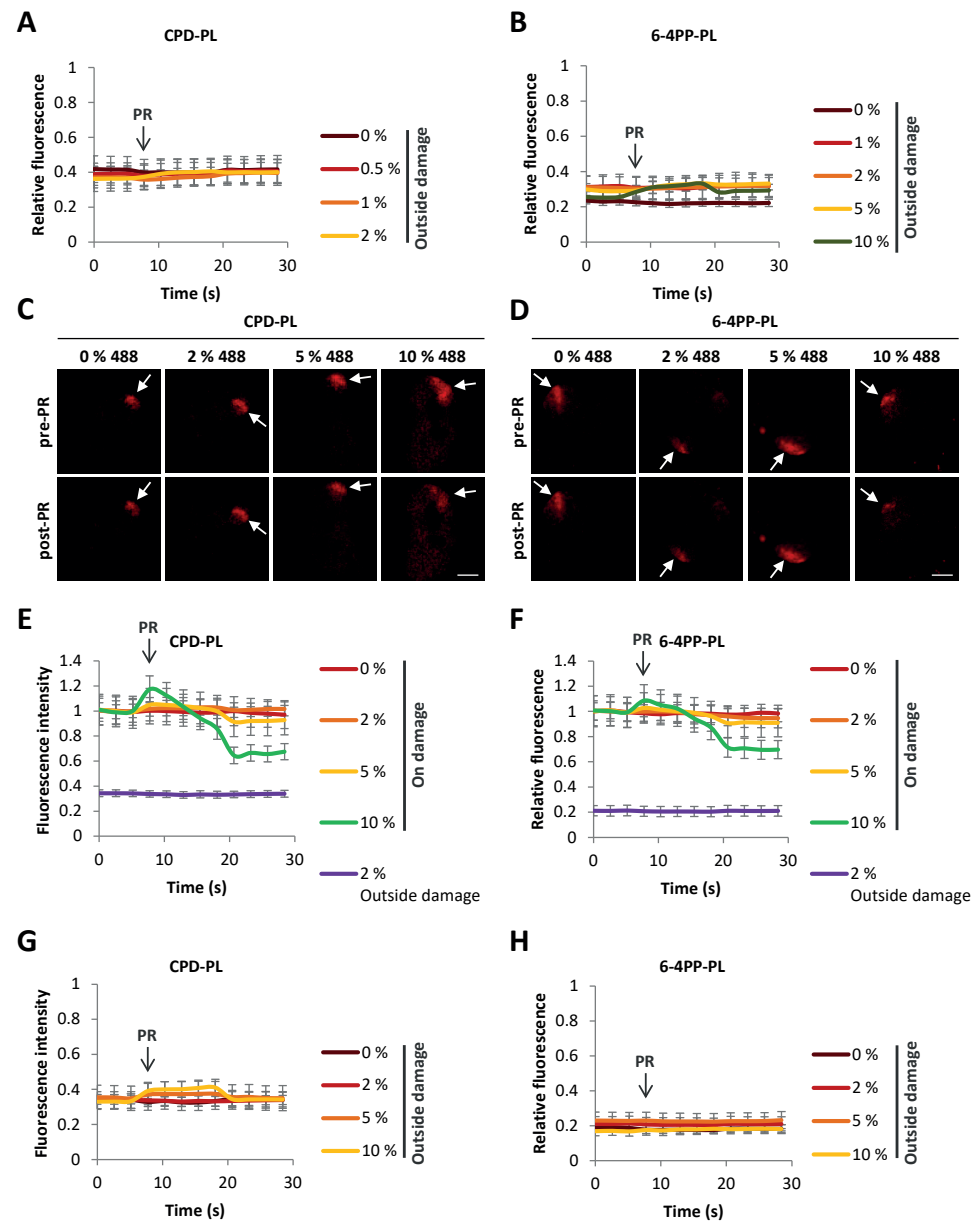
◀ **Supplementary Figure 5.** (**A** and **B**) Relative mCherry fluorescence signal of CPD-PL (**A**) and 6-4PP-PL (**B**) in a non-damaged nuclear region following PR normalized to pre-PR intensities at the local damage (n = 8 cells, mean ± SEM). Cells were locally UV-irradiated (60 J/m^2), then a non-damaged nuclear region was exposed after 7.5 sec (indicated by arrow and PR) to the indicated intensities of 405 nm laser for 13 s. (n = 8 cells, mean ± SEM). (**C** and **D**) Representative images of CPD-PL (**C**) and 6-4PP-PL (**D**) expressing VH10 cells before and 13 s after PR using 488nm laser at the indicated intensity. Arrows indicate local UV damages. Scale bar: 5 μm. (**E, F, G and H**) Cells were locally UV-irradiated (60 J/m^2), the local DNA damage spot and a region of the exact same size outside the damage within the nucleus were exposed after 7.5 s (indicated by arrow and PR) to the indicated intensities of 488 nm laser for 13 s. Relative fluorescence signal normalized to pre-PR intensities at the local damage of the mCherry-tagged PLs was quantified inside (**E and F**) and outside (**G and H**) the DNA damage within the nucleus. (n = 8 cells, mean ± SEM).
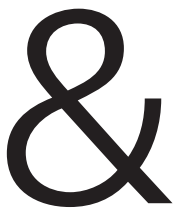
Appendix &

SUMMARY

The genetic information of all living organisms is stored in the nucleotide sequence of their DNA. Preserving the integrity of the DNA is a prerequisite for accurate gene expression and for faithful transmission of the genetic information to subsequent generations. DNA insults caused by endogenous metabolites or by environmental agents can damage DNA and can disrupt the essential cellular processes of gene transcription and DNA replication by blocking the forward translocation of the respective molecular machineries. If unrepaired, persistent DNA damage may eventually lead to malignant transformation or contribute to accelerated aging. To overcome these severe consequences, cells have evolved a DNA damage response network that consists of different dedicated DNA repair systems and damage-signaling pathways.

One of these repair pathways is Transcription-coupled Nucleotide excision repair (TC-NER), which specifically repairs DNA damage in the transcribed strand of active genes. TC-NER is initiated when RNA polymerase II (Pol II) is stalled during transcript elongation. DNA lesions that hinder or even fully block Pol II elongation are referred to as transcription-blocking DNA lesions (TBLs). The stalling of Pol II on TBLs triggers a series of highly regulated events, including Pol II processing to make the lesion accessible for DNA repair, R-loop-mediated DNA damage signaling, and the initiation of TC-NER. This multifaceted cellular response is reviewed in **chapter 1**. Furthermore, we discuss multiple aspects that thus far remained unclear in this field of research, including how cells coordinate transcription arrest, TBL repair and the subsequent restart of mRNA synthesis, and highlight research directions that may help answering these open questions.

In **chapter 2** we describe how the development of GFP-RPB1 knock-in cells, intended as a new live-cell imaging tool to study the consequences of DNA damage-induced transcription stress on Pol II, led to the disclosure of a critical finding about the steady state kinetics of Pol II. By photo-bleaching of GFP-RPB1 we identified three kinetically distinct fractions of chromatin-bound Pol II and using dedicated small molecule inhibitors to arrest Pol II at defined stages of the transcription cycle we allocated them to free and initiating, promoter-paused, and elongating Pol II. Monte Carlo-based modeling of Pol II kinetics allowed assessing a quantitative framework of the Pol II transcription cycle. This revealed that on average 7% of all Pol II complexes are freely diffusing, while 10% are chromatin-bound for 2.4 s during initiation, and 23% are promoter-paused for only 42 s. This unexpectedly high turnover of Pol II at promoters is most likely caused by premature termination of initiating and promoter-paused Pol II, and is in sharp contrast to the 23 min that elongating Pol II resides on chromatin. Our findings are highly supportive of a model in which Pol II initiation and promoter pausing are highly dynamic events of iterative cycles of Pol II chromatin binding and release. As a consequence of the frequent abortive release after initiation and pausing only one mRNA transcript will be synthesized per 100 Pol II initiation events, highlighting the importance of initiation and promoter pausing as key regulatory control events. In conclusion our study suggests that the iterative release

&

and re-initiation of promoter-bound Pol II is an important component of transcriptional regulation and supported by multiple independent studies, we suggest that this dynamic behavior creates an integrative hub that permits the rapid adaptation of transcription to various stimuli.

In **chapter 5** we demonstrate how UV-lesion-specific photolyases can be used to locate, quantify and instantaneously repair UV-induced photo-lesions in living cells, providing a highly sensitive new tool to study the UV-induced DNA damage response in real time in living cells.

Finally, in **chapter 6,** the main conclusions derived from the data presented in this thesis are summarized, and their implications and contributions to our understanding of DNA-damage induced cellular response to transcription stress are emphasized.

SAMENVATTING

De genetische informatie van alle levende organismen wordt opgeslagen in de nucleotidesequentie van hun DNA. Het behouden van de integriteit van het DNA is een voorwaarde voor correcte genexpressie en voor een goede overdracht van de genetische informatie naar volgende generaties. DNA-beschadigingen veroorzaakt door endogene metabolieten of omgevingsfactoren kunnen de essentiële cellulaire processen van gentranscriptie en DNA-replicatie verstoren door de progressie van deze moleculaire mechanismes te blokkeren. Indien niet hersteld, kan langaanhoudende DNA-schade uiteindelijk leiden tot kwaadaardige transformatie of bijdragen aan versnelde veroudering. Om deze ernstige gevolgen te voorkomen, hebben cellen een DNA-schade reactienetwerk ontwikkeld dat bestaat uit verschillende specifieke DNA-reparatiesystemen en schadesignaleringsroutes.

Een van deze reparatieroutes is transcriptie-gekoppeld nucleotide excisie herstel (TC-NER), wat specifiek DNA-schade in de getranscribeerde streng van actieve genen herstelt. TC-NER wordt geïnitieerd wanneer RNA polymerase II (Pol II) vastloopt tijdens transcriptie elongatie. DNA-laesies die Pol II elongatie belemmeren of zelfs volledig blokkeren, worden transcriptie-blokkerende DNA-laesies (TBL's) genoemd. Het vastlopen van Pol II op TBL's start een reeks sterk gereguleerde gebeurtenissen, waaronder het verwerken van Pol II om de laesie toegankelijk te maken voor DNA-reparatie, R-loop-gemedieerde DNA-schadesignalering en het starten van TC-NER. Deze veelzijdige, cellulaire respons wordt besproken in **hoofdstuk 1**. Verder bespreken we meerdere aspecten die tot nu toe onduidelijk bleven in dit onderzoeksveld, waaronder hoe cellen het vastlopen van transcriptie coördineren, TBL reparatie en de daaropvolgende herstart van de mRNA synthese, en belichten we onderzoeksrichtingen die mogelijk helpen bij het beantwoorden van deze open vragen.

In **hoofdstuk 2** beschrijven we hoe de ontwikkeling van GFP-RPB1 knock-in cellen, bedoeld als nieuw hulpmiddel voor het bestuderen van de gevolgen van door DNA-schade geïnduceerde transcriptie stress op Pol II in levende cellen, leidde tot de onthulling van een belangrijke bevinding over de steady-state kinetiek van Pol II. Door middel van het fotobleken van GFP-RPB1 hebben we drie kinetisch verschillende fracties van chromatine-gebonden Pol II geïdentificeerd en met behulp van specifieke klein molecuulremmers, die Pol II in verschillende stadia van de transcriptiecyclus houden, hebben we de fracties ingedeeld als vrij en initiërend, promotor-gepauzeerd en elongerend Pol II. Het bekijken van de Pol II kinetiek met een op Monte Carlo modellering gebaseerd model gaf een kwantitatieve basis voor het beoordelen van de transcriptie cyclus. Hieruit bleek dat gemiddeld 7% van alle Pol II complexen vrij diffundeert, terwijl 10% chromatine-gebonden is gedurende 2,4 seconden tijdens de initiatie en 23% pauzeert op de promotor gedurende slechts 42 seconden. Deze onverwacht hoge omzet van Pol II op promotors wordt hoogstwaarschijnlijk veroorzaakt door voortijdige beëindiging van initiërend en promotor-gepauzeerd Pol II, en staat in schril contrast met de 23 minuten dat elongerend Pol II.

Pol II op chromatine aanwezig is. Onze bevindingen ondersteunen een model waarin Pol II initiatie en promotor-pauzering zeer dynamische gebeurtenissen zijn van opeenvolgende cycli van het binden en loslaten van chromatine door Pol II. Als gevolg van het frequente afbreken van transcriptie en het loslaten van Pol II na initiatie en pauzering, zal slechts één mRNA-transcript worden gesynthetiseerd per 100 Pol II initiatie gebeurtenissen, wat het belang van initiatie en promotor-pauzering als regulerend controlepunt benadrukt. In conclusie suggereert onze studie dat de opeenvolgende loslating en nieuwe binding van promotor-gebonden Pol II een belangrijk onderdeel is van de transcriptie regulatie en, ondersteund door meerdere onafhankelijke studies, suggereren we dat dit dynamische gedrag een centrum creëert die de snelle aanpassing van transcriptie mogelijk maakt op verschillende stimuli.

In **hoofdstuk 5** laten we zien hoe een UV-laesie specifieke fotolyase gebruikt kan worden voor het lokaliseren, kwantificeren en direct repareren van een UV-geïnduceerde fotolaesie in levende cellen, dit geeft een erg gevoelig, nieuw mechanisme voor het bestuderen van de UV-geïnduceerde DNA-schade reactie in levende cellen in tijd.

Tot slot, in **hoofdstuk 6**, worden de belangrijkste conclusies samengevat die worden getrokken uit de data in dit proefschrift, en worden de implicaties en toevoegingen aan ons begrip van de DNA-schade geïnduceerde cellulaire reactie op transcriptie stress benadrukt.

&

&

DEUTSCHE ZUSAMMENFASSUNG

Die genetische Information aller lebenden Organismen ist in der Nukleotidsequenz ihrer DNA gespeichert. Die Erhaltung der Integrität der DNA ist eine Voraussetzung für eine genaue Genexpression und für eine getreue Übertragung der genetischen Informationen an nachfolgende Generationen. Umwelteinflüsse oder endogene Metaboliten können die DNA schädigen und die wesentlichen zellulären Prozesse der Gentranskription sowie DNA-Replikation stören, indem sie die Vorwärtsverlagerung der jeweiligen molekularen Maschinen blockieren. Wenn solche DNA-Schäden nicht repariert werden, können sie zu bösartigen Zellveränderungen führen oder den Alterungsprozess beschleunigen. Um diese schwerwiegenden Folgen zu überwinden, haben Zellen ein DNA-Schadensantwort-Netzwerk entwickelt, welches aus verschiedenen speziellen DNA-Reparatursystemen und Schadenssignalwegen besteht.

Eines dieser Reparatursysteme ist die Transkriptionsgekoppelte Nucleotid Excision Repair (TC-NER), die speziell DNA-Schäden im transkribierten Strang aktiver Gene repariert. TC-NER wird eingeleitet, wenn die RNA-Polymerase II (Pol II) während der Transkription blockiert wird. DNA-Schäden, die die Pol-II-Transkription behindern oder sogar vollständig blockieren, werden als transkriptionsblockierende DNA-Läsionen (TBLs) bezeichnet. Der Stillstand von Pol II auf TBLs löst eine Reihe von stark regulierten Ereignissen aus, darunter die Pol II-Verarbeitung um die Läsion für die DNA-Reparatur zugänglich zu machen, die R-Schleifen-vermittelte DNA-Schadenssignalisierung und die Einleitung von TC-NER. Diese vielschichtige zelluläre Reaktion wird in **Kapitel 1** behandelt. Darüber hinaus diskutieren wir mehrere Aspekte, die in diesem Forschungsgebiet bisher unklar geblieben sind, darunter die Frage, wie Zellen den Transkriptionsstillstand, die TBL-Reparatur und den anschließenden Neustart der mRNA-Synthese koordinieren, und zeigen Forschungsrichtungen auf, die bei der Beantwortung dieser offenen Fragen helfen können.

In **Kapitel 2** beschreiben wir wie GFP-RPB1-Knock-In-Zellen, die als neues „live-cell-Imaging" Tool zur Untersuchung der Folgen von DNA-Schaden-induziertem Transkriptionsstress auf Pol II entwickelt wurden, zu neuen Erkenntnissen über die stationäre Kinetik von Pol II führten. Durch die „Fotobleichung"-Technik von GFP-RPB1 identifizierten wir drei kinetisch unterschiedliche Fraktionen von chromatingebundenen Pol II. Zusätzlich dazu hielten wir Pol II mithilfe von speziellen „small-molecule" Inhibitoren in definierten Phasen des Transkriptionszyklus an, und konnten sie somit dem freien und initiierenden, promotorpausenden und verlängerten Pol II zuordnen. Die Monte-Carlo-basierte Modellierung der Pol-II-Kinetik ermöglichte die Bewertung eines quantitativen Rahmens des Pol-II-Transkriptionszyklus. Dabei stellte sich heraus, dass durchschnittlich 7% aller Pol-II-Komplexe frei diffundierend sind, während 10% während der Initiierung für 2,4 Sekunden lang chromatingebunden sind und 23% nur 42 Sekunden lang promotorpausiert sind. Dieser unerwartet hohe Umsatz von Pol II bei Promotoren wird höchstwahrscheinlich durch die vorzeitige Beendigung der

Initiierung und der promotorpausierten Pol II verursacht, und steht in scharfem Kontrast zu den 23 Minuten, in denen sich das verlängerte Pol II auf Chromatin befindet. Unsere Ergebnisse unterstützen ein Modell, bei dem die Pol-II-Initiation und die Promotorpause hochdynamische Ereignisse iterativer Zyklen der Pol-II-Chromatinbindung und -freigabe sind. Als Folge des häufigen Abbruchs nach der Initiierung und Pause, wird nur ein mRNA-Transkript pro 100 Pol II-Initiationsereignisse synthetisiert, was die Bedeutung von Initiierung und Promotorpause als wichtige regulatorische Kontrollereignisse deutlich unterstreicht. Abschließend schlägt unsere Studie vor, dass die iterative Freisetzung und Reinitiierung von Promoter-gebundenem Pol II ein wichtiger Bestandteil der transkriptionellen Regulierung ist und, durch mehrere unabhängige Studien unterstützt, schlagen wir vor dass dieses dynamische Verhalten eine integrative Drehscheibe schafft, die die rasche Anpassung der Transkription an verschiedene Reize ermöglicht.

.

In **Kapitel 5** zeigen wir, wie UV-läsionsspezifische Photolyasen verwendet werden können, um UV-induzierte Photoläsionen in lebenden Zellen zu lokalisieren, zu quantifizieren und sofort zu reparieren. Damit stellen wir ein hochempfindliches neues Werkzeug zur Verfügung, um die UV-induzierte DNA-Schadenreaktion in Echtzeit in lebenden Zellen zu untersuchen.

Abschließend werden in **Kapitel 6** die wichtigsten Schlussfolgerungen aus den Daten, die in dieser Arbeit zusammengefasst sind, sowie ihre Auswirkungen und Beiträge zu unserem Verständnis von DNA-Schäden induzierte zelluläre Reaktionen auf Transkriptionsstress erläutert und betont.

CURRICULUM VITAE

Barbara Steurer

Date of birth: 18.05.1988
Place of birth: Bregenz, Austria
Nationality: Austrian

Work experience

2018 – Present Post Doctoral Researcher, Department of Molecular Genetics, Erasmus MC Rotterdam
2013 – 2018 PhD Candidate Biomedical Sciences, Erasmus MC Rotterdam

Education

2011 – 2013 MSc in Molecular Medicine, Erasmus MC Rotterdam
2007 – 2011 BSc in Chemistry, University of Innsbruck
1998 – 2006 Bundesgymnasium Bregenz Blumenstrasse

&

&

&

LIST OF PUBLICATIONS

Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. Steurer B, Janssens RC, Geverts B, Geijer ME, Wienholz F, Theil AF, Chang J, Dealy S, Pothof J, van Cappellen WA, Houtsmuller AB, Marteijn JA. Proc Natl Acad Sci U S A. 2018 Apr 9. PMID:29632207

Traveling Rocky Roads: The Consequences of Transcription-Blocking DNA Lesions on RNA Polymerase II. Steurer B, Marteijn JA. J Mol Biol. 2017 Oct 27;429(21):3146-3155. Review, PMID:27851891

Fluorescently-labelled CPD and 6-4PP photolyases: new tools for live-cell DNA damage quantification and laser-assisted repair. Barbara Steurer, Yasemin Turkyilmaz, Marvin van Toorn, Wessel van Leeuwen, Paula Escudero-Ferruz, Jurgen A. Marteijn. Nucleic Acids Res.2019 Apr 23;47(7):3536-3549.

Ultra-soft X-ray system for imaging the early cellular responses to X-ray induced DNA damage. Kochan JA, van den Belt M, von der Lippe J, Desclos ECB, Steurer B, Hoebe RA, Scutigliani EM, Verhoeven J, Stap J, Bosch R, Rijpkema M, van Oven C, van Veen HA, Stellingwerf I, Vriend LEM, Marteijn JA, Aten JA, Krawczyk PM. Nucleic Acids Research, 2019 Sep 26;47(17):e100.

Disruption of TTDA results in complete nucleotide excision repair deficiency and embryonic lethality. Theil AF, Nonnekens J, Steurer B, Mari PO, de Wit J, Lemaitre C, Marteijn JA, Raams A, Maas A, Vermeij M, Essers J, Hoeijmakers JH, Giglia-Mari G, Vermeulen W. PLoS Genet, 2013 Apr;9(4):e1003431.

C8-glycosphingolipids preferentially insert into tumor cell membranes and promote chemotherapeutic drug uptake. Cordeiro Pedrosa LR, van Cappellen WA, Steurer B, Ciceri D, ten Hagen TL, Eggermont AM, Verheij M, Goñi FM, Koning GA, Contreras FX. Biochim Biophys Acta. 2015 Aug;1848(8):1656-70.

PHD PORTFOLIO

Name: Barbara Steurer
PhD period: October 2013 – April 2018
Promotor: W. Vermeulen
Copromoter: J.A.F. Marteijn
Erasmus MC Department: Molecular Genetics
Research School: Biomedical Sciences, MGC

General Courses

Safely Working in the laboratory,	2013
Research Integrity,	2014
Genetics,	2014
Cell and Developmental Biology,	2015
Biochemistry and Biophysics,	2015
Biostatistical Methods I: Basic Principles,	2016
Biomedical English Writing and Communication,	2017

Specialized Courses

Leica Confocal Introduction Course,	2013
Genome maintenance and cancer,	2014
Technology facilities: Proteomics,	2014
The microscopic image analysis: from theory to practice,	2014
In vivo imaging from molecule to organism,	2015
Workshop Microsoft Excel Basic,	2016
Workshop Microsoft Excel Advanced,	2016

Teaching Assistance and Project Supervision

BSc Nanobiology, Genetics,	2014-2018
MSc project, Paula Escudero Ferruz,	2015
MSc project, Wessel van Leeuwen,	2016

Meetings, Workshops, Conferences

Vermeulen group discussions, oral presentations,	2013-2018
Molecular Genetics workdiscussions, oral presentations,	2013-2018
Vermeulen Literature discussions,	2013-2018
MGC DNA repair meeting, Rotterdam/Leiden,	2013-2018
Cancer Genomics meeting, Amsterdam,	2013-2018
ZonMW project progress meetings,	2013-2017
MGC PhD Workshops,	2013-2015

PTMs in Cell Signaling, 2014
Copenhagen Bioscience Cluster, Copenhagen, Denmark ,
Responses to DNA damage: 2016
from molecule to disease, Egmond aan zee, The Netherlands,
Mechanisms of Eukaryotic Transcription, 2017
Cold Spring Harbor Laboratories, New York

ACKNOWLEDGEMENTS

You never know what's around the corner.
It could be everything. Or it could be nothing.
You keep putting one foot in front of the other,
and then one day you look back,
and you've climbed a mountain.

Tom Hiddleston

I'd like to finalize this thesis by acknowledging everyone who contributed to it, who helped me climb this mountain.

Dear Professor Vermeulen, dear **Wim**, thank you for giving me the opportunity to do my Masters internship in your lab, which opened the doors for this PhD journey. I feel truly grateful to have you as a promoter almost 8 years later. Your calm, positive, and always supportive mentorship in this hectic academic world has always been inspiring to me. Thank you for your guidance through all these years.

Dear **Jurgen**, without you I wouldn't be where I am today. Thank you for offering me this PhD position, and for your continuous support ever since. Your trust and belief in me and my work have always kept me motivated. Thank you for your encouraging words whenever I needed to hear them. I am also very grateful for the opportunity to finish the work on chapter 3 and to start new projects together. Last but not least, thank you for all the (bad) jokes. Laughing is a great way to deal with failed experiments.

Arjan, your supervision during my first internship has shaped me as a scientist in more ways than you can imagine. Thank you for your always being straight forward, your willingness to help out whenever needed and your collaboration on almost all my projects. Working with you has always been fun! **Hannes**, since I sit downstairs you almost never annoy me anymore, yeay! **Karen**, thank you for all your help throughout the years, and especially the moral support towards the end. **Maria,** thank you for always greeting me with a big smile and your interest and critical feedback on my work. **Roel**, without you this thesis would only be half as thick. Sharing a project is not always easy, but I think we made it work. Thank you for sharing the frustrations and successes along the way.

Cristina, we've been on this academic journey together for a while now and I'm very grateful our office has always been a place to share what's really going on - the good and the bad. Thank you for your open heart and ears, and for your enthusiastic interest in my work.

Marit, working with you has been fun, efficient and uncomplicated from the very beginning. Thank you for always brightening up the lab, meetings, and conferences with your contagious smile. I'm very grateful to you for your enthusiasm to be my paranymph, I couldn't wish a stronger woman to support me on that day.

Alex, thank you for making every hallway encounter an obstacle course, it has put a smile on my face many times and few things are worth more during busy lab days. Nils, I truly appreciate the support and friendship you have provided over the years. Thank you for helping, listening, and the refreshing conversations about everything and nothing! Di, thank for your cheerfulness, the honesty, and the hugs. Keep up the good work and the smile, all these challenges will make you grow even stronger. Dear Anja, thank you for your friendliness and for always having a spare tube of EU to borrow. Marvin, thank you for always using deodorant before you wear my lab coat, and for always sharing your excitement about RPB1. To all my new lab mates Diana, Alba, Carlota, Melanie, Maarten, Irene. Thank you for all the fresh winds and new motivation you're bringing to the lab. I wish you all the best of luck with your projects and a smooth ride through your PhD/ postdoc journeys.

To those who are continuing their journey elsewhere. Imke, Mariangela, Jana, Yasemin, Masaki. Thank you for great times during lunch and coffee breaks, and all the fun that was had in between the benches.

I am deeply thankful to all staff of the Biomedical science theme. Jasperina, Marieke, Christine, Leo, Koos, Anette, Sjozef, Ton, and the kitchen ladies. Thank you for keeping this place up and running!

My acknowledgements to everyone I have had the pleasure of collaborating with: Joris, Akos, Jiang, Gert, Martin, Bart, Adriaan. Thank you for all your help and your contributions to this thesis.

Dear Paula and Wessel, thank you for your great efforts during your internships in our lab, without your help chapter 5 probably wouldn't exist. I hope you'll have the pleasure of working with equally motivated students during your own PhDs.

Angela, du bist zum perfekten Zeitpunkt gekommen, um mir vorzuleben, wie man den Endspurt am besten meistert. Deine Gelassenheit und Fröhlickeit sind bewundernswert. Danke fürs Zuhören und all die guten Ratschläge, danke für die Freunschaft im und ausserhalb vom Labor.

Nesrin, thank you for your friendship during this PhD and beyond. Nobody could better understand and relate to the doubts and worries about life in and outside academia like you. Thank you for always listening and understanding, for our raw and honest conversations, for all the hugs, and for bringing me to Bikram Yoga with you, it has opened up a whole new world for me.

Dear BYR team, a special heartfelt thank you for all the lessons you have taught, and all the lessons I have learned along the way. The hot room has been a safe space to forget about experiments, presentations, and unfinished papers for a little while. Without the peace of mind, resilience and determination I have gained during practice, this thesis might have never been finished. I am also truly grateful for the beautiful friendships that developed from back-bending and sweating together. Paula, Stephanie, Richard, Johan, and Sevda... you inspire me in your own unique ways. Thank you for all the honest

conversations over juice, coffee and wine. Thank you for your interest in my work, and above all for pulling me out of the science bubble every now and then.

Frank, Rod, Jon. Climbing with you guys has been a weekly highlight during the past few months. Thank you for taking my mind of this thesis for a few hours and thank you for all the safe belays!

Caterina, our paths could not have been more different after the Masters, but look at us and how far we've come, I'm so proud of us! Even though we're far apart most of the time, it feels like you're right around the corner whenever I need you. Thank you for always being there for me, for supporting me through all highs and lows of (PhD) life. Your friendship means the world to me!

Dear Fred, the last parts are always the hardest, with you by my side it didn't feel that way. I am beyond grateful for your love, the laughs, the snuggles, and all the beta! Thank you for all the new things you teach me and for changing my perspectives in so many ways.

Meine Lieben daheim, Katharina, Stephan, Lea und Pius. Danke fuer die wohltuende Ablenkung während eurer Holland Besuche, und die vielen gemeinsamen Wanderungen inklusive Gipfelgummibärle während der Heimaturlaube.

Liebste Mama, bester Papa. So fern, und im Herzen doch so nah. Ohne euch wäre es nicht gegangen. Ihr habt mir immer die Freiheit gelassen meinen eigenen Weg zu finden, und habt mich gleichzeitig auf allen Wegen unterstützt. Zu wissen, dass Ihr immer für mich da seid, auch wenn ich weit weg bin, hat mir die notwendige Sicherheit für dieses PhD Abenteuer gegeben. Aus den zahlreichen Heimaturlauben konnte ich immer wieder neue Kraft, Ruhe und Motivation schöpfen um am nächsten Kapitel weiterzuarbeiten. Ich bin euch unbeschreiblich dankbar für jegliche Unterstützung, euer Vertrauen und eure Liebe.