

DISSECTING THE DNA DAMAGE-INDUCED TRANSCRIPTION STRESS RESPONSE

Marit Geijer

Cover design: Niek Hoogenboom & Marit Geijer
Illustration 'DNA': Flaticon.com
Layout: Chris Driessen Desktopping
Printed by: ProefschriftMaken | www.proefschriftmaken.nl

ISBN: 978-94-6423-131-1

Copyright © 2020 Marit Geijer

All rights reserved. No part of this thesis may be reprinted, reproduced, or transmitted in any form or by any means, without prior written consent of the author.

**DISSECTING THE DNA DAMAGE-INDUCED
TRANSCRIPTION STRESS RESPONSE**

**HET ONTLEDEN VAN DE DNA-SCHADE GEÏNDUCEERDE
TRANSCRIPTIE STRESS REACTIE**

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 31 maart 2021 om 15.30 uur

door

Marit Esmée Geijer
geboren te Leiden

Promotiecommissie

Promotor Prof.dr. W. Vermeulen

Overige leden Dr. J.A.A. Demmers
 Dr. J. Lebbink
 Prof.dr. M.A.T.M. van Vugt

Co-promotor Prof.dr. J.A.F. Martijn

Table of contents

Scope of the thesis	6
Chapter 1 General introduction	9
Chapter 2 Quantitative Pol II interaction proteomics to identify new factors in the UV-induced transcription stress response	41
Chapter 3 Elongation factor ELOF1 drives transcription-coupled repair and prevents genome instability	99
Chapter 4 DNA damage-induced replication stress results in PA200 proteasome-mediated degradation of acetylated histones	155
Chapter 5 General discussion	185
Appendix Summary Samenvatting Curriculum Vitae List of publications PhD portfolio Acknowledgements	197

Scope of this thesis

Faithful transcription of our genome is essential to maintain proper cell function. Therefore, transcription is tightly regulated at the different steps of the transcription cycle. Transcription is threatened by DNA lesions caused by endogenous and exogenous factors. These so-called transcription-blocking lesions (TBLs) can partially or completely block transcription or may result in the onset of mutant transcripts. Impaired transcription can severely reduce cell function and eventually may result in senescence or cell death. Luckily, cells have a dedicated repair pathway called transcription-coupled nucleotide excision repair (TC-NER) that removes TBLs. TC-NER is initiated upon stalling of RNA polymerase II (Pol II) upon a TBL and promotes the recruitment of downstream TC-NER factors. These repair factors help to recognize the TBL, verify the lesion and excise the damaged DNA after which the resulting single strand gap is filled by DNA synthesis. After damage removal, transcription has to be resumed to maintain cell function. In addition to this direct (*in cis*) consequence of Pol II stalling, transcription is also regulated in a genome-wide manner (*in trans*). This includes a block in transcription initiation, release of promoter-paused Pol II and alternative splicing. Together, these mechanisms attempt to diminish the severe consequences of TBL-induced transcriptional interference and promote lesion-recognition and repair to maintain genome integrity. The direct and indirect consequences of TBLs are called the transcription stress response to DNA damage and this response is evaluated in detail in **Chapter 1**.

Although the TC-NER pathway has been discovered decades ago and most key factors have been studied extensively, the exact molecular mechanism is not yet known. For example, Pol II is a tightly regulated protein that needs to be resolved upon stalling to allow repair factors to access the TBL. However, the exact factors regulating Pol II upon DNA damage have yet to be determined. Therefore, in **Chapter 2**, we described different SILAC-based Pol II interaction proteomics experiments to systematically identify, potentially novel, UV-specific interacting proteins. We used different isolation and extraction methods to extract Pol II and its interactors from the chromatin, immunoprecipitate the Pol II complex and identified its interactors. Moreover, we used proximity labeling to covalently tag proteins that are located in close proximity of Pol II in response to UV irradiation. Combining the analysis of the different isolation methods identified new UV-specific Pol II interactors that need to be further evaluated to provide insight whether and how they regulate Pol II during the transcription stress response. Our study compared different methods that could be used to identify Pol II-interacting proteins and could be used to study its interactors at different time points after UV.

Additionally, we have used a different approach to identify factors involved in the transcription stress response based on their sensitivity to TBLs, by using a genome-wide CRISPR/Cas9 sgRNA-based genetic screen (**Chapter 3**). We used UV irradiation as a selective pressure to identify genes involved in the transcription stress response. Analysis of the guide RNAs that were lost upon UV irradiation identified transcription elongation factor 1 homolog (ELOF1) as a new UV-sensitive gene, amongst many other UV-sensitive genes. In this study, we further investigated the role of ELOF1 and identified this protein as a component of the Pol II complex where it functioned as an elongation factor. We showed that ELOF1 depletion resulted in hypersensitivity to different types of TBLs and

that it was an important factor for the TC-NER pathway as repair was severely impaired due to improper recruitment of UVSSA and other downstream repair factors. Moreover, we showed that the function of ELOF1 in the transcription stress response was highly evolutionary conserved. Furthermore, we uncovered a role for ELOF1 in preventing transcription-replication conflicts in addition to its function in TC-NER.

So far, we mainly focused on proteins involved in repairing DNA and regulating the transcription stress response to DNA damage. However, transcription, replication and DNA damage repair take place at the chromatin, in which the DNA is tightly wrapped around nucleosomes, which consists of histone octamers. To regulate the access for DNA transacting processes to the DNA, the chromatin can be remodeled. This can be regulated by for example post-translational modifications (PTMs) of histones, that can directly regulate chromatin compaction or provide a binding platform for chromatin remodelers. One of the first PTMs that was shown to regulate chromatin compaction during transcription and DNA damage repair was histone acetylation. This was proposed to be involved in increasing the accessibility of repair factors to DNA lesions and to restore compaction after repair. However, a complete overview of the dynamic acetylation events after UV-induced DNA damage is lacking.

Therefore, in **Chapter 4**, we focused on chromatin reorganization via post-translational modifications and specifically focused on acetylation levels in response to UV-induced DNA damage. We used SILAC-based interaction proteomics and identified a genome-wide histone deacetylation in response to UV irradiation. This deacetylation was independent of transcription but dependent on replication. Interestingly, we showed that the replication stress-induced deacetylation was a result of degradation of acetylated histones. The degradation was independent of ubiquitin and functioned via the PA200 proteasome complex. Degradation of acetylated histones likely inhibited replication to prevent transcription-replication conflicts and safeguard genome stability.

Finally, in **Chapter 5**, we evaluate the main findings obtained in the experimental work described in this thesis and discuss their implications on the transcription stress response. Moreover, some future experiments are proposed that will help to further understand the molecular mechanism of the DNA damage-induced transcription stress response.



GENERAL INTRODUCTION

What happens at the lesion does not stay at the lesion:

Transcription-coupled nucleotide excision repair and the effects of DNA damage on transcription in cis and trans

Marit E. Geijer¹, Jurgen A. Marteijn¹

¹ Department of Molecular Genetics, Oncode Institute, Erasmus MC, Rotterdam, The Netherlands

Abstract

Unperturbed transcription of eukaryotic genes by RNA polymerase II (Pol II) is crucial for proper cell function and tissue homeostasis. However, the DNA template of Pol II is continuously challenged by damaging agents that can result in transcription impediment. Stalling of Pol II on transcription-blocking lesions triggers a highly orchestrated cellular response to cope with these cytotoxic lesions. One of the first lines of defense is the transcription-coupled nucleotide excision repair (TC-NER) pathway that specifically removes transcription-blocking lesions thereby safeguarding unperturbed gene expression. In this perspective, we outline recent data on how lesion-stalled Pol II initiates TC-NER and we discuss new mechanistic insights in the TC-NER reaction, which have resulted in a better understanding of the causative-linked Cockayne syndrome and UV sensitive syndrome. In addition to these direct effects on lesion-stalled Pol II (effects *in cis*), accumulating evidence shows that transcription, and particularly Pol II, is also affected in a genome-wide manner (effects *in trans*). We will summarize the diverse consequences of DNA damage on transcription, including transcription inhibition, induction of specific transcriptional programs and regulation of alternative splicing. Finally, we will discuss the function of these diverse cellular responses to transcription-blocking lesions and their consequences on the process of transcription restart. This resumption of transcription, which takes place either directly at the lesion or is reinitiated from the transcription start site, is crucial to maintain proper gene expression following removal of the DNA damage.

Keywords

Transcription-coupled nucleotide excision repair, RNA polymerase II, DNA damage, Transcription inhibition, Transcription restart.

Abbreviations

Pol II, RNA polymerase II; TC-NER, transcription-coupled nucleotide excision repair; TBLs, transcription-blocking lesions; CPDs, cyclobutane-pyrimidine dimers; CSA, Cockayne syndrome protein A; CSB, Cockayne syndrome protein B; UVSSA, UV-stimulated scaffold protein A; CRL4, cullin-RING ubiquitin ligases 4; TFIIH, transcription factor II H; ALE, alternative last exon; AS, alternative splicing.

Introduction

The eukaryotic genome is transcribed by different RNA polymerases. These polymerases consist of multiple subunits, are structurally alike and function in a similar manner although they all transcribe a different part of the genome¹. In this perspective, we will focus on RNA polymerase II (Pol II), the RNA polymerase responsible for transcription of protein-coding genes and synthesis of most non-coding snRNAs and miRNAs¹. Correct temporal and spatial regulation of gene expression is crucial for proper cell function and homeostasis. To safeguard this, transcription is tightly controlled at almost each step of the dynamic transcription cycle, ranging from initiation, promoter proximal pausing, productive elongation to transcription termination^{2,3}. However, the DNA template transcribed by Pol II is compromised on a daily basis by numerous types of DNA damaging factors. Several types of these DNA lesions can block or strongly impede progression of Pol II and are therefore referred to as transcription-blocking lesions (TBLs). If TBLs are not resolved properly, prolonged stalling of Pol II can lead to severely disrupted cellular homeostasis due to absence of newly synthesized RNA molecules or the appearance of mutant RNA molecules^{4,5}. In addition, prolonged stalled Pol II induces R-loops and may result in collisions with advancing replication forks⁶. Altogether, these TBLs may result in genome instability, severe cellular dysfunction, premature cell death and senescence^{7,8} which finally may result in DNA damage-induced, accelerated aging⁹.

The effects of different transcription blocking lesions on Pol II

TBLs can originate from both endogenous and exogenous sources. The main examples of DNA damage of endogenous origin are by-products of metabolic processes in mitochondria. These reactive oxygen species can for example generate 8,5'-Cyclopurine-2'-deoxynucleosides or 8-oxo-7,8-dihydroguanine (8-oxo-G) lesions¹⁰⁻¹². Damage from exogenous sources which cause TBLs include clinically used chemotherapeutics like cisplatin which causes inter- and intrastrand crosslinks, or ultraviolet (UV) radiation mainly causing 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs) and cyclobutane-pyrimidine dimers (CPDs)¹³. While all these structurally different types of damage can interfere with elongating Pol II^{9,14}, the stalling is mechanistically different for these diverse types of lesions¹⁵. For example, Pol II most likely does not stall at 8-oxo-G lesions itself but is affected by base excision repair intermediates after the action of lesion-specific DNA glycosylase, such as OGG1¹⁶⁻¹⁸. Cisplatin-induced interstrand crosslinks do not pass through the so-called Pol II translocation barrier. This impairs delivery of this bulky lesion to the active site of Pol II and stalls Pol II in front of the lesion¹⁹. In contrast, UV-induced CPD lesions can pass the translocation barrier and therefore enter the active site of Pol II. This results in direct stalling of Pol II, thereby completely covering the lesion with a 35-nucleotide footprint; 10 nucleotides downstream and 25 nucleotides upstream of the CPD lesion^{13,20,21}. These UV-induced CPDs form a stable road block for Pol II, as shown by extreme stability of CPD-stalled Pol II complexes, with half-lives of approximately 20 hours *in vitro*²². As a consequence of the different effects on elongating Pol II, the diverse types of lesions also trigger different response mechanisms such as transcriptional bypass for cyclopurines or transcription-coupled nucleotide excision repair (TC-NER) for CPDs (figure 1)^{13,23,24}.

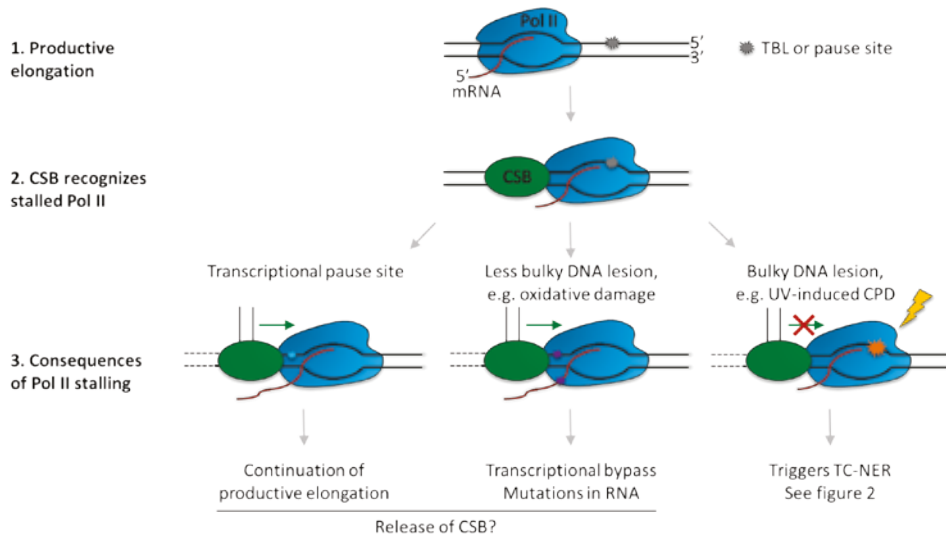


Figure 1. CSB probes Pol II for lesion recognition. (1) During productive elongation Pol II may run into DNA damage or natural occurring pause sequences that consequently impede Pol II forward translocation. **(2)** CSB recognizes and binds stalled Pol II. **(3)** Binding of CSB will result in an 80 degree bending of the DNA. Consequently, the ATPase activity of CSB is suggested to pull on the template DNA thereby mediating Pol II forward translocation. **(3, left and middle panel)** Translocation can be successful in case of natural occurring pause sites and less bulky DNA lesions, including oxidative damage, resulting in continuation of productive elongation or transcriptional bypass respectively. After the CSB-mediated forward movement, CSB might be released from Pol II. **(3, right panel)** Pol II cannot be translocated over bulky, transcription-blocking DNA lesions like UV-induced CPDs. This unsuccessful forward translocation of Pol II likely increases the residence time of CSB and thereby functions as a trigger for the initiation of TC-NER as illustrated in figure 2.

In this perspective we will give an overview on how cells cope with TBLs and provide insight in the cell-wide consequences of DNA damage on Pol II and transcription. First we will discuss how cells can efficiently remove TBLs by using the dedicated TC-NER pathway. Furthermore, we will discuss new insights on the consequences of TBLs *in trans*. These effects include both TBL-induced signaling events but also effects of DNA damage on non-lesion stalled Pol II by regulating the transcription cycle or affecting splicing. Finally, we will discuss factors and mechanisms involved in the last crucial step of overcoming the cytotoxic effects of TBLs; restart of transcription.

Transcription-coupled nucleotide excision repair

The concept of TC-NER was discovered almost three decades ago by the observation that UV-induced DNA damage in an actively transcribed gene was removed faster compared to damage in a non-transcribed genomic region^{25,26}. Follow-up studies showed that preferential repair of active genes was specifically observed in the transcribed strand^{25,27}. Since the discovery of TC-NER, many factors have been identified that play an important role in removal of TBLs and their discovery resulted in a better understanding of the molecular mechanism of TC-NER.

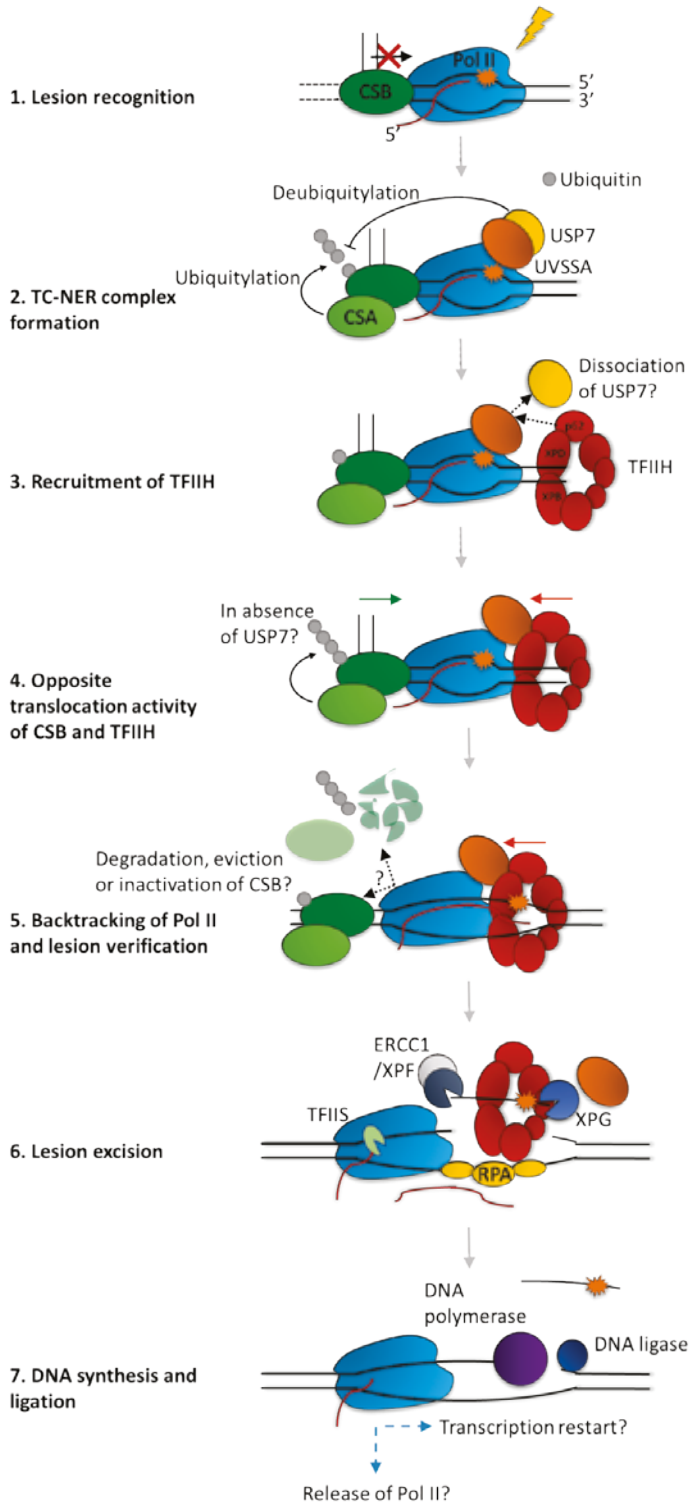
CSB senses lesion-stalled Pol II and initiates repair

TC-NER is initiated by recognition of lesion-stalled Pol II by the three main TC-NER factors: cockayne syndrome protein A and B (CSA and CSB) and UV-stimulated scaffold protein A (UVSSA) (figure 2)²⁸⁻³⁰. CSB is considered to be a master regulator of TC-NER as it plays a key role in recruitment of several proteins to the TC-NER complex. For example, CSB is essential for translocation of CSA to the nuclear matrix^{31,32} and recruitment to the TC-NER complex³³. In addition, CSB, together with CSA, recruits other factors to the TC-NER complex including the pre-mRNA splicing involved protein XAB2, nucleosome binding protein HMG1 and p300 histone-acetyl transferase³³. However, the exact mechanism of how this 1493 amino acid long multifunctional CSB protein contributes to the repair of TBLs remained elusive for a long period. While the C-terminal domain of CSB is required for its interaction with Pol II and translocation of CSA to the nuclear matrix³⁴, the SWI2/SNF2 DNA-dependent ATPase activity located in the central region of CSB was expected to play a crucial role during TC-NER-mediated TBL removal³⁵ and was shown to mediate chromatin remodeling after UV-induced DNA damage³⁵⁻³⁷. Interestingly, its key function during TC-NER was only recently disclosed by unraveling the structure of the yeast homolog of CSB, Rad26, in complex with lesion-stalled Pol II using cryo-EM studies³⁸.

This lesion-stalled structure showed that Rad26 binds to DNA upstream of Pol II and the TBL, causing a CSB-mediated 80 degree bending of the extruding DNA. Importantly, the 3'-5' ATP-dependent translocase activity of RAD26 pulls the DNA away from Pol II in a similar manner as Snf2, another member of the SWI2/SNF2 ATPase family, pulls DNA from nucleosomes³⁹. This "DNA Pulling" is suggested to stimulate forward translocation of Pol II over for example naturally occurring pause sites or small blocking lesions (figure 1). However, Rad26 cannot translocate Pol II over bulky DNA lesions that lead to a transcription block, like CPDs³⁸. This study provides important new insights in the long lasting question how TC-NER, and CSB specifically, could discriminate between normal paused Pol II and TBL-stalled Pol II in such a way that TC-NER is only initiated when needed. Since Rad26 is highly homologous to mammalian CSB, a similar key role for CSB in sensing damage-paused Pol II in mammalian cells is proposed. Interestingly, this model suggests that CSB is constantly probing Pol II to sense for an obstruction, indicating that CSB interacts both with lesion-stalled as well as unperturbed elongating Pol II. In line with this, it was observed that CSB transiently interacts with chromatin in a transcription-dependent manner in non-damaged conditions, most likely by binding to Pol II. A larger fraction of CSB was bound when cells were challenged with Pol II-stalling agents including UV-induced DNA damage or actinomycin D⁴⁰. Interestingly, the binding affinity of Rad26 to Pol II upon stalling at TBLs was not increased *in vitro*³⁸. This suggests that the observed increased CSB binding in cells might be caused by its prolonged binding to lesion-stalled Pol II, which could be the trigger for TC-NER complex assembly and subsequent repair of the TBL (figure 1)^{22,38}.

The role of CSA and UVSSA in TC-NER

A key protein which is recruited by CSB to the TC-NER complex is DDB1- and Cul4-associated factor (DCAF) CSA³³ which forms the CRL4^{CSA} E3 ligase complex together with DDB1, Cul4A and Roc1⁴¹. In this CRL4^{CSA} complex, CSA is a dedicated substrate receptor, giving the complex its essential target specificity⁴². CSA contains a seven-bladed



WD40 propeller and binds via its helix-loop-helix motif to DDB1^{28,42-44}. A recent study uncovered that the chaperonin TCP-1 ring complex (TRiC) interacts with CSA, especially those CSA proteins that are not incorporated in CRL4^{CSA} complexes. TRiC interaction is important for stability of CSA and most likely mediates CSA handover to DDB1 to form properly functioning CRL4^{CSA}. In unperturbed conditions, the ligase activity of CRL4^{CSA} is inhibited by binding of the COP9 signalosome^{41,46}. However, upon DNA damage, the COP9 signalosome dissociates, resulting in activation of CRL4^{CSA} ligase activity. Consequently, CSB, the substrate of the CRL4^{CSA} ligase, is ubiquitylated and targeted for proteasomal degradation (figure 2)^{47,48}. However, also other, yet unknown, targets of the CRL4^{CSA} complex may exist and function during TC-NER.

The third main factor in TC-NER initiation is UVSSA, which interacts with both CSA and Pol II^{30,49}. One of the roles of UVSSA during the TC-NER reaction is to specifically recruit ubiquitin-specific protease 7 (USP7) (figure 2)^{30,49,50}. USP7 was, just as UVSSA, shown to be crucial for TC-NER³⁰. The deubiquitylating activity of USP7 plays an important role in several different pathways following DNA damage. For example, Mdm2, XPC, p53, ALKBH3⁵¹⁻⁵⁴ but also many factors outside the DNA damage response are described as USP7 substrates⁵⁵. In a large number of these processes, USP7 is recruited to its substrate as part of a relatively stable complex. This seems to be a common mode of action for USP7 and might be essential to specifically target this abundant and pleiotropic deubiquitylating enzyme to its substrate^{56,57}. In addition, complex formation of USP7 stimulates its deubiquitylating activity. This can be facilitated by remodelling its structure to a more competent state mediated by its binding partner^{56,58-60}. Strikingly, in contrast to stimulation of USP7 activity following complex formation, the UVSSA-USP7 interaction was reported to inhibit the deubiquitylating activity of USP7⁶¹. Following induction of TBLs, USP7 is specifically recruited via UVSSA to the TC-NER complex, where it counteracts CRL4^{CSA}-mediated CSB ubiquitylation, thereby most likely increasing the half-life of CSB in the TC-NER complex (figure 2)^{30,50}. In addition, recently it was shown that USP7 also deubiquitylates UVSSA, which is important for efficient TC-NER⁶².

◀ **Figure 2. A model how transcription-coupled nucleotide excision repair (TC-NER) removes transcription-blocking lesions (TBLs).** **Step 1**, UV-induced TBLs result in stalling of Pol II and the inability of CSB to forward translocate Pol II. This results in an increased residence time for CSB on Pol II and most likely results in recruitment of the TC-NER factors CSA, UVSSA and USP7. **Step 2**, CSA is part of the CRL4^{CSA} E3 ligase complex and is recruited to the stalled Pol II complex via CSB resulting in polyubiquitylation of CSB. The deubiquitylating enzyme USP7 is recruited to the TC-NER complex by UVSSA and stabilizes CSB via deubiquitylation. **Step 3**, subsequently, TFIIH is recruited, most likely via a direct interaction with UVSSA. As USP7 and TFIIH bind the same domain of UVSSA, TFIIH binding might compete with USP7, resulting in the loss of USP7 activity in the TC-NER complex. **Step 4**, the forward translocating activity of CSB might counteract the reverse translocating helicase activity of the XPB and XPD subunits of TFIIH. To overcome these counteracting forces, either the ATPase activity of CSB needs to be inhibited, or CSB can be degraded or evicted from the TC-NER complex. Increased CSB ubiquitylation due to absence of USP7 might contribute to one of these processes. **Step 5**, due to the loss of CSB or its ATPase activity, TFIIH is able to efficiently reverse translocate (backtrack) Pol II, while at the same time the TFIIH complex verifies the DNA lesion. **Step 6**, following successful Pol II backtracking and damage verification, the DNA lesion is efficiently removed by dual incision of the damaged strand by the ERCC1/XPF and XPG nucleases. RPA binds the undamaged strand. **Step 7**, the ssDNA gap is filled by DNA synthesis and ligated to finalize repair. Subsequently, transcription has to be restarted. This can be initiated either directly at the lesion or by new Pol II initiation at the transcription start site when lesion-stalled Pol II is removed from the template DNA during the repair reaction.

Even though UVSSA has high affinity for USP7, it also has functions during TC-NER for which thus far no role of USP7 is observed. For example, the largest subunit of Pol II, RPB1, was shown to be ubiquitylated in a UVSSA-dependent manner but this modification does not lead to proteasomal degradation of Pol II⁶³. However, thus far it remains unclear whether this is a direct or indirect consequence of UVSSA. In line with previously observed interactions between UVSSA and subunits of transcription factor II H (TFIIH)⁶³, it was recently shown that UVSSA directly interacts with the PH domain of TFIIH subunit p62, via a short, highly conserved, acidic region in the central part of UVSSA⁶⁴. Interestingly, this acidic region in UVSSA was identified as it was highly similar to the p62 PH-binding region of global genome nucleotide excision repair (GG-NER) DNA damage sensor XPC^{64,65} (figure 2). This observation suggests that TFIIH is recruited to TBLs by UVSSA following TC-NER initiation via a similar mechanism as used by XPC to recruit TFIIH during GG-NER initiation^{63,64}. TFIIH, which is normally involved in transcription initiation, is a stable complex of ten subunits, including the helicases XPB and XPD⁶⁶. The ATPase activity of XPB was suggested to recruit TFIIH to damage and initiate opening of DNA around the lesion. The helicase activity of XPD helps to extend unwinding of the DNA in a 5' to 3' direction and is thought to verify the lesion with help of the weaker helicase activity of XPB⁶⁷⁻⁷¹. XPB is suggested to co-translocate with XPD and helps to scan the non-damaged complementary DNA strand since it has opposite directionality of XPD and thereby stimulates unwinding and lesion verification (figure 2)^{68,70}. XPA promotes lesion recognition by enhancing stalling of XPB and XPD and helps to detect chemically altered nucleotides^{68,72}. After proper lesion verification, TFIIH together with replication protein A (RPA) recruits the structure-specific endonucleases ERCC1/XPF and XPG in the correct orientation to excise the damaged strand^{73,74}. Repair is finished by refilling the gap with DNA synthesis and ligation (figure 2)^{75,76}.

New insights in the TC-NER pathway

The recently acquired insights in the mode of action of these TC-NER initiation factors have implications for current TC-NER models. One of the most important findings is that CSB, mediated by its ATP-dependent translocase activity, can discriminate between Pol II stalled at a DNA lesion or at pause sites by constantly probing the Pol II complex for its ability to forward translocate³⁸. This indirect recognition of lesion-stalled Pol II instead of detection of the DNA lesion itself allows detection of a large spectrum of structurally different types of DNA damage. However, at the same time this mechanism may result in different outcomes for lesion-stalled Pol II, depending on the type of DNA damage. For example, CSB forward translocation of Pol II could promote transcriptional bypass of less bulky, oxidative damage like 8-oxo-G lesions (figure 1)^{24,38,77}. In contrast, upon stalling at bulky lesions like CPDs, Pol II cannot be translocated by CSB, resulting in a longer residence time of Pol II and CSB at the lesion, which eventually results in initiation of TC-NER (figure 1)³⁸. The increased residence time of CSB suggests that CSB has to be stabilized, since it is normally targeted for proteasomal degradation by the CRL4^{CSA} complex following UV-induced DNA damage⁴⁸. Protection from degradation could be mediated by concerted action of UVSSA and deubiquitylating enzyme USP7 and thereby likely provides time for CSB to recruit downstream TC-NER machinery (figure 2)^{30,49,50,63,78}.

UVSSA was recently suggested to be involved in recruitment of TFIIH to Pol II^{63,64}. TFIIH, with its XPD 5'-3' and XPB 3'-5' helicase activity⁶⁸, is hypothesized to bind downstream of Pol II for DNA damage verification. This not only suggests that UVSSA and CSB bind to opposite sides of Pol II, but also suggests that the helicase activity of TFIIH might be involved in reverse translocation (backtracking) of Pol II^{38,70}. This would indicate that proofreading of the lesion and backtracking of Pol II is mediated by the exact same complex, namely TFIIH, assuring efficient subsequent removal of the TBL by excision. Of note, in this model TFIIH-mediated backtracking of Pol II is counteracted by the Pol II forward translocating property of CSB (figure 2). It is therefore tempting to speculate that after TFIIH recruitment, CSB needs to be removed in order for Pol II to be efficiently backtracked by TFIIH. Eviction of CSB from the TC-NER complex might be mediated by ubiquitylation, as CSB was shown to be degraded following CRL4^{CSA}-mediated ubiquitylation⁴⁸. However, this may also be mediated by other E3 ligases like BRCA1-BARD1 as this heterodimer was previously implicated to ubiquitylate CSB and target it for proteasomal degradation⁷⁹. Interestingly, USP7 and TFIIH are described to bind a similar region of UVSSA^{61,64}, which might suggest competitive binding of these proteins. This putative mutual exclusive binding of either TFIIH or USP7 to UVSSA might suggest that USP7-mediated CSB deubiquitylation activity is lost following TFIIH recruitment. This will result in increased CSB polyubiquitylation, which subsequently might result in removal of CSB from the TC-NER complex by proteasomal degradation. This may enable TFIIH to reverse translocate Pol II and verify the DNA damage^{68,70}.

Although it is tempting to speculate that CSB needs to be degraded in order to allow Pol II backtracking and damage verification by TFIIH, it cannot be excluded that the ATPase activity of CSB is inhibited, or that CSB is evicted from the chromatin without being degraded. In line with this last possibility, the C-terminal ubiquitin-associated (UBA) domain of CSB, which interacts with ubiquitin chains, plays a role in eviction of CSB from the TC-NER complex. CSB mutants lacking this UBA domain remain trapped at TC-NER complexes, resulting in increased UV-sensitivity and reduced transcription restart, indicating that removal of CSB is a crucial step during TC-NER^{34,80,81}. In line with a specific role in eviction of CSB at later stages in the TC-NER reaction, deletion of the UBA domain does not interfere with TC-NER complex assembly or ATPase activity of CSB^{80,81}, although CSA translocation to the nuclear matrix was affected³⁴. However, thus far the exact mechanism of CSB eviction and the ubiquitylated substrates which are recognized by the UBA domain of CSB remain elusive. Future experiments are necessary to test the above described model.

Pol II degradation

In addition to backtracking and lesion bypass, cells have evolved an additional mechanism to ensure clearance of lesion-stalled Pol II. The largest Pol II subunit, RPB1, was shown to be degraded by the proteasome following UV-induced DNA damage. This degradation is hypothesized to be a 'last resort' pathway, which only happens in conditions where TC-NER fails or when the damage load is too high⁸²⁻⁸⁴. Pol II degradation is a highly inefficient process, as cells have to generate new elongating Pol II complexes to restart transcription. In addition, degradation may result in a global decrease of the total Pol II pool that will likely influence transcription in general. However, degradation of Pol II prevents severe cytotoxic effects caused by persistent Pol II-stalling at DNA damage which

forms genomic roadblocks for advancing replication forks and other chromatin involved processes and may induce the formation of R-loops^{6-9,85}. In addition, degradation of Pol II will make the DNA lesion accessible for another round of TC-NER or for additional repair pathways to remove the TBL, for example by GG-NER.

During the last resort pathway, RPB1 is polyubiquitylated by the E3 ubiquitin ligase NEDD4, which generates lysine 63-linked polyubiquitin chains⁸⁶. These chains are subsequently trimmed down by deubiquitylating enzymes until a monoubiquitin modification remains on RPB1. This monoubiquitin can be extended with lysine 48-linked polyubiquitin chains by the Elc1/Cul3 ligase complex. Next, this chromatin bound, ubiquitylated RPB1 is recognized by the ubiquitin-selective segregase valosin-containing protein (VCP/p97) which removes RPB1 from the stalled complex and results in its proteasomal degradation^{46,82,87}.

Clinical consequences and phenotypical differences

The above described new insights in TC-NER may also have implications for understanding TC-NER-linked disorders. The importance of functional TC-NER is clearly illustrated by the Cockayne syndrome (CS), a human disorder with defective TC-NER, caused predominantly by mutations in CSA and CSB⁸⁸. CS is characterized by sensitivity to UV light, progressive neurodevelopmental symptoms, growth and developmental problems, mental retardation and severe premature aging⁸⁹⁻⁹¹. Strikingly, UV-sensitive syndrome (UV^SS), which is mainly caused by mutations in UVSSA, is also characterized by absence of TC-NER-mediated removal of UV-induced TBLs⁹², but displays only mild cutaneous UV-sensitivity in sharp contrast to the premature and developmental features observed in CS^{74,78,93}. The phenotypical differences between these two syndromes may partially be explained by additional functions of CS proteins compared to UVSSA. For example, CS proteins were proposed to be implicated in specific transcriptional programs⁹⁴, transcription initiation⁹⁵, redox balance⁹⁶, repair of double strand breaks^{97,98} and maintenance of mitochondrial DNA stability^{99,100}, while thus far no such roles are described for UVSSA. Importantly, it was also suggested that the additional CS features may be derived from a defect in repair of (endogenously produced) oxidative DNA damage interfering with transcription^{25,101} as it was shown that CS cells but not UV^SS cells are sensitive to oxidative DNA damage (figure 3)^{102,103}. CSB deficient cells lack the ability to forward translocate Pol II, which in case of an oxidative lesion will result in the inability to bypass this lesion³⁸. This might result in persistent stalling of Pol II, thereby preventing access for base excision repair proteins to remove the lesion (figure 3). On top of that, persistent stalling of Pol II can eventually lead to collisions with replication forks or result in onset of R-loops^{7,104} which may contribute to the CS phenotype. Recently, it was indicated that also UVSSA is involved in repair of oxidative lesions¹⁰⁵. However, in absence of UVSSA, CSB is expected to still induce transcriptional bypass of oxidative lesions which might suggest that the reduced repair rate of oxidative lesions due to loss of UVSSA might be less cytotoxic than persistently stalled Pol II complexes (figure 3)^{24,38,77,78,90}. It has to be noted that mutations in CSA and CSB result in similar phenotypes⁹⁰, however thus far no role for CSA in lesion bypass of oxidative damage has been reported. In addition, loss of UVSSA might result in destabilization of CSB, due to impaired recruitment of USP7 to the TC-NER complex. Stabilization of CSB by UVSSA/USP7 might only be essential when the ATPase activity of CSB is needed for prolonged time, for example

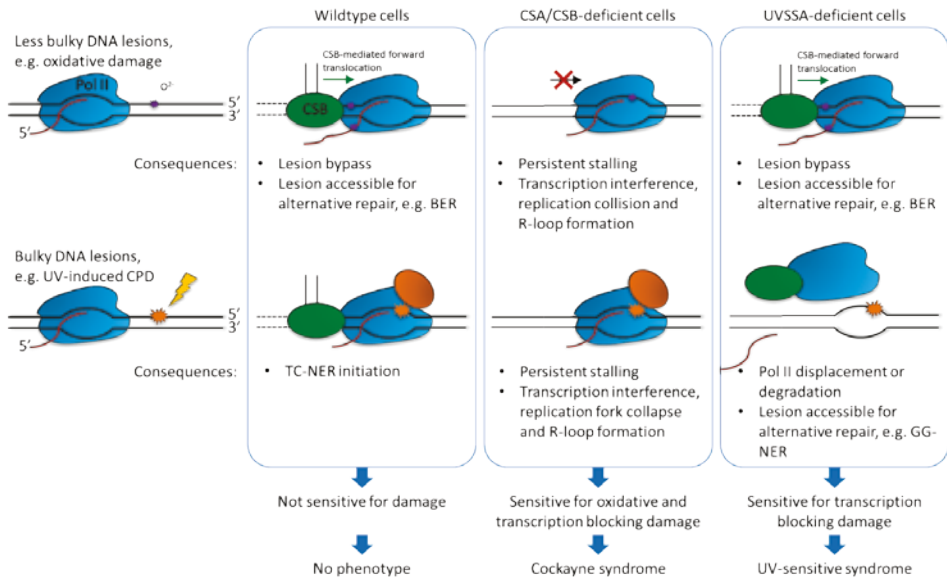


Figure 3. Differential processing of TBLs may explain differences in Cockayne Syndrome and UV sensitive syndrome. Less bulky lesions like oxidative damage (**top row**) and transcription-blocking damage including UV-induced CPDs (**middle row**) will lead to stalling of Pol II. In **wildtype cells**, CSB will bind TBL-stalled Pol II and induce its forward translocation. In case of oxidative lesions, translocation of Pol II by CSB can successfully bypass the lesion making it accessible for alternative repair, e.g. by BER. In contrast, CSB cannot translocate Pol II over TBLs and this will initiate TC-NER to remove the lesion and restart transcription. In **absence of CSB**, TBL-stalled Pol II cannot be forward translocated. Therefore, Pol II may not be able to bypass oxidative damage nor trigger TC-NER and will most likely remain stalled on the lesion. Persistent stalling can lead to transcriptional interference, cause replication-transcription collisions and form R-loops. CSB-deficient cells are sensitive to both oxidative damage as well as TBLs most likely causing the severe Cockayne Syndrome phenotype. In **absence of UVSSA**, CSB is still able to bind the DNA and probe Pol II for translocation. This will lead to successful bypass of oxidative damage similar as in wildtype cells. However, most likely due to absence of UVSSA, CSB will be degraded more rapidly. However, as CSB-induced lesion bypass of oxidative damage might be a swift process, it is therefore not expected to be influenced by the decreased CSB half-life. Since there is still CSA and CSB present in absence of UVSSA, lesion-stalled Pol II might be degraded or displaced from the lesion. This will prevent persistent stalling of Pol II and makes the lesion accessible for repair by alternative repair pathways like GG-NER. As a consequence, UVSSA-deficient cells are sensitive for TBLs but not oxidative damage, which might result in the milder UV-sensitive syndrome.

on UV-induced lesions to initiate TC-NER^{38,78}. However this stabilization might not be essential during the most likely more rapid lesion bypass of oxidative lesions (figure 3). A different but not mutually exclusive explanation for observed phenotypical differences between the TC-NER disorders might be that CSA and CSB are involved in degradation or removal of Pol II¹⁰⁶. Absence of CS proteins might lead to persistent stalling of Pol II with the lesion trapped in the active site thereby completely covering the lesion and preventing repair. In contrast, UV^S cells might still be able to remove Pol II from the lesion and repair UV-induced transcription blocking lesions with alternative repair pathways. In this scenario, 6-4PP lesions will be quickly repaired by GG-NER. However, CPD lesions are less efficiently repaired by GG-NER and their prevalence might explain the UV-sensitivity and failure to restart transcription as observed in UV^S cells (figure 3)⁷⁴.

To obtain a better understanding of the contribution of TBLs to the phenotypes of TC-NER linked disorders, genome instability and DNA damage-induced aging, it is of great importance to understand the exact mode of action of the various TC-NER factors during the different stages of this repair pathway. New insights in the spatio-temporal TC-NER complex composition might give important indications when the different activities of TC-NER factors are crucial, or alternatively, when specific factors need to be evicted from damaged chromatin to allow progression to subsequent reaction-steps. In addition, more in-depth experiments are needed to unravel the exact functions of TC-NER factors. For example, even though the E3 ligase activity of the CRL4^{CSA} complex is crucial for TC-NER, its exact role and substrates remain elusive. In addition to this important ubiquitin-mediated regulation⁴⁶, other post-translational modifications are expected to provide additional layers of control to allow efficient damage recognition and removal. For example, SUMOylation has recently been shown to target CSB and this modification is essential for efficient TC-NER³⁴. To be able to explain the striking differences observed in TC-NER phenotypes^{74,78,93}, it is important to identify putative differential activities of TC-NER factors or changes in the TC-NER complex composition following exposure to different types of TBLs. For example, recently it was shown that CSB is differentially ubiquitinated following UV-induced or oxidative lesions⁸¹, indicating that depending on the type of lesion, the TC-NER complex is regulated in a different manner.

Genome-wide consequences of transcription blocking lesions on transcription

Since TC-NER is initiated via recognition of lesion-stalled Pol II, most research on the effects of transcription-blocking DNA damage has been focused on this specific, damage-engaged subset of elongating polymerases^{25,90}. However, in addition to direct consequences of DNA lesions that impede Pol II forward progression - effects *in cis* -, accumulating evidence shows that several important regulatory mechanisms exist that also affect non-lesion stalled Pol II in a genome-wide manner - effects *in trans* -. Both these *cis* and *trans*-effects are expected to be vital for cells to cope with the severe consequences of TBLs^{85,107}. In this section, we will focus on genome-wide regulation of Pol II following DNA damage. We will discuss examples of TBL-induced effects that result in specific transcriptional programs, induced by either targeting the transcription cycle or by affecting mRNA splicing. Especially the highly regulated process of Pol II-mediated transcription offers several important control steps that can be targeted to regulate transcription, prevent genome instability and reduce cytotoxicity following exposure to TBLs.

Pol II-mediated transcription is initiated by general transcription factors that facilitate recruitment of Pol II and assembly of the preinitiation complex (figure 4)¹⁰⁸. During initiation, the CDK7 kinase activity of TFIIH phosphorylates serine 5 of the C-terminal domain (CTD) of RPB1, the core catalytic subunit of Pol II. This allows Pol II to engage the DNA template and start transcribing a short stretch of RNA followed by a transient pause ~60 bp downstream of the transcription start site (TSS) (figure 4)^{3,109}. The release of Pol II from promoter proximal pause sites into productive elongation is mediated by the CDK9 kinase activity of the positive transcription elongation factor b complex (p-TEFb). CDK9-mediated phosphorylation converts the DRB sensitivity inducing factor (DSIF) into a positive elongation factor, facilitates the release of the Negative

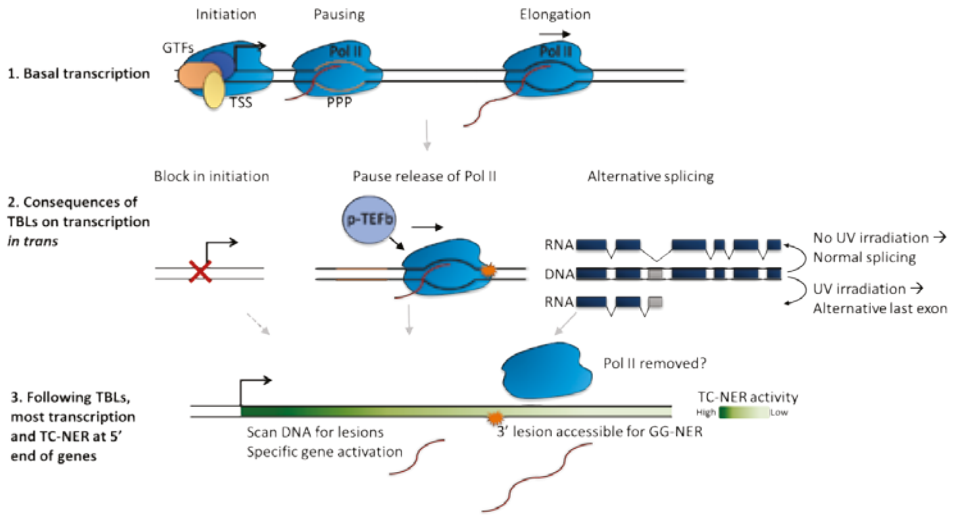


Figure 4. Effects of transcription blocking lesions *in trans*. (1) Basal transcription consists of initiation, pausing and elongation. General transcription factors help Pol II to initiate transcription from the transcription start site (TSS). Pol II pauses on the promoter proximal pause site (orange, PPP) before being released into productive elongation. (2) TBLs have different effects on transcription in a genome-wide manner. (2, left panel) Upon UV irradiation, transcription initiation is inhibited, decreasing the formation of new pre-initiation complexes, resulting in transcription inhibition. (2, middle panel) In addition, p-TEFb is activated independent of transcription. Activated p-TEFb stimulates release of Pol II from the PPP into productive elongation. This might result in increased sensing of TBLs and increased TC-NER. (2, right panel) Following UV-induced DNA damage, transcription switches to expression of shorter mRNA isoforms by using alternative last exons. (3) Both increased pause release of Pol II from the PPP into productive elongation, in combination with Pol II release from the DNA template following stalling at TBLs, as well as spatial restriction of transcription by the expression of shorter isoforms with alternative last exons, results in preferential transcription and TC-NER in the 5' of genes. This might leave the distal part of genes accessible for repair by GG-NER.

elongation factor (NELF) complex, and phosphorylates the CTD of RPB1 on serine 2³. The different phosphorylation statuses of Pol II also mediate the binding and release of splicing factors as previously reviewed¹¹⁰. The CTD serves as a 'landing path' for the spliceosome and mediates co-transcriptional splicing¹¹⁰.

Effects on Pol II transcription upon UV damage

One of the first indications of TBL-induced genome-wide effects on transcription was the observation that TATA-binding protein (TBP) is sequestered at cisplatin- and UV-damaged DNA. This results in a reduced availability of TBP to bind at promoter regions, subsequently leading to less transcription initiation¹¹¹. This observation was followed by the discovery of another *in trans* effect, that results in a massive depletion of the hypophosphorylated initiating form of Pol II upon UV irradiation, with a concomitant increase in the hyperphosphorylated elongating Pol II¹¹². Additional research indicated that this shift could be explained by the inhibition of transcription initiation that was detected in UV-treated cell extracts using *in vitro* assays¹¹². Although this indicates that DNA damage directly interferes with the transcription cycle, the loss of hypophosphorylated Pol II can also partially be explained by stalling of elongating Pol II at TBLs,

thereby increasing the fraction of hyperphosphorylated Pol II. In line with these early observations, more recently published genome-wide Pol II ChIP-seq data showed that directly following UV-induced DNA damage, Pol II was cleared from the promoter¹¹³. The loss of Pol II ChIP-seq reads near the promoter, which most likely represent promoter paused Pol II, can be explained in different ways: Either (1) promoter paused Pol II is released into the gene body, (2) paused Pol II is specifically evicted from the chromatin or (3) transcription initiation is inhibited.

As this study showed that in addition to the loss of Pol II, TFIID promoter-binding was reduced following UV irradiation, it was concluded that there is less TFIID available for transcription initiation¹¹³. Reduced availability of TFIID during transcription initiation might be caused by the involvement of this general transcription factor in the TC-NER reaction following DNA damage. In line with this hypothesis, reduced promoter binding of TFIID and Pol II following UV-induced DNA damage could be rescued by depletion of CSB¹¹³. This is an intriguing finding since the vast majority of repair-associated TFIID is active in GG-NER which makes up the bulk portion of NER¹¹⁴. This may suggest that specifically occupation of TC-NER-associated TFIID is causing inhibition of transcription initiation (figure 4). It is interesting to note that, although most of the genes show a decrease in promoter bound Pol II as a result of DNA damage, a specific subset is shown to be regulated differently. Interestingly, this set of genes is shown to have strongly increased binding of Pol II and consists of genes mainly involved in the p53 response, DNA damage response and apoptosis¹¹³, indicating that this mechanism can stimulate the expression of specific genes following TBL induction.

p-TEFb activation stimulates transcription of the 5' end of genes

Another genome-wide effect on the transcription cycle following UV-induced DNA damage is caused by activation of p-TEFb¹¹⁵. In unperturbed conditions, a pool of p-TEFb is kept in an inactive state in the 7SK snRNP complex¹¹⁶⁻¹²⁰. Upon UV-irradiation, most likely not as a direct effect of DNA damage but via the damage-induced Ca²⁺/calmodulin signaling pathway, p-TEFb is released from its inhibitory complex^{116,121}. It was recently shown that a rapid increase of active p-TEFb levels following UV exposure resulted in a wave-like release of promoter paused polymerases into productive elongation on almost all active genes (figure 4)¹²². This will increase the likelihood of Pol II encountering a lesion, thereby promoting lesion-recognition and initiation of TC-NER. Such a wave-like release might also partially explain the loss of promoter bound Pol II directly after UVB irradiation as described above¹¹³. Interestingly, such a *de novo* wave-release of Pol II into the gene body¹²² in combination with inhibited transcription initiation¹¹¹⁻¹¹³ might be indicative for a "final" round of transcription to 'sense' TBLs and swiftly initiate repair. Recently it was suggested that a significant population of lesion-stalled Pol II is released from the DNA template during TC-NER¹²³. This suggests that individual elongation complexes will not engage in multiple rounds of TC-NER on successive lesions and indicates that only the first encountered TBL in a gene will be recognized. Such a scenario would result in preferential repair of TBLs close to the transcription start site (figure 4). This hypothesis was supported by a meta-analysis of previously performed excision repair sequencing (XR-seq) data showing that most TC-NER is executed in the beginning of genes^{122,124}. This is an intriguing finding since TBLs located more downstream will still inhibit gene expression when not repaired properly.

Preferential repair of the 5' end of genes might suggest that the genetic information encoded in this region is especially important to preserve and might play an important role in the DNA damage response.

The role of gene size in response to DNA damage

The idea that the 5' end of genes is important is supported by the finding that transcription might be spatially restricted to the first 20 to 25 kb of a gene in response to UV¹²⁵. TBL-induced expression of shorter mRNAs is associated with a shift from expression of long mRNAs to shorter isoforms, thereby incorporating alternative last exons (ALEs) that are located closer to 5' end of the gene (figure 4). Expression of shorter isoforms will therefore result in damage-induced, altered gene expression¹²⁵. A key example of altered gene expression due to switching to short isoforms is the Activating Signal Cointegrator 1 Complex Subunit 3 (ASCC3). Following UV-damage, the long ASCC3 isoform (over 370 kb, 42 exons) is replaced by expression of the short ASCC3 isoform (25 kb) which only shares three exons with the long isoform and has a unique terminal exon. Interestingly, the long ASCC3 isoform encodes the ASCC3 protein that is involved in repression of transcription^{125,126}. In contrast, the short isoform is functioning as a non-coding RNA which is essential for proper transcription restart¹²⁵. These two isoforms have opposing effects on transcription and the changed balance upon UV irradiation can regulate transcription inhibition and restart. The importance of these shorter transcripts following TBL-induction might explain the observed preferential repair near 5' end of genes.

The concept that short genes are less susceptible to DNA damage is more commonly observed in the DNA damage response. The human genome consists of over 20.000 genes which vary greatly in gene size. Assuming that TBLs are genome-wide dispersed in a stochastic manner, long genes are more susceptible to gene inactivation due to direct stalling of Pol II on a TBL than shorter genes^{127,128}. A higher chance for TBLs in long genes compared to short genes, which will be especially relevant under physiologically relevant low damage loads, was linked to a shift in gene expression in favor of small genes upon DNA damage exposure^{127,129}. Possibly, gene size has been selected during evolution to maintain proficient expression of genes that are important for proper cellular responses following exposure to transcription-blocking DNA damage¹²⁹.

A clear example of a small gene that plays an important role following UV-induced DNA damage is the immediate early gene (IEG) activating transcription factor 3 (ATF3)¹³⁰⁻¹³³. IEGs are in general short genes and their expression can be rapidly induced following cellular stress¹³⁰. In line, also expression of ATF3 is strongly increased following UV-induced DNA damage¹³¹⁻¹³³. Upon expression, ATF3 is targeted to CRE/ATF-binding sites which are located near promoters of specific genes, thereby inhibiting specifically the expression of these genes^{131,132}. Importantly, this ATF3-mediated transcription inhibition *in trans* needs to be resolved to allow proper transcription restart as discussed below in the transcription restart section. In addition to ATF3, also BMI1 together with the E3 ligase UBR5 was recently shown to repress Pol II elongation and nascent RNA synthesis at UV-induced DNA lesions¹³⁴. Together, the above discussed examples show that different factors actively repress transcription *in trans* following DNA damage, either in general or by targeting a specific subset of genes.

Alternative splicing enhances DNA damage signaling

In addition to the above described selection of alternative last exons, which results in expression of smaller mRNAs following TBL induction, also other alternative splicing (AS) mechanisms play an important role. Damage-induced effects on splicing can induce specific gene expression programs or may contribute to damage-involved signaling. AS-induced expression of different isoforms following DNA damage, has thus far been attributed to changes either in Pol II elongation rate (kinetic coupling)^{135,136} or in the interaction between Pol II and the core spliceosome or splicing regulators (recruitment coupling)^{104,137,138}.

An example of the kinetic coupling model is nicely illustrated by the observed increase in phosphorylation of the C-terminal domain (CTD) of Pol II, which slows down transcription elongation¹³⁵. A reduced elongation rate can result in exon inclusion^{135,136} or skipping^{139,140}. UV-induced AS is a general mechanism regulating expression of many genes, including several genes specifically involved in regulation of survival and apoptosis. For example, the ratio between anti-apoptotic Bcl-xL and pro-apoptotic Bcl-xS, both isoforms of the Bcl-x gene, is shifted towards the Bcl-xS isoform contributing to a higher UV-induced cell death¹³⁵. Interestingly, it was shown that UV-induced AS also happens on non-damaged genes, indicative for an effect *in trans* on Pol II and co-transcriptional splicing¹³⁵.

In line with the finding that regulation of AS by kinetic coupling is regulated *in trans*, it was recently shown that ataxia telangiectasia and Rad3-related (ATR) signaling plays an important role in UV-induced hyperphosphorylation of the CTD of Pol II¹⁴¹. ATR is known to be activated upon UV-irradiation in non-cycling cells, following excision of the lesion-containing DNA during the XPF/ERCC1- and XPG-mediated excision step of NER. The residual repair-intermediate, containing a 22-35 nucleotide long piece of ssDNA, is bound by RPA which is subsequently recognized by ATR interacting protein (ATRIP)/ATR complexes thereby triggering ATR signaling^{74,142,143}. Of note, the study on ATR signaling used keratinocytes in which the majority of ssDNA intermediates will be generated via the much more active GG-NER sub-pathway¹¹⁴. This suggests that stalling of Pol II itself may not be the initiating event, but that Pol II is a target of the ATR-induced effect on elongation speed¹⁴¹.

Together these studies show that in addition to direct impediment of the forward translocation of elongating Pol II by the lesion^{13,16-21}, the elongation rate of Pol II is also affected *in trans* by ATR. A reduced transcription elongation rate will subsequently lead to AS events¹³⁶, resulting in induction of specific isoforms following DNA damage^{135,141}. Whether and how this ATR and DNA damage-induced Pol II hyperphosphorylation is different from the canonical hyperphosphorylated (Pol IIO) remains currently unknown. Of note, most likely ATR affects Pol II in an indirect manner as there are no target sequences known in the CTD for this DNA-damage kinase. Interestingly, in addition to reducing Pol II elongation rate, ATR activation may also affect different axillary factors involved in AS¹⁴⁴⁻¹⁴⁷.

In addition to the above described AS events according to the kinetic coupling model, TBLs also result in AS following the recruitment model^{104,137,138}. For example, UV

irradiation- or camptothecin-induced TBLs induce co-transcriptional exon skipping of for example the *MDM2*, *CHEK2* and *MAP4K2* transcripts^{137,138}. This damage-induced AS is linked to the loss of interaction between EWS, a member of the TET family of RNA and DNA-binding proteins, and its target RNAs¹³⁷. The loss of interaction between these co-transcriptional binding partners might be mediated by the lost interaction between EWS and the spliceosome-associated factor YB-1 upon DNA damage¹³⁸. Also in *Drosophila* cells, camptothecin induces AS which is mediated by ATR activation and results in proteasome-mediated degradation of splicing regulator Tra2¹⁴⁶. Interestingly, in addition to damage-induced AS by targeting splicing regulators, it was recently shown that also the core spliceosome is affected following the induction of TBLs. Pol II stalling on TBLs promotes chromatin displacement of late-stage spliceosomes¹⁰⁴, composed of U2, U5 and U6 small nuclear ribonucleoproteins¹⁴⁸, and initiate a positive feedback loop centered on the signaling kinase ATM. The initial spliceosome displacement results in an increased R-loop formation through hybridization of pre-mRNA with template DNA^{104,149}. Interestingly, R-loop formation near the TBL leads to a non-canonical activation of the protein kinase ATM, which signals to impede core spliceosome organization even further, consequently resulting in increased intron retention and altered splicing in a genome-wide manner^{104,150}.

Regulation of transcription restart

Even when TBLs are successfully repaired, their induction is expected to remain highly cytotoxic if transcription is not properly resumed. Therefore, transcription restart is important to assure proper *de novo* mRNA production and to maintain cellular homeostasis⁸⁵. Transcription restart was thus far mostly assumed to reconvene at the lesion where Pol II was stalled, as soon as the TBL is removed by TC-NER^{25,151}. However, as discussed above, also a significant part of TBL-induced transcription inhibition is caused by effects *in trans* that have direct consequences for the mode of transcription restart. For example, if the TBL-induced block in transcription initiation is reversed¹¹¹⁻¹¹³, this will result in transcription restart by new initiation events at the promoter. In line with these findings, genome-wide analyses of nascent RNA sequencing data¹⁵² showed that transcription recovery of RNA synthesis occurred as a wave in the 5'-3' direction following UV- or camptothecin-induced TBLs. This indicates that a significant part of transcription restarts at the beginning of genes^{127,128}, which could provide time for the GG-NER pathway to remove TBLs at the more distal parts of genes before the transcription machinery encounters these lesions¹²⁷. This could result in a smooth progression of transcription once restarted. Transcription restart at the promoter might be caused by recovery of *in trans*-mediated transcription initiation blockage but may also indicate that transcription does not always resume from the position where it was initially stalled (*in cis*). In line with the latter, it was recently suggested that a significant population of lesion-stalled Pol II is released from the DNA template during the TC-NER reaction¹²³, indicating that the observed transcription resumption from the 5' end of genes might also be a common mechanism for transcription inhibition *in cis*.

Of note, these findings do not exclude that transcription can also be restarted at the site of the lesion itself which is in line with several factors suggested to be involved in backtracking or 3'-end RNA processing^{25,107}. Furthermore, transcription resumption directly at the repaired lesion seems to be the most efficient restart mechanism

since the same Pol II complex is able to continue with transcription of the already partially synthesized mRNA.

Factors involved in transcription restart

Thus far, several factors have been suggested to be involved in backtracking of Pol II to allow repair. To resume transcription from this backtracked position it is crucial that protruding nascent RNA is cleaved so that the 3' end of the RNA is properly realigned with the DNA in the active site of Pol II¹⁵³. This reaction is mediated by transcription factor II S (TFIIS) which is suggested to stimulate the intrinsic 3'–5' exonuclease activity of Pol II^{21,153–155}. TFIIS is recruited in a CSA- and CSB-dependent manner³³. The Ccr4-Not complex further supports recruitment of TFIIS and enhances its cleaving activity thereby suggesting that TFIIS and Ccr4–Not jointly reactivate arrested Pol II¹⁵⁶. Absence of TFIIS results in significantly decreased, but not completely absent, transcription resumption. This can be explained by the fact that remaining intrinsic cleavage activity of Pol II is sufficient to cleave the RNA and restart transcription. Even though this restart will happen in a less efficient manner, it is most likely sufficient to prevent increased UV-sensitivity upon depletion of TFIIS^{153,157,158}. The intrinsic cleavage activity of Pol II might also explain the observed differences in the role of TFIIS in transcription resumption^{157,158}, but this may also be caused by the presence of the redundant TFIIS paralogue TCEA2, which is not solely expressed in the testis as originally described¹⁵⁹. In addition to its stimulating function on TFIIS, Ccr4-Not was suggested to directly promote transcription elongation by binding to the emerging transcript, thereby stimulating Pol II to resume transcription after repair of a lesion¹⁶⁰. Despite that TFIIS and Ccr4-Not are implied in Pol II backtracking to allow resumption of transcription, it cannot be excluded that their role in facilitating RNA cleavage of the protruding RNA may also be involved in release of Pol II from chromatin following TC-NER initiation¹²³.

Another factor which was shown to stimulate transcription restart is elongation factor ELL which binds to TFIIF via the CDK7 subunit of the CDK-activating kinase (CAK) complex. Thus far, the exact function of ELL during transcription restart remains unknown. However, it was hypothesized that ELL functions as a docking protein, thereby enabling other proteins to bind and stimulate transcription resumption. ELL was shown to be specifically involved in transcription resumption since depletion of ELL does not affect TC-NER¹⁶¹. This uncoupling of repair and transcription restart might also indicate that ELL is involved in transcription restart at the TSS as a consequence of damage-induced transcription inhibition *in trans*¹⁶¹.

Another important process during transcription restart is resolving ATF3-mediated transcription inhibition, which has been shown to affect approximately 5000 genes^{131,132}. Interestingly, CSB and the CRL4^{CSA} E3 ligase complex, possibly together with the E3 ubiquitin ligase Mdm2, were shown to ubiquitylate ATF3¹³¹. The subsequent proteasomal degradation of ATF3 at CRE/ATF sites relieves the ATF3-mediated transcriptional repression. This finding also suggests that on top of absence of functional TC-NER in CSA and CSB-deficient cells, loss of transcription restart might be partially explained by maintained transcriptional repression of ATF3-regulated genes and consequently contributes to the severe phenotype observed in CS patients^{131,132}. Interestingly, in addition to the well-described function of CSB on stalled Pol II, this finding suggests

that CSB has an additional function at the promoter. This might be in line earlier with observations that CSB is involved in regulating gene expression⁹⁴.

Chromatin remodeling factors involved in Pol II restart

Histone chaperones and ATP-dependent chromatin remodelers are responsible for histone sliding, eviction and insertion to remodel chromatin and facilitate different DNA transacting processes. Several remodeling factors were identified to be specifically involved in NER-mediated repair¹⁶². In addition to their role in repair, several of these chromatin involved factors were also shown to play a key role in the restart of transcription^{163,164}. For example, nucleosome binding protein HMGN1 and p300 histone-acetyl transferase were shown to be recruited to the TC-NER complex³³. These factors are hypothesized to induce sliding of upstream nucleosomes resulting in a more open chromatin structure, which might facilitate Pol II backtracking^{25,165,166}.

In addition, the histone chaperones Histone regulator A (HIRA) and Facilitating Chromatin Transcription (FACT) were identified to remodel histones following TBL induction, a process that was shown to be essential for transcription restart^{167,168}. HIRA is recruited to sites of DNA damage where this histone chaperone deposits histone variant H3.3 near the damage¹⁶⁷. Histone 3.3 is normally involved in promoting transcription or removing inhibitory factors via specific marks¹⁶⁹ and might promote transcription restart via this mechanism. The other chaperone, FACT, consists of the SPT16 and SSRP1 subunits and was shown to exchange histone H2A and H2B which stimulates Pol II transcription along chromatin by destabilizing nucleosomes¹⁷⁰. Interestingly, only the SPT16 subunit of the FACT complex was shown to be important for transcription restart, in line with SPT16-dependent accelerated H2A/H2B exchange at the site of damage¹⁶⁸. This suggests that SPT16 increases plasticity of chromatin via enhanced incorporation of histone H2A/H2B¹⁷¹, thereby promoting translocation of Pol II either to enhance repair, reverse translocate Pol II from the lesion or restart transcription after repair¹⁶⁸. In addition to these two histone chaperones, proteins involved in post-translational modifications of histones may promote transcriptional restart. For example, the lysine methyltransferase DOT1L normally methylates H3K79, a histone mark that regulates transcription^{172,173}. DOT1L knock-out cells show increased UV-sensitivity coupled to a deficient recovery of transcription restart following TBL induction without affecting TC-NER¹⁷⁴. These DOT1L effects can be rescued by treatment with Trichostatin A, which relaxes the chromatin structure, suggesting that DOT1L promotes transcription initiation by opening up chromatin of UV-repressed genes and is therefore essential for transcription restart¹⁷⁴. Interestingly, HIRA and DOT1L are not necessary for transcription restart in response to the reversible transcription inhibitor DRB^{167,174}, indicative for a specific regulation of transcription restart following removal of TBLs. In line with this additional layer of control during the cellular response to TBLs, it was observed that transcription restart can be regulated in a gene-specific manner¹²⁷.

Outlook

The development of new sequencing approaches, including nascent RNA-seq¹²⁷, XR-seq¹²⁴ and ChIP-seq^{113,122}, has resulted in important new insights in the underlying mechanism of transcription inhibition in response to UV-induced DNA damage. In addition to the direct physical block of Pol II once encountering a TBL (inhibition *in cis*), accumulating

evidence shows that the highly regulated transcription cycle is targeted at different key steps to efficiently induce transcription inhibition in a genome-wide manner (inhibition *in trans*). Examples are inhibition of transcription initiation^{112,113} or induction of immediate response genes (IEGs) like ATF3 that inhibit transcription by binding to their response elements near promoters¹³¹⁻¹³³. Most likely many more cellular processes that are involved in transcription inhibition following TBL induction await their discovery. One of the main questions that remains, is why transcription is inhibited in a genome-wide manner, while cells are equipped with a highly efficient TC-NER pathway that can directly resolve lesion-stalled Pol II. Apparently, additional back-up mechanisms are required for proper cell survival following DNA damage. It is tempting to speculate that these mechanisms have evolved to prevent persistent Pol II stalling, as lesion-stalled Pol II repair intermediates might be more toxic for a cell than the actual TBL itself, for example due to the induction of R-loops^{7,8} or transcription-replication collisions⁶. Importantly, by affecting a specific subset of genes, these genome-wide regulatory systems allow cells to regulate transcription of genes essential for cells to cope with TBLs. As a consequence, TC-NER activity is focussed on repair of these important genes that are not inhibited *in trans*.

It has been shown that different types of TBLs have different outcomes on the impediment of Pol II *in cis*¹⁵. Therefore, it seems logical that also transcription inhibition *in trans* could be differentially affected depending on the type of damage. In line with this idea, recently a rapid accumulation of Pol II near promoters and enhancers was observed following oxidative damage¹⁷⁵. This is in sharp contrast to rapid release of paused Pol II into productive elongation observed following UV-induced DNA damage¹²². The differential response between these types of DNA damage could be explained by differences in the activation mechanisms of these pathways. Some of the *in trans* effects are induced by the initial stalling of Pol II at a lesion, including release of the core spliceosome and subsequent ATM activation¹⁰⁴. In contrast, UV-induced p-TEFb activation by release from its inhibitory complex is a direct consequence of the UV-damage and is thus activated independent of transcription¹²². To obtain a better insight in the biological relevance of the mechanism of transcription inhibition, it is important to study whether different types of TBLs have different outcomes on transcription inhibition *in trans*. In addition, TBLs may also have strikingly different outcomes in different cell types and organs^{74,176,177}. This is clearly illustrated by extreme damage-sensitivity of photoreceptor cells in retinas of TC-NER-deficient mice or neurodegeneration in CS patients^{178,179}. These differential cellular outcomes to TBLs may be explained by differences in transcription levels, replication or activity of DNA-repair pathways. However, also the presence of multiple mechanisms to inhibit transcription might explain the differential response and sensitivity of different tissues to TBLs.

Improved insights in different modes of transcription inhibition, both *in cis* and *in trans*, is crucial to understand the molecular mechanism of transcription restart. For example, a large contribution of transcription inhibition *in trans* will most likely result in restart from the TSS, while stalling of Pol II *in cis* might result in transcripton resumption by the same Pol II and RNA molecule that was stalled at the TBL. Even though transcription will be restored in both cases, the molecular mechanisms to restart transcription and the involved factors will be completely different for restart at the beginning of genes compared to transcription resumption at the TBL.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgements

This work is part of the Oncode Institute which is partly financed by the Dutch Cancer Society and was funded by grant 10506/2016-1 from the Dutch Cancer Society. Dutch Organization for Scientific Research ZonMW TOP Grant [912.12.132]; VIDI ALW [864.13.004]; Erasmus MC fellowship.

References

1. Arimbasseri, A.G., Rijal, K. & Maraia, R.J. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* **5**, e27639 (2014).
2. Shandilya, J. & Roberts, S.G.E. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1819**, 391-400 (2012).
3. Jonkers, I. & Lis, J.T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-77 (2015).
4. Doetsch, P.W. Translesion synthesis by RNA polymerases: occurrence and biological implications for transcriptional mutagenesis. *Mutat Res* **510**, 131-40 (2002).
5. Marietta, C. & Brooks, P.J. Transcriptional bypass of bulky DNA lesions causes new mutant RNA transcripts in human cells. *EMBO Rep* **8**, 388-93 (2007).
6. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell* **46**, 115-24 (2012).
7. Helmrich, A., Ballarino, M., Nudler, E. & Tora, L. Transcription-replication encounters, consequences and genomic instability. *Nat Struct Mol Biol* **20**, 412-8 (2013).
8. Ljungman, M. & Zhang, F. Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis. *Oncogene* **13**, 823-31 (1996).
9. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-74 (2001).
10. Barnes, D.E. & Lindahl, T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet* **38**, 445-76 (2004).
11. Dizdaroglu, M. Chemical determination of free radical-induced damage to DNA. *Free Radical Biology and Medicine* **10**, 225-242 (1991).
12. Wang, Y. Bulky DNA lesions induced by reactive oxygen species. *Chem Res Toxicol* **21**, 276-81 (2008).
13. Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. *Science* **315**, 859-62 (2007).
14. Gates, K.S. An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol* **22**, 1747-60 (2009).
15. Shin, J.H., Xu, L. & Wang, D. Mechanism of transcription-coupled DNA modification recognition. *Cell Biosci* **7**, 9 (2017).
16. 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* **39**, 5926-34 (2011).
17. Kathe, S.D., Shen, G.P. & Wallace, S.S. Single-stranded breaks in DNA but not oxidative DNA base damages block transcriptional elongation by RNA polymerase II in HeLa cell nuclear extracts. *J Biol Chem* **279**, 18511-20 (2004).
18. Yanamadala, S. & Ljungman, M. Potential role of MLH1 in the induction of p53 and apoptosis by blocking transcription on damaged DNA templates. *Mol Cancer Res* **1**, 747-54 (2003).
19. Damsma, G.E., Alt, A., Brueckner, F., Carell, T. & Cramer, P. Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* **14**, 1127-33 (2007).
20. Tornaletti, S. Transcription arrest at DNA damage sites. *Mutat Res* **577**, 131-45 (2005).
21. Tornaletti, S., Reines, D. & Hanawalt, P.C. Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. *J Biol Chem* **274**, 24124-30 (1999).

22. Selby, C.P. & Sancar, A. Cockayne syndrome group B protein enhances elongation by RNA polymerase II. *Proc Natl Acad Sci U S A* **94**, 11205-9 (1997).
23. Walmacq, C. *et al.* Mechanism of RNA polymerase II bypass of oxidative cyclopurine DNA lesions. *Proc Natl Acad Sci U S A* **112**, E410-9 (2015).
24. Charlet-Berguerand, N. *et al.* RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *Embo J* **25**, 5481-91 (2006).
25. Hanawalt, P.C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol* **9**, 958-70 (2008).
26. Bohr, V.A., Smith, C.A., Okumoto, D.S. & Hanawalt, P.C. 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* **40**, 359-69 (1985).
27. Mellon, I., Spivak, G. & Hanawalt, P.C. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**, 241-9 (1987).
28. Henning, K.A. *et al.* The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. *Cell* **82**, 555-64 (1995).
29. Troelstra, C. *et al.* ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell* **71**, 939-53 (1992).
30. Schwertman, P. *et al.* UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat Genet* **44**, 598-602 (2012).
31. Kamiuchi, S. *et al.* Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proc Natl Acad Sci U S A* **99**, 201-6 (2002).
32. Saijo, M. *et al.* Functional TFIIH is required for UV-induced translocation of CSA to the nuclear matrix. *Mol Cell Biol* **27**, 2538-47 (2007).
33. Foustieri, M., Vermeulen, W., van Zeeland, A.A. & Mullenders, L.H. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol Cell* **23**, 471-82 (2006).
34. Sin, Y., Tanaka, K. & Saijo, M. The C-terminal Region and SUMOylation of Cockayne Syndrome Group B Protein Play Critical Roles in Transcription-coupled Nucleotide Excision Repair. *J Biol Chem* **291**, 1387-97 (2016).
35. Lake, R.J., Geyko, A., Hemashettar, G., Zhao, Y. & Fan, H.Y. UV-induced association of the CSB remodeling protein with chromatin requires ATP-dependent relief of N-terminal autorepression. *Mol Cell* **37**, 235-46 (2010).
36. Citterio, E. *et al.* Biochemical and biological characterization of wild-type and ATPase-deficient Cockayne syndrome B repair protein. *J Biol Chem* **273**, 11844-51 (1998).
37. Muftuoglu, M., Selzer, R., Tuo, J., Brosh, R.M., Jr. & Bohr, V.A. Phenotypic consequences of mutations in the conserved motifs of the putative helicase domain of the human Cockayne syndrome group B gene. *Gene* **283**, 27-40 (2002).
38. Xu, J. *et al.* Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653-657 (2017).
39. Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. **544**, 440 (2017).
40. van den Boom, V. *et al.* DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J Cell Biol* **166**, 27-36 (2004).
41. Groisman, R. *et al.* The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**, 357-67 (2003).

42. Fischer, E.S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024-39 (2011).
43. Zhou, H.-X. & Wang, G. Predicted structures of two proteins involved in human diseases. *Cell Biochemistry and Biophysics* **35**, 35-47 (2001).
44. Stirnimann, C.U., Petsalaki, E., Russell, R.B. & Muller, C.W. WD40 proteins propel cellular networks. *Trends Biochem Sci* **35**, 565-74 (2010).
45. Pines, A., *et al.* (Nature Communications, 2018).
46. van Cuijk, L., Vermeulen, W. & Marteijn, J.A. Ubiquitin at work: the ubiquitous regulation of the damage recognition step of NER. *Exp Cell Res* **329**, 101-9 (2014).
47. Jin, J. & Li, J.-M. CRL Ubiquitin Ligases and DNA Damage Response. *Frontiers in Oncology* **2**(2012).
48. Groisman, R. *et al.* CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes Dev* **20**, 1429-34 (2006).
49. Fei, J. & Chen, J. KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). *J Biol Chem* **287**, 35118-26 (2012).
50. Zhang, X. *et al.* Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nat Genet* **44**, 593-7 (2012).
51. Meulmeester, E. *et al.* Loss of HAUSP-mediated deubiquitination contributes to DNA damage-induced destabilization of Hdmx and Hdm2. *Mol Cell* **18**, 565-76 (2005).
52. He, J. *et al.* Ubiquitin-specific protease 7 regulates nucleotide excision repair through deubiquitinating XPC protein and preventing XPC protein from undergoing ultraviolet light-induced and VCP/p97 protein-regulated proteolysis. *J Biol Chem* **289**, 27278-89 (2014).
53. Li, M. *et al.* Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* **416**, 648-53 (2002).
54. Zhao, Y. *et al.* Noncanonical regulation of alkylation damage resistance by the OTUD4 deubiquitinase. *Embo J* **34**, 1687-703 (2015).
55. Nicholson, B. & Suresh Kumar, K.G. The multifaceted roles of USP7: new therapeutic opportunities. *Cell Biochem Biophys* **60**, 61-8 (2011).
56. Kim, R.Q. & Sixma, T.K. Regulation of USP7: A High Incidence of E3 Complexes. *Journal of Molecular Biology* **429**, 3395-3408 (2017).
57. Nijman, S.M. *et al.* A genomic and functional inventory of deubiquitinating enzymes. *Cell* **123**, 773-86 (2005).
58. Hu, M. *et al.* Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**, 1041-54 (2002).
59. Faesen, A.C. *et al.* Mechanism of USP7/HAUSP activation by its C-terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Mol Cell* **44**, 147-59 (2011).
60. Kim, R.Q., van Dijk, W.J. & Sixma, T.K. Structure of USP7 catalytic domain and three Ubl-domains reveals a connector alpha-helix with regulatory role. *J Struct Biol* **195**, 11-8 (2016).
61. Higa, M., Zhang, X., Tanaka, K. & Saijo, M. Stabilization of Ultraviolet (UV)-stimulated Scaffold Protein A by Interaction with Ubiquitin-specific Peptidase 7 Is Essential for Transcription-coupled Nucleotide Excision Repair. *J Biol Chem* **291**, 13771-9 (2016).
62. Higa, M., Tanaka, K. & Saijo, M. Inhibition of UVSSA ubiquitination suppresses transcription-coupled nucleotide excision repair deficiency caused by dissociation from USP7. *Febs J* (2018).
63. 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* **44**, 586-92 (2012).

64. Okuda, M., Nakazawa, Y., Guo, C., Ogi, T. & Nishimura, Y. Common TFIIH recruitment mechanism in global genome and transcription-coupled repair subpathways. *Nucleic Acids Res* **45**, 13043-13055 (2017).
65. Lafrance-Vanasse, J., Arseneault, G., Cappadocia, L., Legault, P. & Omichinski, J.G. Structural and functional evidence that Rad4 competes with Rad2 for binding to the Tfb1 subunit of TFIIH in NER. *Nucleic Acids Res* **41**, 2736-45 (2013).
66. Compe, E. & Egly, J.M. TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* **13**, 343-54 (2012).
67. Winkler, G.S. *et al.* TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *J Biol Chem* **275**, 4258-66 (2000).
68. Li, C.-L. *et al.* Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Molecular Cell* **59**, 1025-1034 (2015).
69. Oksenyich, V., Bernardes de Jesus, B., Zhovmer, A., Egly, J.M. & Coin, F. Molecular insights into the recruitment of TFIIH to sites of DNA damage. *Embo J* **28**, 2971-80 (2009).
70. Marteijn, Jurgen A., Hoeijmakers, Jan H.J. & Vermeulen, W. Check, Check & Triple Check: Multi-Step DNA Lesion Identification by Nucleotide Excision Repair. *Molecular Cell* **59**, 885-886.
71. Sugawara, K., Akagi, J.-i., Nishi, R., Iwai, S. & Hanaoka, F. Two-Step Recognition of DNA Damage for Mammalian Nucleotide Excision Repair: Directional Binding of the XPC Complex and DNA Strand Scanning. *Molecular Cell* **36**, 642-653 (2009).
72. Camenisch, U., Dip, R., Schumacher, S.B., Schuler, B. & Naegeli, H. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nat Struct Mol Biol* **13**, 278-84 (2006).
73. de Laat, W.L. *et al.* DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* **12**, 2598-609 (1998).
74. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* **15**, 465-81 (2014).
75. Riedl, T., Hanaoka, F. & Egly, J.M. The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J* **22**, 5293-303 (2003).
76. Volker, M. *et al.* Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* **8**, 213-24 (2001).
77. Damsma, G.E. & Cramer, P. Molecular basis of transcriptional mutagenesis at 8-oxoguanine. *J Biol Chem* **284**, 31658-63 (2009).
78. Schwertman, P., Vermeulen, W. & Marteijn, J.A. UVSSA and USP7, a new couple in transcription-coupled DNA repair. *Chromosoma* **122**, 275-84 (2013).
79. Wei, L. *et al.* BRCA1 contributes to transcription-coupled repair of DNA damage through polyubiquitination and degradation of Cockayne syndrome B protein. *Cancer Sci* **102**, 1840-7 (2011).
80. Anindya, R. *et al.* A ubiquitin-binding domain in Cockayne syndrome B required for transcription-coupled nucleotide excision repair. *Mol Cell* **38**, 637-48 (2010).
81. Ranes, M. *et al.* A ubiquitylation site in Cockayne syndrome B required for repair of oxidative DNA damage, but not for transcription-coupled nucleotide excision repair. *Nucleic Acids Research* **44**, 5246-5255 (2016).
82. Wilson, M.D., Harreman, M. & Svejstrup, J.Q. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim Biophys Acta* **1829**, 151-7 (2013).
83. Woudstra, E.C. *et al.* A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature* **415**, 929-33 (2002).

84. Svejstrup, J.Q. Rescue of arrested RNA polymerase II complexes. *J Cell Sci* **116**, 447-51 (2003).
85. Steurer, B. & Marteijn, J.A. Traveling Rocky Roads: The Consequences of Transcription-Blocking DNA Lesions on RNA Polymerase II. *J Mol Biol* **429**, 3146-3155 (2017).
86. Anindya, R., Aygun, O. & Svejstrup, J.Q. Damage-induced ubiquitylation of human RNA polymerase II by the ubiquitin ligase Nedd4, but not Cockayne syndrome proteins or BRCA1. *Mol Cell* **28**, 386-97 (2007).
87. Harreman, M. *et al.* Distinct ubiquitin ligases act sequentially for RNA polymerase II polyubiquitylation. *Proc Natl Acad Sci U S A* **106**, 20705-10 (2009).
88. Laugel, V. Cockayne syndrome: the expanding clinical and mutational spectrum. *Mech Ageing Dev* **134**, 161-70 (2013).
89. Nance, M.A. & Berry, S.A. Cockayne syndrome: review of 140 cases. *Am J Med Genet* **42**, 68-84 (1992).
90. Vermeulen, W. & Foustieri, M. Mammalian transcription-coupled excision repair. *Cold Spring Harb Perspect Biol* **5**, a012625 (2013).
91. Karikkineth, A.C., Scheibye-Knudsen, M., Fivenson, E., Croteau, D.L. & Bohr, V.A. Cockayne syndrome: Clinical features, model systems and pathways. *Ageing Res Rev* **33**, 3-17 (2017).
92. Spivak, G. *et al.* Ultraviolet-sensitive syndrome cells are defective in transcription-coupled repair of cyclobutane pyrimidine dimers. *DNA Repair (Amst)* **1**, 629-43 (2002).
93. Spivak, G. UV-sensitive syndrome. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **577**, 162-169 (2005).
94. Wang, Y. *et al.* Dysregulation of gene expression as a cause of Cockayne syndrome neurological disease. *Proc Natl Acad Sci U S A* **111**, 14454-9 (2014).
95. Proietti-De-Santis, L., Drane, P. & Egly, J.M. Cockayne syndrome B protein regulates the transcriptional program after UV irradiation. *EMBO J* **25**, 1915-23 (2006).
96. Pascucci, B. *et al.* An altered redox balance mediates the hypersensitivity of Cockayne syndrome primary fibroblasts to oxidative stress. *Aging Cell* **11**, 520-9 (2012).
97. Sakai, A., Sakasai, R., Kakeji, Y., Kitao, H. & Maehara, Y. PARP and CSB modulate the processing of transcription-mediated DNA strand breaks. *Genes Genet Syst* **87**, 265-72 (2012).
98. Batenburg, N.L., Thompson, E.L., Hendrickson, E.A. & Zhu, X.D. Cockayne syndrome group B protein regulates DNA double-strand break repair and checkpoint activation. *EMBO J* **34**, 1399-416 (2015).
99. Aamann, M.D. *et al.* Cockayne syndrome group B protein promotes mitochondrial DNA stability by supporting the DNA repair association with the mitochondrial membrane. *The FASEB Journal* **24**, 2334-2346 (2010).
100. 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* **207**, 379-90 (2010).
101. Hoeijmakers, J.H. DNA damage, aging, and cancer. *N Engl J Med* **361**, 1475-85 (2009).
102. Stevnsner, T., Muftuoglu, M., Aamann, M.D. & Bohr, V.A. The role of Cockayne Syndrome group B (CSB) protein in base excision repair and aging. *Mech Ageing Dev* **129**, 441-8 (2008).
103. Spivak, G. & Hanawalt, P.C. Host cell reactivation of plasmids containing oxidative DNA lesions is defective in Cockayne syndrome but normal in UV-sensitive syndrome fibroblasts. *DNA Repair (Amst)* **5**, 13-22 (2006).
104. Tresini, M. *et al.* The core spliceosome as target and effector of non-canonical ATM signalling. *Nature* **523**, 53-8 (2015).
105. Guo, J., Hanawalt, P.C. & Spivak, G. Comet-FISH with strand-specific probes reveals transcription-coupled repair of 8-oxoGuanine in human cells. *Nucleic Acids Research* **41**, 7700-7712 (2013).

106. Bregman, D.B. *et al.* UV-induced ubiquitination of RNA polymerase II: a novel modification deficient in Cockayne syndrome cells. *Proc Natl Acad Sci U S A* **93**, 11586-90 (1996).
107. Mullenders, L. DNA damage mediated transcription arrest: Step back to go forward. *DNA Repair (Amst)* **36**, 28-35 (2015).
108. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 129-43 (2015).
109. Adelman, K. & Lis, J.T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13**, 720-31 (2012).
110. Bentley, D.L. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* **15**, 163-75 (2014).
111. Vichi, P. *et al.* Cisplatin- and UV-damaged DNA lure the basal transcription factor TFIID/TBP. *EMBO J* **16**, 7444-56 (1997).
112. Rockx, D.A. *et al.* UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II. *Proc Natl Acad Sci U S A* **97**, 10503-8 (2000).
113. 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* **10**, e1004483 (2014).
114. D'Errico, M., Lemma, T., Calcagnile, A., Proietti De Santis, L. & Dogliotti, E. Cell type and DNA damage specific response of human skin cells to environmental agents. *Mutat Res* **614**, 37-47 (2007).
115. Marshall, N.F., Peng, J., Xie, Z. & Price, D.H. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem* **271**, 27176-83 (1996).
116. Nguyen, V.T., Kiss, T., Michels, A.A. & Bensaude, O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**, 322-5 (2001).
117. Yang, Z., Zhu, Q., Luo, K. & Zhou, Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**, 317-22 (2001).
118. Michels, A.A. *et al.* Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J* **23**, 2608-19 (2004).
119. Yik, J.H. *et al.* Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell* **12**, 971-82 (2003).
120. Byers, S.A., Price, J.P., Cooper, J.J., Li, Q. & Price, D.H. HEXIM2, a HEXIM1-related protein, regulates positive transcription elongation factor b through association with 7SK. *J Biol Chem* **280**, 16360-7 (2005).
121. Chen, R. *et al.* PP2B and PP1 α cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca(2+) signaling. *Genes & Development* **22**, 1356-1368 (2008).
122. Lavigne, M.D., Konstantopoulos, D., Ntakou-Zamplara, K.Z., Liakos, A. & Foustieri, M. Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate. *Nat Commun* **8**, 2076 (2017).
123. Chiou, Y.Y., Hu, J., Sancar, A. & Selby, C.P. RNA polymerase II is released from the DNA template during transcription-coupled repair in mammalian cells. *J Biol Chem* **293**, 2476-2486 (2018).
124. Hu, J., Adar, S., Selby, C.P., Lieb, J.D. & Sancar, A. Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. *Genes Dev* **29**, 948-60 (2015).
125. Williamson, L. *et al.* UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. *Cell* **168**, 843-855 e13 (2017).
126. Li, J. *et al.* A short hairpin RNA screen of interferon-stimulated genes identifies a novel negative regulator of the cellular antiviral response. *MBio* **4**, e00385-13 (2013).

127. Andrade-Lima, L.C., Veloso, A., Paulsen, M.T., Menck, C.F. & Ljungman, M. DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes. *Nucleic Acids Res* **43**, 2744-56 (2015).
128. Veloso, A. *et al.* Genome-wide transcriptional effects of the anti-cancer agent camptothecin. *PLoS One* **8**, e78190 (2013).
129. McKay, B.C. *et al.* Regulation of ultraviolet light-induced gene expression by gene size. *Proc Natl Acad Sci U S A* **101**, 6582-6 (2004).
130. Bahrami, S. & Drabløs, F. Gene regulation in the immediate-early response process. *Advances in Biological Regulation* **62**, 37-49 (2016).
131. Epanchintsev, A. *et al.* Cockayne's Syndrome A and B Proteins Regulate Transcription Arrest after Genotoxic Stress by Promoting ATF3 Degradation. *Mol Cell* **68**, 1054-1066 e6 (2017).
132. 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* **110**, E2261-70 (2013).
133. Cui, H. *et al.* The Stress-responsive Gene ATF3 Mediates Dichotomous UV Responses by Regulating the Tip60 and p53 Proteins. *J Biol Chem* **291**, 10847-57 (2016).
134. Sanchez, A. *et al.* BMI1-UBR5 axis regulates transcriptional repression at damaged chromatin. *Proc Natl Acad Sci U S A* **113**, 11243-11248 (2016).
135. Muñoz, M.J. *et al.* DNA Damage Regulates Alternative Splicing through Inhibition of RNA Polymerase II Elongation. *Cell* **137**, 708-720 (2009).
136. de la Mata, M. *et al.* A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* **12**, 525-32 (2003).
137. Paronetto, M.P., Minana, B. & Valcarcel, J. The Ewing sarcoma protein regulates DNA damage-induced alternative splicing. *Mol Cell* **43**, 353-68 (2011).
138. Dutertre, M. *et al.* Cotranscriptional exon skipping in the genotoxic stress response. *Nat Struct Mol Biol* **17**, 1358-66 (2010).
139. Dujardin, G. *et al.* How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Molecular Cell* **54**, 683-690 (2014).
140. Fong, N. *et al.* Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev* **28**, 2663-76 (2014).
141. Munoz, M.J. *et al.* Major Roles for Pyrimidine Dimers, Nucleotide Excision Repair, and ATR in the Alternative Splicing Response to UV Irradiation. *Cell Rep* **18**, 2868-2879 (2017).
142. Hanasoge, S. & Ljungman, M. H2AX phosphorylation after UV irradiation is triggered by DNA repair intermediates and is mediated by the ATR kinase. *Carcinogenesis* **28**, 2298-304 (2007).
143. 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* **186**, 835-47 (2009).
144. Wickramasinghe, V.O. & Venkitaraman, A.R. RNA Processing and Genome Stability: Cause and Consequence. *Mol Cell* **61**, 496-505 (2016).
145. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160-6 (2007).
146. Katzenberger, R.J., Marengo, M.S. & Wassarman, D.A. Control of alternative splicing by signal-dependent degradation of splicing-regulatory proteins. *J Biol Chem* **284**, 10737-46 (2009).
147. Katzenberger, R.J., Marengo, M.S. & Wassarman, D.A. ATM and ATR pathways signal alternative splicing of Drosophila TAF1 pre-mRNA in response to DNA damage. *Mol Cell Biol* **26**, 9256-67 (2006).
148. Valadkhan, S. & Jaladat, Y. The spliceosomal proteome: at the heart of the largest cellular ribonucleoprotein machine. *Proteomics* **10**, 4128-41 (2010).

149. Sordet, O. *et al.* Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks. *EMBO Rep* **10**, 887-93 (2009).
150. Tresini, M., Marteijn, J.A. & Vermeulen, W. Bidirectional coupling of splicing and ATM signaling in response to transcription-blocking DNA damage. *RNA Biol* **13**, 272-8 (2016).
151. Donahue, B.A., Yin, S., Taylor, J.S., Reines, D. & Hanawalt, P.C. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc Natl Acad Sci U S A* **91**, 8502-6 (1994).
152. Paulsen, M.T. *et al.* Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods* **67**, 45-54 (2014).
153. Sigurdsson, S., Dirac-Svejstrup, A.B. & Svejstrup, J.Q. Evidence that Transcript Cleavage Is Essential for RNA Polymerase II Transcription and Cell Viability. *Molecular Cell* **38**, 202-210 (2010).
154. Kettenberger, H., Armache, K.J. & Cramer, P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**, 347-57 (2003).
155. Reines, D. Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J Biol Chem* **267**, 3795-800 (1992).
156. Dutta, A. *et al.* Ccr4-Not and TFIIS Function Cooperatively To Rescue Arrested RNA Polymerase II. *Mol Cell Biol* **35**, 1915-25 (2015).
157. Jensen, A. & Mullenders, L.H. Transcription factor IIS impacts UV-inhibited transcription. *DNA Repair (Amst)* **9**, 1142-50 (2010).
158. Mackinnon-Roy, C., Stubbert, L.J. & McKay, B.C. RNA interference against transcription elongation factor SII does not support its role in transcription-coupled nucleotide excision repair. *Mutat Res* **706**, 53-8 (2011).
159. Shema, E., Kim, J., Roeder, R.G. & Oren, M. RNF20 inhibits TFIIS-facilitated transcriptional elongation to suppress pro-oncogenic gene expression. *Mol Cell* **42**, 477-88 (2011).
160. Kruk, J.A., Dutta, A., Fu, J., Gilmour, D.S. & Reese, J.C. The multifunctional Ccr4-Not complex directly promotes transcription elongation. *Genes Dev* **25**, 581-93 (2011).
161. Mourgues, S. *et al.* ELL, a novel TFIIF partner, is involved in transcription restart after DNA repair. *Proc Natl Acad Sci U S A* **110**, 17927-32 (2013).
162. Lans, H., Marteijn, J.A. & Vermeulen, W. ATP-dependent chromatin remodeling in the DNA-damage response. *Epigenetics Chromatin* **5**, 4 (2012).
163. Mandemaker, I.K., Vermeulen, W. & Marteijn, J.A. Gearing up chromatin: A role for chromatin remodeling during the transcriptional restart upon DNA damage. *Nucleus* **5**, 203-10 (2014).
164. Dabin, J., Fortuny, A. & Polo, S.E. Epigenome Maintenance in Response to DNA Damage. *Mol Cell* **62**, 712-27 (2016).
165. Lim, J.H. *et al.* Chromosomal protein HMG1 enhances the acetylation of lysine 14 in histone H3. *EMBO J* **24**, 3038-48 (2005).
166. Shiama, N. The p300/CBP family: integrating signals with transcription factors and chromatin. *Trends Cell Biol* **7**, 230-6 (1997).
167. Adam, S., Polo, S.E. & Almouzni, G. Transcription recovery after DNA damage requires chromatin priming by the H3.3 histone chaperone HIRA. *Cell* **155**, 94-106 (2013).
168. Dinant, C. *et al.* Enhanced chromatin dynamics by FACT promotes transcriptional restart after UV-induced DNA damage. *Mol Cell* **51**, 469-79 (2013).
169. Szenker, E., Lacoste, N. & Almouzni, G. A developmental requirement for HIRA-dependent H3.3 deposition revealed at gastrulation in *Xenopus*. *Cell Rep* **1**, 730-40 (2012).
170. Reinberg, D. & Sims, R.J., 3rd. de FACTo nucleosome dynamics. *J Biol Chem* **281**, 23297-301 (2006).

171. Keller, D.M. & Lu, H. p53 serine 392 phosphorylation increases after UV through induction of the assembly of the CK2.hSPT16.SSRP1 complex. *J Biol Chem* **277**, 50206-13 (2002).
172. Nguyen, A.T. & Zhang, Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev* **25**, 1345-58 (2011).
173. Feng, Q. *et al.* Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* **12**, 1052-8 (2002).
174. Oksenyich, V. *et al.* Histone methyltransferase DOT1L drives recovery of gene expression after a genotoxic attack. *PLoS Genet* **9**, e1003611 (2013).
175. Nilson, K.A. *et al.* Oxidative stress rapidly stabilizes promoter-proximal paused Pol II across the human genome. *Nucleic Acids Res* **45**, 11088-11105 (2017).
176. Lans, H. & Vermeulen, W. Tissue specific response to DNA damage: C. elegans as role model. *DNA Repair (Amst)* **32**, 141-148 (2015).
177. Nospikel, T.P., Hyka-Nospikel, N. & Hanawalt, P.C. Transcription domain-associated repair in human cells. *Mol Cell Biol* **26**, 8722-30 (2006).
178. Gorgels, T.G. *et al.* Retinal degeneration and ionizing radiation hypersensitivity in a mouse model for Cockayne syndrome. *Mol Cell Biol* **27**, 1433-41 (2007).
179. Jaarsma, D., van der Pluijm, I., van der Horst, G.T. & Hoeijmakers, J.H. Cockayne syndrome pathogenesis: lessons from mouse models. *Mech Ageing Dev* **134**, 180-95 (2013).



2

Quantitative Pol II interaction proteomics to identify new factors in the UV-induced transcription stress response

Marit E Geijer¹, Roel Janssens¹, Dick Dekkers²,
Jeroen AA Demmers², Jurgen A Marteijn¹

¹ Department of Molecular Genetics, Oncode Institute, Erasmus University Medical Centre, Rotterdam, the Netherlands.

² Proteomics Center, Erasmus University Medical Center, Rotterdam, the Netherlands.

Abstract

Accurate transcription is essential for correct cellular function. However, the transcription process is threatened on a daily basis by DNA damage that may severely impede progression of elongating RNA polymerase II (Pol II). Lesion-stalled Pol II may result in transcription inhibition and in the formation of R-loops and transcription-replication conflicts that can induce genome instability. Moreover, transcription-blocking DNA damage has genome-wide effects on the transcription process, in addition to its direct effects on lesion-stalled Pol II. This DNA damage-induced transcription-stress response is a dedicated cellular response to allow cells to cope with transcription-blocking DNA damage. Pol II is the central factor in this highly regulated response and is involved in most of the reaction steps, ranging from lesion recognition and repair to transcription inhibition and restart. However, the exact molecular mechanism how cells cope with TBLs remains unknown. To identify new factors involved in repairing TBLs, we performed quantitative Pol II-interaction proteomics. For this purpose, we compared different Pol II extraction and immunoprecipitation methods and set up a proximity labeling interaction proteomics approach. The different methods each identified the Pol II basal complex. Surprisingly, only CSB was identified as a UV-specific interactor with all methods, in addition to a rather limited set of unknown UV-specific interacting proteins. Therefore, to obtain a better overview of the TBL-induced Pol II interactome, we combined the data of the different approaches and discussed the most identified and promising candidates.

Introduction

Faithful transcription by RNA polymerase II (Pol II) is crucial for proper cell function and homeostasis and is therefore tightly regulated from transcription initiation and promoter-proximal pausing to productive elongation and finally termination^{1,2}. However, correct transcription is daily threatened by endogenous metabolic processes causing oxidative lesions via reactive-oxygen-species (ROS)³ or exogenous factors. The latter includes chemotherapeutics like cisplatin which causes crosslinks or ultraviolet light (UV) that induces 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs), cyclobutane-pyrimidine dimers (CPDs) and many other bulky chemical DNA adducts⁴. Many of these DNA lesions strongly impede transcription elongation and are therefore called transcription-blocking lesions (TBLs). For example, 6-4PPs and CPDs induce helix distortions that make the DNA less flexible and prevent continuation of elongating Pol II, thereby forming a direct transcription block⁴⁻⁷. Impeded transcription can lead to the expression of mutant mRNAs due to decreased transcription fidelity during transcription bypass or can result in complete absence of *de novo* synthesis of mRNAs due to blocked transcription^{8,9}. The impact of stalled Pol II may contribute to severe cellular dysfunction, senescence and premature cell death and thereby lead to accelerated aging¹⁰⁻¹². Moreover, lesion-stalled Pol II can lead to the formation of R-loops^{13,14}. R-loops are RNA-DNA hybrids that leave the unpaired DNA strand more susceptible to DNA damage, leading to single-strand breaks and damaged bases¹⁵. In addition, the R-loop, as well as the stalled Pol II complex, can form physical road blocks for the replication machinery thereby leading to transcription-replication conflicts¹⁴. Transcription-replication conflicts can lead to double-strand breaks and strongly impair DNA replication leading to severely decreased genome stability that eventually can lead to the development of cancer^{15,16}.

In addition to the direct, *in cis* effects of TBLs on elongating Pol II, there are several *in trans* effects comprising genome-wide effects on Pol II that is not stalled on a lesion^{10,17}. One of these consequences includes a block in transcription initiation^{18,19} and a genome-wide clearance of Pol II from the promoter²⁰. In addition, there is an increased release of Pol II from the promoter-proximal pause site²¹⁻²³. This leads to increased transcription elongation and likely increases the ability to detect TBLs. Moreover, TBLs result in a decreased elongation rate^{24,25} that, together with an altered spliceosome composition, can lead to alternative splicing events^{13,26}. In addition, cells can switch to inclusion of an alternative last exon, thereby promoting shorter transcripts that are less susceptible to transcription interference by TBLs²⁵. These *in trans* responses to TBLs, the direct effects of Pol II stalling at a TBL, and repair of TBLs, together form the DNA damage-induced transcription stress response and help cells to cope with the severe consequence of TBLs.

A crucial part of the transcription stress response is a dedicated repair mechanism that prevents the deleterious effects of prolonged stalling and counteracts the genome-wide consequences of transcription stress. Transcription-coupled nucleotide excision repair (TC-NER) repairs TBLs in actively transcribed genes²⁷. The biological importance of removing TBLs is exemplified by the severe consequences of Cockayne syndrome (CS). CS patients suffer from progressive neurodegeneration, photosensitivity and premature ageing caused by inactivating mutations in TC-NER genes²⁸.

TC-NER is initiated upon Pol II stalling at a TBL, which stabilizes the transient interaction of the SNF2-like DNA-translocase CSB (ERCC6) with Pol II²⁹. The forward translocating ability of CSB helps to discriminate between Pol II stalled at natural occurring pause sites or at TBLs³⁰. Once the lesion is recognized, CSB initiates further complex assembly by recruiting CSA (ERCC8)^{31,32}. CSA is part of a Cullin 4-RING E3 ubiquitin ligase complex (CRL4^{CSA}), which acts as substrate receptor likely targeting CSB for ubiquitylation and subsequent proteasome-mediated degradation^{33,34}. Next, the scaffold protein UVSSA is recruited and brings its binding partner USP7 to the complex, which likely counteracts CSB degradation via its deubiquitylation activity³⁵⁻³⁷. In turn, UVSSA recruits TFIIH to the complex via a direct interaction^{38,39}. TFIIH might be involved in the reverse translocation of Pol II¹⁷. The latter is necessary because Pol II may shield the lesion from repair factors and its backtracking will provide access^{4,30}. In addition, TFIIH forms the core pre-incision complex together with RPA and XPA. This complex verifies the presence of lesions, and recruits ERCC1/XPF and XPG. These endonucleases excise the TBL and the resulting ssDNA gap is refilled by DNA polymerases and ligated to finalize repair^{40,41}. If the damage load is too high for swift removal of TBLs by TC-NER, lesion-stalled Pol II can be degraded via proteasome-dependent degradation of RPB1, the largest catalytic subunit of Pol II^{42,43}. This so-called 'last-resort' pathway does not initiate TC-NER, but makes the TBL accessible for different repair pathways and most likely prevents the detrimental effects of prolonged stalled Pol II, like transcription-replication conflicts.

After successful resolving the TBL, it is essential to resume transcription to maintain proper cellular homeostasis. Several factors have been suggested that enable transcription restart directly at the lesion, most likely via cleavage of the extruding mRNA, thereby aligning the backtracked Pol II to continue transcription¹⁷. Although it seems most efficient to restart transcription directly at the lesion, this appeared not the only way to resume transcription after repair. Transcription can also restart at the beginning of genes as shown by the wave-like transcription restart from promoters^{44,45}. In line with this, it has been suggested that Pol II can be evicted from the chromatin upon stalling⁴⁶. However, restart from the promoter could also be explained by inhibition of transcription initiation upon TBL exposure leading to restart after release of this inhibition²⁰.

Although many processes have been linked to Pol II-mediated transcription regulation and TC-NER following DNA damage, the exact molecular mechanism and the factors involved have yet to be determined. As Pol II is the central player in the cellular reaction to transcription stress both *in cis* and *in trans*, we aimed to investigate which factors are involved in regulating Pol II during stalling, eviction, repair, and restart. The identification of new factors in the DNA damage-induced transcription stress response will provide new insights in its molecular mechanism.

In this study, we compared different methods to identify Pol II complex interactors. We applied and optimized different native and cross-linked immunoprecipitation methods and set up proximity-labeling approaches and used these methods to identify Pol II-interacting proteins and UV-specific interactors using SILAC-based quantitative interaction proteomics. All these approaches were able to identify established Pol II complex subunits and elongation factors in undamaged conditions demonstrating the successful isolation

of elongating Pol II together with its interaction partners. Following UV damage, we could identify the established damage-specific Pol II interactor CSB, showing proof of principle of this approach. However, only a few other TBL-induced interactors were identified. The most effective identification method was the immunoprecipitation of P-Ser2-modified elongating Pol II. This approach identified more damage-specific interactors such as CSA although it did not directly result in the identification of new interactors. Finally, we combined the different immunoprecipitation approaches to reveal interactors that are identified in multiple approaches. This combined analysis identified several putative TBL-induced Pol II interactors that will be subject to follow-up studies to determine their function during the transcription stress response.

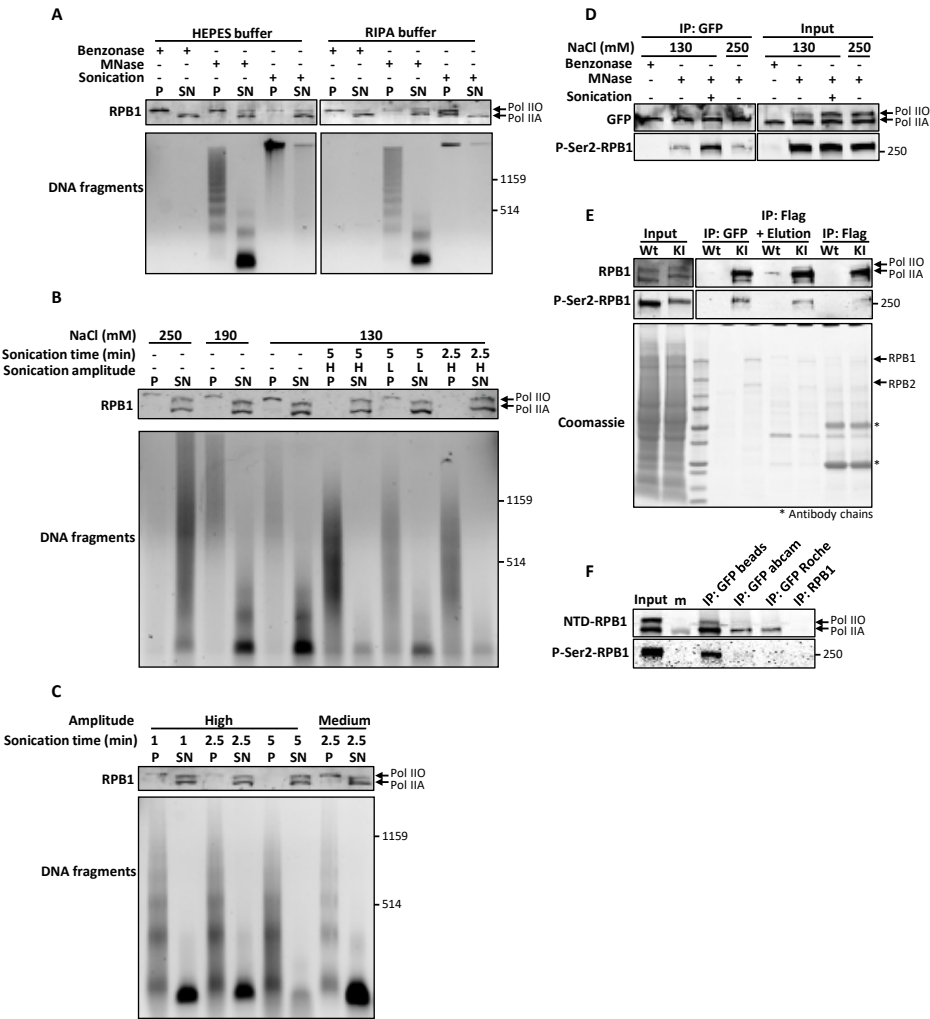
Results and Discussion

Optimization of extraction of initiating and elongating Pol II from the chromatin using a native immunoprecipitation

While initiating and promoter-paused Pol II are only transiently bound to DNA⁴⁷, elongating Pol II is tightly bound to the chromatin. Since we aimed to identify Pol II interaction partners during the *in cis* transcription block but also during *in trans* regulation of promoter-bound Pol II, we aimed to extract all fractions of Pol II. To isolate factors bound to elongating Pol II, it is necessary to extract chromatin-bound Pol II from the DNA before pulling down the complex. Therefore, we started with assessing a native extraction method using RIPA or milder HEPES buffers to lyse cells and determined the optimal conditions for extracting Pol II from the nucleoplasm and chromatin (Fig. 1A). To efficiently extract chromatin-bound Pol II we compared different DNA fragmentation methods. Elongating Pol II is hyperphosphorylated and can be visualized with a Pol II antibody on immunoblot as a higher migrating band (Pol IIO, Fig. 1A) compared to the lower migrating hypophosphorylated band that mainly represents initiating and promoter-paused Pol II (Pol IIA, Fig. 1A). As elongating Pol II is characterized by phosphorylation of the Serine at position 2 (Ser2) in the tandem repeats of a heptapeptide sequence present in the C-terminal domain (CTD) of Pol II, Pol IIO can also be visualized using an antibody recognizing the P-Ser2-specific form of Pol II. Fractionation assays help to discriminate between chromatin-bound proteins in the pellet fraction (P) or solubilized proteins in the supernatant fraction (SN). Proteins that cannot be solubilized due to firm chromatin binding reside in the pellet fraction. To solubilize chromatin-bound proteins, the DNA can be cut into smaller pieces using different enzymatic or mechanic approaches. For example, benzonase can be used to digest both DNA and RNA to single nucleotides and micrococcal nuclease (MNase) cuts the naked DNA in between the nucleosomes resulting in DNA in fragments of approximately 150 nucleotides. However, with sonication ultrasonic frequencies are used to shear the DNA.

To assess which treatment successfully solubilized elongating Pol II, we compared the different digestion methods using a cellular fractionation assay. This showed that benzonase treatment did not properly extract the higher migrating elongating fraction of Pol II from the chromatin as this remained present in the pellet fraction (P), but that only the lower migrating initiating and promoter-bound pool of Pol II was solubilized under these conditions (Fig. 1A, supernatant, SN). Both digestion with MNase or

shearing the DNA for 5 minutes using sonication extracted a small fraction of elongating Pol II from the chromatin although the majority of elongating Pol II remained in the chromatin (Fig. 1A). We assessed the size of the digested DNA to determine if the DNA fragments were properly broken into smaller fragments that could be immuno-precipitated. Upon benzonase treatment we lost all the DNA. This means that either the DNA was complete digested, risking loss of interactors bound via DNA and RNA, or the DNA was not digested leaving DNA fragments that are too large for detection (Fig. 1A, DNA fragments). Shearing the DNA using sonication left rather large fragments resulting in additional pull-down of interactors bound to the DNA further away from the Pol II bait. In the supernatant, MNase treatment digested the DNA in fragments of approximately 150 nucleotides, which most likely resulted in the efficient extraction of elongating Pol II from the pellet fraction. The presence of larger DNA fragments in the pellet fraction



might explain why not all elongating Pol II is solubilized. The small DNA fragments in the supernatant fraction, consisting of 1-2 nucleosomes, provided an optimal size to pull down proteins in close proximity of Pol II that were interacting via the DNA.

To assess the effect of buffer content on extracting Pol II from the chromatin, we compared RIPA buffer to milder HEPES buffer. There is little difference between the two buffers but using HEPES buffer showed a slightly higher yield of Pol II in the supernatant (SN) fraction compared to RIPA buffer (Fig. 1A).

Since MNase digestion and sonication both extracted some fraction of elongating Pol II, we combined both treatments to further optimize Pol II extraction (Fig. 1B). MNase treatment in combination with only 2.5 minutes of sonication already extracted all initiating and elongating Pol II and generated optimal DNA fragments that were smaller than 300 nucleotides. As an alternative, we tested if high salt concentrations could break protein interactions and solubilize Pol II. Therefore, we used high salt buffers in combination with MNase digestion to extract Pol II without shearing the DNA. This extracted both initiating as well as elongating Pol II although there was still elongating Pol II in the pellet fraction and compared to the sonication treatment, high salt left bigger sizes of DNA (Fig. 1B). Since sonication reduced elongating Pol II levels in the pellet fraction but did not increase the levels in the supernatant, this could suggest that Pol II is degraded during the procedure. Therefore, we further decreased the sonication time to maximize Pol II extraction but limit its degradation and chose 1.5 minutes of sonication in combination with MNase digestion for our immunoprecipitation (IP) method (Fig. 1C). Of note, decreasing the amplitude also extracted less Pol II from the chromatin, making it less suitable for IP (Fig. 1B-C).

As combining MNase digestion with shearing the DNA with mild sonication seemed to be most efficient to extract Pol II from the chromatin, we assessed if these were also the best conditions to pull down Pol II. Therefore, we used MRC-5 Flag-GFP-RPB1-KI

- ◀ **Figure 1. Optimization steps for extraction of Pol II and immunoprecipitation using a native protocol.** (A) Immunoblot stained for RPB1 (top) or gel electrophoresis of DNA fragments (bottom) of pellet (P) or supernatant (SN) after chromatin digestion with benzonase (1 hour, 1250 U), MNase (1 hour, 2 U) or shearing using sonication (5 minutes, high amplitude, 15 sec on/15 sec off). (B) Immunoblot stained for RPB1 (top) or gel electrophoresis of DNA fragments (bottom) for pellet (P) or supernatant (SN) after chromatin digestion with the indicated salt concentrations, sonication times and sonication intensities. All conditions were treated with MNase (1 hour, 2 U). (C) Immunoblot stained for RPB1 (top) or gel electrophoresis of DNA fragments (bottom) for pellet (P) or supernatant (SN) after chromatin digestion with the indicated sonication times and intensities. Sonication cycle consisted of 15 sec on/15 sec off, all conditions were treated with MNase (1 hour, 2 U). (D) Immunoprecipitation of Flag-GFP-RPB1 using GFP-beads followed by immunoblot analysis of the indicated proteins. Chromatin was digested using benzonase (1 hour, 625 U), MNase (1 hour, 1 U), MNase with 1.5 minutes of sonication (high, 15sec on/15sec off) or MNase with 250 mM NaCl. (E) Immunoprecipitation of Flag-GFP-RPB1 using indicated antibody-coupled beads, followed by immunoblot analysis of the indicated proteins (top) or coomassie protein staining (bottom). Elution: elution using Flag peptide instead of boiling to dissociate protein complexes. Chromatin was digested using MNase (1 hour, 1 U) with 1.5 minutes of sonication (high, 15 sec on/15 sec off). (F) Immunoprecipitation of Flag-GFP-RPB1 using indicated beads or antibody-coupled beads, followed by immunoblot analysis of the indicated proteins. Chromatin was digested using MNase (1 hour, 1 U) with 1.5 minutes of sonication (high, 15 sec on/15 sec off).

cells in which RPB1, the largest subunit of Pol II, is GFP-tagged and expressed from its endogenous locus⁴⁷. We compared the different extraction methods followed by IP of GFP-RPB1 using GFP-beads (Fig. 1D). Indeed, combining MNase with sonication was the most efficient extraction method to pull down Pol II. Even more efficient than using a high salt concentration which also seemed to extract both fractions of Pol II in fractionation experiments (see input). To compare the efficiency of different pull-down approaches, we compared GFP-beads to Flag-beads followed by boiling or elution (Fig. 1E). Using GFP-beads or Flag-beads was equally efficient when comparing the amount of pulled down Pol II (Fig. 1E). However, since GFP-beads do not contain heavy and light antibody chains, this makes them more suitable than Flag beads. When we eluted the protein complexes from the Flag-beads, the antibody chains could not be detected but the elution had a higher background compared to GFP-beads (Fig. 1E, coomassie). Next, we compared GFP-beads to GFP-specific antibodies or an RPB1-specific antibody (Fig. 1F). Immunoprecipitations using GFP-beads were more efficient than using GFP antibody-coupled beads (Fig. 1F). Using the RPB1-specific antibody barely immunoprecipitated Pol II (Fig. 1F). Therefore, we continued to use GFP-beads for immunoprecipitation. It is interesting to note that although equal amounts of initiating and elongating Pol II were extracted, as seen in the input sample, less elongating Pol II was immunoprecipitated. Possibly, this fraction was degraded during the pulldown or was bound to strands of DNA that were lost during the procedure.

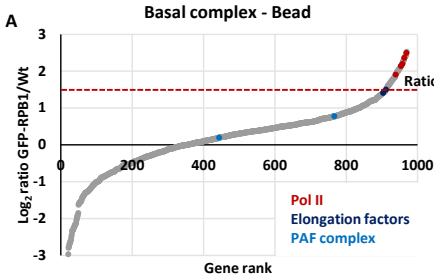
Native immunoprecipitation of Pol II identified components of the elongating Pol II complex in basal conditions and in response to UV irradiation

To test whether we were able to successfully extract and IP both initiating and elongating Pol II, we assessed the complex composition of Pol II using SILAC-based mass spectrometry in undamaged conditions. In addition, we compared the efficiency of two digestion methods used for preparing the samples for mass spectrometry: on-bead digestion and digestion after gel-separation (Fig. 2). First, we performed GFP-RPB1 immunoprecipitations in light labeled MRC-5 Wt and heavy labeled MRC-5 Flag-GFP-RPB1 KI cells and determined the protein interactors using on-bead digestion (Fig. 2A-C). The Log₂ SILAC ratio of identified proteins are ranked and all interactors with a Log₂ ratio of 1.5 or bigger were considered specific Pol II-interactors (Fig. 2A). There was a minor skew towards a SILAC ratio of 1 for the majority of proteins identified. This could be explained by better incorporation of the heavy amino acids, more identified Pol II interactors in the heavy labeled cells or a severe imbalance in the protein content in one of the samples that results in abnormalities in the normalization by MaxQuant. Regardless, we quantified approximately 1000 proteins and among these were 59 proteins with a SILAC ratio bigger than 1.5. These included Pol II subunits⁴⁸, several transcription elongation factors such as SPT6^{49,50} (Fig. 2A-B). Some components of the PAF complex were also isolated with lower SILAC ratios⁵¹ (Fig. 2A-B). GO term analysis indicated that we mainly identified proteins involved in Pol II transcription, mRNA splicing and processing (Fig. 2C), confirming we specifically isolated the Pol II complex. Performing the same analysis for Pol II-specific interactors with mass spectrometry on gel-separated samples identified approximately 1200 Pol II-interactors (Fig. 2D), slightly more than the on-bead digestion (Fig. 2A). With this procedure, we also quantified Pol II subunits, elongation factors and one subunit of the PAF complex (Fig. 2E). GO term

analysis identified transcription-related processes (Fig. 2F), similar to the on-bead digestion. Overall, the number of peptides found to identify the specific interactors was slightly higher in the gel-separated sample (Fig. 2E). This suggests, together with the higher number of proteins identified in general, that gel-separation was a slightly better method to efficiently detect interactors and therefore we used this digestion in combination with mass spectrometry.

We identified few RNA-binding proteins using mass spectrometry, such as spliceosome components. Thus far, we used MNase to digest the DNA but this is also known to cut RNA. Since we were also interested in interactors that were bound via interaction with the nascent RNA, we searched for an alternative DNA-specific nuclease. DNase I is known to specifically cut DNA and not RNA and successfully extracted initiating Pol II from the chromatin, which was not influenced by inhibiting the activity using EDTA (Fig. 2G). However, to extract elongating Pol II, the chromatin also needed shearing of the DNA using sonication, similar to MNase-digested samples (Fig. 2G). We successfully immunoprecipitated Pol II after DNase treatment (Fig. 2G) and therefore used this digestion method to detect the UV-specific Pol II complex composition using SILAC-based mass spectrometry. To do so, we labeled mock-treated MRC-5 Flag-GFP-RPB1 KI cells with light amino acids and labeled MRC-5 Flag-GFP-RPB1 cells one hour after irradiation with 8 J/m² UV-C with heavy amino acids. Log₂ SILAC ratios of UV-interactors over mock-treated are plotted and proteins were considered significant interactors with a ratio of 1 or bigger (Fig. 2H). Most subunits of the Pol II complex were identified but their interaction was not changed in response to UV, suggesting that the Pol II complex composition was not changed (Fig. 2H-I). We did not identify CSB or any of the downstream TC-NER subunits¹⁷ but did identify TFIIH. Interestingly, these subunits seemed to bind slightly less to the Pol II complex upon DNA damage which is in contrast to the described enrichment of TFIIH during TC-NER^{32,38}. Previously, a decreased TFIIH-promoter binding in response to UV was reported due to decreased transcription initiation²⁰. Since we immunoprecipitated a large fraction of initiating Pol II (Fig. 2G), this could explain why we noticed a decrease in TFIIH-binding to Pol II. Moreover, only a small fraction of the total pool of TFIIH has been described to be engaged in TC-NER⁵², and as a consequence this fraction was too small to detect the UV-induced increased interaction with lesion-stalled Pol II.

Interestingly, some of the top hits that were identified as UV-specific Pol II interactors were components of the 26S proteasome that is involved in degradation of ubiquitylated proteins (Fig. 2I-J)⁵³. Moreover, we identified SQSTM1 which is a protein that protects the ubiquitin chains from deubiquitylation while transporting proteins to the proteasome⁵⁴. Pol II degradation is normally serving as a last resort pathway when the damage load is too high for cells to overcome^{42,43}. During this pathway, RPB1 is ubiquitylated and degraded via the proteasome^{42,43}. Since multiple proteasome subunits were binding Pol II, this could indicate that Pol II was degraded in response to UV irradiation and might explain why we did not detect known UV-specific interactors. Moreover, it is possible that we did not successfully identify CSB as a UV-specific interactor due to loss of the interactors during extraction of Pol II from the chromatin or during the immunoprecipitation procedure.

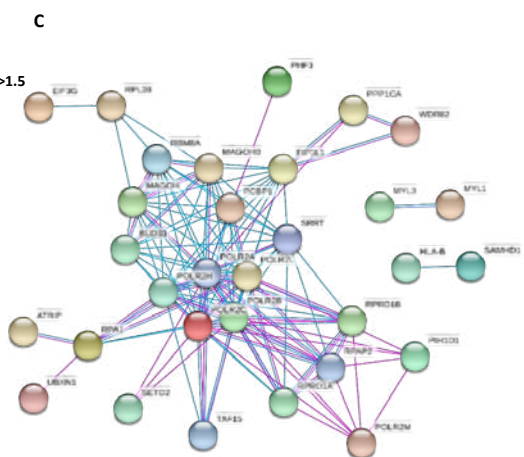


B Native GFP IP

Bead extraction GFP-RPB1/Wt

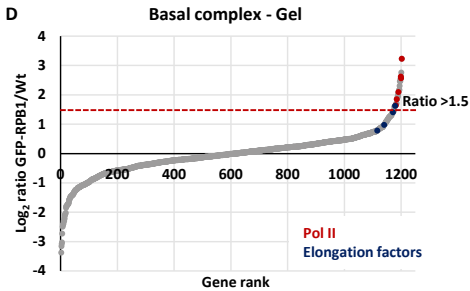
Gene name	# of peptides	SILAC ratio (Log ₂)
POLR2A	79	1.91
POLR2B	41	2.20
POLR2C	10	2.49
Pol II POLR2E	2	1.44
POLR2H	6	2.14
POLR2J	4	0.94
POLR2L	2	2.36
Elongation factors RECQL5	8	1.50
SUPT16H	17	1.14
SUPT6H	21	1.33
TCEA1	5	1.41
PAF Cpl. CTR9	5	0.77
PAF1	5	0.19

0 3



GO Biological processes. Top terms of n=59 (Log₂ SILAC ratio > 1.5)

Pathway description	Gene count	Total gene count	FDR
mRNA processing	16	456	1.52E-09
RNA splicing	15	391	1.52E-09
snRNA transcription by RNA polymerase II	9	71	1.75E-09
mRNA splicing, via spliceosome	13	284	2.26E-09
snRNA metabolic process	9	88	3.88E-09

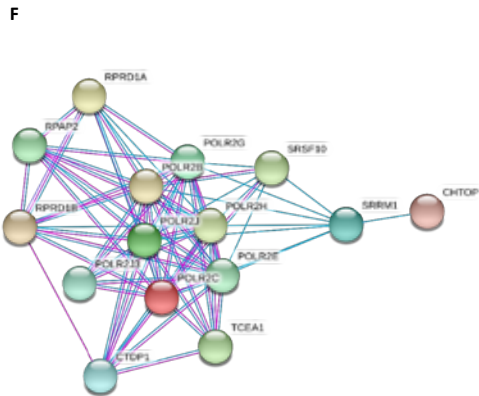


E Native GFP IP

Gel extraction GFP-RPB1/Wt

Gene name	# of peptides	SILAC ratio (Log ₂)
POLR2A	77	1.45
POLR2B	56	1.85
POLR2C	6	2.55
Pol II POLR2E	4	2.09
POLR2G	2	1.64
POLR2H	2	3.23
POLR2J	3	2.61
RECQL5	11	1.40
SUPT16H	20	0.97
SUPT5H	8	0.31
SUPT6H	24	0.77
TCEA1	12	1.62
PAF Cpl. CDC73	6	0.43

0 4



GO Biological processes. Top terms of n=28 (Log₂ SILAC ratio > 1.5)

Pathway description	Gene count	Total gene count	FDR
transcription by RNA polymerase II	16	784	1.62E-12
snRNA transcription by RNA polymerase II	9	71	1.62E-12
snRNA metabolic process	9	88	2.45E-12
mRNA processing	13	456	6.05E-12
transcription elongation from RNA polymerase II promoter	8	75	3.48E-11

Figure 2. (continued)

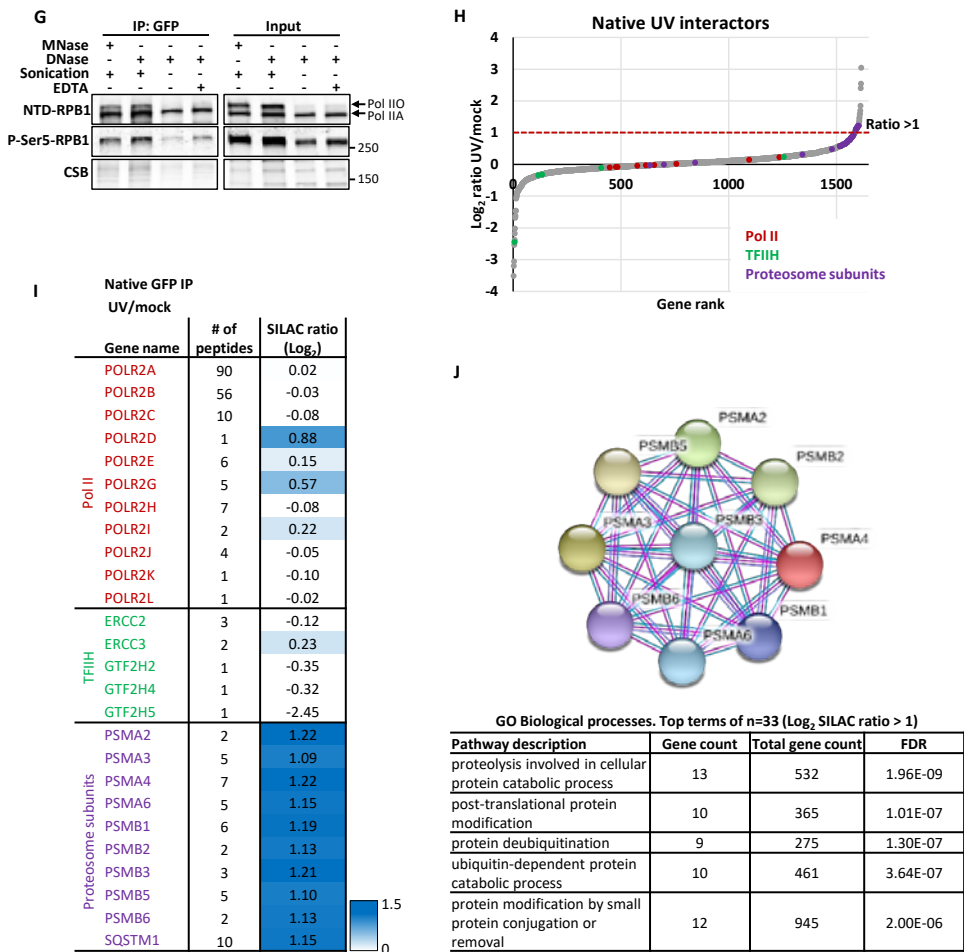


Figure 2. Pol II basal complex composition and UV-specific interactors identified using native IP. (A) Pol II-interacting proteins ranked based on the Log₂ SILAC ratios determined using quantitative interaction proteomics of GFP IP comparing MRC-5 Flag-GFP-RPB1 KI cells to MRC-5 Wt cells after on-bead digestion. Dotted line indicates SILAC ratio threshold (> 1.5) for identified interactors. Pol II subunits, elongation factors and PAF complex subunits are color-coded. **(B)** Interaction heat map of the Log₂ SILAC ratios of Pol II-interacting proteins of Flag-GFP-RPB1 cells compared to Wt cells. Positive values indicate increased interaction. Second column indicates the number of unique peptides quantified. **(C)** Top: protein-protein interactions as determined by STRING protein interaction analysis of genes with Log₂ SILAC ratio > 1.5. Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of Pol II-specific interactors with Log₂ SILAC ratio > 1.5. **(D-F)** Same as (A-C) but for quantitative interaction proteomics of MRC-5 Flag-GFP-RPB1 KI cells compared to MRC-5 Wt cells after gel separation and digestion. **(G)** Immunoprecipitation of Flag-GFP-RPB1 using GFP beads followed by immunoblot analysis of the indicated proteins. Chromatin was digested using MNase (1 hour, 1 U) or DNase I (1 hour, 8 U), in combination with 1.5 minute of sonication (high, 15 sec on/15 sec off) where indicated. EDTA: after 1 hour, 5 mM EDTA was added to stop DNase activity. **(H)** UV-specific, Pol II-interacting proteins ranked based on the Log₂ SILAC ratios as determined using quantitative interaction proteomics of GFP IP

- comparing MRC-5 Flag-GFP-RPB1 KI cells after 8 J/m² UV-C to mock-treated cells. Dotted line indicates threshold (> 1) for identified interactors. Pol II subunits, TFIIH complex subunits and proteasome subunits are color-coded. **(I)** Interaction heat map of the Log₂ SILAC ratios of UV-specific Pol II-interacting proteins of UV-irradiated compared to mock-treated cells. Positive values indicate increased interaction, negative values indicate decreased interaction. Second column indicates the number of peptides identified. **(J)** Top: protein-protein interactions as determined by STRING protein interaction analysis of genes with Log₂ SILAC ratio > 1 . Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of UV-specific interactors with a Log₂ SILAC ratio > 1 .

Crosslinked immunoprecipitation approach to study interactors of Pol II elongation complex in basal or DNA damage conditions

To prevent loss of interactors during the relative harsh extraction methods and to be able to identify transient interactors⁵⁵, we adapted a crosslinked immunoprecipitation (XIP) protocol using paraformaldehyde (PFA) fixation. Chemical crosslinking with PFA crosslinks DNA-protein and protein-protein interactions that are in close proximity since it has a reach of approximately 2 Å⁵⁵⁻⁵⁷. Therefore, this XIP method allows harsh chromatin digestion to efficiently extract elongating Pol II while preserving protein interactions.

We started with optimizing the conditions for optimal extraction of Pol II using this crosslinked approach (Fig. 3A-B). Extending the crosslink time showed that there was more Pol II trapped in the pellet fraction after 10 minutes of crosslinking, suggesting that Pol II could not be properly extracted from the chromatin anymore (Fig. 3A). This was not due to improper reversal of the crosslink bonds because also after 30 minutes of reverse crosslinking, Pol II could not be extracted from the chromatin (Fig. 3A). Further extending the crosslink time to 15 minutes showed that we even lost Pol II and this also seemed to be the case after 30 minutes of reverse crosslinking. However, since we aimed to reach maximal reversal of induced crosslinks to be able to perform SILAC-based mass spectrometry, we continued to reverse crosslink for 30 minutes^{58,59}. Shearing the DNA using sonication promoted extraction of Pol II from the chromatin but extensive sonication induced Pol II degradation (Fig. 3B). We decided to crosslink the cells for 7 minutes, shear the DNA for 7 minutes and reverse crosslink for 30 minutes, as under these conditions almost all elongating Pol II was extracted from the chromatin. As crosslinking might affect lysine-residues present in the epitopes recognized by the antibodies used for IP, we checked which antibody most efficiently isolated GFP-RPB1 from crosslinked samples. We compared pull-down efficiency of Flag-GFP-RPB1 by GFP beads, GFP antibody-coupled beads and Flag M2 beads. Flag M2 beads were most efficient in pulling down both initiating and elongating Pol II (Fig. 3C).

To validate our crosslinked immunoprecipitation approach, we set out to identify the Pol II complex composition in undamaged conditions using SILAC-based mass spectrometry. Wt cells were labeled with light amino acids and KI cells were labeled with heavy amino acids in the forward experiment and we performed a label swap in the reverse experiment. The log₂ SILAC ratios of KI/Wt cells of both experiments are plotted and proteins with a ratio bigger than 2 were considered to be specific Pol II-interactors (Fig. 3D). This resulted in the detection of RPB1 interaction with most of the Pol II subunits,

several transcription elongation factors such as SPT4 and SPT5^{49,50} and components of the PAF complex (Fig. 3D-E). GO term analysis confirmed that we specifically pulled down elongating Pol II and transcription-related interactors (Fig. 3F). Compared to the native IP, this method identified more Pol II-interacting proteins and identified them with more unique peptides suggesting a more efficient extraction of Pol II and its interactors (Fig. 2). Moreover, this indicated that crosslinking and reverse crosslinking did not interfere with efficient detection using mass spectrometry because more peptides were identified compared to the native IP method. Therefore, we used this method to also identify UV-specific Pol II interactors. We harvested heavy-labeled MRC-5 Flag-GFP-RPB1 KI cells 1 hour after irradiation with 16 J/m² UV-C and compared them to mock-treated, light-labeled MRC-5 Flag-GFP-RPB1 KI cells and performed a label swap. The log₂ SILAC ratios are depicted in a scatterplot (Fig. 3G) or the average log₂ SILAC ratios of selected proteins are depicted in an interaction heatmap (Fig. 3H). Proteins were considered significant UV-specific interactors when the Log₂ SILAC ratio was 0.75 or bigger. The most significant interactor identified was CSB, which is the first protein of the TC-NER reaction and its interaction with Pol II is stabilized upon UV irradiation²⁹. This confirmed that our approach could identify UV-specific interactors. In addition, we identified WDR83. WDR83 is also known as MORG1 and functions as a scaffold protein in the response to hypoxia⁶⁰ and the MAPK/ERK pathway⁶¹, which was also identified as a biological process using GO term analysis (Fig. 3I). Moreover, it has previously been shown to be isolated as a part of the spliceosome^{62,63}. However, thus far a role for WDR83 during DNA damage-induced transcription stress is unknown and thus requires further investigation.

Pol II interaction proteomics using crosslinked chromatin in the presence of VCP inhibitors

Although our approach seemed to be successful since we identified CSB as a UV-induced Pol II interactor, the number of UV-induced interactions remained low. The absence of these interactors could be explained by degradation of lesion-stalled Pol II in response to UV irradiation, for example during the last resort pathway^{42,43}. During this last resort pathway, ubiquitylated Pol II is extracted from the chromatin by the VCP/p97 ubiquitin-selective segregase resulting in its proteasomal degradation⁶⁴⁻⁶⁶. To inhibit degradation of lesion-stalled Pol II and to keep it bound to the chromatin, we determined the UV-induced Pol II interactors in presence of VCP inhibitor (VCPi). We compared VCPi-treated cells 1 hour after irradiation with 16 J/m² UV-C to cells treated with VCPi and mock-irradiation. The log₂ SILAC ratios are depicted in a scatterplot (Fig. 3J) or the average log₂ SILAC ratios of selected proteins are depicted in an interaction heatmap (Fig. 3K). Proteins were considered specific interactors when the Log₂ SILAC ratio was 0.65 or bigger. This again identified CSB as a UV-specific interactor with a similar ratio (Fig. 3J-K). This suggests that lesion-stalled Pol II was not massively degraded under the used conditions, otherwise the CSB-Pol II interaction was most likely increased upon VCPi. Moreover, as we still did not identify additional UV-specific interactors (Fig. 3J-L), this suggests that proteasomal degradation of Pol II is not the cause for the limited set of UV-induced Pol II interactors observed.

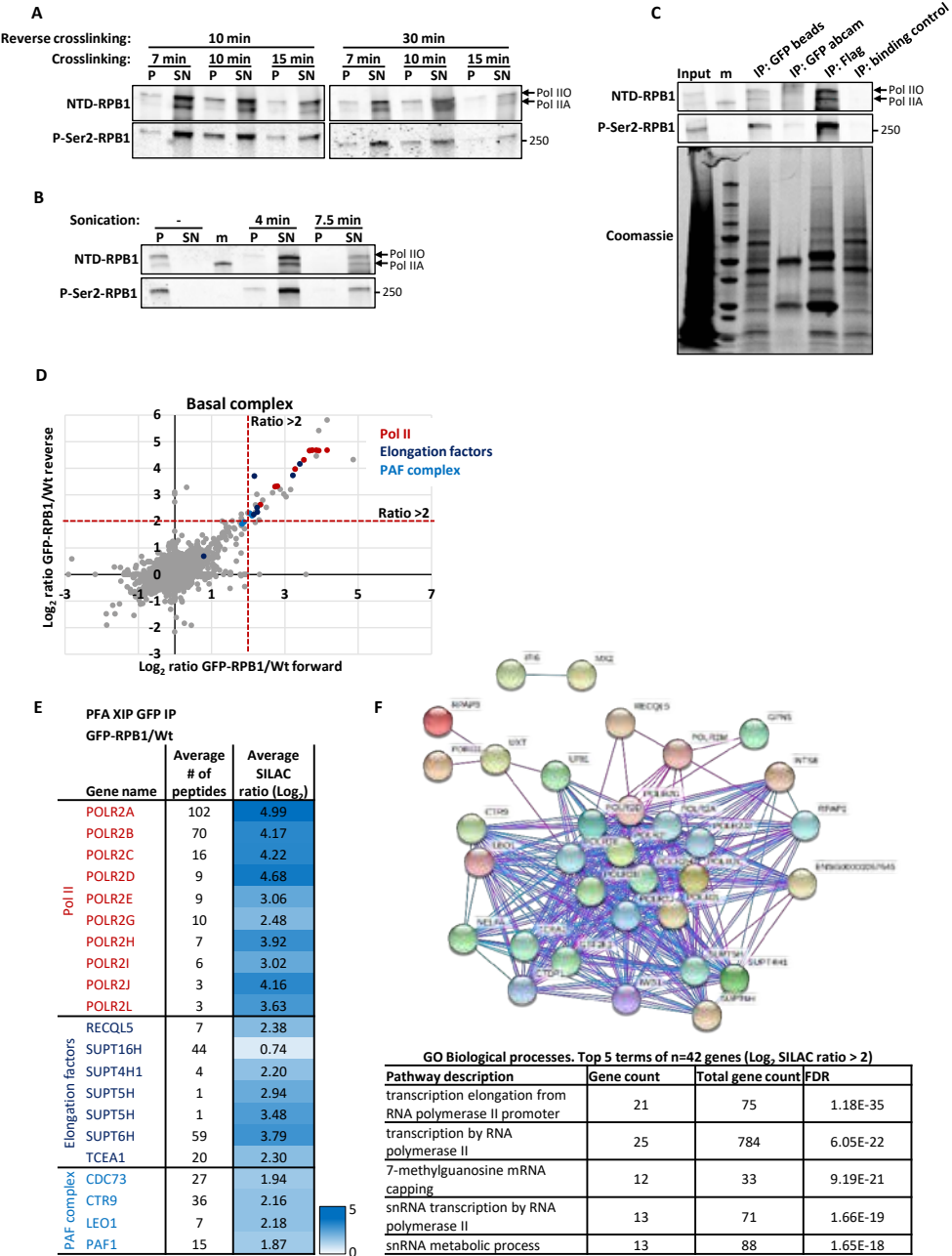
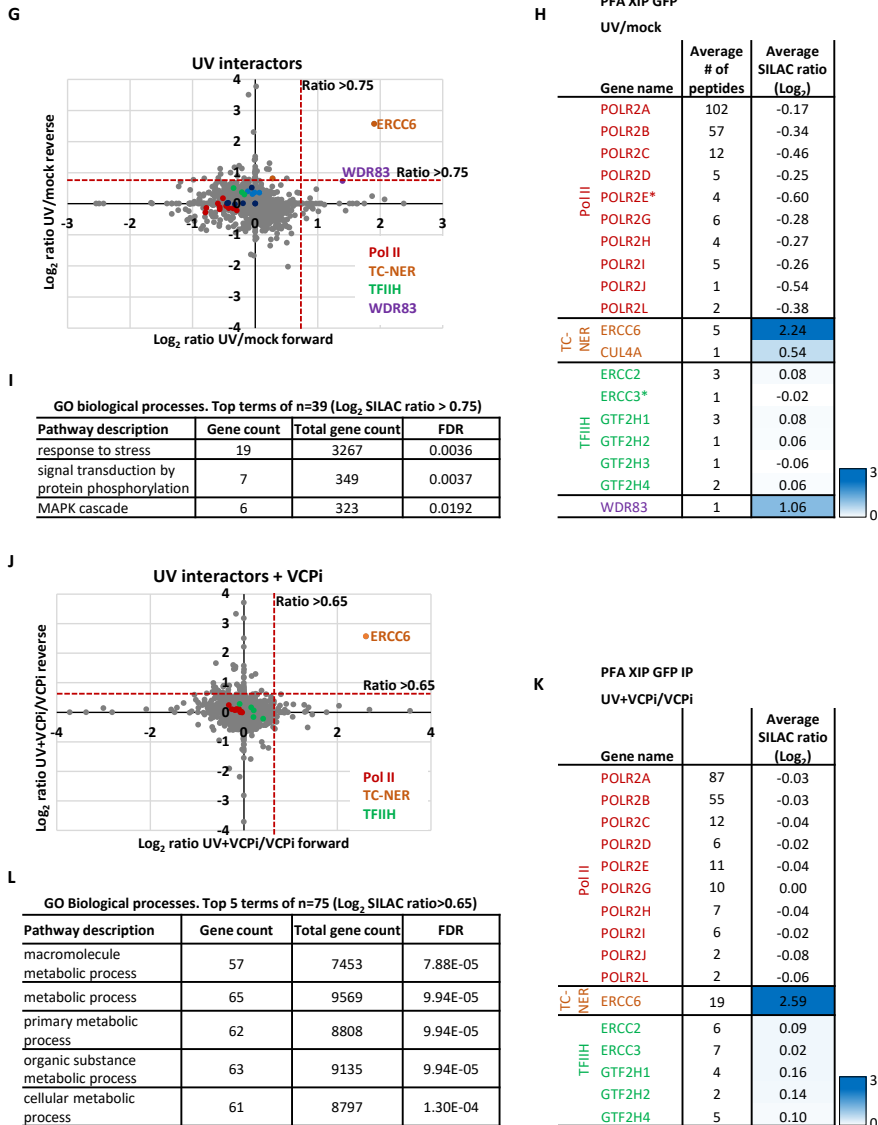


Figure 3. XIP optimization and identification of Pol II basal complex composition and UV-specific interactors using XIP. (A) Immunoblot stained for indicated proteins in pellet (P) or supernatant (SN) after shearing the DNA using sonication (7.5 minutes at high, 15 sec on/15 sec off). Crosslink (1% PFA) and reverse crosslink (boiling at 95°C) times were used as indicated. **(B)** Immunoblot stained for indicated proteins in pellet (P) or supernatant (SN) after shearing the chromatin. Sonication times were varied as indicated (high, 15 sec on/15 sec off). Cells were crosslinked for 7 minutes (1% PFA) and reverse crosslinking for 30 minutes

Figure 3. (continued)



(boiling at 95°C). (C) Immunoprecipitation of Flag-GFP-RPB1 using indicated beads or antibody-coupled beads, followed by immunoblot analysis of the indicated proteins (top) or coomassie protein staining (bottom). Cells were crosslinked for 7 minutes (1% PFA), DNA was sheared with sonication for 7 minutes (high, 15 sec on/15 sec off), and crosslinking was reversed for 30 minutes (boiling at 95°C). (D) Pol II-interacting proteins as determined using quantitative interaction proteomics of MRC-5 Flag-GFP-RPB1 KI cells compared to MRC-5 Wt cells after XIP using Flag-beads. The log₂ ratios of KI/Wt from forward experiment are plotted against the log₂ ratios of KI/Wt from reverse experiment. Dotted lines indicate threshold (> 2) for identified interactors. Pol II subunits, elongation factors and PAF complex subunits are color-coded. (E) Interaction heat map of the average Log₂ SILAC ratios of Pol II-interacting proteins determined by comparing Flag-GFP-RPB1 to Wt cells. Positive values indicate increased interaction. Second column ►

- indicates the number of peptides quantified. **(F)** Top: protein-protein interactions as determined by STRING protein interaction analysis of genes with Log_2 SILAC ratio > 2 identified in at least 1 of the experiments. Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of Pol II-specific interactors with a Log_2 SILAC ratio > 2 identified in at least 1 of the experiments. **(G)** UV-specific Pol II interactors identified using quantitative interaction proteomics of MRC-5 Flag-GFP-RPB1 KI cells 1 hour after 16 J/m^2 UV-C compared to mock-treated cells. The log_2 SILAC ratios of UV/mock-treated from forward experiment are plotted against the log_2 ratios of UV/mock-treated from reverse experiment. Dotted line indicates threshold (> 0.75) for identified interactors. Pol II subunits, TFIIH complex subunits, TC-NER subunits, and WDR83 are color-coded. **(H)** Interaction heat map of the average log_2 SILAC ratios of UV-specific Pol II-interacting proteins 1 hour after 16 J/m^2 UV-C compared to mock-treated. Positive values indicate increased interaction, negative values indicate decreased interaction. Second column indicates the number of peptides quantified. **(I)** Top enriched gene ontology (GO) terms (biological processes) of UV-specific interactors with a Log_2 SILAC ratio > 0.75 in at least one of the experiments (forward, reverse or both). **(J-L)** Same as (G-I) but after pre-treatment with and in presence of VCPI (1 hour, $5 \mu\text{M}$). Threshold for specific interactors is Log_2 SILAC ratio > 0.65 .

Using an alternative protein crosslinker for Pol II interaction proteomics

PFA induces more protein-DNA than protein-protein crosslinks due to lysine-specific bonds^{55,67,68}. The absence of numerous UV-specific interactors, might suggest that DNA-protein crosslinking is not sufficient. Therefore, we changed the crosslinking agent to dithiobis(succinimidyl propionate) (DSP), which mainly crosslinks protein-protein interactions⁵⁷. Moreover, DSP has a bigger reach of about 12 \AA , thanks to the size of its linker⁶⁹, enabling the crosslinking of interactors that are further away from Pol II. Finally, DSP also has the advantage of containing an easily cleavable disulfide bond making it very easy to reverse crosslinking without the need to boil the samples for 30 minutes and risking degradation⁵⁷.

We first compared the efficiency of extracting Pol II after 30 minutes of crosslinking with 1 mM DSP to 7 minutes of crosslinking with 1% or 0.5% PFA. Crosslinking using DSP successfully extracted both fractions of Pol II although there seemed to be slightly less Pol II, especially less initiating Pol II, and the DNA fragments were slightly larger (Fig. 4A). Immunoprecipitating Pol II after DSP crosslinking successfully pulled down Pol II, particularly elongating Pol II (Fig. 4B). Next, we investigated if we could decrease the sonication time or if we could decrease the crosslinking time by increasing the concentration of DSP (Fig. 4C). Decreasing the sonication time from 7 to 3.5 minutes did yield more immunoprecipitated Pol II (Fig. 4C). However, maintaining 7 minutes of shearing the DNA while crosslinking using 2 mM DSP for 15 minutes was even more efficient to extract and pulldown Pol II. Therefore, we used these conditions to perform SILAC-based interaction proteomics to identify UV-specific Pol II interactors.

To study the UV-specific Pol II complex composition and if this composition changes over time, we compared light-labeled mock-treated cells to heavy-labeled UV-irradiated cells, 15 or 75 minutes after 16 J/m^2 UV-C. We plotted the Log_2 SILAC ratios 15 minutes (Fig. 4D, left) or 75 minutes (Fig. 4D, right) after UV over mock-treated. Proteins were identified as significant interactors with a ratio of 1 or bigger. 15 minutes after UV-C irradiation, we identified CSB as a UV-specific interactor (Fig. 4D, left, 4E), similar as in the XIP experiments with PFA. In addition to CSB, we identified DDB1 (Fig. 4D, left,

4E), which is recruited during TC-NER complex assembly as part of the CRL4^{CSA} E3 ligase complex⁷⁰. Moreover, we identified several subunits of the TFIIH complex, possibly because we extracted more elongating Pol II (Fig. 4D, left, 4E). 75 minutes after UV-C irradiation, we also identified CSA (Fig. 4D, right, 4E). DSP XIP identified more known UV-specific interactors and moreover, the proteins were identified with more peptides (Fig. 4E) compared to the PFA XIP (Fig. 3H+K). Interestingly, ATM was identified as a UV-specific interactor 75 minutes after UV-C irradiation (Fig. 4E-F), suggesting that ATM is interacting with Pol II. ATM has previously been shown to be activated by the formation of R-loops in response to lesion-stalled Pol II, resulting in displacement of the spliceosome¹³. Our results might indicate that ATM is directly binding to the lesion-stalled Pol II complex or ATM binds in close proximity of Pol II and was identified due to the relatively big reach of DSP. Surprisingly, assessment of the biological processes identified in either 15 minutes or 75 minutes after UV or that were present in both time points, revealed proteins involved in cell cycle regulation and not damage repair (Fig. 4F). This observation might suggest that under these conditions, lesion-stalled Pol II may control cell cycle regulation in response to UV.

Performing Pol II interaction proteomics in TC-NER-deficient cells

Although DSP XIP identified more UV-interactors and the interactors found were identified with more peptides, we still did not discover many previously unknown interactors. This could be explained by the fact that the fraction of lesion-stalled Pol II was too small to clearly identify new interactors. To overcome this, we used XPA KO cells to accumulate lesion-stalled Pol II complexes and set out to investigate Pol II complex composition. XPA assists TFIIH during lesion verification of the lesion¹⁷ and is essential for the downstream NER-complex assembly. In the absence of XPA, Pol II is expected to remain stalled on the lesion. While the TC-NER complex assembly is initiated, TC-NER cannot progress beyond the verification step and therefore likely accumulates repair complexes, thereby enriching for lesion-stalled Pol II. We compared MRC-5 XPA KO Flag-GFP-RPB1 KI cells after mock-treatment to MRC-5 XPA KO Flag-GFP-RPB1 KI cells 15 or 75 minutes after irradiation with 16 J/m² UV-C. The Log₂ SILAC ratios (Fig. 4G) 15 minutes (left) or 75 minutes (right) after UV compared to mock-treated are plotted and proteins were considered UV-specific interactors if the ratio was 1.25 or higher. CSB was identified as a UV-specific interactor, similar as in the XPA-proficient DSP XIP, but other components of the TC-NER complex were not identified anymore (Fig. 4G-H).

Interestingly, we identified FANCM as a Pol II-interactor, especially 15 minutes after UV and with a more subtle increased log₂ SILAC ratio 75 minutes after UV. FANCM has previously been shown to recruit the core fanconi anemia complex in response to DNA damage such as interstrand crosslinks but also replication fork stalling⁷¹. FANCM is phosphorylated to stimulate its chromatin-binding and activity and in turn monoubiquitylates FANCD2 which is necessary for functional DNA crosslink repair and homologous recombination⁷¹. Moreover, FANCM has been shown to be involved in resolving R-loops, thereby preventing genome instability⁷². Failure to repair TBLs by TC-NER might result in prolonged stalled of Pol II at lesions, which can subsequently induce the formation of R-loops^{13,14}. Therefore, in the absence of TC-NER, FANCM binding to Pol II might help to resolve these R-loops. In addition to FANCM, we identified subunits of the 26S

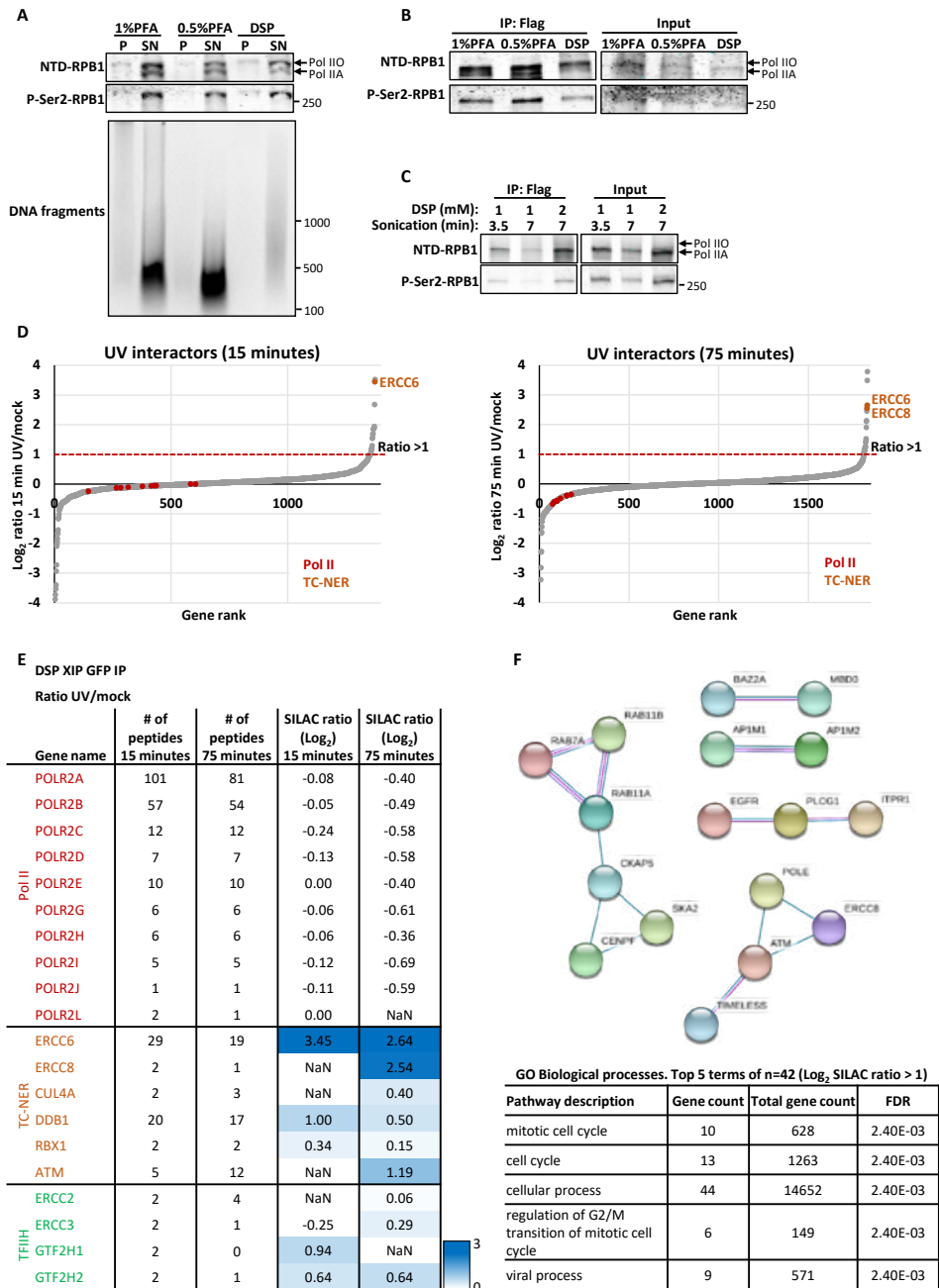
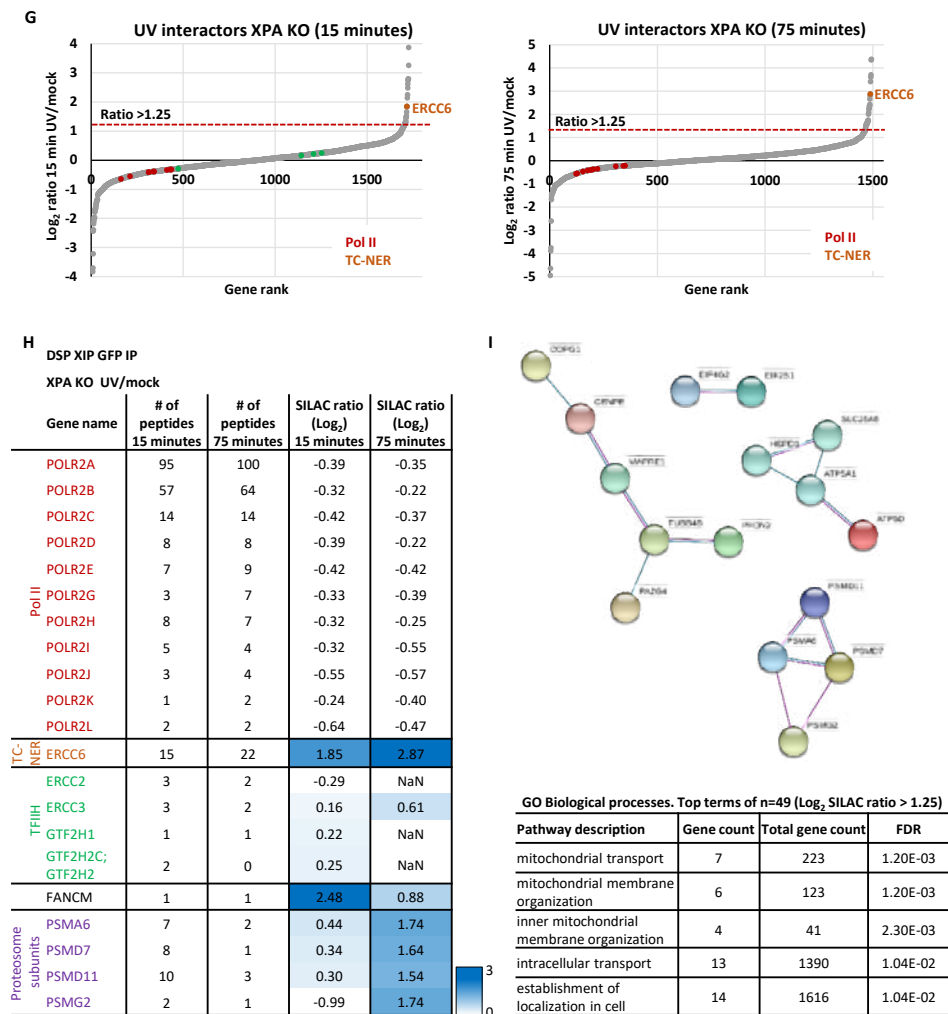


Figure 4. Optimization of DSP XIP and identification of UV-specific Pol II-interacting proteins using DSP XIP. (A) Immunoblot stained for RPB1 (top) or gel electrophoresis of DNA fragments (bottom) for pellet (P) or supernatant (SN) after shearing the DNA using 7 minutes of sonication (high, 15 sec on/15 sec off). Samples were crosslinked with PFA (7 minutes, 0.5 or 1%) or DSP (30 minutes, 1 mM). (B) Immunoprecipitation of Flag-GFP-RPB1 using Flag beads with crosslinking conditions as indicated (PFA: 7 minutes, DSP: 30 minutes, 1 mM), followed by immunoblot analysis of indicated proteins. (C) Immunoprecipitation of Flag-GFP-RPB1

Figure 4. (continued)



using Flag beads after crosslinking with DSP (30 minutes with 1 mM or 15 minutes with 2 mM), followed by immunoblot of indicated proteins. DNA was sheared for indicated time with cycles of 15 sec on/15 sec off on high amplitude. **(D)** UV-specific proteins ranked based on the Log₂ SILAC ratios as determined using quantitative interaction proteomics of GFP IP comparing MRC-5 Flag-GFP-RPB1 KI cells 15 minutes (left) or 75 minutes (right) after 16 J/m² UV-C to mock-treated cells. Dotted line indicates threshold (> 1) for identified interactors. Pol II subunits and TC-NER factors are color-coded. **(E)** Interaction heat map of the Log₂ SILAC ratios of UV-specific proteins 15 and 75 minutes after UV-C compared to mock-treated. Positive values indicate increased interaction, negative values indicate decreased interaction. Second and third columns indicate the number of peptides identified. **(F)** Top: Protein-protein interactions as determined by STRING protein interaction analysis of genes with Log₂ SILAC ratio > 1. Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of UV-specific interactors with Log₂ SILAC ratio > 1 identified 15 or 75 minutes after UV-C. **(G-I)** Same as (D-F) but for MRC-5 XPA KO Flag-GFP-RPB1 KI cells. Interactors were considered identified with a Log₂ SILAC ratio of > 1.25.

proteasome (Fig. 4H-I)⁵³, suggesting that Pol II is being degraded. Likely, prolonged stalling of Pol II due to deficient TC-NER, initiated the last resort pathway. Degrading Pol II provides access for alternative repair pathways and prevents that the damage load becomes too high for cells to overcome^{42,43}.

Using proximity labeling in Apex2-RPB1 knock-in cells to identify Pol II interactors

So far, using the different Pol II extraction and immunoprecipitation methods, we revealed only a few known UV-specific Pol II-interacting proteins, while many more were expected. As crosslinking protein interactions, inhibiting Pol II degradation or inhibiting the TC-NER reaction did not result in more interactions, we set out to study the interactions with Pol II in living cells using proximity labeling proteomics. Proximity labeling methods use a biotin-tag to label proteins in close proximity of the bait protein⁷³. Biotin is covalently-bound to proteins⁷³ and consequently biotin-tagged proteins can be isolated using streptavidin, making it easy to identify interactors while using harsh extraction methods. We set up two different methods to biotinylate proteins either by using an APEX2 or a BioID2 tag. APEX2 is an engineered ascorbate peroxidase that uses hydrogen peroxide (H_2O_2) to convert biotin-phenol (BP) into highly reactive biotin-phenoxy radicals (BP[•]) (Fig. 5A)^{73,74}. These radicals are covalently attached as biotin to tyrosine residues on proteins in close proximity. APEX2 has the advantage that it uses short labeling times. BioID2 is a biotin ligase that activates biotin and covalently binds it to lysine residues of proteins in close proximity (Fig. 6A)^{73,75}. BioID2 is smaller than its predecessor BioID, making the targeting more specific. In addition, it has higher efficacy and requires less biotin⁷⁵. BioID2 has the advantage over APEX2 that it does not require addition of H_2O_2 , which can induce oxidative stress and thereby affect Pol II-mediated transcription⁷⁶. Moreover, BioID2 biotinylates lysines that are found more frequent throughout the human genome compared to tyrosines that are biotinylated by APEX2⁷³. Finally, biotin is actively transported into the cell, making it easily accessible⁷³ but requires long incubation times to reach maximum biotinylation.

We first assessed the APEX2-tag to identify Pol II-interactors that are in close proximity. Therefore, we endogenously tagged the largest subunit of Pol II and generated a homozygous Flag-APEX2-RPB1 KI cell line in MRC-5 cells, similar as described previously for GFP-RPB1 (Fig. 5B-C)⁴⁷. Next, we set out to determine the optimal conditions for successful biotinylation using these APEX2-KI cells. The cells did not need to be cultured in biotin-depleted media since normal media induced better biotinylation of proteins than biotin-depleted media, as determined by the streptavidin-positive smear (Fig. 5D). We started using 500 μ M biotin-phenol for 30 minutes as increasing the incubation time to 60 minutes did not increase biotinylation (Fig. 5D). Comparing different incubation times and concentrations of H_2O_2 showed that increasing the incubation time did not increase successful biotinylation, but increasing the of concentration H_2O_2 did enhance protein biotinylation (Fig. 5D). Of note, the three clearly biotinylated proteins are most likely endogenously biotinylated proteins as these were detected in all conditions, even in cells that do not express APEX2.

Next, we tested if we could decrease the concentration of biotin-phenol and maintain similar levels of biotinylation. This was not the case, so we kept on using 500 μM biotin-phenol (Fig. 5E). However, when we isolated biotin-tagged proteins using streptavidin beads, we did not enrich for these proteins, likely because of competition between the free reactive biotin-phenoxyl species and biotin-tagged proteins (Fig. 5E). To prevent competition and remove free biotin radicals, we increased the number of streptavidin beads during IP or we filtered the lysates before IP using a 3K filter that filters out all proteins smaller than 3 kDa. Both methods efficiently removed free biotin and successfully pulled down biotinylated proteins and we continued to use more beads to prevent competition (Fig. 5F). Finally, we compared RIPA buffer to different concentrations of Urea buffer to extract biotinylated proteins. Urea buffer more efficiently isolated biotinylated proteins than RIPA buffer but there was no clear difference between 2 or 4 M urea buffer (Fig. 5G). We wanted to be sure that the high concentration of urea did not influence identification of proteins using mass spectrometry, and therefore continued to use 2 M urea for protein extraction.

Apex2-mediated biotinylation successfully identified transcription elongation factors in basal conditions

To validate the proximity labeling method, we assessed the Pol II complex composition using SILAC-based mass spectrometry in undamaged conditions. Therefore, we compared heavy-labeled MRC-5 APEX2-KI cells to light-labeled MRC-5 Wt cells. We plotted the Log_2 SILAC ratios of APEX2-KI over Wt (Fig. 5H). Similar to the basal complex composition after native IP, there was a slight skew towards a Log_2 SILAC ratio of 1 (Fig. 2A, 5H). Many Pol II-interactors were identified with a Log_2 SILAC ratio bigger than 1, including most of the Pol II subunits, several elongation factors and multiple components of the PAF complex (Fig. 5I). To limit the selection of Pol II-specific interactors, we considered proteins to be significant interactors with an arbitrary SILAC ratio of 3.75 or bigger. In addition to the known Pol II-interacting proteins, the proteins identified as specific interactors were mainly involved in mRNA processing and splicing, as determined by GO term analysis of the biological process (Fig. 5J). For example, several heterogeneous nuclear ribonucleoproteins (hnRNPs) were identified (Fig. 5I-J). These proteins are involved in many different processes of the nucleic acid metabolism and for example help to regulate mRNA stability⁷⁷. This confirmed that we mainly identified proteins involved in transcription.

Since proximity labeling seemed to be successful in identifying Pol II-specific interactors, we continued to identify UV-specific interactors. Therefore, we light-labeled mock-treated APEX2-KI cells and compared them to heavy-labeled APEX2-KI cells that were harvested 20 minutes after irradiation with 20 J/m^2 UV-C. Directly before harvest, H_2O_2 was added for 1 minute to biotinylate UV-specific interactors and followed by quenching to stop biotinylation. We plotted the Log_2 SILAC ratios of UV-irradiated samples over mock-treated and identified significant interactors with a Log_2 SILAC ratio 0.5 or bigger (Fig. 5K). Similar to previous approaches and in line with literature, we identified CSB as a UV-specific Pol II interactor but only with few peptides (Fig. 5L). Besides CSB, no known TC-NER factors were identified. Interestingly, RPB5 (POLR2E), one of the Pol II subunits, was identified to be more biotinylated. This could indicate that RPB5 is more stably

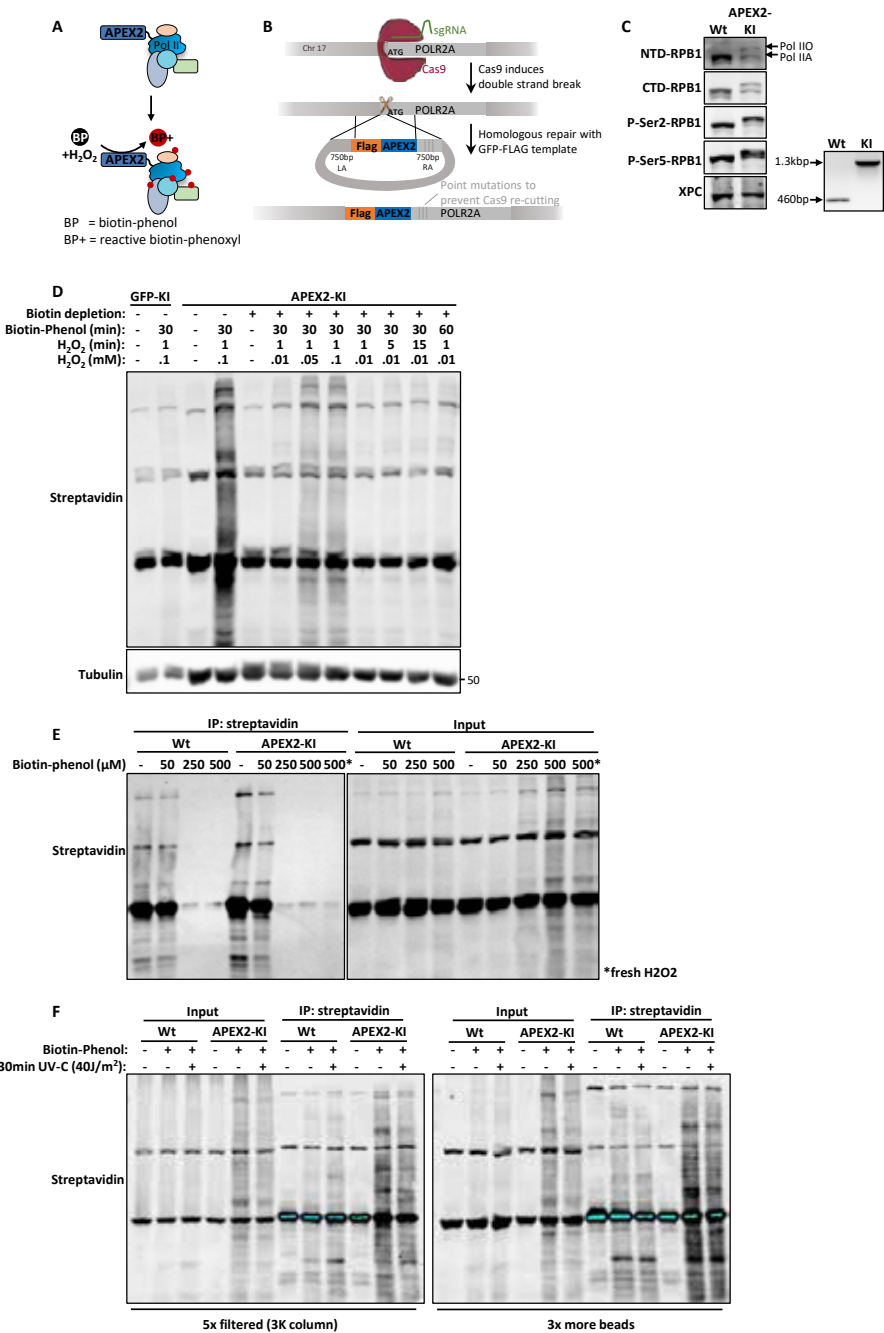
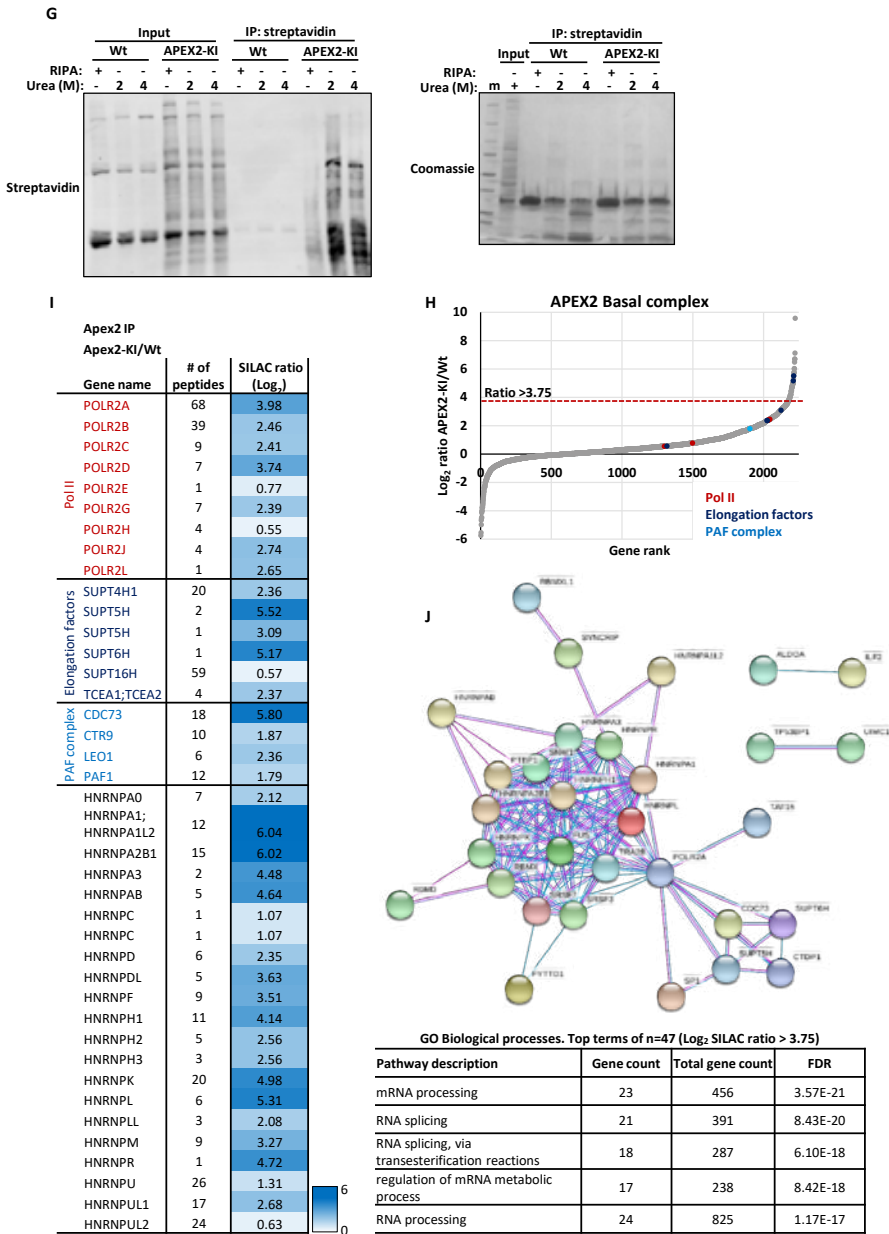


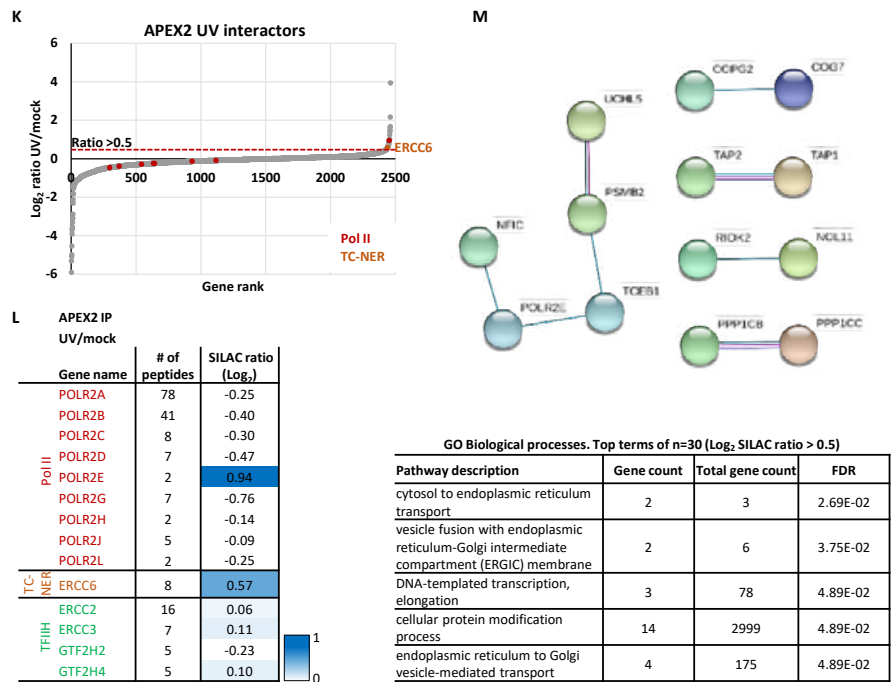
Figure 5. Optimization of proximity labeling and identification of Pol II basal complex and UV-specific interactors using a Flag-Apex2-RPB1 KI cell line. (A) Schematic for biotinylation using Flag-Apex2-RPB1 KI cells. **(B)** Schematic for generating MRC-5 Flag-Apex2-RPB1 KI cells. **(C)** Immunoblot for indicated proteins (left) and genotyping PCR (right) for MRC-5 Flag-Apex2-RPB1 KI cells. XPC was used as loading control. **(D)** Immunoblot stained for streptavidin in Flag-GFP-RPB1 or Flag-Apex2-RPB1 KI cells with indicated

Figure 5. (continued)



conditions for biotin depletion, incubation time with Biotin-Phenol (500 μ M), incubation time with H_2O_2 and concentration of H_2O_2 . Tubulin was used as loading control. (E) Top: Immunoblot stained for indicated proteins after adding Biotin-Phenol (500 μ M) and H_2O_2 (1 minute, indicated concentrations). Bottom: Immunofluorescence images showing indicated stainings after treatment with indicated conditions. Biotin-Phenol: 30 minutes, 500 μ M. H_2O_2 : 1 minute. (F) Immunoprecipitation of biotinylated proteins using streptavidin beads in Wt or Apex2-KI cells, followed by immunoblot analysis. H_2O_2 : 1 minute, 0.1 mM. Biotin-Phenol was added for 30 minutes with indicated concentrations. (G) Left: Immunoprecipitation of

Figure 5. (continued)



► biotinylated proteins using streptavidin beads in Wt or Apex2-KI cells, 30 minutes after mock-treatment or irradiation with 40 J/m² where indicated. Input was filtered 5 times with a 3K filter to remove free biotin. IP was followed by immunoblot analysis. Biotin-Phenol (500 μM) was added for 30 minutes directly after UV. Right: same as left but instead of using a filter, 3 times more streptavidin-beads were added during IP. (H) Immunoprecipitation using streptavidin beads in Wt or Apex2-KI cells followed by immunoblot analysis (top) or coomassie protein staining (bottom). After pull-down, beads were washed with RIPA or urea buffer as indicated. (I) Pol II-interacting proteins are ranked based on the Log₂ SILAC ratio determined using quantitative interaction proteomics comparing streptavidin IP of biotinylated proteins from MRC-5 Flag-Apex2-RPB1 KI cells to MRC-5 Wt cells. Dotted line indicates the threshold (> 3.75) for identified interactors. Pol II subunits, elongation factors and PAF complex subunits are color-coded. (J) Interaction heat map of the Log₂ SILAC ratios of Pol II-interacting proteins comparing biotinylated proteins from Apex2-RPB1 KI cells to Wt cells. Positive values indicate increased interaction. Second column indicates the number of peptides identified. (K) Top: protein-protein interactions as determined by STRING protein interaction analysis of genes with Log₂ SILAC ratio > 3.75. Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of Pol II-specific interactors with Log₂ SILAC ratio > 3.75. (L-N) as (I-K) but for UV-specific interactors by comparing MRC-5 Flag-Apex2-RPB1 KI cells 20 minutes after 20 J/m² UV-C to mock-treated cells and using a threshold of > 0.5 for identified interactors.

interacting with RPB1 or RPB5 comes in close proximity of the APEX2 tag due to structural changes. So far, no role for RPB5 has been identified in the UV-induced transcription stress response. In addition, we identified TCEB1 or Elongin C as a UV-induced interactor. Elongin C forms the heterotrimeric Elongin complex together with Elongin A (TCEB3) and B (TCEB2) that functions as a transcription elongation factor⁴⁹. Besides RPB5 and TCEB1, several hits were identified that do not have a clear correlation with Pol II. Therefore, these hits were considered to be less interesting interactors and were not followed

up. The top involved biological processes, as determined using GO term analysis, only revealed transcription-related and vesicle-related transport processes (Fig. 5M).

BioID2 proximity labeling to identify Pol II interactors

To further assess the efficiency of identifying Pol II interactors using proximity, we used the biotin ligase BioID2. Therefore, we tagged RPB1 on its endogenous locus, similar as for the APEX2-KI cells, and generated a homozygous MRC-5 Flag-BioID2-RPB1 KI cell line (Fig. 6B-C). First, we determined the optimal concentration of and incubation time with biotin for successful biotinylation. BioID2-KI cells have to be cultured in biotin-free media for several days before adding biotin because biotin from the culture media already biotinylated most proteins, as determined by the streptavidin-positive smear in absence of biotin depletion (Fig. 6D). Increasing the concentration of biotin did not increase efficient biotinylation but increasing the incubation time did increase biotinylation (Fig. 6D). Already after 2 hours, there was a clear increase in the streptavidin signal that further increased over time and seemed to be maximal after 24 hours, suggesting proteins kept on being biotinylated (Fig. 6D-E). The biotinylation was specific for Pol II because we could see a clear overlap in the streptavidin signal with the P-Ser5-RPB1 signal in immunofluorescence experiments, suggesting that Pol II was specifically biotinylated (Fig. 6E). Since increasing the biotin concentration did not improve biotinylation of target proteins, we investigated if it was possible to decrease the concentration. Indeed, lowering the concentration to 1 μM still induced successful biotinylation (Fig. 6F). With 50 μM biotin we biotinylated proteins but did not precipitate biotin-tagged proteins, suggesting that free biotin competed with biotin-tagged proteins thereby preventing specific pull-down (Fig. 6F). With the 1 μM concentration, we specifically isolated biotinylated proteins using streptavidin beads in BioID2-KI cells but not Wt cells and used this concentration for future experiments (Fig. 6F). Filtering the lysates before IP using a 3K filter to remove free biotin indeed successfully removed excess biotin and enabled pull-down of biotin-tagged proteins (Fig. 6G). Therefore, we included the filtering step to prevent competition.

To identify Pol II-specific interactors, we performed SILAC-based interaction proteomics in undamaged conditions. Therefore, we labeled MRC-5 Wt cells with light amino acids and MRC-5 BioID2-KI cells with heavy amino acids. The Log_2 SILAC ratio of BioID2-KI/Wt cells are plotted (Fig. 6H). A clear skew towards a Log_2 SILAC ratio of 1 was shown for the interactors. This suggests that we identified many specific Pol II interactors. Many interactors were identified with a Log_2 SILAC ratio bigger than 1, including some of the Pol II subunits, transcription elongation factors and subunits of the PAF complex (Fig. 6I). To limit the number of hits, all interactors with an arbitrary Log_2 SILAC ratio of 2.25 or bigger were considered to be Pol II-specific interactors. However, this excluded some of the known Pol II interactors. Possibly, 4 hours labeling was too short to identify Pol II-specific interactors and we needed to extend the labeling with biotin up to 24 hours. Alternatively, during the 4-hour labeling, a lot of non-specific interactors were biotinylated.

GO term analysis of the interactors with a ratio bigger than 2.25 did show mRNA metabolic processes as one of the top biological processes (Fig. 6J). In addition, the

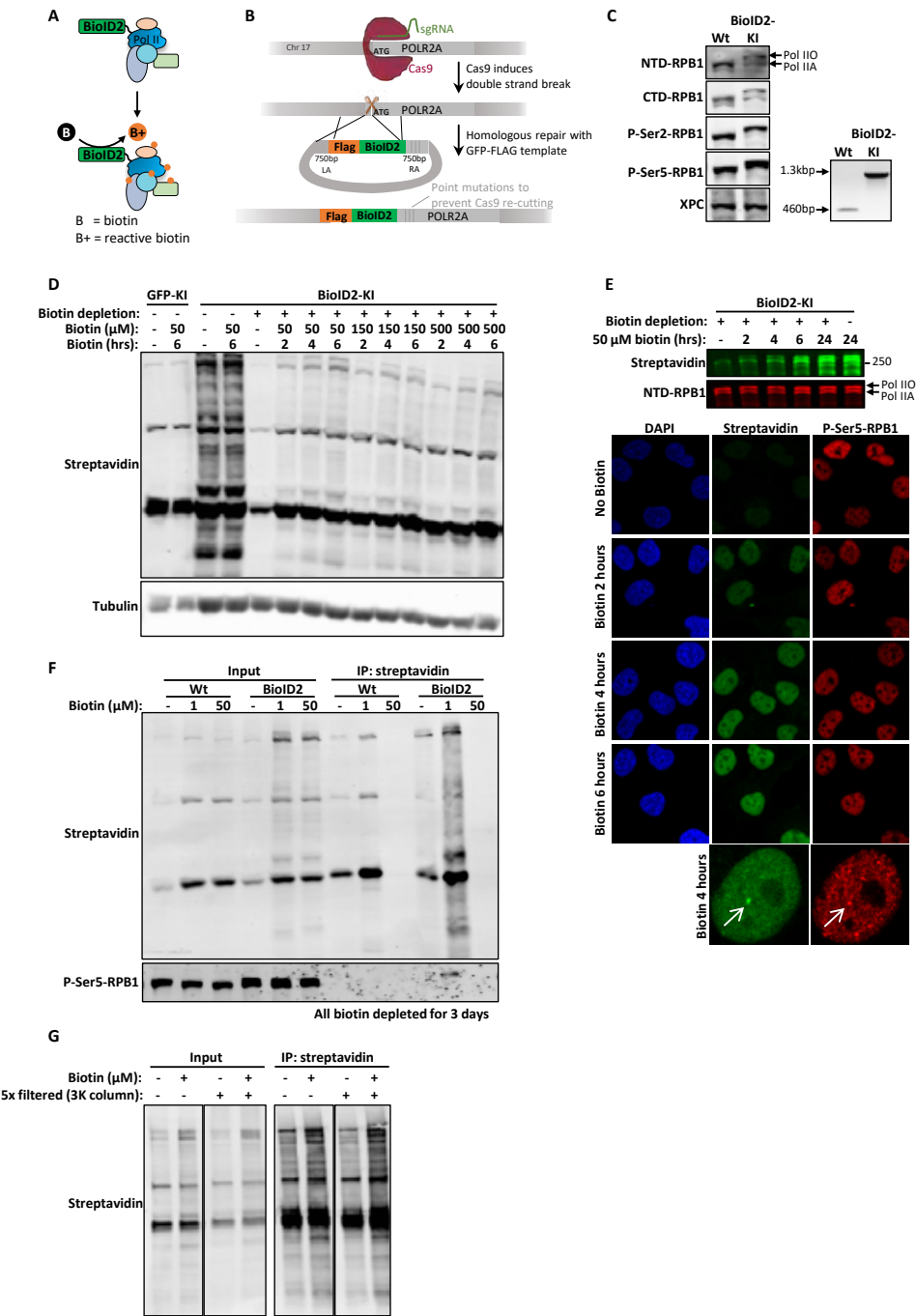


Figure 6. (continued)

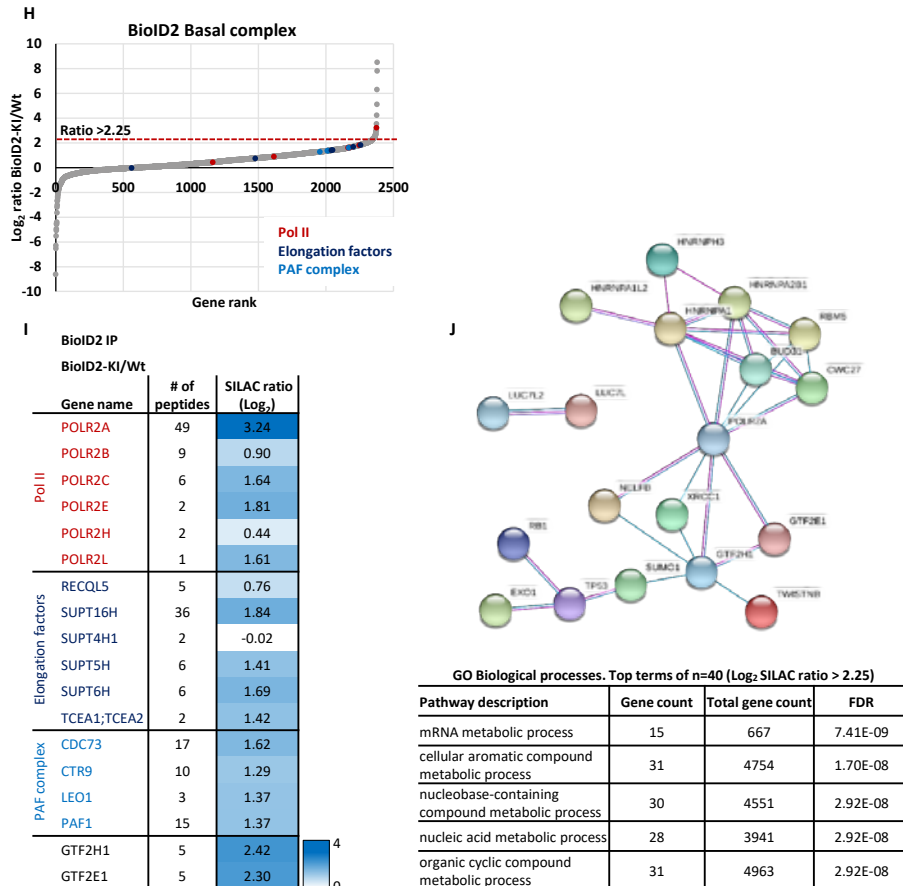
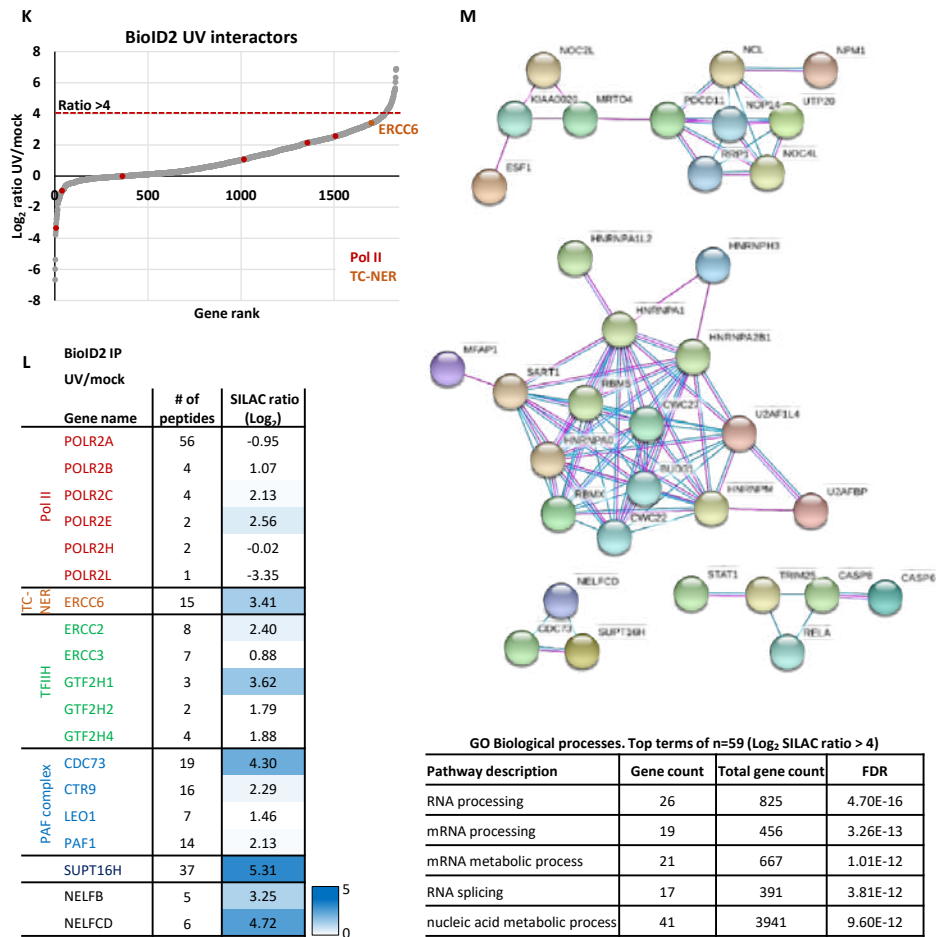


Figure 6. Optimization of proximity labeling and identification of Pol II basal complex and UV-specific interactors using a Flag-BioID2-RPB1 KI cell line. (A) Schematic for biotinylation using Flag-BioID2-KI cells. **(B)** Schematic for generating MRC-5 Flag-BioID2-RPB1 KI cells. **(C)** Immunoblot for indicated proteins (left) and genotyping PCR (right) for MRC-5 Flag-BioID2-RPB1 KI cells. XPC was used as loading control. **(D)** Immunoblot for streptavidin in Flag-GFP-RPB1 or Flag-BioID2-RPB1 KI cells after biotin treatment as indicated. Cells were biotin depleted for 24 hours where indicated. Tubulin was used as loading control. **(E)** Top: Immunoblot for streptavidin in BioID2-KI cells after indicated incubation times with biotin (50 μ M) and biotin depletion. Bottom: Immunofluorescence images of streptavidin-Alexa 488 or P-Ser5-RPB1 staining after indicated incubations with biotin. **(F)** Immunoprecipitation of biotinylated proteins using streptavidin-beads in BioID2-KI or Wt cells after incubating for 20 hours with indicated concentrations of biotin. Cells were biotin depleted for 72 hours prior to IP. **(G)** Immunoprecipitation using streptavidin-beads in BioID2-KI cells after 24 hours incubation with indicated concentrations of biotin. Lysates were filtered 5 times with 3K column prior to IP where indicated. **(H)** Pol II-interacting proteins were ranked based on the Log₂ SILAC ratios determined using quantitative interaction proteomics comparing streptavidin IP of biotinylated proteins from MRC-5 Flag-BioID2-RPB1 KI cells to MRC-5 Wt cells. Dotted line indicates threshold (> 2.25) for identified interactors. Pol II subunits, elongation factors and PAF complex subunits are color-coded. **(I)** Interaction heat map of the Log₂ SILAC ratios of Pol II-interacting proteins as determined by quantitative interaction proteomics following streptavidin-IP of biotinylated proteins from Flag-BioID2-RPB1 KI cells compared to Wt cells. Positive values indicate increased interaction. Second column indicates

Figure 6. (continued)



► the number of peptides identified. (**J**) Top: protein-protein interactions as determined by STRING protein interaction analysis of genes with Log₂ SILAC ratio > 2.25. Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of Pol II-specific interactors with Log₂ SILAC ratio > 2.25. (**K-M**) Same as (H-J) but for UV-specific interactors comparing BioID2-KI cells 4 hours after 20 J/m² to mock-treated and using a threshold of > 4 for identified interactors.

protein interaction analysis revealed interaction of Pol II with TFIH subunit p62 (GTF2H1) and with TFIIE (GTF2E1) (Fig. 6I-J). The known interactors were identified with much less peptides when compared to the number of peptides identified using APEX2 (Fig. 5J, 6I). This suggested that less peptides were biotinylated using the BioID2 compared to using APEX2. This could be explained by the short labeling using BioID2, which normally needs 24 hours labeling to reach maximum biotinylation whereas APEX2 works quickly and thus suggests a much less efficient or slower BioID2-mediated biotinylation.

Next, we used the BioID2-KI cell line to assess UV-specific Pol II interactors as identified using SILAC-based interaction proteomics. Since the biotin labelling takes four hours, this could provide a complete overview of the UV-induced Pol II-interacting proteins during the hours ranging from UV irradiation until the moment of harvesting the cells. To assess all the UV-induced interactors, we mock-treated BioID2-KI cells labeled with light amino acids and irradiated BioID2-KI cells labeled with heavy amino acids with 20 J/m² of UV-C 4 hours before harvest and directly added biotin to initiate biotinylation. We plotted the Log₂ SILAC ratios of UV-irradiated over mock-treated cells (Fig. 6K). Plotting the SILAC ratios showed a clear skew in the ratios from zero towards two, suggesting there was an imbalance in the incorporation of amino acids, an imbalance between labeled proteins during analysis, more proteins were bound to Pol II after UV in general or non-specific interactors were identified during the biotin labeling. This made it more challenging to identify UV-specific interactors, as many proteins have a ratio bigger than 1. For example, CSB, one of the interactors of lesion-stalled Pol II, was detected with a SILAC ratio of 3.41 (Fig. 6L). To limit the pool and identify new UV-specific interactors, we considered proteins to be UV-specific Pol II interactors with a ratio of 4 or higher. Interestingly, most identified biological processes were involved in RNA processing (Fig. 6M). This suggested that in response to UV, many RNAs were regulated and processed, which is in line with the highly regulated transcription-stress response as described in our introduction. Moreover, pathway analysis identified CDC73 and SPT16, both components involved in inducing promoter proximal pausing and release into transcription elongation¹. This is in line with the described increase in transcription elongation in response to UV¹⁸⁻²². Among the UV-specific interactors, few interactors were identified that could be associated with regulating Pol II and most interactors were considered to be non-specifically binding to Pol II.

Native immunoprecipitation approach to isolate elongating Pol II and identify its interactors in basal conditions and after UV irradiation

Thus far, we have used different extraction and immunoprecipitation methods to isolate all forms of Pol II, including initiating, promoter-paused and elongating Pol II. These methods enabled the identification of Pol II complex components in undamaged conditions, thereby proving the functionality of these methods. When we assessed UV-specific interactors, we identified the most prominent TC-NER factor CSB, but we did not identify previously unidentified UV-induced interactors. Perhaps searching for UV-specific interactors in the complete pool of Pol II dilutes the subset of lesion-engaged Pol II proteins and thereby makes the identification of new interactors difficult. To more specifically identify lesion-stalled Pol II, we used a different native approach that could specifically isolate and IP P-Ser2-modified Pol II. This subset of elongating Pol II recognizes the lesion in response to UV-induced DNA damage. However, the downside of this approach is that factors involved in the regulation of Pol II *in trans* will not be identified.

First, we compared immunoprecipitation using GFP beads to using a P-Ser2-modified Pol II-specific antibody (Fig. 7A). As seen in the input sample, this method extracted both initiating as well as elongating Pol II. After immunoprecipitation using the GFP-beads, both initiating and elongating Pol II were pulled down whereas using the P-Ser2-specific antibody, only elongating Pol II was pulled down as expected (Fig. 7A). Both IP methods

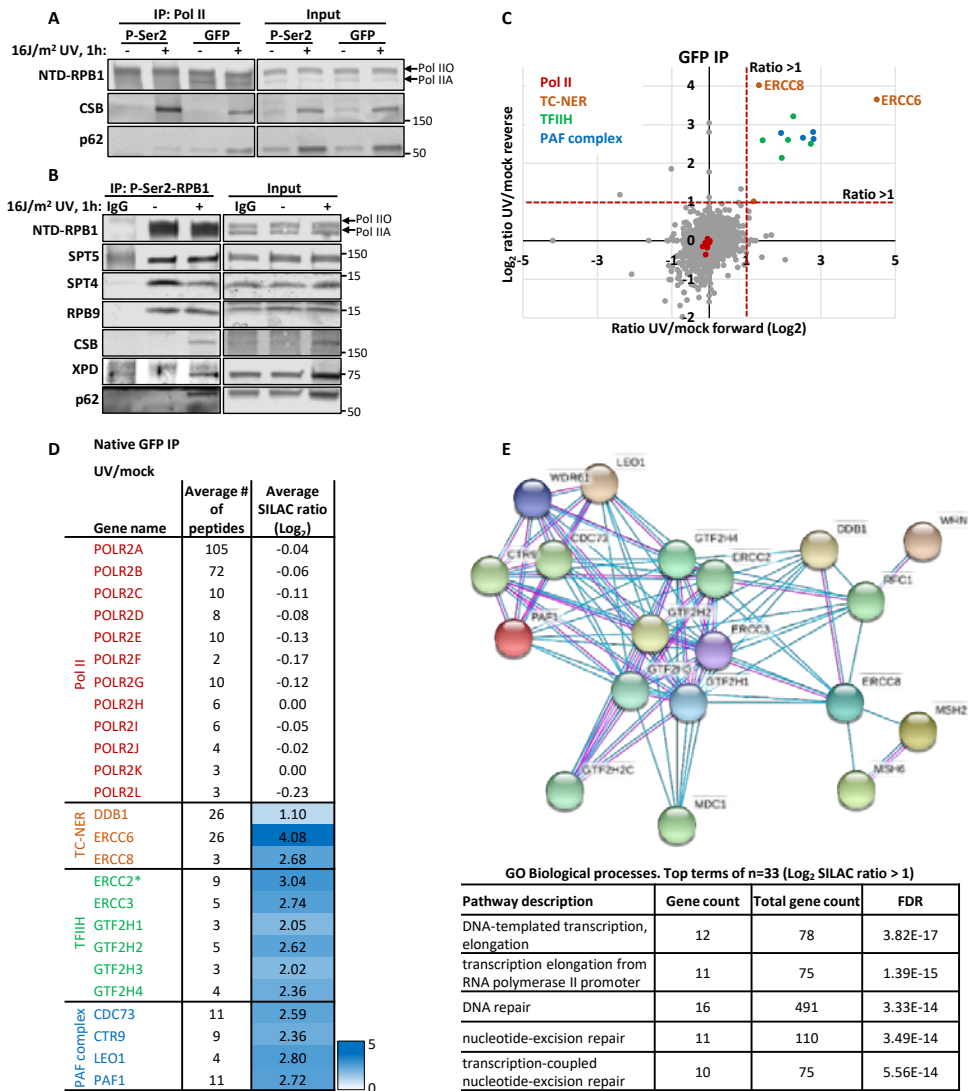
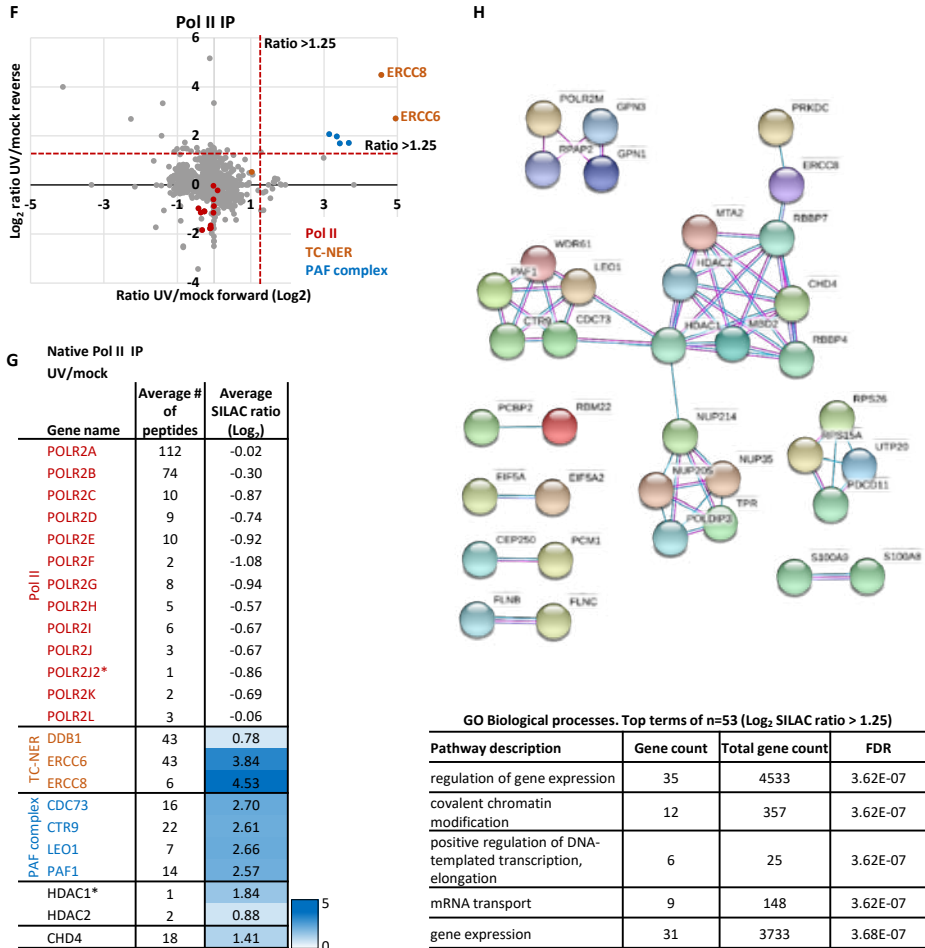


Figure 7. Native fractionation IP identifies UV-specific interactors. (A) Immunoprecipitation of Flag-GFP-RPB1 using a P-Ser2-RPB1-specific antibody or GFP-beads in MRC-5 Flag-GFP-RPB1 KI cells followed by immunoblot analysis of the indicated proteins. Cells were harvested 1 hour after 16 J/m² UV-C or mock-treatment. (B) Immunoprecipitation of P-Ser2-RPB1 in HCT116 cells followed by immunoblot analysis of the indicated proteins. Cells were harvested 1 hour after 16 J/m² UV-C or mock-treatment. IgG was used as binding control. (C) UV-specific Pol II-interacting proteins as determined using quantitative interaction proteomics of MRC-5 Flag-GFP-RPB1 KI cells 1 hour after 16 J/m² UV-C compared to mock-treated cells using GFP-beds. The Log₂ SILAC ratios of UV/mock-treated from the forward experiment are plotted against the Log₂ SILAC ratios of UV/mock-treated from the reverse experiment. Dotted lines indicate threshold (> 1) for identified interactors. Pol II subunits, TC-NER factors, TFIIH complex subunits and PAF complex are color-coded. (D) Interaction heat map of the average Log₂ SILAC ratios of UV-treated compared to mock-treated. Positive values indicate increased interaction, negative values indicate decreased interaction. Second column indicates the number of peptides identified. (E) Top: protein-protein interactions as

Figure 7. (continued)



determined by STRING protein interaction analysis of genes with Log₂ SILAC ratios > 1, identified in at least 1 of the experiments. Minimum required interaction score was set to high confidence (0.700) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of UV-specific interactors with Log₂ SILAC ratios > 1 identified in at least 1 of the experiments. (F-H) Same as (C-E) but using P-Ser2-RPB1 IP and using a threshold of > 1.25 for identified interactors.

also showed a UV-specific interaction with CSB and p62, one of the subunits of the TFIIH complex. Interestingly, p62 also interacted with Pol II in undamaged conditions using the GFP pull-down. This could be explained by the co-immunoprecipitation of initiating Pol II that interacts with TFIIH during transcription initiation. Further assessing the P-Ser2-specific pull-down of Pol II showed clear, UV-independent interactions with Pol II subunit RPB9 and elongation factors SPT4 and SPT5 (Fig. 7B). Moreover, we confirmed the UV-induced interaction with CSB and TFIIH subunits p62 and XPD (Fig. 7B).

Next, we performed SILAC-based interaction proteomics to determine new UV-induced Pol II interactors. To be able to compare the identified interactors to the previous experiments, we first assessed UV-specific interactors using the GFP IP to identify interactors from both initiating and elongating Pol II. Therefore, we harvested mock-treated light-labeled MRC-5 Flag-GFP-RPB1 KI cells and compared them to heavy-labeled MRC-5 Flag-GFP-RPB1 KI cells harvested one hour after 16 J/m² UV-C. We performed a label swap and plotted the Log₂ SILAC ratios of UV-irradiated over mock-treated from both experiments (Fig. 7C). Proteins identified with a Log₂ SILAC ratio of 1 or bigger were considered to be UV-specific Pol II interactors. The average SILAC ratios of both experiments were plotted in an interaction heatmap (Fig. 7D). We identified all 12 Pol II subunits with more peptides than with any of the previous methods and none of them had a changed interaction as expected (Fig. 7D). We identified CSB, CSA and DDB1, one of the components of the CRL4^{CSA} E3 ubiquitin ligase (Fig. 7C-D)⁷⁰, similar as with the DSP XIP (Fig. 4E). However, in contrast to the DSP XIP, we also identified several subunits of the TFIIH complex as UV-specific interactors. TFIIH is one of the downstream factors during TC-NER, involved in verification of the lesion¹⁷. The fact that we identified TFIIH as a UV-specific interactor suggested that the subset of lesion-stalled Pol II, involved in TC-NER, is bigger in this IP procedure than in the previous experiments (Fig. 7E) or the conditions used during this Pol II isolation procedure were more favorable to maintain protein-protein interactions. Further assessing the UV-specific interactors revealed several subunits of the PAF complex (Fig. 7C, 7E). The PAF complex has been shown to be involved in transcription elongation and has previously been shown to bind to Pol II in response to UV damage⁷⁸.

Immunoprecipitation of the elongating Pol II identified several known UV-induced interactors

Since this extraction method in combination with pulling down both initiating and elongating Pol II already successfully identified known repair-engaged proteins, we wondered if immunoprecipitation of the lesion-bound subset of Pol II further increased the specificity of identifying UV-induced interactors. Therefore, we performed the same interaction proteomics on P-Ser2-modified Pol II and identified UV-specific interactors by comparing mock-treated MRC-5 Flag-GFP-RPB1 KI cells to MRC-5 Flag-GFP-RPB1 KI cells harvested 1 hour after irradiation with 16 J/m² UV-C. We performed a label swap and plotted the Log₂ SILAC ratios of both experiments (Fig. 7F) and the average Log₂ SILAC ratios in an interaction heatmap (Fig. 7G). Proteins were identified as significant interactors with a Log₂ SILAC ratio of 1.25 or bigger. Once more we identified all the Pol II subunits and showed that their interaction was not changed upon DNA damage (Fig. 7F-G). We identified CSA, CSB and DDB1, similar as with the GFP IP but now with many more peptides for CSB and DDB1, suggesting we enriched for lesion-stalled Pol II. Interestingly, we did not detect peptides from any of the TFIIH subunits so we could not determine if their interaction was changed upon UV-induced damage. However, we did show a UV-induced interaction of TFIIH with Pol II using the P-Ser2-modified Pol II IP followed by immunoblot analysis.

Protein interaction analysis of interactors with a ratio bigger than 1.25 in one or both experiments revealed enrichment for the PAF complex, similar to the GFP IP, chromatin remodeler CHD4, and histone deacetylases HDAC1 and 2 (Fig. 7H). CHD4 has been

shown to be involved in enabling homologous recombination to remove double strand breaks (DSBs) together with HDAC1 and 2 and other proteins from their cluster (Fig. 7H)⁷⁹. Moreover, the two HDACs have been shown to be important for non-homologous end-joining, another mechanism to repair DSBs⁸⁰. Possibly, these HDACs are recruited to stalled Pol II in response to a conflict with a progressing replication fork that may finally result in a DSB. The chromatin remodeling ability of the HDACs is likely needed to enable DSB repair. Alternatively, the recruitment of these factors to lesion-stalled Pol II could also suggest a yet unknown role for this complex in remodeling the chromatin to facilitate efficient TC-NER but this needs further investigation. GO term analysis indicated that we mainly identified proteins involved in the regulation of transcription and gene expression (Fig. 7H). This suggests that we isolated lesion-stalled Pol II and were able to look at genome-wide *in trans* effects on transcription although we did not identify a clear UV-induced Pol II-interactor involved in regulating transcription.

Combined analysis of all UV-induced Pol II interaction experiments

The in this chapter described different Pol II isolation procedures resulted in the identification of a few UV-induced Pol II-interactors. However, most screens resulted in the identification of different interactors that could be identified as a UV-induced Pol II interactor due to the chosen stringent SILAC ratios. To obtain a more complete overview of DNA damage-induced interactors, we combined all hits for the described proteomic screens. With a lower threshold, some factors will be identified in several experiments and could be potential candidates. Subsequently, for such an overall analysis we combined the identified UV-specific interactors of the individual experiments into one analysis. Since we were interested in interactors that were found with multiple approaches, we loosened the threshold for each experiment to detect more interacting proteins. Using the lower thresholds, we determined for each protein in how many experiments it was identified as an interactor. Proteins were considered UV-specific Pol II interactors when they were identified in four or more different methods (Table 1). Most of the methods contained a label swap or two different time points, while for the proximity labeling methods there was only one experiment to identify UV-interactors. To compensate for this, we decided to weigh the identification of a hit in the APEX2-KI cells or the BioID2-KI cells twice.

GO term analysis of the biological processes of the identified UV-induced interactors revealed that most were involved in transcription elongation (Fig. 8A). More directly looking at the protein-protein interactions of these top hits using STRING protein-protein analysis revealed that we identified several components of the TC-NER complex, including CSA and other components of the CRL4^{CSA} E3 ubiquitin ligase complex such as DDB1, Cullin 4A (CUL4A) and NEDD8, a ubiquitin-like molecule that covalently binds CUL4A to activate the ubiquitin ligase complex⁴¹. Moreover, we identified several subunits of the TFIIF complex (ERCC2, GTF2H1) confirming that TFIIF was recruited to lesion-stalled Pol II¹⁷. This indicates that combining several proteomic experiments using lower thresholds is a promising approach to identify UV-specific interactors. For example, the TC-NER involved Cul4A/B was not recognized in individual experiments with a stringent threshold, but is identified as a bona-fide hit using this combined approach. Interestingly, we did not identify the endonucleases ERCC1/XPF and XPG. This suggests that Pol II

Pathway description	Gene count	Total gene count	FDR
DNA-templated transcription, elongation	11	78	5.64E-11
regulation of DNA-templated transcription, elongation	9	44	3.50E-10
transcription elongation from RNA polymerase II promoter	10	75	4.98E-10
protein modification by small protein conjugation or removal	22	945	2.07E-09
positive regulation of DNA-templated transcription, elongation	7	25	9.80E-09

Pathway description	Gene count	Total gene count	FDR
snRNA metabolic process	18	88	1.72E-21
snRNA transcription by RNA polymerase II	17	71	1.77E-21
ncRNA metabolic process	26	497	5.39E-19
nucleic acid metabolic process	56	3941	7.41E-19
RNA processing	30	825	3.22E-18

needs to be displaced or translocated from the lesion before the endonucleases can bind the DNA and excise the lesion. Interaction with the TC-NER complex clearly illustrates that repair was initiated upon stalling of Pol II on a lesion and suggests that we extracted lesion-stalled Pol II in response to UV. Of note, CSB was not identified using STRING protein-protein analysis but was identified as one of the top hits in almost all proteomics approaches (Table 1).

In addition to the components of the TC-NER pathway, we also identified several subunits of the 26S proteasome, which is involved in degradation of proteins⁵³. The identification of these proteasome subunits could indicate that a part of Pol II, or RPB1 specifically, is degraded as part of the so called last resort pathway, for examples as a consequence of an insurmountable DNA damage load^{42,43}. This is surprising, since we also identified TC-NER components, indicating that repair is correctly initiated. This could indicate that TC-NER and the last resort pathway can be active at the same time, most likely at different lesion-stalled Pol II complexes. Alternatively, Pol II was degraded to provide access for the repair proteins to repair the lesion¹⁷. In line with the identification of the proteasome and Pol II degradation, we also identified Elongin A (TCEB3) and B (TCEB2). These components of the heterotrimeric Elongin complex have been shown to form an E3 ubiquitin ligase complex together with RBX2 and Cullin 5 upon induction of genotoxic stress⁴⁹. Consequently, this complex ubiquitylates RPB1 and thereby targets it for proteasomal degradation as part of the last resort pathway^{42,49}. The interaction of Cullin 5 to Pol II was enriched in three of the experiments, but RBX2 was not identified in any of the experiments (Table 1). Identification of these factors supported the idea that a subset of Pol II was degraded. However, we could not exclude that the increased interaction of the elongation factor Elongin to Pol II was a result of a larger fraction of elongating Pol II complexes.

Identification of transcription-regulating and RNA-regulating and -processing proteins suggests a highly regulated UV-induced transcription response

The idea that transcription elongation is stimulated to regulate transcription levels in response to UV-irradiation²¹⁻²³ is supported by the UV-induced identification of the PAF complex, which plays an important role during transcription elongation⁵¹. As recently shown, the PAF complex was not essential for repair of UV-induced lesions but promoted restart of transcription after repair, likely via releasing Pol II from the promoter-proximal pause site into transcription elongation⁸¹. Interestingly, we already identified the interaction of the PAF complex with Pol II one hour after UV. However, at this time point, transcription is still inhibited and transcription only partially recovers after six hours⁸¹.

- ◀ **Figure 8. Combined analysis of the different approaches identified multiple UV-specific Pol II interactors.** (A) Top: protein-protein interactions as determined by STRING protein interaction analysis of upregulated UV-specific Pol II interactors identified with a significant Log_2 SILAC ratio in at least 4 experiments. Threshold for significant SILAC ratios was determined for each experiment separately. Minimum required interaction score was set to medium confidence (0.400) and only connected nodes based on experimental data (pink) or curated databases (blue) were shown. Bottom: Top 5 enriched gene ontology (GO) terms (biological processes) of UV-specific interactors identified with a significant Log_2 SILAC ratio in at least 4 experiments. (B) Same as (A) but for genes that were significantly released from Pol II upon UV irradiation in at least 4 experiments. Minimum required interaction score was set to high confidence (0.700).

Table 1. UV-specific Pol II interacting proteins identified using combined analysis of the different approaches. Highlighted values are considered significant and specific interactors are ranked based on the

Log ₂ Ratio H/L normalized																
Gene names	PFA XIP				Native GFP IP	DSP XIP				APEX2 IP	BioID2 IP	Native fractionation				# of times identified:
	Forward	VCPI		15 minutes		XPA KO		GFP IP Forward	GFP IP Reverse			Pol II IP Forward	Pol II IP Reverse			
		Reverse	Forward			Reverse	Forward							Reverse	Forward	
ERCC6	2.74	3.20	2.53	2.63	NaN	3.44	2.68	2.92	1.96	0.60	3.45	4.56	3.77	4.94	2.59	16
RECQL5	NaN	0.76	0.52	-0.27	0.17	0.86	0.66	0.51	NaN	-0.91	4.27	0.74	0.47	1.15	-0.08	9
DDB1	0.16	0.94	0.31	0.42	0.31	0.97	0.51	0.92	0.59	-0.15	0.37	1.20	1.04	1.06	0.59	9
CDC73	-0.08	0.37	0.05	0.17	0.53	0.18	-0.04	0.25	-0.22	-0.15	4.36	2.50	2.64	3.64	1.69	7
CUL4A	0.24	0.82	0.20	0.18	NaN	0.51	0.41	NaN	NaN	0.17	NaN	0.52	0.53	0.51	-0.62	6
HSPB1	0.04	-0.26	0.26	-0.09	0.17	0.57	-0.08	0.92	-0.05	-0.03	3.56	1.57	0.15	0.87	0.12	6
ERCC8	NaN	NaN	NaN	NaN	NaN	NaN	2.51	NaN	3.17	NaN	NaN	1.32	3.99	4.65	2.40	6
PSMD8	0.28	-0.22	0.15	-0.06	0.81	1.74	0.10	0.63	0.16	0.25	NaN	0.55	0.17	NaN	NaN	6
CUL4B	0.20	0.14	0.29	0.27	0.17	0.72	0.40	0.70	0.46	-0.26	NaN	0.74	0.28	0.72	-0.41	5
NEDD8-MDP1;																
NEDD8	NaN	NaN	1.49	0.18	-0.81	NaN	0.42	0.26	NaN	NaN	NaN	0.88	0.58	0.69	-0.31	5
RPS21	0.10	0.02	0.08	-0.28	-0.02	NaN	0.37	0.94	0.90	0.82	NaN	NaN	NaN	NaN	-0.08	5
PELP1	-0.21	0.60	-0.25	0.31	NaN	-0.09	-0.30	-0.11	-0.36	-0.25	3.11	0.41	NaN	-0.38	0.67	5
NCL	0.02	0.10	-0.01	0.35	-0.27	-0.01	-0.17	0.03	-0.20	-0.69	6.02	0.64	-0.31	0.47	1.11	5
MKI67	0.02	0.65	0.05	0.24	0.19	0.04	1.05	-0.23	-0.27	-1.11	3.18	-0.15	0.03	-0.28	1.24	5
GTF2H1	-0.29	0.40	-0.05	0.07	NaN	0.92	NaN	NaN	0.11	-1.52	3.54	2.24	2.19	NaN	NaN	5
TCEB3	0.57	0.41	0.38	0.15	NaN	0.48	-0.19	-0.17	-0.27	-0.74	3.25	NaN	NaN	NaN	NaN	5
ATPSA1	0.69	-0.50	-0.52	-0.19	0.04	NaN	0.91	0.97	1.41	0.04	0.05	0.07	-0.20	-0.22	-1.69	4
TCEB2	0.12	-0.36	0.12	0.24	0.51	0.11	0.13	0.59	0.19	0.36	1.38	0.06	0.44	0.16	-0.57	4
HSPA9	0.50	-0.75	0.04	0.21	-0.10	0.04	0.39	0.32	0.46	0.23	0.25	0.24	-0.01	0.44	-0.53	4
KIAA0020	1.04	0.41	NaN	0.43	NaN	0.08	NaN	-0.05	-0.01	0.07	4.38	-0.40	-0.27	NaN	-0.51	4
ZNF326	-0.23	0.61	0.09	0.05	-0.26	-0.02	-0.02	-0.17	-0.23	-0.41	3.45	-0.10	-0.35	0.58	-0.16	4
RRP1	NaN	NaN	NaN	0.30	NaN	0.44	-0.30	-0.01	-0.46	-0.03	4.35	-0.47	0.01	-0.55	-0.09	4
PPFIBP1	-0.18	0.02	-0.03	-0.08	NaN	NaN	NaN	NaN	1.69	NaN	3.06	NaN	0.24	NaN	-0.06	4
NOC4L	0.45	0.94	0.51	0.05	NaN	-0.01	NaN	-0.86	NaN	0.15	4.81	-0.17	-1.46	-0.72	0.02	4
PHIP	1.15	-0.34	-0.15	0.09	NaN	-0.77	0.39	0.24	-0.11	0.14	-0.09	0.74	0.20	0.53	0.64	4
PFN1	0.24	-0.36	0.04	0.06	0.05	0.30	-0.12	0.66	0.34	0.20	NaN	NaN	NaN	NaN	0.80	4

number of times they were identified in the different experiments. Threshold for significant Log₂ SILAC ratios was determined for each experiment separately. Complete table can be found in supplemental table 2.

Log ₂ Ratio H/L normalized																
Gene names	PFA XIP				Native GFP IP	DSP XIP				APEX2 IP	BioID2 IP	Native fractionation				# of times identified:
	Forward	VCPi		15 minutes		75 minutes	XPA KO		GFP IP Forward			GFP IP Reverse	Pol II IP Forward	Pol II IP Reverse		
		Reverse	Forward				Reverse	Forward							Reverse	
DNTTIP2	NaN	NaN	0.39	0.00	NaN	0.09	NaN	-0.15	-0.40	0.05	3.20	-0.05	0.21	-0.28	0.83	4
RPS10;																
RPS10P5	0.02	0.03	0.02	-0.04	0.00	NaN	0.08	0.62	0.44	0.07	3.56	0.16	-0.46	-0.05	0.89	4
WDR61	0.10	0.08	0.02	0.35	0.01	0.26	0.08	0.57	0.04	-0.14	NaN	NaN	2.78	3.07	0.96	4
FASN	0.12	-0.19	0.44	-0.18	0.92	-0.04	-0.03	0.46	0.31	-0.06	0.04	0.14	0.48	NaN	0.96	4
UTP20	-1.03	NaN	-0.65	0.97	NaN	0.14	-0.26	-0.12	-0.59	0.01	4.19	-1.04	NaN	-1.03	1.52	4
PAF1	-0.02	0.35	0.14	0.13	0.28	0.10	0.03	0.08	-0.16	-0.28	2.09	2.83	2.75	3.46	1.81	4
LEO1	0.02	0.33	0.10	0.20	0.01	0.07	0.07	-0.02	-0.07	-0.68	1.37	2.82	2.76	3.37	2.10	4
CTR9	0.01	0.37	-0.04	0.11	0.06	0.23	0.03	0.28	-0.13	-0.24	2.03	1.91	2.82	3.15	2.11	4
DDX47	NaN	NaN	0.35	0.66	0.44	-0.12	-0.02	-0.19	-0.49	0.01	3.13	-0.25	-0.12	-0.66	NaN	4
NONO	-0.23	0.35	0.12	0.14	0.43	-0.06	-0.12	-0.06	-0.18	-0.13	3.80	0.00	0.31	2.10	NaN	4
ZGPAT	NaN	NaN	0.44	-0.47	-0.15	NaN	NaN	1.11	-0.25	NaN	3.27	0.19	-0.34	NaN	NaN	4
NOSIP	-0.27	0.46	0.31	-0.19	0.64	-1.19	-0.26	NaN	0.11	NaN	5.02	-0.09	-0.18	NaN	NaN	4
OPA1	NaN	-0.16	0.03	NaN	0.46	NaN	0.45	1.69	0.85	0.18	-0.04	0.53	-0.04	NaN	NaN	4
TMEM33	NaN	NaN	0.11	0.38	0.05	NaN	0.20	NaN	NaN	0.25	0.02	0.45	0.05	NaN	NaN	4
MAP2K3	0.23	0.58	-0.42	0.04	-0.05	NaN	0.00	-0.03	-0.17	-0.16	3.16	0.49	0.07	NaN	NaN	4
TSEN34	-0.20	0.49	0.06	-0.62	NaN	0.21	0.06	NaN	-0.55	0.29	3.91	-0.07	0.10	NaN	NaN	4
ASNS	1.00	-0.78	0.02	-0.08	-1.00	NaN	0.53	NaN	0.36	-0.14	3.88	NaN	0.17	NaN	NaN	4
PSMA7	0.52	NaN	0.02	-0.06	0.92	NaN	0.36	0.63	0.21	-0.31	NaN	-0.24	0.29	NaN	NaN	4
ACACA	0.09	-0.23	0.26	-0.38	0.55	NaN	0.26	NaN	0.61	0.45	-0.41	0.48	0.30	NaN	NaN	4
VRK3	NaN	NaN	0.97	NaN	NaN	NaN	NaN	NaN	NaN	0.23	NaN	-0.01	0.30	NaN	NaN	4
CASP6	NaN	NaN	0.56	-0.33	NaN	NaN	NaN	NaN	NaN	NaN	5.08	NaN	0.33	NaN	NaN	4
FMNL3	-0.08	NaN	-0.16	-0.14	NaN	0.46	0.01	-0.56	-0.67	0.35	0.42	NaN	0.39	NaN	NaN	4
ECI2	0.94	-1.32	-1.43	0.38	0.57	NaN	NaN	NaN	NaN	-0.01	-0.05	0.30	0.45	NaN	NaN	4
CDKN2A	0.10	-0.31	0.71	-0.60	NaN	NaN	-0.76	-0.27	-0.43	NaN	3.12	0.17	0.48	NaN	NaN	4
PSMC1	0.30	-0.06	-0.12	-0.14	0.68	0.25	0.11	0.73	0.32	-0.01	0.62	0.78	0.51	NaN	NaN	4
DHCR24	NaN	NaN	0.38	-0.39	0.63	-0.26	0.20	0.91	0.33	0.19	-0.15	0.12	0.52	NaN	NaN	4

Log ₂ Ratio H/L normalized																
Gene names	PFA XIP				DSP XIP				Native fractionation							
	Forward	VCPI		Native GFP IP	15 minutes	75 minutes	XPA KO		APEX2 IP	BioID2 IP	GFP IP		Pol II IP Forward	Pol II IP Reverse	# of times identified:	
		Reverse	Forward				15 minutes	75 minutes			Forward	Reverse				
PUM1	0.05	0.27	0.11	0.02	-0.09	0.02	NaN	NaN	NaN	4.11	1.03	0.52	NaN	NaN	4	
RPL38	0.15	0.07	0.03	-0.12	0.02	0.32	0.40	0.12	0.27	0.26	0.33	0.56	NaN	NaN	4	
HLA-H	-0.21	0.51	NaN	0.44	0.32	0.45	0.45	0.12	-0.27	-0.29	NaN	0.64	NaN	NaN	4	
PFDN2	0.68	-0.98	0.08	-0.49	-0.73	NaN	0.36	0.36	NaN	NaN	-0.10	0.65	NaN	NaN	4	
G3BP1	-0.21	0.25	0.37	-0.06	0.17	NaN	-0.36	0.13	0.06	4.48	NaN	0.68	NaN	NaN	4	
PSMD1	0.51	-0.44	-0.04	-0.07	0.75	-0.10	0.20	0.93	-0.15	0.03	NaN	0.78	NaN	NaN	4	
MRE11A	-0.05	-0.12	0.06	0.04	0.04	0.20	0.07	0.60	-0.21	3.22	-0.04	0.85	NaN	NaN	4	
EIF3G	-0.21	0.27	0.09	-0.13	0.06	0.30	-0.10	0.76	0.25	2.23	NaN	0.94	NaN	NaN	4	
MSH6	-0.14	0.09	0.02	0.01	0.22	0.20	0.07	0.41	-0.02	3.00	1.21	1.97	NaN	NaN	4	
GTF2H2C;																
GTF2H2	-0.42	0.26	0.05	0.16	-0.32	0.65	0.63	NaN	-0.16	NaN	2.47	2.54	NaN	NaN	4	
ERCC2	-0.19	0.51	0.27	-0.05	-0.09	0.69	0.03	NaN	0.08	2.41	2.17	3.17	NaN	NaN	4	
SLC3A2	3.28	-0.16	0.43	-0.40	-0.72	NaN	0.64	0.65	0.12	0.13	NaN	NaN	NaN	NaN	4	
RTF1	0.17	0.71	0.07	0.05	NaN	0.45	0.47	0.87	0.11	0.54	NaN	NaN	NaN	NaN	4	
VCPIP1	0.08	-0.20	0.06	-0.01	NaN	NaN	NaN	NaN	0.25	3.02	NaN	NaN	NaN	NaN	4	
PEG10	NaN	NaN	0.30	-0.21	1.26	NaN	NaN	NaN	NaN	3.20	NaN	NaN	NaN	NaN	4	
FANCG	NaN	NaN	NaN	2.97	NaN	NaN	0.12	NaN	-0.40	3.82	NaN	NaN	NaN	NaN	4	
KRIT1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.97	3.83	NaN	NaN	NaN	NaN	4	
USP48	NaN	NaN	0.43	0.12	0.90	NaN	NaN	-0.19	-0.86	4.00	NaN	NaN	NaN	NaN	4	
DIS3	-0.42	0.06	-0.03	0.09	0.14	NaN	0.17	1.29	-0.03	4.42	NaN	NaN	NaN	NaN	4	
TIMELESS	NaN	-0.38	-0.26	0.32	NaN	1.12	NaN	NaN	-0.44	5.44	NaN	NaN	NaN	NaN	4	
PSMD12	0.21	-0.85	0.04	-0.03	0.67	-0.24	0.31	0.94	0.30	NaN	NaN	NaN	NaN	NaN	4	
PSMB2	0.25	-0.01	0.29	-0.05	1.12	-0.32	0.26	0.64	0.51	NaN	NaN	NaN	NaN	NaN	4	
PSMD7	0.26	-0.36	0.01	0.05	0.79	-0.02	0.20	1.53	0.72	NaN	NaN	NaN	NaN	NaN	4	
COPG2	0.69	NaN	0.14	0.00	0.79	NaN	-0.59	NaN	1.90	NaN	NaN	NaN	NaN	NaN	4	

Possibly, this interaction is not only important for transcription restart, but might also contribute to promoter clearance²⁰ or an increased release from the promoter-proximal pause site to stimulate transcription elongation and sense for TBLs to initiate repair²¹⁻²³.

In addition to the identification of the PAF complex, we also identified RECQL5 as a UV-specific interactor of Pol II. RECQL5 has previously been shown to be a direct interactor of Pol II⁸² and inhibit transcription elongation⁸³. Recently it was shown that RECQL5 indeed inhibits transcription to preserve a more dispersed transcription elongation. This prevents interrupted transcription and transcription stress and thereby prevents genome instability⁸⁴. This might indicate that the increased interaction of RECQL5 with Pol II in response to UV (Fig. 8A, Table 1), slows down transcription elongation, perhaps to prevent transcription stress-induced genome instability. Interestingly, we also identified MRE11 as a UV-induced Pol II interactor, which is part of the MRN complex and has been shown to direct recruitment RECQL5 to DSBs to regulate repair⁸⁵. This might suggest that RECQL5 could be recruited to lesion-stalled Pol II complexes near DSBs via the MRN complex. However, this recruitment can be independent of transcription, as suggested previously⁸⁶. For example, if RECQL5 is recruited to DSBs induced by transcription-replication conflicts, which happen in close proximity of lesion-stalled Pol II. The latter hypothesis is supported by the identification of TIMELESS (Fig. 8A, Table 1) that normally binds stalled replication forks to maintain their stability and promote restart⁸⁷.

We also identified several RNA regulating and processing factors such as DDX47, DIS3 and PUM1. DDX47 has been shown to be recruited to process RNA transcripts to relieve R-loops and prevent transcription-replication conflicts⁸⁸. DIS3 (or RRP44) is a component of the exosome complex involved in regulating decay of aberrant RNAs⁸⁹. PUM1 is one of the pumilio proteins that contain an RNA binding domain and are involved in promoting mRNA decay and inhibiting gene expression by preventing translation⁹⁰. In some cases, they can also regulate expression of non-coding RNAs, which might be an important action in the UV-induced transcription stress response⁹⁰. Regulating or even degrading mRNAs might be a crucial step in the transcription stress response, for example, when Pol II is removed from the DNA template⁴⁶ or degraded upon DNA damage. In these scenarios, the RNA transcripts might be aberrant or incomplete and therefore need to be processed. The identification of these RNA processing factors suggests that there are different processing mechanisms for mRNAs, depending on the fate of lesion-stalled Pol II.

Factors that are released from Pol II upon UV-induced DNA damage

By combining the results of the different approaches, we also identified several proteins that lost their interaction with Pol II upon UV irradiation in several experiments using lower thresholds. Similar as for the UV-specific interactors, we counted the proximity labeling approaches twice to have a more balanced count based on the experimental background. Proteins that released from Pol II in four or more experiments were considered UV-specifically downregulated for this analysis (Table 2).

Interestingly, we identified RPB4, RPB5, RPB7 and RPB10, all subunits of the Pol II complex, to be released from Pol II in response to UV-induced DNA damage. (Fig. 8B, Table 2).

Table 2. Proteins releasing from Pol II in response to UV identified using combined analysis of the different approaches. Highlighted values are considered significant and specific releasing proteins are ranked based

Gene names	Log ₂ Ratio H/L normalized															
	PFA XIP				DSP XIP					Native fractionation						
		VCPI	VCPI	Native	15	75	XPA KO	XPA KO			BioID2	GFP IP	GFP IP	Pol II IP	Pol II IP	# of times
	Forward	Reverse	Forward	GFP IP	minutes	minutes	minutes	minutes	APEX2	IP	IP	Forward	Reverse	Forward	Reverse	Identified:
PHF3	-0.74	-0.30	-0.43	-0.31	0.85	-0.19	-1.00	-0.91	-0.44	-0.74	-1.34	NaN	-1.68	0.86	0.69	10
NELFE	-0.98	-0.65	-0.50	-0.10	NaN	-0.50	-1.48	NaN	-1.88	-0.96	0.98	NaN	NaN	NaN	NaN	8
SNW1	-0.16	0.13	0.12	0.01	0.15	-0.13	-0.41	-0.55	-0.80	-1.13	NaN	-0.58	-0.61	0.19	-0.76	7
COIL	-0.63	0.33	0.09	-0.13	-0.46	-0.98	-0.38	-0.66	-0.73	-0.65	2.07	0.01	-0.24	-0.19	0.39	7
NELFB	-0.69	-0.56	-0.36	-0.01	-0.02	-0.63	-0.98	-1.01	-0.85	-0.06	3.67	NaN	NaN	NaN	NaN	7
S100A8	NaN	NaN	NaN	-1.49	-1.43	-3.72	NaN	NaN	NaN	-1.47	1.80	NaN	NaN	-1.48	0.31	6
RGPD3	-0.62	0.36	-0.09	0.19	-0.18	0.24	-0.05	-0.66	NaN	0.03	-4.03	-6.01	1.74	-4.51	0.41	6
FAM208B	-1.61	0.79	-0.51	-0.32	NaN	-1.68	-0.18	-0.12	0.30	-0.89	NaN	NaN	NaN	0.52	1.19	6
INTS1	-1.07	-0.24	-0.10	-0.37	NaN	-0.06	-0.87	-0.76	-0.52	-0.32	-0.93	NaN	-0.15	NaN	NaN	6
RM11	-0.57	0.33	-0.34	NaN	NaN	-0.59	0.11	-0.89	NaN	-1.35	NaN	NaN	0.77	NaN	NaN	6
NELFA	-0.76	-0.86	-0.29	-0.28	NaN	-0.19	-0.37	-1.39	-1.77	-0.92	0.02	NaN	NaN	NaN	NaN	6
PCIF1	-0.85	-0.79	-0.42	0.01	NaN	-0.57	-0.44	-1.17	-0.66	-0.29	NaN	NaN	NaN	NaN	NaN	6
IGKV2D-24	-6.33	6.20	4.04	-5.00	NaN	-8.45	-8.06	-7.77	-7.49	NaN	NaN	NaN	NaN	NaN	NaN	6
IGHG2;IGHG4	-5.12	5.15	5.40	-5.47	NaN	-5.87	-5.97	-2.33	-5.31	NaN	NaN	NaN	NaN	NaN	NaN	6
IGKV A18; IGKV2D-26; IGKV2D-29	-4.12	4.22	4.19	-4.10	NaN	-3.37	-2.97	-3.62	-3.65	NaN	NaN	NaN	NaN	NaN	NaN	6
SPIN2B; SPIN2A	-0.61	NaN	NaN	-0.64	NaN	-0.40	-0.57	-0.82	-1.03	NaN	NaN	NaN	NaN	NaN	NaN	6
POLR2E	-0.70	NaN	0.07	-0.24	0.19	-0.13	-0.55	-0.45	-0.46	-0.57	NaN	-0.06	-0.17	-0.08	-1.75	5
ZMYM4	-0.66	NaN	-0.04	-0.01	NaN	-0.01	NaN	0.25	0.01	-0.73	1.43	-0.68	NaN	-0.46	-1.68	5
IK	-0.28	0.38	0.07	0.15	0.52	0.09	-0.24	-0.10	-0.64	-0.73	3.31	-0.77	-0.59	-0.10	-0.68	5
RPRD2	-0.09	0.38	0.23	0.08	0.53	0.09	0.27	-0.29	-0.53	-0.58	0.07	-0.56	-0.55	-0.65	-0.62	5
DDX56	NaN	-0.52	NaN	NaN	NaN	-0.52	-0.73	-0.20	-0.61	0.04	0.18	-0.44	-0.45	-0.76	-0.17	5
BANF1	0.10	0.02	NaN	NaN	NaN	NaN	NaN	-0.12	-0.71	-0.53	-0.57	NaN	NaN	NaN	0.14	5
NOP10	NaN	NaN	0.37	-0.07	NaN	NaN	-0.36	-0.51	-1.00	0.03	-0.74	-0.96	-0.10	-1.36	0.46	5
NUP153	-0.21	0.35	0.20	-0.48	0.28	-1.03	NaN	-0.66	-0.20	-0.98	0.06	0.18	-0.11	-0.77	0.53	5
ICE1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-1.25	NaN	-0.82	-0.98	NaN	NaN	NaN	0.64	5
INTS6	-1.05	-0.26	-0.04	-0.44	NaN	-0.15	-0.98	-0.76	-0.55	-0.38	NaN	-0.24	-0.65	NaN	NaN	5

on the number of times they were identified in the different experiments. Threshold for significant Log_2 SILAC ratios was determined for each experiment separately. Complete table can be found in supplemental table 2.

Log ₂ Ratio H/L normalized																
Gene names	PFA XIP				DSP XIP				Native fractionation				# of times Identified:			
	Forward	Reverse	VCPI		Native GFP IP	15 minutes	XPA KO		APEX2 IP	BioID2 IP	GFP IP			Pol II IP Forward	Pol II IP Reverse	
			Forward	Reverse			75 minutes	75 minutes			Forward	Reverse				
CDKN2AIP	-0.58	NaN	-0.23	0.23	NaN	-0.86	NaN	NaN	-0.66	NaN	NaN	-0.52	NaN	NaN	5	
DIDO1	-0.46	0.23	0.17	-0.20	NaN	-0.07	-0.72	-0.45	-0.64	-1.43	NaN	NaN	NaN	NaN	5	
CTDP1	-0.68	NaN	0.12	-0.40	-0.20	-1.05	-0.86	-1.01	-0.48	0.22	NaN	NaN	NaN	NaN	5	
NELFCD	-0.77	-0.55	-0.65	0.05	NaN	-0.54	-1.25	-0.27	NaN	NaN	NaN	NaN	NaN	NaN	5	
IGKV4-1	-4.42	2.40	1.90	-1.88	NaN	-7.88	-5.49	-4.66	NaN	NaN	NaN	NaN	NaN	NaN	5	
EHMT1	NaN	0.34	0.09	-0.15	NaN	0.12	0.00	-0.51	-0.77	0.68	-0.16	-0.33	NaN	-2.44	4	
POLR2G	-0.48	-0.11	0.05	-0.09	0.59	-0.06	-0.66	-0.39	-0.80	NaN	-0.07	-0.19	-0.07	-1.68	4	
SNRPB2	-0.31	0.58	-0.12	0.28	-0.01	NaN	-0.62	0.00	-0.22	-0.33	-0.82	-0.36	NaN	-1.16	4	
POLR2D	-0.40	-0.19	0.04	0.00	0.85	-0.14	-0.60	-0.31	-0.42	NaN	-0.11	-0.11	-0.29	-1.11	4	
FLOT2	-0.51	NaN	-0.31	0.32	NaN	0.00	NaN	-1.13	0.33	NaN	NaN	NaN	NaN	-0.65	4	
CDYL	0.01	NaN	-0.31	-0.01	0.33	-0.57	-0.71	0.37	NaN	0.12	0.36	NaN	0.23	0.14	4	
RPAP2	-0.38	-0.49	-0.04	-0.51	0.30	NaN	-1.15	-0.76	-0.41	NaN	0.84	0.34	1.63	-0.58	4	
HNRNPR	-0.10	0.24	-0.05	0.13	-0.25	0.06	0.02	0.14	-0.19	-0.84	3.07	-0.44	-0.33	NaN	4	
SF3A1	-0.20	0.32	-0.16	0.08	-0.16	0.04	0.04	0.05	-0.26	-1.14	NaN	-0.45	-0.42	0.06	4	
EMG1	NaN	NaN	0.31	-0.04	NaN	-0.45	-0.16	-0.76	NaN	NaN	-1.03	-0.66	-0.78	-0.16	4	
POLR2L	-0.63	-0.18	0.10	-0.18	-0.03	-0.13	-0.59	-0.34	-0.57	-0.33	-3.42	-0.01	-0.07	-0.24	4	
TAF15	-0.24	0.08	0.06	0.00	NaN	0.04	-0.08	-0.32	-0.33	-0.55	NaN	-0.75	-1.90	-0.15	4	
MATR3	-0.17	0.43	0.02	0.08	0.13	-0.02	0.07	-0.42	-0.86	-0.86	3.50	-0.56	-0.40	-0.19	4	
RACGAP1	-0.21	0.05	-0.22	0.11	NaN	-0.57	0.36	-0.01	-0.20	-0.70	2.93	0.02	-0.44	-0.04	4	
BMS1	NaN	NaN	0.47	0.69	NaN	0.03	-0.29	-0.35	-0.73	-0.01	2.11	-0.44	-0.38	-0.99	4	
NOP14	-1.03	NaN	NaN	NaN	NaN	-0.12	-0.64	0.05	-0.36	0.01	4.42	-1.49	NaN	-1.11	4	
UTP14A	0.23	NaN	NaN	NaN	NaN	0.15	-2.33	-0.19	-1.08	-0.34	3.98	-0.27	-0.76	-0.91	4	
FN1	-0.28	0.16	0.07	-0.36	0.15	-0.31	-0.40	-0.49	-0.67	0.47	0.41	-2.04	-0.81	-0.69	4	
HEATR1	NaN	0.50	0.08	0.33	-0.14	0.07	-0.55	-0.14	-0.55	0.03	2.94	-0.86	-1.19	-0.93	4	
UTP20	-1.03	NaN	-0.65	0.97	NaN	0.14	-0.26	-0.12	-0.59	0.01	4.19	-1.04	NaN	-1.03	4	
TMEM201	-0.82	NaN	NaN	NaN	NaN	NaN	-0.47	-1.16	NaN	0.02	NaN	NaN	NaN	-1.05	4	
CMTR1	NaN	NaN	-0.09	-0.12	NaN	0.29	-0.27	-0.25	-0.10	-0.56	NaN	-0.90	-1.06	NaN	4	
DDX49	NaN	NaN	NaN	NaN	0.01	-0.61	NaN	-0.52	-0.72	-0.12	1.88	-0.45	-0.48	NaN	4	

Log ₂ Ratio H/L normalized																
Gene names	PFA XIP				Native GFP IP	DSP XIP				APEX2 IP	BioID2 IP	Native fractionation				# of times Identified:
	Forward	Reverse	VCPI Forward	VCPI Reverse		15 minutes	75 minutes	XPA KO 15 minutes	XPA KO 75 minutes			GFP IP Forward	GFP IP Reverse	Pol II IP Forward	Pol II IP Reverse	
CPSF3L	NaN	NaN	0.02	-0.33	0.07	-0.05	-0.89	-0.68	-0.60	0.09	1.20	NaN	-0.38	NaN	NaN	4
TRIM33	-0.27	0.19	0.15	0.02	-0.17	NaN	0.34	-0.04	0.07	-1.07	-1.51	NaN	-0.32	NaN	NaN	4
SF1	-0.29	0.14	0.20	0.05	0.13	0.06	-0.16	-0.21	-0.53	-1.02	-1.62	0.27	-0.29	NaN	NaN	4
CDK9	-0.15	NaN	-0.07	-0.37	0.63	0.26	0.05	NaN	-0.95	-0.53	3.13	0.02	-0.25	NaN	NaN	4
INTS3	-1.12	-0.17	0.04	-0.26	NaN	-0.04	-0.94	-0.80	-0.58	-0.01	2.88	-0.51	-0.18	NaN	NaN	4
HCFC1	-0.42	0.19	-0.10	0.12	0.01	-0.04	-0.25	-0.56	-0.05	-0.81	-0.48	0.51	-0.02	NaN	NaN	4
TSN	0.26	-0.76	-0.32	0.34	NaN	NaN	NaN	NaN	NaN	-0.53	NaN	0.21	0.13	NaN	NaN	4
TXNDC5	0.21	-0.35	2.89	-1.81	NaN	NaN	-0.75	NaN	0.76	-1.14	1.16	0.23	0.28	NaN	NaN	4
AHCY	0.17	-0.35	-0.03	-0.01	-0.30	-0.21	0.10	0.07	0.42	-0.86	-0.40	0.14	0.29	NaN	NaN	4
PHF8	NaN	-0.76	-0.52	0.05	NaN	-0.53	-0.83	-0.37	-0.64	-0.27	3.28	NaN	0.35	NaN	NaN	4
EIF1AD	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-5.11	-2.43	-0.12	0.47	NaN	NaN	4
HMGCS1	-0.09	0.20	0.76	-0.40	0.94	NaN	-0.38	NaN	NaN	-0.12	-0.62	-0.45	0.77	NaN	NaN	4
MAPK8	-0.58	-0.09	-0.24	-0.31	NaN	NaN	NaN	NaN	NaN	-0.59	2.39	NaN	2.10	NaN	NaN	4
INIP	NaN	NaN	NaN	NaN	NaN	NaN	-1.11	-0.66	-0.77	NaN	NaN	-0.62	NaN	NaN	NaN	4
DRD5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-0.87	-1.49	NaN	NaN	NaN	NaN	4
DPYSL3	0.40	-0.61	0.41	-0.01	0.06	NaN	-0.54	NaN	0.08	0.00	-1.13	NaN	NaN	NaN	NaN	4
SEC13	-0.10	0.00	-0.04	-0.50	0.42	-0.47	-0.03	0.31	0.28	NaN	-1.12	NaN	NaN	NaN	NaN	4
PP1P5K1	NaN	NaN	1.06	NaN	NaN	NaN	NaN	NaN	NaN	-0.63	-0.79	NaN	NaN	NaN	NaN	4
GATAD2B	-0.31	0.33	-0.03	-0.16	0.64	-0.13	-0.29	0.22	0.01	-0.54	-0.78	NaN	NaN	NaN	NaN	4
INTS2	-0.53	-0.06	-0.05	-0.33	NaN	-0.02	-0.93	-0.97	-0.39	-0.48	-0.31	NaN	NaN	NaN	NaN	4
RBM27	-0.15	NaN	0.27	0.16	-2.52	-0.10	-0.16	-0.12	-0.85	-1.26	-0.14	NaN	NaN	NaN	NaN	4
INTS12	-1.28	-0.22	0.00	-0.41	NaN	-0.07	-1.11	-1.44	-0.58	NaN	0.15	NaN	NaN	NaN	NaN	4
ELL	-0.53	NaN	0.32	-0.76	NaN	0.74	NaN	NaN	-0.84	0.79	0.79	NaN	NaN	NaN	NaN	4
KDM3B	-0.49	NaN	0.07	-0.31	-0.27	-0.74	-0.26	-0.11	-0.48	-0.98	3.22	NaN	NaN	NaN	NaN	4
ZCCHC8	-0.67	0.07	-0.25	-0.32	NaN	NaN	NaN	NaN	-0.60	-0.98	NaN	NaN	NaN	NaN	NaN	4
DCTN4	0.02	-0.12	0.25	-0.26	0.15	NaN	-0.70	NaN	-3.19	-0.83	NaN	NaN	NaN	NaN	NaN	4
CENPF	2.37	NaN	-0.33	0.16	NaN	1.64	NaN	-0.87	0.05	-0.65	NaN	NaN	NaN	NaN	NaN	4
GTF2B	-0.82	-0.17	0.18	-0.14	NaN	-0.05	-0.80	-0.15	-0.43	-0.64	NaN	NaN	NaN	NaN	NaN	4

Log ₂ Ratio H/L normalized															
Gene names	PFA XIP				DSP XIP				Native fractionation						
			VCPI		Native		15 minutes		75 minutes		XPA KO 15 minutes		XPA KO 75 minutes		BioID2 IP
	Forward	Reverse	Forward	Reverse	GFP IP		minutes	minutes	minutes	minutes	minutes	minutes	minutes	minutes	
CSTF3	-0.02	-4.75	-0.43	0.12	NaN		-0.08	-0.03	-0.16	-0.09	-0.50	NaN	NaN	NaN	NaN
ZC3H4	-0.69	0.14	-0.14	-0.33	-0.35		0.21	-0.49	-0.30	-0.78	-0.50	NaN	NaN	NaN	NaN
INTS5	-1.10	-0.30	0.21	-0.33	NaN		-0.07	-1.05	-0.97	-0.60	NaN	NaN	NaN	NaN	NaN
# of times identified:															
															4
															4
															4

Thus far, no clear role for RPB5 and RPB10 has been shown in response to UV-induced transcription stress, however, RPB4 and RPB7 have been shown to be involved in the stress response in yeast. RPB4 and RPB7 form a heterodimer that is considered a general transcription factor since the dimer restores transcription initiation^{91,92}. In yeast, RPB4/7 regulates TC-NER by inhibiting the RPB9-dependent TC-NER pathway and stimulating the Rad26-dependent pathway, likely by inducing a conformational change upon binding of the heterodimer⁹³. Moreover, the heterodimer forms a stalk module that extrudes from the Pol II complex and thereby likely serves as a binding platform for CTD-interacting proteins, such as phosphatases⁹². In line with this, RPB4 has been shown to be important in recruiting factors to the CTD of Pol II⁹⁴. CTD¹, also known as FCP1, was also found to be less associated with Pol II after UV and is an example of a phosphatase that is likely recruited by RPB4 to facilitate binding to the CTD^{94,95}. Here, it most likely removes the Ser2 phosphorylation, thereby regulating recycling of Pol II for another round of transcription^{95,96}. Together, this suggests a role for the RPB4/7 heterodimer in regulating transcription levels and directing TC-NER in response to UV-induced transcription stress. However, a role for RPB5 and RPB10 in the DNA damage-induce transcription stress response has yet to be determined.

Among the top hits that were released from Pol II after UV were the different components of the negative elongation factor (NELF) (Fig. 8B, Table 2). NELF is known to inhibit transcription by trapping Pol II on the promoter-proximal pause site with the help of DRB-sensitivity-inducing factor (DSIF)¹. For Pol II release into the gene body, positive transcription elongation factor b complex (P-TEFb) is recruited and phosphorylates DSIF, NELF and the CTD of Pol II. Upon phosphorylation, DSIF turns into a positive elongation factor and phosphorylation of NELF induces its release from the paused Pol II complex, resulting in the release of paused Pol II into productive elongation^{1,97}. Release of NELF from Pol II as identified (Fig. 8B, Table 2) thus suggests pause-release and stimulation of transcription elongation. This is in line with the previously observed pause release upon UV damage, caused by p38-MK2-dependent phosphorylation of the NELF complex or increased levels of active p-TEFb²¹⁻²³. However, loss of NELF could also indicate that the promoter-paused Pol II fraction is lost, for example by degradation, which should not necessarily result in more elongating Pol II.

Furthermore, we identified multiple components of the 14-subunit integrator complex (INT) (Fig. 8B, Table 2). INT has been shown to be involved in transcription pause release where it binds the CTD of Pol II and recruits P-TEFb⁹⁸. As described above, this leads to release of NELF and release of Pol II into productive elongation. It is interesting to note that during this process, INT remains bound to the elongating complex. Since we identified a decreased interaction of INT to Pol II, this suggested another role than inducing pause release.

Moreover, we identified several other CTD-binding proteins involved in regulating transcription and phosphorylation of the CTD of Pol II such as PCIF1 and PHF3. PCIF1 has been shown to bind to the CTD and regulate the phosphorylation status, thereby inhibiting transcription activation⁹⁹. PHF3 was identified to release from Pol II in response to UV (Table 2)⁷⁸. Previously, the yeast homolog of PHF3, Bye1, was shown to bind Pol II

at the same region as backtracking factor TFIIS¹⁰⁰ and was shown to compete with TFIIS in binding to Pol II¹⁰¹. Possibly, PHF3 is released from Pol II to facilitate TFIIS binding and thereby, mRNA cleavage to realign the mRNA and enable transcription restart¹⁰².

We have successfully set up different extraction and isolation methods to identify Pol II-interacting proteins and to identify UV-induced interactors. We were able to identify known UV-specific interactors, thereby validating our isolation methods. Although the independent experiments did not identify strong new UV-interactors, combining the different approaches revealed many new interactors. Most of these were involved in processes that regulate transcription by modifying Pol II activity, affecting global transcription, regulating backtracking, Pol II degradation, and mRNA levels. We identified several known UV-specific interactors as well as new, interesting factors that need further investigation to unravel their exact function in the UV-induced transcription stress response. In addition, it might be interesting to assess Pol II-interacting proteins at different time points after UV irradiation. Most of our experiments were performed 1 hour after UV irradiation and assessing different time points can provide insight in the timing of the repair reaction but also in the mechanism of the transcription response. Furthermore, it would be interesting to assess later time points after UV, after efficient repair, to identify factors that are involved in transcription restart. Overall, there are many processes involved in regulating Pol II in response to UV to safeguard efficient repair, induce the transcription stress response, protect mRNA and transcription fidelity, and finally restart transcription. The identification of Pol II-interacting proteins will fuel follow-up studies to better understand the mechanisms of the DNA damage-induced transcription stress response.

Materials and methods

Cell culture

MRC-5 SV40 immortalized human lung fibroblast Wt cells, MRC-5 GFP-RPB1 KI cells⁴⁷, MRC-5 XPA KO GFP-RPB1 KI cells and HCT116 colorectal cancer cells were cultured in a 1:1 mixture of DMEM (Gibco) and Ham's F10 (Invitrogen) supplemented with 10% fetal calf serum (FCS, Biowest) and 1% penicillin-streptomycin in a humidified incubator at 37°C and 5% CO₂. For stable isotope labeling of amino acids in culture (SILAC), cells were grown for two weeks (>10 cell doublings) in arginine/lysine-free SILAC DMEM (ThermoFisher) supplemented with 15% dialyzed FCS (Gibco), 1% penicillin-streptomycin, 200 µg/ml proline (Sigma), and either 73 µg/mL light [¹²C₆]-lysine and 42 µg/mL [¹²C₆, ¹⁴N₄]-arginine (Sigma) or heavy [¹³C₆]-lysine and [¹³C₆, ¹⁵N₄]-arginine (Cambridge Isotope Laboratories). For proximity labeling, cells were cultured in biotin-depleted media or in media with dialyzed FCS. Biotin-depleted media was generated by incubating 45 ml media with 125 µl DynabeadsTM MyOneTM Streptavidin (ThermoFisher) at 4°C for three days.

MRC-5 Flag-APEX2-RPB1 KI and Flag-BioID2 KI cells were generated as described previously for generating Flag-GFP-RPB1 KI cells⁴⁷. Briefly, cells were transiently transfected with an sgRNA-containing pLentiCRISPR.v2 plasmid (GCCTGCCTCCGCCATGCACG) targeting the start codon of RPB1 and co-transfecting a homology-directed repair template containing

Flag, APEX2 or BioID2, a Sugaya linker and homology arms for RPB1 (199 bp for APEX2 and 224 bp for BioID2). One day after transfection, cells were selected for successful transfection by culturing them in presence of 1 µg/ml of Puromycin (Invivogen) for three days. Surviving cells were seeded in a low density to allow expansion. Genomic DNA was isolated from single cell clones using a PureLink™ Genomic DNA Mini Kit (Life Technologies) according to manufacturer's protocol. Genotyping PCR was performed with My Tag DNA polymerase (Bioline) according to manufacturer's protocol using the following primers: GACGAGGTCTGAGCACCTAG and GACTCCCTAGGATTCGTCGG. Homozygous KI clones were selected and expanded for further analysis.

For UV-C irradiation, cells were washed with PBS, and placed under a 254 nm germicidal UV-C lamp (Philips). Duration of irradiation was controlled with an air-pressured shutter connected to a timer and cells were irradiated with doses as indicated. Cells were pre-treated 1 hour before UV irradiation with VCP inhibitor (Seleck Chemicals, 5 µM) and maintained in presence of VCP where indicated. For proximity labeling using APEX2, Biotin-phenol (Iris Biotech) was added to culture medium in a final concentration of 500 µM for 30 minutes unless indicated differently and H₂O₂ (Sigma) was added for 1 minute to a final concentration of 0.1 mM unless indicated otherwise. For proximity labeling using BioID2, Biotin (Sigma) was added to culture media to a final concentration of 1 µM for 4 hours unless indicated differently.

Cell lysis and immunoblot

Cells were directly lysed in SDS-PAGE loading buffer (0.125M Tris pH 6.8, 2% sodium dodecyl sulfate (SDS, Sigma), 0.005% bromophenol blue, 21% glycerol, 4% β-mercaptoethanol) and incubated 5 minutes at 95°C. Cell lysates were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad). Proteins were transferred onto PVDF membranes (0.45µm, Merck Millipore) at 4°C, either 1.5h at 90V with 1x transfer buffer (25mM TRIS, 190mM Glycine, 10% methanol) or overnight at 25V in 2x transfer buffer (50mM TRIS, 380mM Glycine). Membranes were blocked with 5% BSA (Sigma) in PBS-tween (0.05%) and probed with primary antibodies (Table 3, see below). Subsequently, membranes were extensively washed with PBS-tween and incubated with secondary antibodies coupled to IRDyes (LI-COR, table 4, see below) to visualize proteins using an Odyssey CLx infrared scanner (LI-COR).

Native immunoprecipitation (IP)

MRC-5 Flag-GFP-RPB1 KI cells, or Wt cells where indicated, were directly harvested, mock-treated or irradiated with 8 J/m² prior to harvest. One confluent 145 cm² dish per condition was used for IP and three confluent 145 cm² dishes per condition for mass spectrometry.

Cells were collected by trypsinization and pelleted in cold PBS using centrifugation at 1000 rpm for 5 minutes at 4°C. After one wash with cold PBS, cell pellets were lysed for 20 minutes at 4°C using HEPES buffer (30 mM HEPES pH 7.6, 1 mM MgCl₂, 130 mM NaCl, 0.5% Triton X-100 (Sigma), 0.5 mM DTT, 1x cComplete™ EDTA-free Protease Inhibitor Cocktail (Roche), 1:100 phosphatase inhibitor cocktail 2 and 3 (Sigma)) or RIPA buffer (10 mM Tris pH 7.5, 150 mM NaCl, 1 mM MgCl₂, 1% Triton X-100, 0.1% sodium deoxycholate,

0.1% SDS, 1x cComplete™ EDTA-free Protease Inhibitor Cocktail, 1:100 phosphatase inhibitor cocktail 2 and 3) containing 50 μ M MG132. Alternatively, cells were directly lysed in the plate on ice. Therefore, plates were washed once with cold PBS and lysed for 10 minutes using HEPES buffer containing 50 μ M MG132 before scraping. Lysed DNA was sheared using sonication at different amplitudes and duration as indicated using cycles of 15 seconds on/15 seconds off or chromatin was digested for 1 hour at 4°C using benzonase (Novagen), Micrococcal Nuclease (MNase, Sigma), or DNase I (New England Biolabs, activated with CaCl_2) as indicated. MNase activity was stopped using 5 mM EGTA and, if indicated, DNase I activity was stopped with 5 mM EDTA. When both shearing and digestion was used, samples were first sonicated and then digested for 1 hour. For high salt concentrations, the complete extraction protocol was performed in HEPES buffer with indicated concentrations of NaCl in mM. After digestion, the lysates were centrifuged twice at maximum speed for 15 minutes at 4°C. A sample from the pellet (P) and supernatant (SN) was taken to analyze if the proteins were extracted from the chromatin and to determine the DNA fragment size. For protein analysis, SDS-PAGE loading buffer was directly added and samples were incubated for 5 minutes at 95°C before immunoblot. For DNA fragment analysis, the pellet and supernatant samples were treated with RNase (0.2 μ g/ μ l) for 30 minutes at 37°C and Proteinase K (1 μ g/ μ l) for 2-3 hours at 55°C. Loading dye (1x orange G) was added to the remaining DNA and loaded on 1.5% agarose gel in 1x TBE buffer with Ethidium bromide to visualize DNA.

For immunoprecipitation of Flag-GFP-RPB1, supernatant of the centrifuged lysates was incubated overnight with rotation at 4°C with the following beads or antibodies (2.5 μ g): GFP-Trap Agarose beads (Chromotek, 20 μ l slurry per sample), Flag M2 agarose beads (Sigma, 20 μ l slurry per sample), GFP antibody (Abcam, ab290), GFP antibody (Roche) or RPB1 antibody (Santa Cruz, N-20, sc899). Next day, antibody-bound complexes were incubated with Protein G plus Agarose beads (ThermoFisher) for 1.5 hours with rotation at 4°C. Finally, non-specific interactors were removed by washing five times. For mass spectrometry, beads of light and heavy labeled cells were combined during the last wash. Proteins were eluted by incubating 5 minutes at 95°C in SDS-PAGE loading buffer. Samples were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad) and processed for immunoblotting or fixed and stained for analyzing protein content or mass spectrometry using Imperial protein stain (Pierce) according to manufacturer's protocol. For On-bead digestion, beads were directly analyzed for mass spectrometry.

Cross-linked immunoprecipitation (XIP)

MRC-5 Flag-GFP-RPB1 KI cells were directly harvested, mock-treated or irradiated with 16 J/m² UV-C one hour prior to cell harvest. Cell pellets were prepared from 3 confluent 145 cm² dishes per condition for IP or 8 confluent 145 cm² dishes per condition for mass spectrometry.

Crosslinked IP was performed as described previously¹⁰³ with modifications as indicated. Cells were crosslinked with 1% paraformaldehyde (PFA) in serum-free DMEM for 7 minutes at room temperature (RT), or as indicated during optimization, with constant shaking before quenching the reaction for 5 minutes with glycine (final concentration of 0.125 M) at RT. Cells were collected by scraping in 10% glycerol in PBS with 1 mM PMSF

and pelleted at maximum speed for 15 minutes at 4°C. Consequently, chromatin was purified by washing the pellets for 30 minutes with rotation in buffer 1 (50 mM HEPES, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100, 0.5% NP-40, 10% glycerol) and pelleting the cells for 10 minutes at 1300 rpm. For mass spectrometry analysis, light and heavy labeled samples were pooled. Pellets were washed twice with buffer 2 (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) and finally the chromatin was pelleted. Chromatin was stored at -80°C and upon thawing, sonicated in RIPA buffer (10 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% SDS, 1% Sodium Deoxycholate and 0.5 mM EGTA) using the Bioruptor Sonicator (Diagenode) with cycles of 15 seconds on/15 seconds off using the highest amplitude for 7 minutes or as indicated. Extracted chromatin was collected by spinning 15 minutes at maximum speed and supernatant was used for immunoprecipitation. A sample from the pellet (P) and supernatant (SN) fractions was taken to analyze if the proteins were extracted from the chromatin and to determine the DNA fragment size. For protein analysis, SDS-PAGE loading buffer was directly added and samples were incubated for 5 minutes at 95°C before immunoblot. For DNA fragment analysis, the pellet and supernatant samples were reverse crosslinked overnight at 65°C in presence of 0.8 M NaCl. Next, samples were treated with RNase (0.2 µg/µl) for 30 minutes at 37°C and Proteinase K (1 µg/µl) for 2-3 hours at 55°C. Loading dye (1x orange G) was added to the remaining DNA and loaded on 1.5% agarose gel in 1x TBE buffer with Ethidium bromide to visualize DNA. All steps were performed at 4°C unless indicated differently and to all buffers, 1 mM PMSF, 0.5 mM Na₂VO₄, 5 mM NaF, 5 mM NaPPi, 10 mM β-glycerol and cOmplete™ EDTA-free Protease Inhibitor Cocktail was added.

For IP, lysates were pre-cleared for 30 minutes with Protein G agarose beads (Pierce) at 4°C. IP was performed by incubating 4 hours with rotation at 4°C with 30 µl slurry Flag M2 agarose beads, 30 µl slurry GFP-Trap Agarose beads (Chromotek), 30 µg GFP antibody (Abcam, ab290) or 30 µl slurry Pierce prot G agarose beads. After 4 hours, 30 µl slurry Pierce prot G agarose was added to the GFP antibody-bound complexes and incubated for another hour. Finally, non-specific interactors were removed by washing five times with RIPA buffer and proteins were eluted and crosslinking was reversed by incubating 30 minutes, or 10 minutes where indicated, at 95°C in SDS-PAGE loading buffer. Samples were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad) and processed for immunoblotting or fixed and stained to determine protein content or for mass spectrometry using Imperial protein stain (Pierce) according to manufacturer's protocol.

For DSP XIP, the same protocol was followed with minor alterations. Cells were crosslinked in 1 mM dithiobis(succinimidyl propionate) (DSP, Pierce) in PBS for 30 minutes and quenched by adding Tris pH 7.5 to a final concentration of 25 mM for 10 minutes. During optimization, 2 mM of DSP for 15 minutes was used and sonication time was reduced to 3.5 minutes. Beads were incubated for 5 minutes at 95°C to elute and reverse cross-link immunocomplexes.

Proximity labeling using APEX2

MRC-5 Flag-APEX2-RPB1 KI cells were directly harvested, mock-treated or irradiated with UV-C prior to harvest (30 minutes after 40 J/m² or 20 minutes after 20 J/m²). One confluent 145 cm² dish was used per condition for IP or four confluent 145 cm² dishes were used for mass spectrometry. Cells were cultured with dialyzed FCS or in biotin-depleted medium if indicated. Biotinylation was initiated as described above and reaction was quenched by washing three times with quenching solution (5 mM Trolox (Santa Cruz), 10 mM Sodium Ascorbate and 10 mM Sodium Azide (both sigma) in Hank's Balanced Salt Solution (HBSS, Thermofisher)). Cells were harvested by scraping in 5 ml PBS per plate containing 0.5x cOmplete™ EDTA-free Protease Inhibitor Cocktail. Cells were centrifuged at 2000 rpm for 15 minutes and pellets were snap frozen and stored at -80°C until IP.

For immunoprecipitation, thawed pellets were resuspended in 8 M Urea buffer pH 8 (300 mM NaCl, 50 mM Na₂HPO₄, 0.5% NP-40 (all Sigma), and 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail in 8 M Urea). DNA was sheared using a Bioruptor Sonicator with 3 cycles of 15 seconds on/15 seconds off on high amplitude and samples were centrifuged at 13,000 rpm for 15 minutes at 4°C. For mass spectrometry, supernatant of light- and heavy-labeled samples were pooled. Concentration of supernatant was diluted to 2 M urea by adding urea-free buffer (300 mM NaCl, 50 mM Na₂HPO₄, 0.5% NP-40, and 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail in dH₂O). For protein analysis, SDS-PAGE loading buffer was added and samples were incubated at 95°C for 5 minutes. To continue with the IP, lysates were pre-cleared by incubating for 1 hour with Prot G beads (Thermofisher) with rotation at room temperature (RT). Protein complexes were immunoprecipitated by incubating 1 hour at RT with Dynabeads™ MyOne™ Streptavidin (Thermofisher). Non-specific interactors were removed by washing 5 times with 2 M Urea buffer (300 mM NaCl, 50 mM Na₂HPO₄, 0.5% NP-40 (all Sigma), and 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail 2 M Urea). Immunoprecipitated complexes were eluted by incubating in SDS-PAGE loading buffer for 5 minutes at 95°C. Samples were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad) and processed for immunoblotting or fixed and stained to determine protein content or for mass spectrometry using Imperial protein stain (Pierce) according to manufacturer's protocol. Where indicated, buffers were replaced by RIPA buffer.

Proximity labeling using BioID2

MRC-5 Flag-BioID2-RPB1 KI cells were directly harvested, mock-treated or irradiated with UV-C prior to harvest (4 hours after 20 J/m², or time as indicated). One confluent 145 cm² dish was used per condition for IP or four confluent 145 cm² dishes were used for mass spectrometry. Biotinylation was initiated as described above immediately after UV irradiation or mock-treatment. For mass spectrometry, cells were cultured in media with dialyzed serum for two weeks and with additional biotin-depletion for the last 4 days.

Cells were washed twice in ice cold PBS and harvested by scraping in 5 ml PBS per plate containing 0.5x cOmplete™ EDTA-free Protease Inhibitor Cocktail. Cells were centrifuged at 2000 rpm for 15 minutes and pellets were resuspended in 8 M Urea buffer pH 8. For mass spectrometry, light- and heavy-labeled samples were pooled. DNA

was sheared with a Bioruptor Sonicator using 3 cycles of 15 seconds on/15 seconds off on high amplitude and samples were centrifuged at 13,000 rpm for 15 minutes at 4°C. Concentration of supernatant was diluted to 2 M urea by adding urea-free buffer. For protein analysis, lysates were filtered or SDS-PAGE loading buffer was directly added and samples were incubated at 95°C for 5 minutes.

To continue with IP, lysates were filtered using 10K or 3K Amicon® Ultra-15 Centrifugal Filter Units (Merck Millipore) to remove free biotin by centrifuging at 5,000 rpm for 10 minutes at 4°C. Filters were washed twice by adding 2 M Urea buffer and centrifuging at 5,000 rpm for 10 minutes at 4°C. Next, protein complexes were released with 2 M Urea buffer by spinning at 4,000 g for 10 minutes at 4°C and pre-cleared by incubating 1 hour with Prot G beads (ThermoFisher) with rotation at 4°C. Protein complexes were immunoprecipitated by incubating overnight at 4°C with Dynabeads™ MyOne™ Streptavidin (ThermoFisher). Non-specific interactors were removed by washing 5 times with 2 M Urea buffer and once with PBS. Immunoprecipitated complexes were eluted by incubating in SDS-PAGE loading buffer for 4 minutes at 95°C. Samples were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad) and processed for immunoblotting or fixed and stained to determine protein content or for mass spectrometry using Imperial protein stain (Pierce) according to manufacturer's protocol.

Native fractionation IP

Cells were mock-treated or irradiated with 16 J/m² UV-C 1 hour prior to harvest. Cell pellets were prepared from three confluent 145 cm² dishes per condition for IP or eight confluent 145 cm² dishes per condition for mass spectrometry. Cells were collected by trypsinization and pelleted in cold PBS using centrifugation for 5 minutes at 1500 rpm. After one wash with cold PBS, cell pellets were stored at -80°C until immunoprecipitation.

For IP, pellets were thawed on ice and lysed for 20 minutes at 4°C in HEPES buffer containing 30 mM HEPES pH 7.6, 1 mM MgCl₂, 150 mM NaCl, 0.5% NP-40, and 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche). Chromatin was pelleted by spinning 5 minutes at 10,000 g at 4°C and subsequently incubated for 1 hour at 4°C in HEPES buffer containing 500 units of benzonase (Millipore) to digest the chromatin. 2 µg Pol II antibody (P-Ser2-RPB1, ab5095, abcam) or IgG (sc2027, Santacruz) was directly added where indicated. After 1 hour, the NaCl concentration was increased to 300 mM to inactivate benzonase and incubation was continued for another 30 minutes. The undigested fraction was pelleted at 13,200 rpm for 10 minutes at 4°C and the soluble fraction was immunoprecipitated for 90 minutes with rotation at 4°C using 25 µL slurry salmon sperm protein A agarose beads (Millipore) or GFP-Trap Agarose beads (Chromotek). Unbound proteins were removed by washing the beads 5 times in wash buffer (30 mM HEPES pH 7.6, 150 mM NaCl, 1mM EDTA, 0.5% NP-40, and 0.2x cOmplete™ EDTA-free Protease Inhibitor Cocktail). For mass spectrometry, light- and heavy-labeled samples were pooled during last wash step. Bound proteins were eluted in SDS-PAGE loading buffer and separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad). Samples were processed for immunoblotting or fixed and stained for mass spectrometry using Imperial protein stain (Pierce) according to manufacturer's protocol.

Mass spectrometry

For gel-separation digestion, SDS-PAGE gel lanes were cut into slices and subjected to in-gel reduction with dithiothreitol (Sigma, D8255), alkylation with iodoacetamide (Sigma, I6125) and digestion with trypsin (sequencing grade; Promega) as previously described³⁶. For on-bead digestion, beads were directly digested with trypsin. Nanoflow liquid chromatography tandem mass spectrometry (nLC-MS/MS) was performed on an EASY-nLC 1200 coupled to a Lumos Tribid Orbitrap mass spectrometer (ThermoFisher Scientific) operating in positive mode. Peptide mixtures were trapped on a 2 cm x 100 μ m Pepmap C18 column (Thermo Fisher 164564) and then separated on an in-house packed 50 cm x 75 μ m capillary column with 1.9 μ m Reprosil-Pur C18 beads (Dr. Maisch) at a flowrate of 250 nL/min, using a linear gradient of 0–32% acetonitrile (in 0.1% formic acid) during 90 min. The eluate was directly sprayed into the electrospray ionization (ESI) source of the mass spectrometer. Spectra were acquired in continuum mode, fragmentation of the peptides was performed in data-dependent mode by HCD.

Mass spectrometry data were analyzed using MaxQuant software (version 1.6.3.3). The false discovery rate (FDR) of both PSM and protein was set to 0.01 and the minimum ratio count was set to 1. The Andromeda search engine was used to search the MS/MS spectra against the UniProt database (taxonomy: Homo sapiens, release June 2017), concatenated with the reversed versions of all sequences. A maximum of two missed cleavages was allowed. In case the identified peptides of two proteins were the same or the identified peptides of one protein included all peptides of another protein, these proteins were combined by MaxQuant and reported as one protein group. Before further analysis, known contaminants and reverse hits were removed. For each experiment, a threshold was determined for identified interactions. Gene ontology (GO) term enrichment analysis and protein-protein network analysis were performed using the STRING website¹⁰⁴. Genes with a(n) (average) SILAC ratio above the threshold were analyzed and the top 5 biological processes affected were identified or the interaction network was visualized. For the latter, only proteins that were part of an interaction network were displayed and interactions based on experimental data (pink) or curated databases (blue) were shown. Minimum required interaction score was set to high confidence (0.700) and medium confidence (0.400) for combined analysis of increased interactors.

Immunofluorescence

Cells were grown on 24-mm glass coverslips and treated to biotinylate proteins as described. For APEX2-KI cells, cells were fixed for 5 minutes in cold methanol (Sigma) at -20°C and incubated with 3% BSA (Sigma) for 1 hour at 4°C to prevent non-specific staining. For BioID2-KI cells, cells were incubated for 15 minutes with 0.4% Triton X-100 (Sigma) and fixed with 0.5% SDS in PBS for 5 minutes before blocking non-specific staining using PBS+ (0.15% BSA and 0.15% glycine in PBS) for 1 hour at 4°C. For both cell lines, P-Ser5-RPB1 was stained by incubating 2 hours at RT with rat anti-P-Ser5-RPB1 antibody (table 3, see below) in PBS+. Next, cells were washed extensively with 0.1% Triton and once with PBS+ before incubating 1 hour at RT with anti-streptavidin Alexa Fluor 488-conjugated antibody and goat anti-rat Alexa Fluor 594-conjugated antibody (table 4, see below). After extensive washes with 0.1% Triton and once with PBS+,

coverslips were mounted using Vectashield with DAPI (Vector Laboratories). Images were acquired with a Zeiss LSM700 Axio Imager Z2 upright microscope equipped with a 63x Plan-apochromat 1.4 NA oil immersion lens (Carl Zeiss Micro Imaging Inc.).

Table 3. Primary antibodies.

Antibody	Host	Source	Dilutions	
			WB	IF
CSB/ERCC6	G	Santa Cruz, sc10459	1/500	N.A.
CSB/ERCC6	Rb	Antibodies-online, ABIN2855858	1/1000	N.A.
CTD-RPB1 (Pol II)	Ms	Abcam, ab817	1/1000	N.A.
GFP	Ms	Roche, 14314500	1/1000	N.A.
GFP	Rb	Abcam, Ab290	1/1000	N.A.
NTD-RPB1 (Pol II)	Rb	Cell signalling, D8L4Y	1/1000	N.A.
p62/GTF2H1	Ms	Sigma Aldrich, WH0002965M1	1/1000	N.A.
P-Ser2-RPB1	R	Chromotek, 3E10	1/1000	N.A.
P-Ser2-RPB1	Rb	Abcam, ab5095	1/1000	N.A.
P-Ser5-RPB1	R	Chromotek, 3E8	1/1000	1/500
RPB1	Rb	Santa Cruz, sc899	1/500	N.A.
RPB9	Rb	Abcam, ab192407	1/500	N.A.
SPT4/SUPT4H1	Rb	Cell signalling, D3P2W	1/1000	N.A.
SPT5/SUPT5H	Rb	Bethyl, A300-869A	1/500	N.A.
Streptavidin		Li-Cor, 926-32230	1/1000	N.A.
Tubulin	Ms	Sigma Aldrich, B512	1/5000	N.A.
XPC	Rb	Homemade, fraction V	1/1000	N.A.
XPD	Ms	Abcam, ab54676	1/1000	N.A.

Table 4. Secondary antibodies.

Antibody	Host	Source	Dilutions	
			WB	IRDye
Rabbit	goat	Sigma, sab4600215	1/10000	770
Rabbit	goat	Sigma, sab4600200	1/10000	680
Mouse	goat	Sigma, sab4600199	1/10000	680
Mouse	goat	Sigma, sab4600214	1/10000	770
Goat	donkey	Sigma, sab4600375	1/10000	770
Rat	goat	Sigma, sab4600479	1/10000	770
Antibody	Host	Source	IF	AlexaFluor
Rat	Goat	Invitrogen, A-11007	1/5000	594
Streptavidin		Thermofisher, s-32354	1/5000	488

Supplemental table 1. Complete table containing the individual SILAC-based mass spectrometry experiments. <https://1drv.ms/x/s!AvEm1Jm7uzqFnnBfgbP5cbpgQKMD?e=SPVabI>

Supplemental table 2. Complete table containing the combined SILAC-based mass spectrometry approach. <https://1drv.ms/x/s!AvEm1Jm7uzqFnnTCe6BsRK3NlxRr?e=nX0WQP>

References

1. Jonkers, I. & Lis, J.T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-77 (2015).
2. Shandilya, J. & Roberts, S.G.E. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1819**, 391-400 (2012).
3. Wang, Y. Bulky DNA lesions induced by reactive oxygen species. *Chem Res Toxicol* **21**, 276-81 (2008).
4. Wang, W., Xu, J., Chong, J. & Wang, D. Structural basis of DNA lesion recognition for eukaryotic transcription-coupled nucleotide excision repair. *DNA repair* **71**, 43-55 (2018).
5. Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. *Science* **315**, 859-62 (2007).
6. Tornaletti, S., Reines, D. & Hanawalt, P.C. Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. *J Biol Chem* **274**, 24124-30 (1999).
7. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-74 (2001).
8. Doetsch, P.W. Translesion synthesis by RNA polymerases: occurrence and biological implications for transcriptional mutagenesis. *Mutat Res* **510**, 131-40 (2002).
9. Marietta, C. & Brooks, P.J. Transcriptional bypass of bulky DNA lesions causes new mutant RNA transcripts in human cells. *EMBO Rep* **8**, 388-93 (2007).
10. Lans, H., Hoeijmakers, J.H.J., Vermeulen, W. & Marteijn, J.A. The DNA damage response to transcription stress. *Nat Rev Mol Cell Biol* **20**, 766-784 (2019).
11. Gómez-González, B. & Aguilera, A. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes & development* **33**, 1008-1026 (2019).
12. López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194-1217 (2013).
13. Tresini, M. *et al.* The core spliceosome as target and effector of non-canonical ATM signalling. *Nature* **523**, 53-8 (2015).
14. Crossley, M.P., Bocek, M. & Cimprich, K.A. R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell* **73**, 398-411 (2019).
15. Gaillard, H. & Aguilera, A. Transcription as a Threat to Genome Integrity. *Annu Rev Biochem* **85**, 291-317 (2016).
16. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell* **46**, 115-24 (2012).
17. Geijer, M.E. & Marteijn, J.A. What happens at the lesion does not stay at the lesion: Transcription-coupled nucleotide excision repair and the effects of DNA damage on transcription in cis and trans. *DNA Repair (Amst)* (2018).
18. Vichi, P. *et al.* Cisplatin- and UV-damaged DNA lure the basal transcription factor TFIID/TBP. *EMBO J* **16**, 7444-56 (1997).
19. Rockx, D.A. *et al.* UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II. *Proc Natl Acad Sci U S A* **97**, 10503-8 (2000).
20. 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* **10**, e1004483 (2014).
21. Lavigne, M.D., Konstantopoulos, D., Ntakou-Zamplara, K.Z., Liakos, A. & Foustier, M. Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate. *Nat Commun* **8**, 2076 (2017).

22. Bugai, A. *et al.* P-TEFb Activation by RBM7 Shapes a Pro-survival Transcriptional Response to Genotoxic Stress. *Molecular cell* **74**, 254-267.e10 (2019).
23. Borisova, M.E. *et al.* p38-MK2 signaling axis regulates RNA metabolism after UV-light-induced DNA damage. *Nature communications* **9**, 1017-1017 (2018).
24. Muñoz, M.J. *et al.* DNA Damage Regulates Alternative Splicing through Inhibition of RNA Polymerase II Elongation. *Cell* **137**, 708-720 (2009).
25. Williamson, L. *et al.* UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. *Cell* **168**, 843-855 e13 (2017).
26. Tresini, M., Marteijn, J.A. & Vermeulen, W. Bidirectional coupling of splicing and ATM signaling in response to transcription-blocking DNA damage. *RNA Biol* **13**, 272-8 (2016).
27. Mellon, I., Spivak, G. & Hanawalt, P.C. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**, 241-9 (1987).
28. Laugel, V. Cockayne syndrome: the expanding clinical and mutational spectrum. *Mech Ageing Dev* **134**, 161-70 (2013).
29. van den Boom, V. *et al.* DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J Cell Biol* **166**, 27-36 (2004).
30. Xu, J. *et al.* Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653-657 (2017).
31. Kamiuchi, S. *et al.* Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proc Natl Acad Sci U S A* **99**, 201-6 (2002).
32. van der Weegen, Y. *et al.* The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nature communications* **11**, 2104-2104 (2020).
33. Groisman, R. *et al.* CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes Dev* **20**, 1429-34 (2006).
34. Fischer, E.S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024-39 (2011).
35. Fei, J. & Chen, J. KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). *J Biol Chem* **287**, 35118-26 (2012).
36. Schwertman, P. *et al.* UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat Genet* **44**, 598-602 (2012).
37. Zhang, X. *et al.* Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nat Genet* **44**, 593-7 (2012).
38. Okuda, M., Nakazawa, Y., Guo, C., Ogi, T. & Nishimura, Y. Common TFIIH recruitment mechanism in global genome and transcription-coupled repair subpathways. *Nucleic Acids Res* **45**, 13043-13055 (2017).
39. Nakazawa, Y. *et al.* Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair. *Cell* **180**, 1228-1244 e24 (2020).
40. Schärer, O.D. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol* **5**, a012609 (2013).
41. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* **15**, 465-81 (2014).
42. Wilson, M.D., Harreman, M. & Svejstrup, J.Q. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim Biophys Acta* **1829**, 151-7 (2013).
43. Woudstra, E.C. *et al.* A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature* **415**, 929-33 (2002).

44. Paulsen, M.T. *et al.* Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods* **67**, 45-54 (2014).
45. Andrade-Lima, L.C., Veloso, A., Paulsen, M.T., Menck, C.F. & Ljungman, M. DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes. *Nucleic Acids Res* **43**, 2744-56 (2015).
46. Chiou, Y.Y., Hu, J., Sancar, A. & Selby, C.P. RNA polymerase II is released from the DNA template during transcription-coupled repair in mammalian cells. *J Biol Chem* **293**, 2476-2486 (2018).
47. Steurer, B. *et al.* Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. *Proceedings of the National Academy of Sciences* **115**, E4368-E4376 (2018).
48. Armache, K.-J., Kettenberger, H. & Cramer, P. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 6964-6968 (2003).
49. Conaway, R.C. & Conaway, J.W. The hunt for RNA polymerase II elongation factors: a historical perspective. *Nature Structural & Molecular Biology* **26**, 771-776 (2019).
50. Hartzog, G.A., Wada, T., Handa, H. & Winston, F. Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes Dev* **12**, 357-69 (1998).
51. Jaehning, J.A. The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochimica et biophysica acta* **1799**, 379-388 (2010).
52. D'Errico, M., Lemma, T., Calcagnile, A., Proietti De Santis, L. & Dogliotti, E. Cell type and DNA damage specific response of human skin cells to environmental agents. *Mutat Res* **614**, 37-47 (2007).
53. Budenholzer, L., Cheng, C.L., Li, Y. & Hochstrasser, M. Proteasome Structure and Assembly. *Journal of molecular biology* **429**, 3500-3524 (2017).
54. Livnat-Levanon, N. & Glickman, M.H. Ubiquitin-Proteasome System and mitochondria — Reciprocity. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1809**, 80-87 (2011).
55. Hoffman, E.A., Frey, B.L., Smith, L.M. & Auble, D.T. Formaldehyde crosslinking: a tool for the study of chromatin complexes. *The Journal of biological chemistry* **290**, 26404-26411 (2015).
56. Sutherland, B.W., Toews, J. & Kast, J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *Journal of Mass Spectrometry* **43**, 699-715 (2008).
57. Lomant, A.J. & Fairbanks, G. Chemical probes of extended biological structures: Synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis(succinimidyl propionate). *Journal of Molecular Biology* **104**, 243-261 (1976).
58. Kennedy-Darling, J. & Smith, L.M. Measuring the formaldehyde Protein-DNA cross-link reversal rate. *Anal Chem* **86**, 5678-81 (2014).
59. Jackson, V. Studies on histone organization in the nucleosome using formaldehyde as a reversible cross-linking agent. *Cell* **15**, 945-954 (1978).
60. Hopfer, U., Hopfer, H., Jablonski, K., Stahl, R.A.K. & Wolf, G. The Novel WD-repeat Protein Morg1 Acts as a Molecular Scaffold for Hypoxia-inducible Factor Prolyl Hydroxylase 3 (PHD3). *Journal of Biological Chemistry* **281**, 8645-8655 (2006).
61. Vomastek, T. *et al.* Modular construction of a signaling scaffold: MORG1 interacts with components of the ERK cascade and links ERK signaling to specific agonists. *Proc Natl Acad Sci U S A* **101**, 6981-6 (2004).
62. Huttlin, E.L. *et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505-509 (2017).
63. Huttlin, E.L. *et al.* The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425-440 (2015).

64. van den Boom, J. & Meyer, H. VCP/p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling. *Mol Cell* **69**, 182-194 (2018).
65. Verma, R., Oania, R., Fang, R., Smith, G.T. & Deshaies, R.J. Cdc48/p97 mediates UV-dependent turnover of RNA Pol II. *Mol Cell* **41**, 82-92 (2011).
66. He, J., Zhu, Q., Wani, G. & Wani, A.A. UV-induced proteolysis of RNA polymerase II is mediated by VCP/p97 segregase and timely orchestration by Cockayne syndrome B protein. *Oncotarget* **8**, 11004-11019 (2017).
67. Solomon, M.J. & Varshavsky, A. Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 6470-6474 (1985).
68. Rohs, R. *et al.* Origins of specificity in protein-DNA recognition. *Annu Rev Biochem* **79**, 233-69 (2010).
69. David, E.N., Bing, T. & Allan, R.B. Two-step cross-linking method for identification of NF- κ B gene network by chromatin immunoprecipitation. *BioTechniques* **39**, 715-725 (2005).
70. Groisman, R. *et al.* The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**, 357-67 (2003).
71. Whitby, M.C. The FANCM family of DNA helicases/translocases. *DNA Repair* **9**, 224-236 (2010).
72. Schwab, R.A. *et al.* The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol Cell* **60**, 351-61 (2015).
73. Chen, C.L. & Perrimon, N. Proximity-dependent labeling methods for proteomic profiling in living cells. *Wiley Interdiscip Rev Dev Biol* **6**(2017).
74. Lam, S.S. *et al.* Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat Methods* **12**, 51-4 (2015).
75. Kim, D.I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell* **27**, 1188-96 (2016).
76. Mello Filho, A.C., Hoffmann, M.E. & Meneghini, R. Cell killing and DNA damage by hydrogen peroxide are mediated by intracellular iron. *The Biochemical journal* **218**, 273-275 (1984).
77. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Human genetics* **135**, 851-867 (2016).
78. Boeing, S. *et al.* Multiomic Analysis of the UV-Induced DNA Damage Response. *Cell Reports* **15**, 1597-1610 (2016).
79. Stanley, F.K.T., Moore, S. & Goodarzi, A.A. CHD chromatin remodelling enzymes and the DNA damage response. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **750**, 31-44 (2013).
80. Miller, K.M. *et al.* Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining. *Nature structural & molecular biology* **17**, 1144-1151 (2010).
81. van den Heuvel, D. *et al.* A CSB-PAF1C axis restores processive transcription elongation after DNA damage repair. *bioRxiv*, 2020.01.04.894808 (2020).
82. Aygün, O., Svejstrup, J. & Liu, Y. A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proceedings of the National Academy of Sciences* **105**, 8580 (2008).
83. Aygün, O. *et al.* Direct inhibition of RNA polymerase II transcription by RECQL5. *J Biol Chem* **284**, 23197-203 (2009).
84. Saponaro, M. *et al.* RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell* **157**, 1037-49 (2014).
85. Zheng, L. *et al.* MRE11 complex links RECQ5 helicase to sites of DNA damage. *Nucleic Acids Research* **37**, 2645-2657 (2009).

86. Popuri, V. *et al.* Recruitment and retention dynamics of RECQL5 at DNA double strand break sites. *DNA Repair (Amst)* **11**, 624-35 (2012).
87. Errico, A. & Costanzo, V. Mechanisms of replication fork protection: a safeguard for genome stability. *Critical Reviews in Biochemistry and Molecular Biology* **47**, 222-235 (2012).
88. Okamoto, Y. *et al.* FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress. *The FEBS Journal* **286**, 139-150 (2019).
89. Wasmuth, Elizabeth V. & Lima, Christopher D. Exo- and Endoribonucleolytic Activities of Yeast Cytoplasmic and Nuclear RNA Exosomes Are Dependent on the Noncatalytic Core and Central Channel. *Molecular Cell* **48**, 133-144 (2012).
90. Goldstrohm, A.C., Hall, T.M.T. & McKenney, K.M. Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins. *Trends in Genetics* **34**, 972-990 (2018).
91. Edwards, A.M., Kane, C.M., Young, R.A. & Kornberg, R.D. Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *J Biol Chem* **266**, 71-5 (1991).
92. Calvo, O. RNA polymerase II phosphorylation and gene looping: new roles for the Rpb4/7 heterodimer in regulating gene expression. *Current Genetics* (2020).
93. Li, S. & Smerdon, M.J. Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *The EMBO Journal* **21**, 5921-5929 (2002).
94. Allepuz-Fuster, P. *et al.* Rpb4/7 facilitates RNA polymerase II CTD dephosphorylation. *Nucleic Acids Research* **42**, 13674 - 13688 (2014).
95. Cho, H. *et al.* A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev* **13**, 1540-52 (1999).
96. Friedl, E.M., Lane, W.S., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. The C-terminal domain phosphatase and transcription elongation activities of FCP1 are regulated by phosphorylation. *Proceedings of the National Academy of Sciences* **100**, 2328 (2003).
97. Vos, S.M. *et al.* Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature* **560**, 607-612 (2018).
98. Baillat, D. & Wagner, E.J. Integrator: surprisingly diverse functions in gene expression. *Trends in Biochemical Sciences* **40**, 257-264 (2015).
99. Hirose, Y. *et al.* Human phosphorylated CTD-interacting protein, PCIF1, negatively modulates gene expression by RNA polymerase II. *Biochemical and Biophysical Research Communications* **369**, 449-455 (2008).
100. Kinkelin, K. *et al.* Structures of RNA polymerase II complexes with Bye1, a chromatin-binding PHF3/DIDO homologue. *Proc Natl Acad Sci U S A* **110**, 15277-82 (2013).
101. Pinskaya, M. *et al.* PHD and TFIIS-Like domains of the Bye1 transcription factor determine its multivalent genomic distribution. *PLoS One* **9**, e102464 (2014).
102. Sigurdsson, S., Dirac-Svejstrup, A.B. & Svejstrup, J.Q. Evidence that Transcript Cleavage Is Essential for RNA Polymerase II Transcription and Cell Viability. *Molecular Cell* **38**, 202-210 (2010).
103. Wienholz, F. *et al.* FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res* **47**, 4011-4025 (2019).
104. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613 (2019).



Elongation factor ELOF1 drives transcription-coupled repair and prevents genome instability

Marit E Geijer¹, Di Zhou^{1,#}, Kathiresan Selvam^{2,#}, Barbara Steurer^{1,#}, Bastiaan Evers^{3,#}, Chirantani Mukherjee^{1,#}, Simona Cugusi^{4,#}, Marvin van Toorn¹, Melanie van der Woude¹, Wenzhi Gong⁵, Roel Janssens¹, Anja Raams¹, Joyce HG Lebbink^{1,6}, Bart Geverts⁷, Dalton A Plummer², Karel Bezstarosti⁸, Arjan F Theil¹, Richard Mitter⁹, Adriaan B Houtsmuller⁷, Wim Vermeulen¹, Jeroen AA Demmers⁸, Shisheng Li⁵, Hannes Lans¹, René Bernards³, Jesper Q Svejstrup⁴, Arnab Ray Chaudhuri¹, John J Wyrick², Jorgen A Marteijn¹

¹ Department of Molecular Genetics, Oncode Institute, Erasmus University Medical Center, Rotterdam, the Netherlands

² School of Molecular Biosciences, Washington State University, Pullman, Washington 99164, USA.

³ Oncode Institute, Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands.

⁴ Mechanisms of Transcription Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK.

⁵ Department of Comparative Biomedical Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA, 70803, USA

⁶ Department of Radiation Oncology, Erasmus University Medical Center, Rotterdam, The Netherlands.

⁷ Erasmus Optical Imaging Center, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands

⁸ Proteomics Center, Erasmus University Medical Center, Rotterdam, the Netherlands.

⁹ Bioinformatics and Biostatistics, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

#These authors contributed equally

Submitted to Nature Cell Biology

Abstract

Correct transcription is crucial for life. However, DNA damage severely impedes elongating RNA Polymerase II (Pol II), causing transcription inhibition and transcription-replication conflicts. Cells are equipped with intricate mechanisms to counteract the severe consequence of these transcription-blocking lesions (TBLs). However, the exact mechanism and factors involved remain largely unknown. Here, using a genome-wide CRISPR/cas9 screen, we identified elongation factor ELOF1 as an important new factor in the transcription stress response upon DNA damage. We show that ELOF1 has an evolutionary conserved role in Transcription-Coupled Nucleotide Excision Repair (TC-NER), where it promotes recruitment of the TC-NER factors UVSSA and TFIIH to efficiently repair TBLs and resume transcription. Additionally, ELOF1 modulates transcription to protect cells from transcription-mediated replication stress, thereby preserving genome stability. Thus, ELOF1 protects the transcription machinery from DNA damage by two distinct mechanisms.

Introduction

Faithful transcription is essential for proper cell function. However, transcription is continuously threatened by DNA damaging agents, which induce transcription-blocking lesions (TBLs) that strongly impede or completely block forward progression of RNA polymerase II (Pol II)^{1,2}. Impeded transcription elongation by DNA damage can affect transcription fidelity or result in complete absence of newly synthesized mRNA transcripts^{3,4}. This can result in severe cellular dysfunction, senescence and cell death, consequently contributing to aging⁵⁻⁷. Furthermore, prolonged stalling of Pol II at TBLs can form physical road blocks for the replication machinery, thereby giving rise to transcription-replication conflicts. These conflicts are detrimental for cells since they can lead to genome instability and onset of cancer⁷⁻¹⁰. Cells are equipped with an intricately regulated cellular response to overcome the highly toxic consequences of TBLs. This transcription stress response includes repair of TBLs and mechanisms to overcome transcription-replication conflicts^{1,8}.

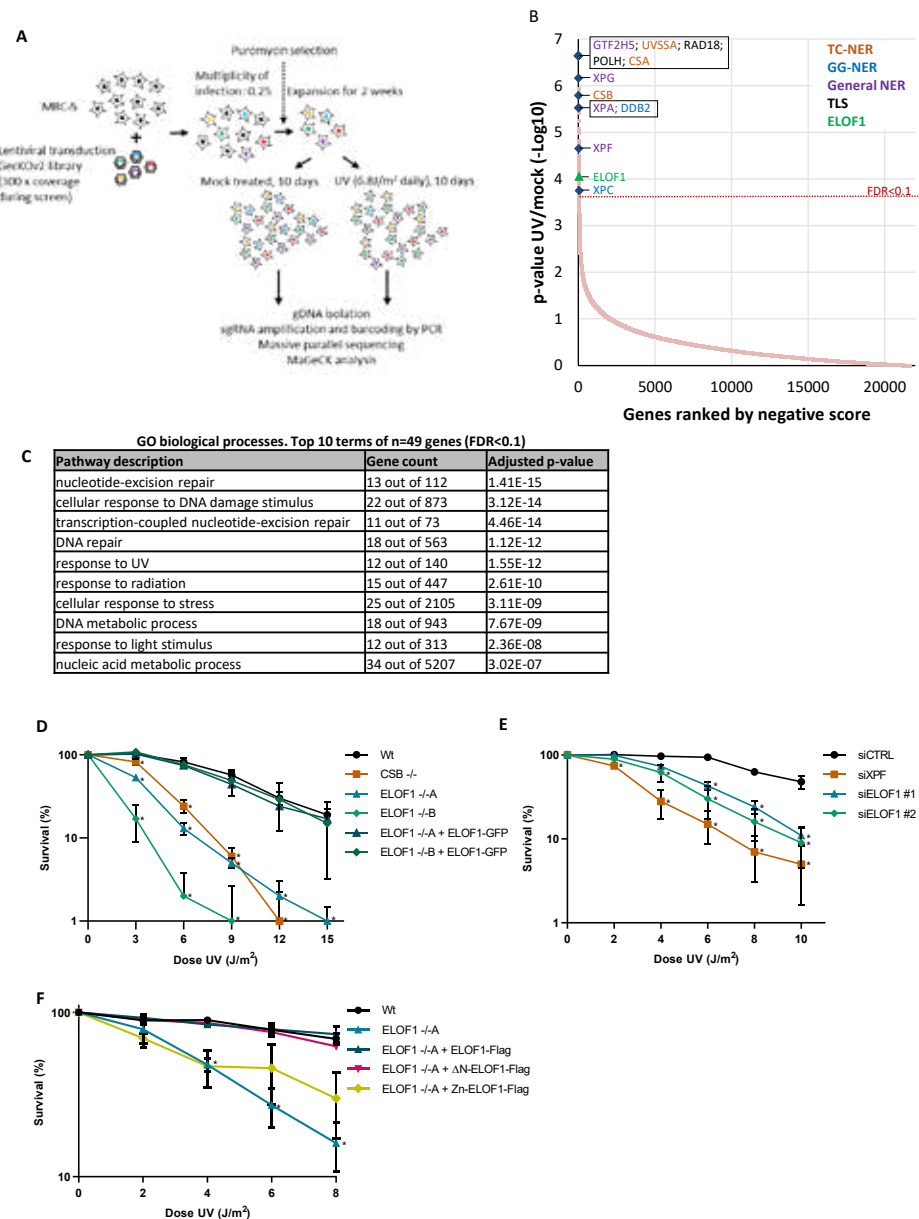
The main mechanism to remove TBLs is transcription-coupled nucleotide excision repair (TC-NER). TC-NER removes a wide spectrum of environmentally or endogenously-induced TBLs, such as UV-light-induced lesions or oxidative damage caused by metabolic processes^{1,11}. The biological consequences of TBLs and relevance of the TC-NER pathway are best illustrated by the fact that inactivating mutations in TC-NER genes can cause Cockayne syndrome (CS), which is characterized by photosensitivity, progressive neurodegeneration and premature aging¹². The TC-NER initiating factor CSB (ERCC6) is recruited upon Pol II stalling. CSB uses its forward translocating ability to discriminate between lesion-stalled and other forms of paused Pol II¹³. When lesion-stalled Pol II is recognized, TC-NER complex assembly is continued by recruitment of CSA (ERCC8)¹⁴⁻¹⁶, which is part of a Cullin 4-RING E3 ubiquitin ligase complex (CRL4^{CSA})¹⁷, and UVSSA. Interestingly, recently it was shown that also the ubiquitylation of lesion-stalled Pol II plays an important role in the transcription stress response^{18,19}. UVSSA subsequently promotes the recruitment of TFIIH^{18,20} which forms the core incision complex with XPA and RPA. The incision complex unwinds the DNA, verifies the lesion, and recruits the endonucleases ERCC1/XPF and XPG to excise the TBL²¹. Repair is finalized by refilling and ligating the gap²², after which transcription can restart. The repair reaction also resolves lesion-stalled Pol II, which helps to lower the frequency of TBL-induced transcription-replication conflicts.

Although several key factors have been identified in the cellular response to DNA damage-induced transcription stress, the exact molecular mechanism to repair TBLs by TC-NER or to avoid collisions of lesion-stalled Pol II with the replication machinery remain largely unknown.

Results

To obtain mechanistic insights in the DNA damage-induced transcription stress response, we performed a genome-wide CRISPR/Cas9 loss-of-function screen to identify novel factors involved in this cellular response. Briefly, fibroblasts were lentivirally transduced with an sgRNA library²³ at low multiplicity-of-infection (<0.25). The resulting pool of gene-edited cells was split into two populations. The control group was mock-treated,

while in the other group TBLs were induced by exposure to a daily UV-C dose of 6.8 J/m² for 10 consecutive days (Fig. 1A). This UV dose maintained a ~50% cell confluency throughout the screen (Suppl. Fig 1A). sgRNA abundance was determined by next generation sequencing of PCR-amplified incorporated sgRNAs from the isolated genomic DNA of surviving cell pools²⁴. sgRNA counts from UV-exposed cells were compared to those from untreated cells and negatively selected genes were identified using MAGeCK²⁵ (Fig. 1B, Suppl. Table 1). Gene ontology analysis among the top hits



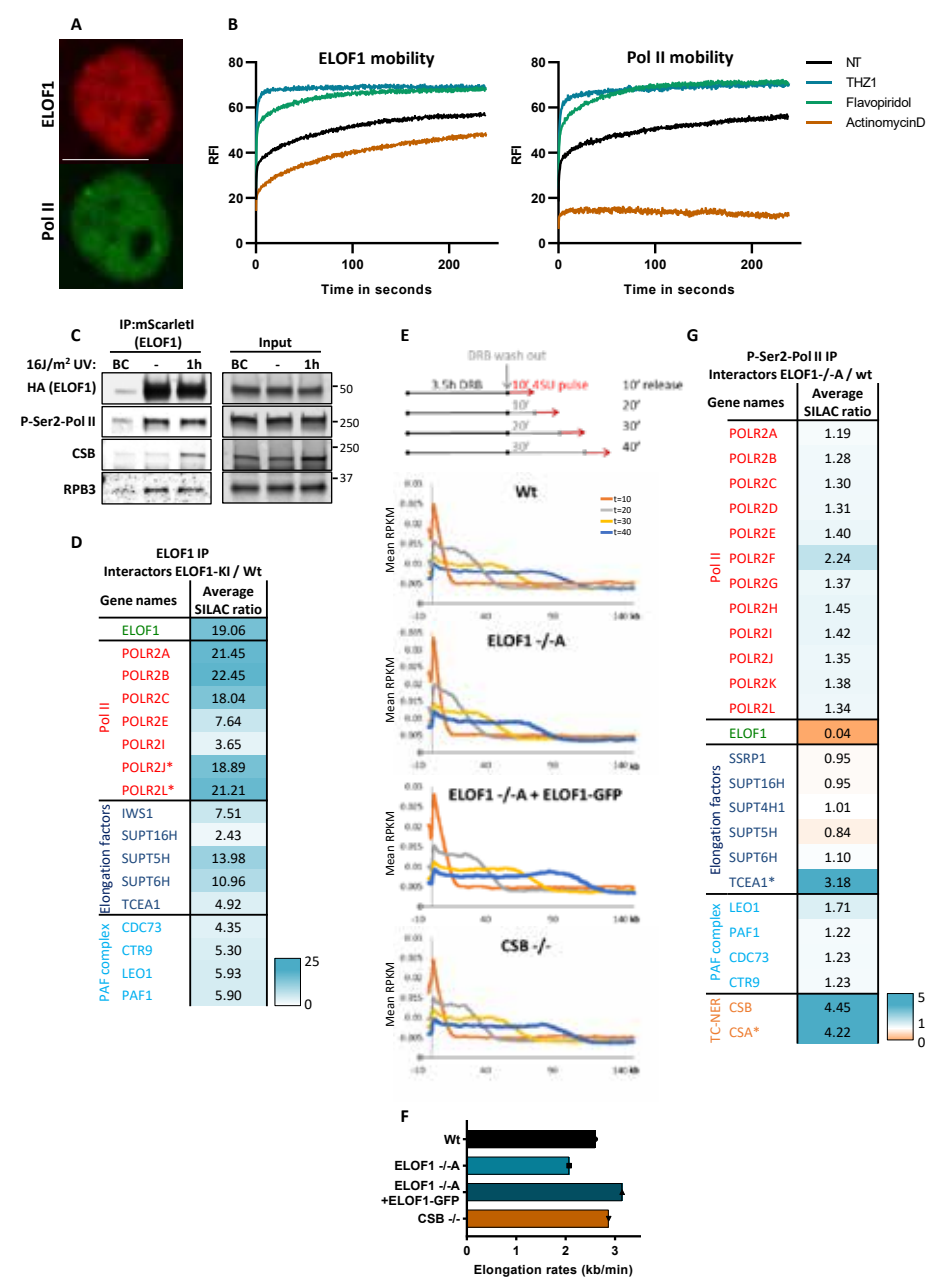
(FDR<0.1), identified many genes involved in the UV-induced DNA damage response (Fig. 1C). These genes included factors involved in translesion synthesis (TLS), like *RAD18* and *POLH*²⁶, and many NER genes including the global genome-NER (GG-NER) damage sensors *DDB2* and *XPC*²¹. Especially the identification of key TC-NER factors *CSA*, *CSB* and *UVSSA* underscored the potential of this screen to identify factors involved in the DNA damage-induced transcription stress response¹ (Fig. 1B).

Interestingly, one of the top hits was Elongation Factor 1 Homolog (*ELOF1*), an evolutionary-conserved small zinc-finger protein (~10 kDa)²⁷. *ELOF1* was identified in budding yeast in which disruption of its orthologue, *Elf1*, was shown to be synthetic lethal with mutation of genes encoding elongation factors such as *SPT6* and *TFIIS*²⁸. Follow-up studies in yeast revealed that Elf1 interacts with the core Pol II elongation complex as shown by proteomics²⁹, Cryo-EM studies³⁰, and by its presence at gene bodies as shown by ChIP^{28,31}. *In vitro* studies showed that Elf1 binds downstream of Pol II at the DNA entry tunnel and promotes elongation through nucleosomes³². However, its exact function as a transcription elongation factor, especially in mammalian cells, and its role in the DNA damage response has thus far remained unknown.

To validate the sensitivity to UV upon *ELOF1* depletion, as determined in our CRISPR/Cas9 screen, we performed clonogenic survival experiments using two independent *ELOF1* knockout (KO) cell lines (Suppl. Fig. 1B-F). *ELOF1* KO resulted in a severe UV hypersensitivity, even slightly higher than observed in TC-NER-deficient *CSB* KO cells (Fig. 1D). Similar results were obtained upon siRNA-mediated depletion of *ELOF1* (Fig. 1E, Suppl. Fig. 1G,H). Re-expression of *ELOF1* in *ELOF1* KO cells fully rescued their UV sensitivity, indicating that the observed effects are specific for *ELOF1*. Although the N-terminal tail of *ELOF1* promotes Pol II progression on the nucleosome³², constructs without this N-terminal tail could still rescue the UV sensitivity in *ELOF1* KO cells (Fig. 1F, Suppl. Fig. 1I). However, the conserved zinc-finger domain of *ELOF1* was crucial for survival upon UV-induced DNA damage. Furthermore, photolyase-mediated reversal of UV-induced cyclobutane pyrimidine dimer (CPD) lesions (33) almost completely rescued the UV sensitivity of *ELOF1* KO cells, showing that this sensitivity is due to induction of DNA damage and not RNA or protein damage (Suppl. Fig. 1J).

◀ **Figure 1. Genome-wide CRISPR/cas9 screen identifies *ELOF1* as a novel factor involved in the UV-induced DNA damage response.** (A) Schematic of the CRISPR/cas9 screen. MRC-5 (SV40) cells infected with a lentiviral sgRNA library were mock-treated or irradiated daily with 6.8 J/m² UV-C for 10 consecutive days. sgRNA abundance was determined by sequencing and UV-sensitive genes were identified by comparing the abundance in UV-irradiated cells over mock-treated cells. The screen was performed in duplicate. (B) UV-sensitive genes were ranked based on the gene-based P-value resulting from MaGeCK analysis of the change in abundance of sgRNAs in UV-treated over mock-treated. Dotted line indicates FDR=0.1. Genes involved in NER or TLS are color-coded. (C) Top 10 enriched GO terms (biological process) identified using g:Profiler of UV-sensitive genes with FDR<0.1 (n=49). (D) Relative colony survival of HCT116 wildtype (Wt) cells, indicated knock-out cells (-/-) or rescued cells exposed to the indicated doses of UV-C. (E) Relative colony survival of MRC-5 cells transfected with indicated siRNAs following exposure to the indicated doses of UV-C. (F) Relative colony survival of HCT116 *ELOF1* KO cells with expression of the indicated *ELOF1* mutants following exposure to the indicated doses of UV-C. Zn: zinc-finger mutant, ΔN: deletion of N-terminus. Plotted curves represent averages of three independent experiments ± SEM. *P≤0.05.

We first tested whether ELOF1 is part of the elongating Pol II complex, as previously observed for Elf1 in yeast^{28,29}. Since we could not obtain antibodies capable of recognizing endogenous ELOF1, we generated homozygous *ELOF1-mScarlet1-HA* knock-in (KI) cells to allow detection of endogenously expressed ELOF1 (Suppl. Fig. 2A). In living cells, ELOF1 was localized strictly to the nucleus, excluded from the nucleoli, and showed high level of co-localization with endogenously expressed GFP-tagged RPB1³³, the largest subunit



of Pol II (Fig. 2A and Suppl. Fig. 2B,C). Previous live-cell imaging studies on GFP-RPB1 mobility showed that fluorescence recovery after photobleaching (FRAP) experiments are a sensitive way to study Pol II-mediated transcription, as the different steps of the transcription cycle are characterized by kinetically distinct Pol II populations³³. Therefore, we compared the mobility of ELOF1 to that of Pol II using FRAP, and observed that it was almost identical in non-treated conditions (Fig. 2B). The large ELOF1 immobilization, which slowly redistributes in time, suggests that the majority of ELOF1 molecules is chromatin-bound, most likely engaged in transcription elongation, similar as observed for Pol II³³. The engagement of ELOF1 in transcription elongation was confirmed by its swift chromatin release, as shown by its strong mobilization upon inhibition of transcription initiation with the CDK7 inhibitor THZ1³⁴, or inhibition of release of promoter-paused Pol II into the gene body by the CDK9 inhibitor Flavopiridol³⁵ (Fig 2B). This almost complete mobilization upon transcription inhibition suggests that ELOF1 is exclusively involved in transcription-related processes. Furthermore, this highly similar dynamic behavior, also upon transcription inhibition, suggests that ELOF1 is closely associated with Pol II. Interestingly, treatment with the DNA intercalator actinomycin D³⁶ resulted in a severe immobilization of Pol II, while ELOF1 was only transiently immobilized. This suggests that ELOF1 can still dissociate from actinomycin D-stalled Pol II complexes while RPB1 remains trapped on the DNA.

To further investigate whether ELOF1 is part of the elongating Pol II complex we immunoprecipitated (IP) ELOF1 and detected its interaction with the RPB1 and RPB3 subunits of Pol II (Fig. 2C). The interaction of ELOF1 with P-Ser2-modified RPB1, which primarily marks productively elongating Pol II, indicates that ELOF1 is present in the elongating Pol II complex. The reciprocal IP of P-Ser2-modified Pol II confirmed that ELOF1 interacts with elongating Pol II (Suppl. Fig. 2D). Moreover, SILAC-based interaction proteomics of endogenously expressed GFP-RPB1³³ identified ELOF1 as a genuine Pol II interactor with similar SILAC ratios as other elongation factors (Suppl. Fig. 2E). To obtain a complete overview of ELOF1 protein interactions, we performed SILAC-based

◀ **Figure 2. ELOF1 is part of the elongating Pol II complex.** (A) Co-localization of ELOF1 and Pol II in HCT116 cells with ELOF1-mScarlet1-HA and GFP-RPB1 during live-cell imaging. Scale bar: 10 μ m. (B) Fluorescence recovery after photobleaching (FRAP) analysis of endogenously expressed ELOF1-mScarlet1 (Left) and GFP-RPB1 (Right). Cells were mock-treated (NT) or inhibited at different steps of the transcription cycle using indicated inhibitors. Relative Fluorescence Intensity (RFI) was measured over time, background-corrected, and normalized to pre-bleach fluorescence intensity. $n \geq 20$ for ELOF1-KI and $n \geq 8$ for RPB1-KI cells. (C) Immunoprecipitation of ELOF1 using RFP beads in ELOF1-KI cells followed by immunoblotting for indicated proteins. Cells were harvested 1 hour after mock treatment or irradiation with 16 J/m² UV-C. BC: binding control. (D) Interaction heat map of the SILAC ratios of ELOF1-interacting proteins as determined by quantitative interaction proteomics following HA-IP of ELOF1. Average SILAC ratios of duplicate experiments are plotted and represent ELOF1-interactors relative to empty beads. SILAC ratio >1 indicate increase in interaction. * indicates proteins quantified in one experiment. (E) Top panel: Schematic of DRB/TT_{chem}-seq to measure Pol II elongation rates. Bottom panel: Metagene profiles of DRB/TT_{chem}-seq in HCT116 Wt or indicated KO (-/-) cells, with ELOF1 re-expression where indicated, 10, 20, 30, or 40 minutes after DRB release. (F) Average elongation rates as determined by DRB/TT_{chem}-seq. (G) Interaction heat map based on the SILAC ratios as determined by quantitative interaction proteomics of P-Ser2-modified Pol II-interacting proteins in ELOF1 -/-A cells relative to Wt cells. Average SILAC ratios of duplicate experiments are plotted. * indicates proteins quantified in one experiment. SILAC ratios <1 indicate loss of interaction, >1 indicate increase in interaction.

interaction proteomics for ELOF1, revealing high SILAC ratios for many Pol II subunits and elongation factors including TFIIS, SPT6, SPT5 and the PAF complex^{37,38} (Fig. 2D and Suppl. Table 2). Gene ontology analysis of the most enriched ELOF1-interactors showed specific involvement in transcription-related processes (Suppl. Fig. 2F). Of note, the ELOF1-Pol II interaction did not change upon UV-induced DNA damage, in contrast to the interaction with CSB³⁹ (Fig. 2C and Suppl. Fig. 2D). Together, these live-cell imaging and interaction data indicated that ELOF1 is an integral component of the transcription elongation complex, independent of DNA damage.

Next, we tested whether ELOF1 acts as a transcription elongation factor by determining its effect on Pol II elongation rates. Therefore, we performed DRB/TT_{chem}-seq⁴⁰, in which nascent RNA is labeled with 4SU to determine the Pol II position in a gene body at different time points after its release from the promoter by DRB washout (Fig. 2E). Single gene profiles (Suppl. Fig. 3A) and metagene analysis (Fig. 2E) showed that ELOF1 KO resulted in a clear decrease in elongation rate, while ~6-fold overexpression of ELOF1 (Suppl. Fig. 1E) resulted in increased elongation rate. Based on metagene analysis, an average decrease in elongation speed from 2.6 kb/min to 2.0 kb/min was observed for ELOF1 KO cells, while an increase to 3.1 kb/min was observed after ELOF1 overexpression (Fig. 2F). In line with this reduced elongation rate, the overall nascent RNA synthesis was also reduced upon ELOF1 depletion (Suppl. Fig. 3B-D). In contrast, loss of CSB had no obvious effect on Pol II elongation rate.

To identify the mechanism for the reduction in elongation speed after loss of ELOF1, we compared the composition of the elongation complex with and without ELOF1. Endogenous P-Ser2-modified Pol II was isolated, and differences in the Pol II interactome were detected using SILAC-based proteomics. Absence of ELOF1 did not affect the presence of the core Pol II subunits or the majority of elongation factors in the elongation complex (Fig. 2G, Suppl. Table 2). For example, presence of the SPT4/5 dimer, which interacts genetically and biochemically with yeast ELOF1^{28,32}, was not changed in the elongation complex. Interestingly, the biggest change in complex composition was found for CSA, CSB and TFIIS, each having a 3- to 4-fold increased interaction with Pol II without ELOF1 (Fig. 2G). As CSB recognizes stalled and paused Pol II complexes, for example at DNA lesions or natural pause sites¹³, the increase in CSB binding might indicate that the forward translocation of Pol II is more frequently perturbed in the absence of ELOF1. Such perturbation can induce Pol II backtracking that is recognized by TFIIS to stimulate subsequent transcript cleavage⁴¹ to allow continued forward translocation. In line with such a model, we observed increased TFIIS binding to elongating Pol II in ELOF1 KO cells (Fig. 2G). In addition, depletion of TFIIS gave rise to synthetic lethality with ELOF1 KO (Suppl. Fig. 3E) as was previously also observed in yeast²⁸.

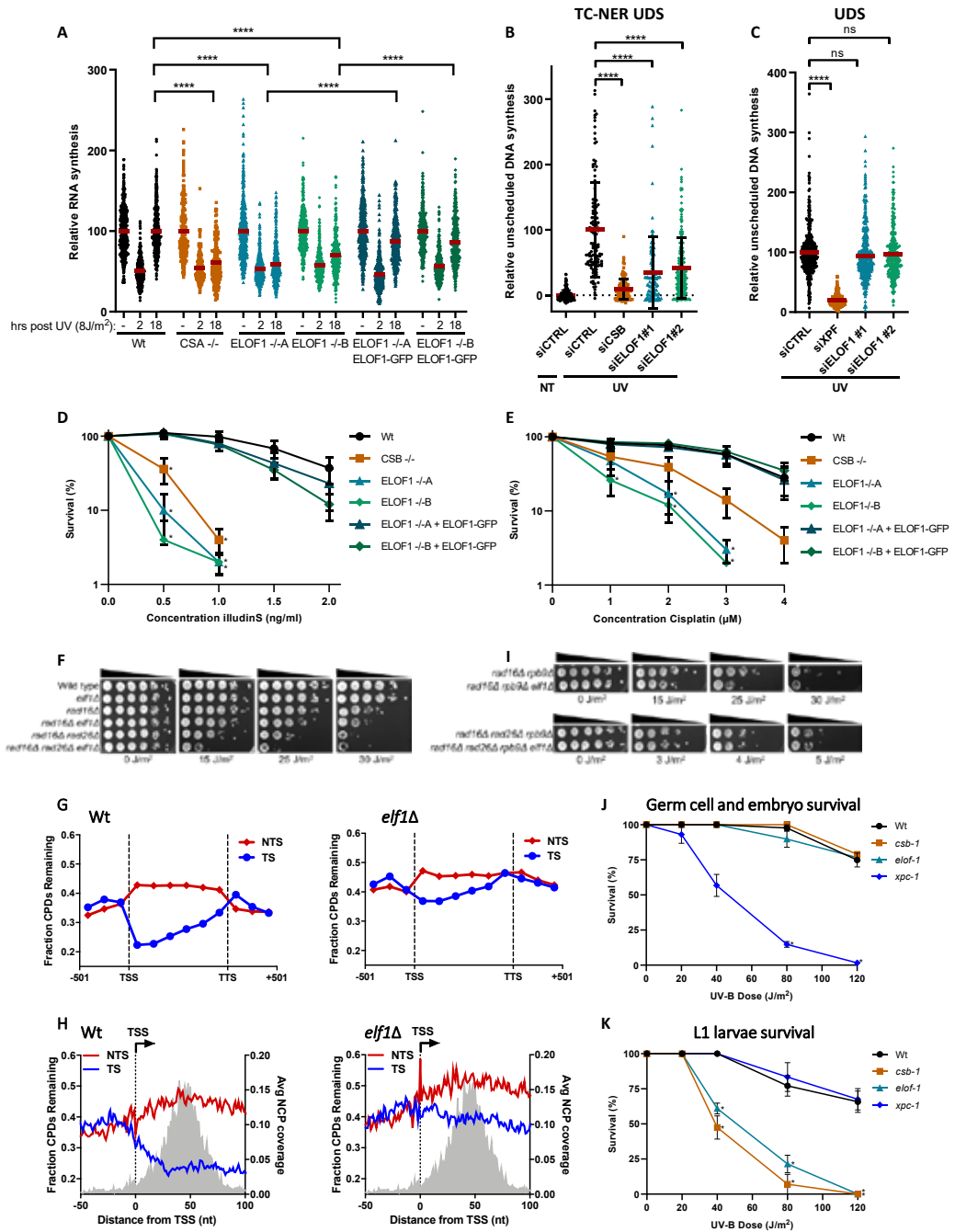
After having established that ELOF1 is a *bona fide* elongation factor, we studied its role in the DNA damage response. Since ELOF1 was shown to be an integral part of the transcription elongation machinery, and ELOF1 KO cells are sensitive for UV-induced DNA damage, which is a potent inhibitor of transcription, we tested whether ELOF1 is needed for recovery of transcription after UV irradiation by quantifying nascent transcription levels by EU incorporation⁴². Transcription was severely reduced 2 hours

after UV damage but fully recovered in Wt cells after 18 hours (Fig. 3A and Suppl. Fig. 4A). Strikingly, the transcription recovery was completely abolished in ELOF1 KO cells, as in TC-NER-deficient CSA KO cells, but could be rescued by re-expression of ELOF1. Similar results were obtained using siRNA-mediated ELOF1 knockdown (Suppl. Fig. 4B,C). This indicates that ELOF1 either has a function in the removal of TBLs, or in the restart of transcription. To distinguish between both possibilities, we measured TC-NER activity by quantifying the gap-filling synthesis using EdU incorporation in non-replicating GG-NER-deficient cells⁴³. Like CSB depletion, loss of ELOF1 severely inhibited the TC-NER activity indicating that ELOF1 has a crucial function in TC-NER (Fig. 3B, Suppl. Fig. 4D-E). The function of ELOF1 was restricted to the TC-NER sub-pathway, since the gap-filling synthesis in GG-NER-proficient cells was not affected (Fig. 3C, Suppl. Fig. 4F-G). Together, this shows that ELOF1 is important for removing UV-induced lesions by TC-NER to subsequently promote transcription recovery.

Next, we tested the sensitivity of ELOF1 KO cells to other types of DNA damage. Interestingly, ELOF1 KO, like CSB KO, resulted in a severe sensitivity to a wide spectrum of genotoxins that cause TBLs, including Illudin S⁴⁴, Cisplatin⁴⁵, Camptothecin⁴⁶ and oxidative lesions⁴⁷ (Fig. 3D-E and Suppl. Fig. 5A-C). However, ELOF1 KO cells were not sensitive to replication stress induced by hydroxyurea (Suppl. Fig. 5D). Importantly, sensitivity to TBLs is generally not observed after depletion of elongation factors since these were not among the top hits of our CRISPR/Cas9 screen for UV-sensitive genes (Suppl. Table 1). In addition, transient knockdown of the core elongation factors SPT4 and SPT5 did not increase UV sensitivity, although this caused a comparable reduction in RNA synthesis, similar as depletion of ELOF1 (Suppl. Fig. 3B-D, 5E).

As *ELOF1* is highly conserved from archaea to mammals²⁷, we tested whether the *ELOF1* orthologues in *Saccharomyces cerevisiae* and *Caenorhabditis elegans* are also involved in repairing TBLs. Similar to mutations in *RAD26*, the budding yeast ortholog of *CSB*, inactivation of *ELF1* (*elf1Δ*) had no effect on the UV sensitivity (Suppl. Figure 6A), which can be explained by the highly efficient GG-NER machinery in budding yeast⁴⁸. To specifically study the effect of *elf1Δ* in TC-NER, we tested the effect of its inactivation on UV survival in GG-NER-deficient *RAD16* mutants (*rad16Δ*). This showed a clearly increased sensitivity to UV for both the *elf1Δ* and the *rad16Δ* mutants, suggesting that Elf1 is involved in TC-NER (Fig. 3F).

To determine if the increased UV sensitivity in the *elf1Δ* mutant is caused by a TC-NER-defect, we analyzed CPD repair profiles in the transcribed strand (TS) and non-transcribed strand (NTS) of yeast genes 2 hours after UV using high-resolution CPD-sequencing⁴⁹ (Suppl. Fig. 6B). This analysis showed that in the *elf1Δ* mutant, GG-NER-mediated repair in the NTS was hardly affected. However, TC-NER-mediated repair in the TS was severely compromised, as shown by meta-analysis of ~4500 genes (Fig. 3G) and by individual genes (Suppl. Fig. 6C). The global repair rate in *elf1Δ* was hardly affected (Suppl. Fig. 6D), which is in agreement with a TC-NER-specific effect that only happens in the TS of active genes. Although Elf1 was described to stimulate Pol II progression on the nucleosome³², no nucleosome-dependent difference in TC-NER efficiency was detected in the TS in *elf1Δ* mutants (Suppl. Fig. 6E). Moreover, deletion of the N-terminus of ELOF1,

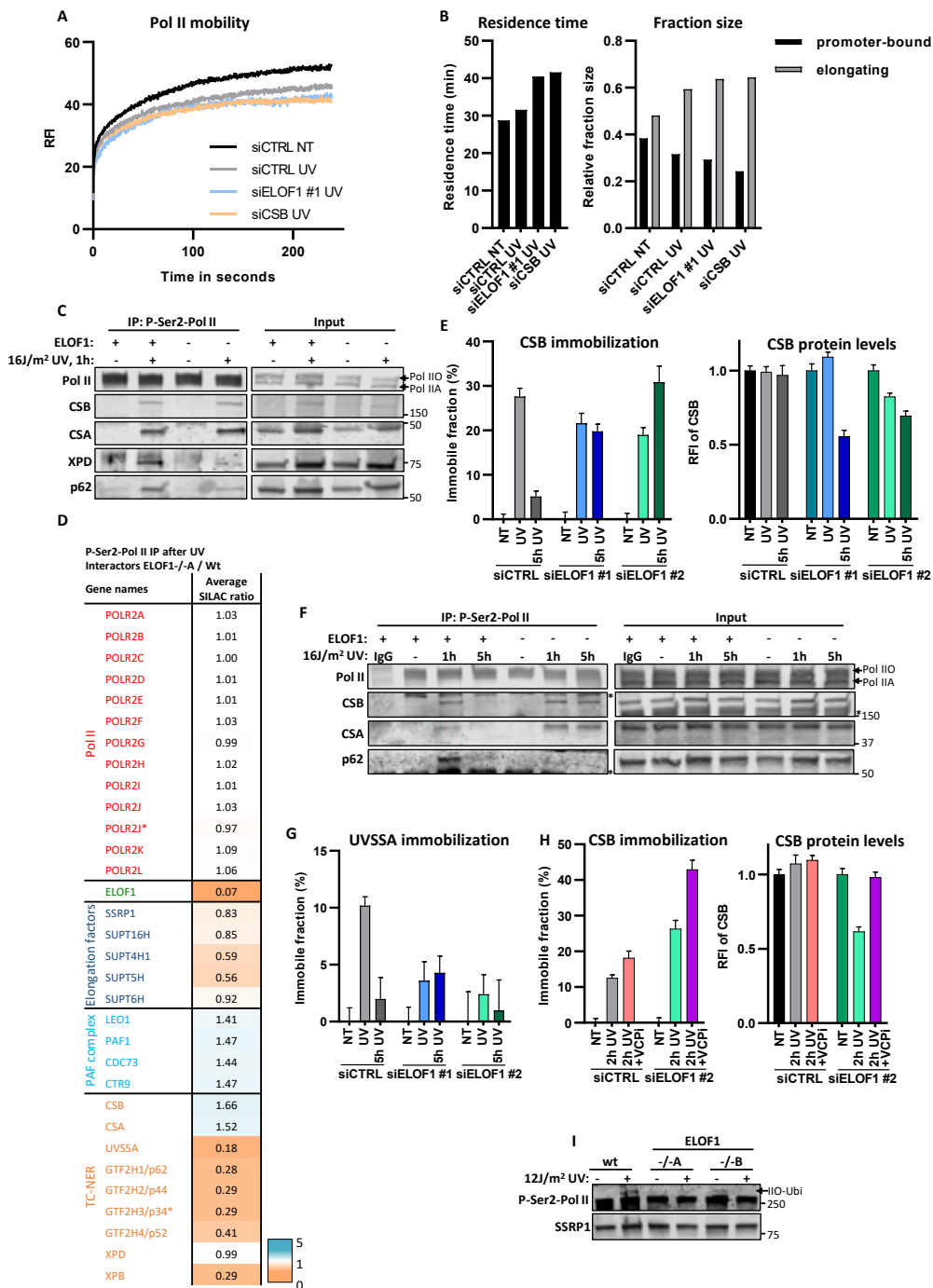


which is involved in transcription processivity at nucleosomes³², had no effect on the UV-survival (Fig. 1F).

Strikingly, the *elf1Δ rad26Δ* double mutant showed an even higher UV sensitivity than the *elf1Δ* and *rad26Δ* single mutants in a *rad16Δ* background, indicating that Elf1 has functions in the UV-induced DNA damage response independent of Rad26 (Fig. 3F). Close-ups of the CPD sequencing data showed that repair in the *elf1Δ* mutant is also compromised immediately downstream of the transcription start site (TSS) (Fig. 3H). This genomic region can be repaired in a Rad26-independent manner⁵⁰ by a Rpb9-mediated transcription-coupled repair mechanism⁵¹. This may suggest that Elf1 also functions in Rpb9-mediated repair, independent of Rad26 (Fig. 3F). Indeed, *elf1Δ* enhances the UV sensitivity in *rad16Δrpb9Δ* mutants, but not in a *rad16Δrpb9Δrad26Δ* mutant (Fig 3I and Suppl Fig. 6F), indicating that Elf1 is involved in both Rad26-dependent and -independent repair. This was confirmed by the finding that deletion of *ELF1* in both *rad16Δrad26Δ* and *rad16Δrpb9Δ* mutants resulted in reduced TC-NER (Suppl. Fig. 6G,H).

To study the role of ELOF1 in a multi-cellular model organism, we made use of the conservation of *ELOF1* in *C. elegans*. We assayed UV-survival of mutant germ and early embryonic cells, which predominantly depends on GG-NER, and of post-mitotic first-stage larvae, which mainly depends on TC-NER⁵². In contrast to inactivation of the GG-NER factor *xpc-1*, did inactivation of *elof-1* not increase UV sensitivity of germ and embryonic cells (Fig. 3J and Suppl. Fig. 6I). However, *elof-1* mutant animals showed a strong UV sensitivity in the first larval stage, similar to TC-NER-deficient *csb-1* animals (Fig. 3K).

◀ **Figure 3. ELOF1 is an evolutionary-conserved core TC-NER factor.** (A) Transcription restart after UV damage as determined by relative EU incorporation in the indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, at the indicated time points after UV-C (8 J/m²). Relative integrated density normalized to mock-treated levels and set to 100%. Red lines indicate average integrated density ± SEM. n≥300 cells from at least three independent experiments. (B) TC-NER-specific UDS as determined by relative EdU incorporation in XP186LV fibroblasts (XP-C) transfected with indicated siRNAs following UV-C-irradiation (7 hours, 8 J/m²). n≥100 cells from two independent experiments. (C) Relative levels of EdU incorporation in C590 (hTert) cells transfected with indicated siRNAs, following UV-C-irradiation (3 hours, 16 J/m²). n≥200 cells from at least two independent experiments (D+E) Relative colony survival of the indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, upon a 24-hour exposure to the indicated concentrations of illudinS (D) or Cisplatin (E). Plotted curves represent average of at least three independent experiments ± SEM. (F) Indicated mutant yeast strains were serially 10-fold diluted, spotted, and exposed to the indicated UV-C doses. (G) CPD-seq analysis of Wt (left) and *elf1Δ* mutant (right) yeast showing the average fraction of unrepaired CPDs remaining on the transcribed strand (TS) and non-transcribed strand (NTS) for ~4500 yeast genes following 2-hour repair relative to no repair. Each gene was divided in 6 equally-sized bins. Repair in flanking DNA upstream of the transcription start site (TSS) and downstream of the transcription termination site (TTS) is also depicted. (H) Close-up of CPD-seq repair data near the TSS in Wt (left) and *elf1Δ* mutant (right) cells. Nucleosome positioning data is shown for reference. (I) Indicated mutant yeast strains were serially 10-fold diluted, spotted, and exposed to the indicated UV-C doses. (J) *C. elegans* germ cell and embryo UV survival assay, measuring GG-NER activity, of wild type, *csb-1*, *xpc-1*, and *elof-1* animals. The percentages of hatched eggs (survival) are plotted against the applied UV-B doses. The mean survival of two replicate experiments each performed in quintuple is depicted. (K) L1 larvae UV survival assay, measuring TC-NER activity, of wildtype, *csb-1*, *xpc-1* and *elof-1* animals. The percentages of animals that developed beyond the L2 stage (survival) are plotted against the applied UV-B doses. The mean survival of three replicate experiments each performed in quintuple is depicted. Error bars represent the SEM. *p<0.05, ****p<0.0001.



Together these data indicate that ELOF1 is an important and highly evolutionary-conserved repair factor, specifically involved in repair of DNA damage in transcribed strands of active genes (Fig. 3). As ELOF1 is an integral part of the elongation complex (Fig. 2), its depletion will most likely affect Pol II forward translocation upon encountering TBLs. To test this, we used GFP-RPB1 KI cells to study Pol II mobility by FRAP, which provides quantitative information on Pol II elongation rates and fraction sizes of elongating and promoter-bound Pol II, i.e. initiating and promoter-paused Pol II³³. UV-induced DNA damage resulted in an increased Pol II immobilization, especially of the long-bound fraction, as evident from the reduced slope of the FRAP curve at time points >100 sec (Fig. 4A), which mainly represents dynamics of elongating Pol II³³. Monte-Carlo-based modeling⁵³ of these FRAP data revealed an increase in the fraction size and residence time of elongating Pol II. This indicates that UV-exposure resulted in more elongating Pol II transcribing with a lower average elongation rate (Fig 4B), most likely caused by Pol II-stalling at TBLs^{54,55}. Interestingly, the long-bound Pol II fraction after UV was further immobilized upon knockdown of ELOF1, to a similar extent as after depletion of CSB. Monte-Carlo-based modeling of these FRAP data showed an approximately 30% increase of the average residence time of elongating Pol II, suggesting that Pol II-stalling at lesions is prolonged in the absence of ELOF1 (Fig. 4A-B, Suppl. Fig. 7C). Similar results were obtained by Pol II ChIP-seq experiments (van der Weegen *et al.* submitted back-to-back). ELOF1 knockdown also resulted in an increased residence time of elongating Pol II in unperturbed conditions, indicative of a reduced elongation rate (Fig. Suppl. 7A-B), in line with our DRB/TT_{chem}-seq data (Fig 2E-F).

Since Pol II elongation was slowed down upon DNA damage induction in the absence of ELOF1, we immunoprecipitated elongating Pol II after UV to study whether specific

◀ **Figure 4. ELOF1 is crucial for proper TC-NER complex assembly.** (A) FRAP analysis of Pol II mobility in MRC-5 *GFP-RPB1* KI cells after depletion of indicated factors in untreated cells (NT) or directly after UV induction (UV, 12 J/m²). Relative Fluorescence Intensity (RFI) was measured over time, background-corrected, and normalized to pre-bleach fluorescence intensity. n≥17 cells. (B) Left panel: residence time of the elongating Pol II fraction. Right panel: relative fraction sizes of promoter-bound or elongating Pol II as determined by Monte-Carlo-based modeling based on the RPB1 mobility shown in (A). (C) Immunoprecipitation of P-Ser2-modified Pol II in Wt and ELOF1^{-/-} cells followed by immunoblotting for indicated proteins. Cells were harvested 1 hour after mock treatment or irradiation with 16 J/m² UV-C. (D) Interaction heat map based on the SILAC ratios as determined by quantitative interaction proteomics of UV-specific Pol II-interacting proteins in ELOF1^{-/-} cells relative to Wt cells. Average SILAC ratios of duplicate experiments are plotted. SILAC ratios <1 indicate loss of interaction, >1 indicate increase in interaction. * indicates proteins quantified in one experiment. (E) Left panel: Relative immobile fraction of CSB in *CSB-mScarlet1* KI cells transfected with indicated siRNAs directly (UV) or 5 hours after UV-C irradiation (5h UV, 4 J/m²) as determined by FRAP analysis (Suppl. Fig. S8E). Right panel: Relative fluorescence intensity of CSB-mScarlet1 in *CSB*-KI cells transfected with indicated siRNAs as determined by live-cell imaging. Plotted values represent mean ± SEM and are normalized to mock-treated. n≥9 cells. (F) Immunoprecipitation of P-Ser2-modified Pol II in Wt and ELOF1^{-/-} cells 1 hour or 5 hours after UV-C (16 J/m²) irradiation followed by immunoblotting for indicated proteins. IgG was used as binding control. *non-specific band. (G) Same as left panel of E but for *UVSSA-mScarlet1* KI cells (Suppl. Fig. 8F-G). n≥16 cells. (H) Relative immobile fraction (left panel) or relative fluorescence intensity (right panel) of *CSB-mScarlet1* in *CSB*-KI cells transfected with indicated siRNAs 2 hours after UV-C irradiation (4 J/m²) as determined by FRAP analysis (Suppl. Fig. 8H). VCPi: treatment with VCP inhibitor. Plotted values represent mean ± SEM and are normalized to mock-treated. n≥10 cells. (I) Immunoblot of chromatin fraction of indicated HCT116 Wt or ELOF1 KO cells 1 hour after 12 J/m² UV-C or mock treatment. SSRP1 is shown as loading control.

reaction steps of TC-NER were compromised. In ELOF1 KO cells, TC-NER initiating factors CSA and CSB were still properly bound to lesion-stalled Pol II. However, the UV-induced Pol II interaction with TFIIF complex subunits XPD and p62 was strongly reduced (Fig. 4C), which was not a consequence of TFIIF degradation (Suppl. Fig. 7D). To obtain a more unbiased overview of the effects of ELOF1 KO on the DNA damage-induced Pol II interactome, we performed SILAC-based interaction proteomics on the elongation complex after UV-induced DNA damage in presence or absence of ELOF1. The interaction of most elongation factors with Pol II was not affected by the absence of ELOF1 (Fig. 4D and Suppl. Fig. 7E). Interestingly, while CSA and CSB could still bind to Pol II in the absence of ELOF1, the proteins most affected in their Pol II binding were UVSSA and TFIIF subunits.

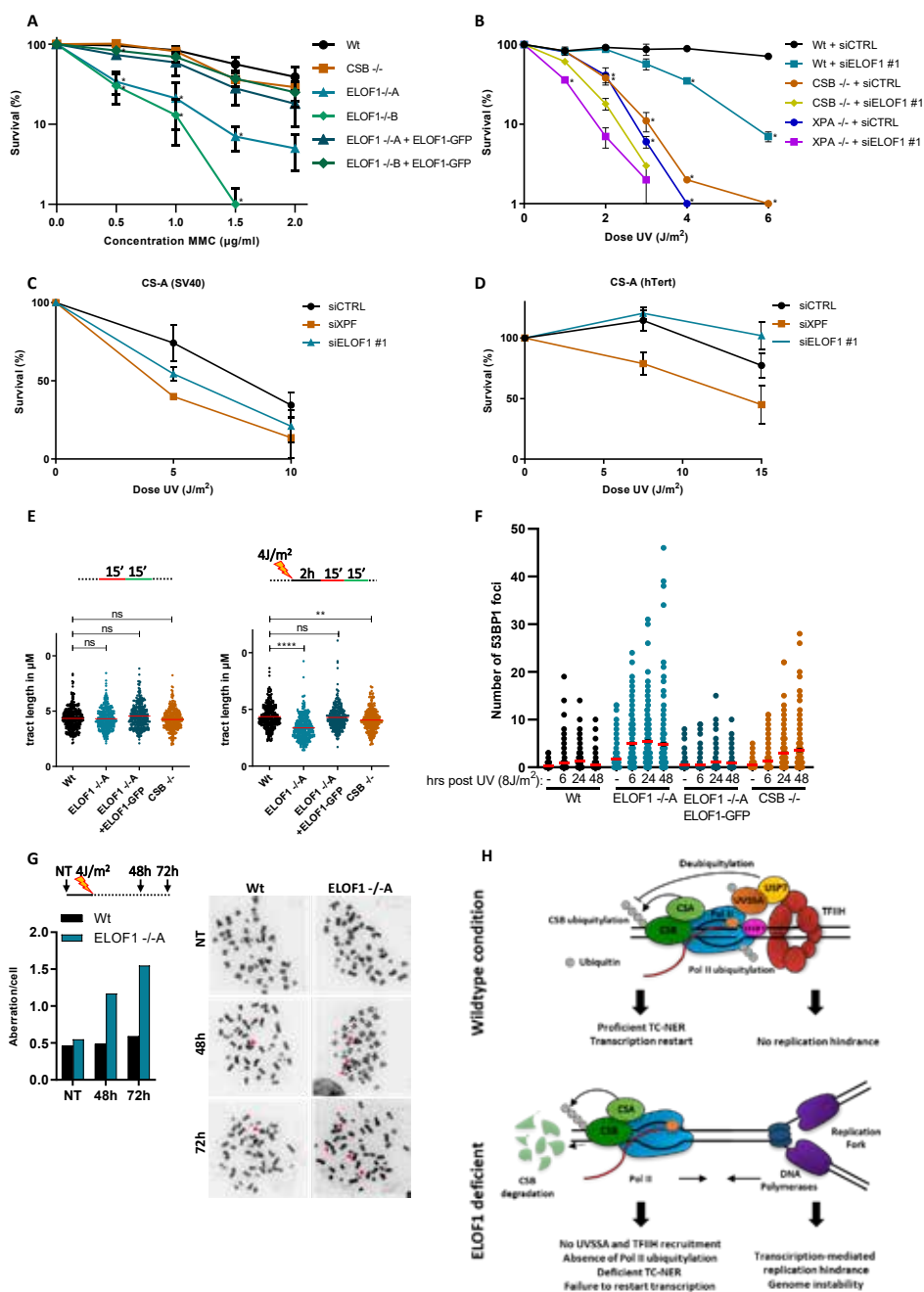
Since UVSSA plays a crucial role in the recruitment of TFIIF to lesion-stalled Pol II^{18,20,56}, the decreased binding of UVSSA is most likely the cause of reduced TFIIF recruitment, and explains the observed TC-NER defects. To confirm these results, we generated CSB and UVSSA knock-in cells (Suppl. Fig. 8A,B), expressing mScarlet1-tagged CSB and UVSSA proteins from their endogenous locus, allowing direct analysis of their quantity and mobility in living cells. TBL-induced immobilization of these TC-NER factors, as determined by FRAP^{39,57}, is an accurate measure for their involvement in TC-NER, as shown by their UV-induced immobilization in a transcription-dependent manner (Suppl. Fig. 8C,D). In line with the IP experiments, did ELOF1 depletion not affect the CSB immobilization 1 hour after UV. CSB remained immobilized up to at least 5 hours after UV (Fig. 4E and Suppl. Fig. 8E). This prolonged binding of CSB to stalled Pol II upon UV was confirmed by IP experiments, which also showed prolonged binding of CSA (Fig. 4F). These observations are in line with a model in which TBLs cannot be removed because of the TC-NER defect caused by ELOF1-deficiency, which will result in prolonged binding of CSB and CSA to lesion-stalled Pol II. In contrast, UVSSA immobilization upon UV damage was severely reduced after ELOF1 depletion (Fig. 4G and Suppl. Fig. 8F-G), further indicating that ELOF1 plays a crucial role in the recruitment of UVSSA to lesion-stalled Pol II. UVSSA recruits the deubiquitylating enzyme USP7, which protects CSB from proteasomal degradation mediated by the ubiquitin-selective segregase VCP/p97⁵⁷⁻⁵⁹. In line with this, we observed that reduced UVSSA recruitment upon ELOF1 depletion, and consequently of USP7, resulted in a UV-induced ~40% decrease of overall CSB levels (Fig. 4E) by VCP-mediated proteasomal degradation (Fig. 4H right panel). Interestingly, FRAP analysis showed an even stronger CSB immobilization upon TBL induction and VCP inhibition in ELOF1-depleted cells (Fig. 4H and Suppl. Fig. 8H), suggesting that chromatin-bound CSB is degraded in the absence of ELOF1, most likely when bound to lesion-stalled Pol II. Therefore, the increase in CSB immobilization upon UV-exposure confirms that in the absence of ELOF1, a larger Pol II fraction remains stalled at the lesion.

Recently, the ubiquitylation of a single lysine mutation in RPB1 (K1268) was described to be an important event in the transcription stress response^{18,19}. Therefore, we tested whether ELOF1 is involved in the UV-induced ubiquitylation of Pol II by means of a slower migrating P-Ser2-modified RPB1 band¹⁸. Interestingly, ELOF1 KO almost completely abolished the UV-induced RPB1 ubiquitylation (Fig. 4I), to the same extent

as CSB KO or inhibiting NEDD8-conjugating enzyme NAE1, which controls the activity of CRL complexes^{18,19} (Suppl. Fig. 8I). Similar results were obtained using siRNA-mediated depletion of ELOF1 (Suppl. Fig. 8J).

Together, our results demonstrate that the presence of ELOF1 in the lesion-stalled Pol II complex is an important determinant for proper Pol II ubiquitylation and for correct assembly of the TC-NER complex. The observed TC-NER defect explains the severe sensitivity of ELOF1 KO cells to different TBLs (Fig. 1D,E, 3D,E and Suppl. Fig. 5A-C). Strikingly, while testing the sensitivity of ELOF1 KO cells for a wide spectrum of DNA lesions, we observed that ELOF1 KO cells were also sensitive to the DNA crosslinker Mitomycin C (MMC) (Fig. 5A). Interestingly, CSB KO cells were not sensitive to MMC suggesting that ELOF1 has an additional function in the DNA damage response, besides canonical TC-NER. The prolonged transcription block in ELOF1 KO cells upon MMC exposure (Suppl. Fig. 9A), which was not observed in CSB KO cells, suggests that this additional role for ELOF1 is also linked to transcription. To investigate this additional function of ELOF1, we depleted ELOF1 in TC-NER-deficient CSB KO or NER-deficient XPA KO cells and strikingly observed that this resulted in increased UV sensitivity (Fig. 5B). Interestingly, CSB has also additional functions to ELOF1 in the response to UV-induced damage (Suppl. Fig. 9B). The additional role of ELOF1 to TC-NER was further confirmed in CS patient cells characterized by inactivating mutations in CSA (CS-A), in which knockdown of ELOF1 also resulted in additional UV sensitivity (Fig. 5C). Remarkably, this additive effect was completely absent in non-cycling CS-A cells (Fig. 5D), indicating that it is dependent on cell proliferation.

This replication-dependent sensitivity, together with the specific role of ELOF1 in transcription (Fig. 2), its additive effect to TC-NER (Fig. 5A and Suppl. Fig. 8A), and the prolonged Pol II-stalling upon ELOF1 knockdown (Fig. 4A,B), opened the possibility that lesion-stalled Pol II collides with incoming replication forks in the absence of ELOF1, thereby causing transcription-replication conflicts. Therefore, we investigated the impact of ELOF1 KO on DNA replication by analyzing the progression rates of individual replication forks by sequentially labelling cells with CldU and IdU. Tract length analysis revealed no significant difference in replication fork progression upon ELOF1 KO in unperturbed conditions, indicating that ELOF1 has no role in fork progression (Fig. 5E). However, 2 hours after UV, the tract length was significantly decreased in ELOF1 KO cells compared to Wt and ELOF1-complemented cells. Also, in CSB KO cells a small effect on fork progression was observed, however not to the same extent as in ELOF1 KO cells. This suggests that loss of the elongation factor ELOF1 results in replication problems upon induction of TBLs, likely due to the transcription-mediated replication blockage. Transcription-replication conflicts have previously been shown to result in under-replicated DNA, which may cause DSBs upon mitotic progression and subsequently give rise to genome instability^{8,60}. In line with this hypothesis, we observed a more pronounced increase in 53BP1-foci upon UV irradiation in ELOF1 KO cells compared to Wt or CSB KO cells (Fig. 5F, Suppl. Fig. 9C). As replication-interference and under-replicated DNA are important drivers of chromosomal aberrations, we assessed this in ELOF1 KO cells. This clearly resulted in an increased number of chromosomal aberrations in ELOF1 KO cells compared to Wt cells (Fig 5G).



Discussion

We unveiled an important role for ELOF1 in the cellular response to DNA damage-induced transcription stress by two independent mechanisms: promoting TC-NER and reducing transcription-mediated replication hindrance (Fig. 5H). First, ELOF1 is crucial for TC-NER. Interestingly, while the interaction of most TC-NER factors with elongating Pol II is strongly increased upon DNA damage^{1,14,39,56-58}, ELOF1 is already an intrinsic part of the elongating complex in unperturbed conditions where it stimulates transcription elongation (Fig. 2). Its dual function as an elongation and repair factor can be the cause of the embryonic lethality observed in ELOF1 KO mice⁶¹ and may explain why thus far no ELOF1 mutations were found in TC-NER related syndromes, something commonly observed for other TC-NER factors¹². The role of ELOF1 in TC-NER is highly conserved, also in yeast in which TC-NER is differently organized. For example, in yeast TBLs can be repaired in a Rad26-independent manner⁵⁰, and no homolog of UVSSA is detected. As ELOF1 is an integral part of the elongation complex we speculate that this is the reason why ELOF1 is crucial in both Rad26-dependent and -independent repair pathways, since its presence in the stalled Pol II complex is not dependent on Rad26.

ELOF1 promoted UVSSA binding to lesion-stalled Pol II, resulting in subsequent TFIIH recruitment, which promotes assembly of the full incision complex to excise the TBL and restart transcription. In the absence of ELOF1, TC-NER can still be initiated since CSB and the CRL4^{CSA} E3 ubiquitin ligase complex are still properly recruited to lesion-stalled Pol II (Fig 5H). Interestingly, although UVSSA was previously shown to be incorporated into the TC-NER complex through a direct interaction with CSA^{56,62}, we found that in the absence of ELOF1, a repair intermediate accumulated that consists of CSA but not of UVSSA. This

◀ **Figure 5. ELOF1 is important for preventing genome instability in addition to its function in TC-NER.** (A) Relative colony survival of indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, upon a 1-hour exposure to the indicated concentrations of mitomycin C. Plotted curves represent averages of three independent experiments \pm SEM. (B) Relative colony survival of MRC-5 Wt or indicated KO (-/-) cell lines, transfected with indicated siRNAs following exposure to the indicated doses of UV. Plotted curves represent averages of at least two independent experiments \pm SEM. * $P \leq 0.05$. (C+D) Viability of replicating CS-A (SV40, C) or non-replicating primary CS-A cells (hTert, D) following exposure to the indicated UV-C doses as determined by AlamarBlue staining. Plotted curves represent averages of at least two experiments \pm SEM. (E) Top panel: Schematic of experimental conditions for fork progression in indicated cell lines labeled with CldU (red) for 15 min followed by IdU (green) for 15 min as indicated. Bottom panel: Fork progression measured by tract lengths of CldU (red) in micrometers (μ M) is depicted for HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, in untreated conditions (left) or 2 hours after 4 J/m² UV-C (right). $n \geq 300$ tracts from three independent experiments. (F) Number of 53BP1 foci in HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, in untreated conditions or the indicated time after UV-C (8 J/m²). Red lines indicate average number of foci \pm SEM. $n \geq 200$ cells from two independent experiments. (G) Left panel: Quantitation of chromosomal aberrations per cell in HCT116 Wt and ELOF1 -/-A cells 48 or 72 hours after irradiation with 4 J/m² UV-C or mock treatment (NT). At least 60 metaphases were analyzed. Right panel: Representative images of metaphase spreads. Arrows indicate chromosomal aberrations. (H) Model showing function of ELOF1. Top panel: wildtype conditions: ELOF1 is an integral part of the elongation complex and binds near the DNA entry tunnel and ubiquitylation site of Pol II to promote TC-NER and subsequent transcription restart. Cells do not have replication problems. Bottom panel: in the absence of ELOF1 are CSA and CSB still recruited to lesion-stalled Pol II, however, UVSSA, TFIIH, and Pol II ubiquitylation are absent. The incomplete assembly of the TC-NER complex prevents functional TC-NER and subsequent transcription restart. In addition, there is an increase in transcription-mediated replication stress leading to genome instability. * $p \leq 0.05$, ** ≤ 0.01 , **** $p \leq 0.0001$.

suggests that more control steps are needed to recruit or stably incorporate UVSSA, and that this is not only mediated via a direct interaction with CSA, which is in line with the previously observed CSA-independent UVSSA recruitment^{57,63}. Such tight regulation will control the subsequent TFIIF recruitment and assembly of the incision complex and may represent an important proof-reading step that prevents build-up of the incision complex on non-lesion stalled Pol II. An example of such a regulatory mechanism is the recently discovered TBL-induced ubiquitylation of a single lysine residue (K1268) of Pol II, that is crucial for Pol II stability and TFIIF recruitment^{18,19}. Interestingly, based on recent structural analysis of the elongation complex in yeast³², the K1268 ubiquitylation site is in close proximity of ELOF1 (Suppl. Fig. 10). In the absence of ELOF1, Pol II ubiquitylation is reduced. We hypothesize that ELOF1 might either stimulate this ubiquitylation by facilitating a correct orientation of the elongation complex, or is involved in recruiting E3 ligases or repair factors that promote Pol II ubiquitylation. As UVSSA was shown to be important for the K1268 ubiquitylation¹⁸ and its recruitment is promoted by ELOF1, this might argue for the latter.

Our data, together with recent cryo-EM studies, indicate that TC-NER factors embrace the complete elongation complex, with CSB binding to upstream DNA extruding from Pol II¹³ and ELOF1 binding to downstream DNA entering Pol II³² (Suppl. Fig. 10). As CSB promotes the forward translocation of Pol II, thereby sensing for a TBL¹³, it is tempting to speculate that the presence of ELOF1 at the opposite site of Pol II might promote Pol II backtracking upon DNA damage, thereby facilitating repair factors to access the TBL. Since TFIIF might be involved in this backtracking process¹, this could explain why TFIIF recruitment is reduced upon ELOF1 depletion.

In addition to its role in TC-NER, our data show that ELOF1 plays an important role in preserving genome stability upon DNA damage, likely by preventing transcription-mediated replication stress (Fig. 5). The chromatin binding of the transcription machinery and the inability to clear Pol II from the DNA is assumed to play an important role in the onset of transcription-replication conflicts^{8,61,64}. Even though CSB and ELOF1 depletion had similar effects on the prolonged binding of Pol II upon DNA damage (Fig. 4A), only ELOF1 KO resulted in a clear replication defect and increased genome instability (Fig. 5E,F). This suggests that Pol II is differently processed in the absence of ELOF1 compared to what happens in CSB-deficient cells. This is most likely not caused by Pol II degradation, as loss of ubiquitylation is observed in the absence of both ELOF1 and CSB (Fig. 4I and Suppl. Fig. 8J). This implies that ELOF1 either has a function in Pol II release upon stalling at a lesion or that, in the absence of ELOF1, Pol II cannot be properly released from the DNA by incoming replication forks, resulting in an increase in transcription-replication conflicts (Fig. 5H). Together, our results show that ELOF1 is an important guardian of elongating Pol II by protecting transcription from the severe consequences of TBLs via two mechanisms; stimulating repair and preventing transcription-replication conflicts.

Acknowledgments

We thank the Optical Imaging Centre and the proteomics center of the Erasmus Medical Center for support with microscopes and mass spectrometry analysis. We thank the Advanced Sequencing Facility of the Francis Crick Institute's for technical assistance on the DRB/TT_{chem}-seq.

Funding

This work is part of the Oncode Institute which is partly financed by the Dutch Cancer Society and was funded by a grant from the Dutch Cancer Society (KWF grant 10506). This work was further funded by the Dutch organization for Scientific Research (NWO-ALW) which awarded a VIDI (864.13.004) and VICI (VI.C.182.025) grant to J.A.M. A.R.C. is supported by the Dutch Cancer Society (KWF grant 11008). S.L. is funded by the National Science Foundation (MCB-1615550). J.J.W. is funded by the National Institute of Environmental Health Sciences (grants R01ES028698, R21ES029655, and R21ES029302). H.L. is funded by The Netherlands Organization for Scientific Research (project nr 711.018.007) and Cancergenomics.nl. J.Q.S. was supported by the Francis Crick Institute (FCI receives funding from Cancer Research UK [FC001166], the UK Medical Research Council [FC001166], and the Wellcome Trust [FC001166]) and by a grant from the European Research Council (Agreement 693327).

Author contributions

M.E.G. performed the majority of the experiments and generated ELOF1 and CSB KO cell lines and ELOF1- and RPB1-KI cell lines. D.Z. generated CSB- and UVSSA-KI cells and performed live-cell imaging experiments and the TCR-UDS. K.S., D.A.P., W.G., S.L., and J.J.W. performed and supervised all experiments in *S.cerevisiae*. B.S. and M.E.G. performed the CRISPR/cas9 screen, and B.E. and R.B. analyzed and supervised the screen. C.M. and A.R.C. performed and supervised the metaphase spread and DNA fiber analysis. S.C., R.M., and J.Q.S. performed and supervised the DRB/TT_{chem}-seq. M.v.T. performed the alamar blue cell viability assay. M.v.d.W. and H.L. performed and supervised the experiments in *C. elegans*. R.J. provided experimental support. J.L. generated images of Pol II structure. B.G. performed Monte-Carlo based-modeling and was supervised by A.H.. K.B. and J.A.A.D. performed and supervised mass spectrometry analysis. A.R. performed UDS experiments and A.F.T. performed FACS sorting, both supervised by W.V.. J.A.M. conceived and supervised the project and together with M.E.G. wrote the manuscript with input from all authors.

Competing interests

Authors declare no competing interests.

Data and materials availability

All DRB/TT_{chem}-seq data used in this study is available under GEO accession: GSE148844. All CPD-seq data is available under GEO accession: GSE149082. Raw mass spectrometry data is available upon request.

Materials and methods

Cell lines and cell culture

MRC-5 (SV40) immortalized human lung fibroblast cells and HCT116 colorectal cancer cells were cultured in a 1:1 mixture of DMEM (Gibco) and Ham's F10 (Invitrogen) supplemented with 10% fetal calf serum (FCS, Biowest) and 1% penicillin-streptomycin in a humidified incubator at 37°C and 5% CO₂. C5RO fibroblasts (hTert), CS3BE (CS-A, SV40), XP186LV (XPC-/-) and CS216LV (CS-A, hTert) cells were maintained in Ham's F10 with 15% FCS and antibiotics.

For stable isotope labeling of amino acids in culture (SILAC), cells were grown for two weeks (>10 cell doublings) in arginine/lysine-free SILAC DMEM (ThermoFisher) supplemented with 15% dialyzed FCS (Gibco), 1% penicillin-streptomycin, 200 µg/ml proline (Sigma), and either 73 µg/mL light [¹²C₆]-lysine and 42 µg/mL [¹²C₆, ¹⁴N₄]-arginine (Sigma) or heavy [¹³C₆]-lysine and [¹³C₆, ¹⁵N₄]-arginine (Cambridge Isotope Laboratories).

HCT116 knock-out cells were generated by transiently transfecting HCT116 cells with a pLentiCRISPR.v2 plasmid²³ containing appropriate sgRNAs. Transfected cells were selected using 1 µg/ml puromycin (Invitrogen) for 2 days and single cells were seeded to allow expansion. Genotyping of single-cell clones was performed by immunoblotting or genomic PCR as indicated. sgRNAs sequences can be found in table 1, see below.

Table 1. Primers

Name	Sequence	function	Gene
sgRNA1	CACGGTACAAGAGATGACTC	sgRNA ELOF1 -/- B	ELOF1
sgRNA2	TCGTCTATAAACACAGAGACA	sgRNA ELOF1 -/- A	ELOF1
sgRNA3	GGGCAGGATGGAGAACAGCG	sgRNA ELOF1 -/- A	ELOF1
sgRNA4	GGTGATGACGCCTTCACCAA	sgRNA ELOF1 -/- A	ELOF1
sgRNA5	AATGAGGGAATCCCCACTC	sgRNA CSB -/-	CSB
sgRNA6	CACCGGCAATCAGTAGCGACACAG	sgRNA ELOF1 KI	ELOF1
sgRNA7	CACCGAATGTTGTTTAGCAGTATTC	sgRNA CSB KI	CSB
sgRNA8	CACCGCTACGCACTGAAGTAGAGAG	sgRNA UVSSA KI	UVSSA
fw1	GCCTCACTATGTTGCCAGG	Genotype KO	ELOF1
rv1	TCCTCTAGGCACACGGTACA	Genotype KO	ELOF1
fw2	CCCTGGTGACAGGGCCAAAGC	Genotype front of KI and complete KI	ELOF1
rv2	GGGCCACCGCTTGATTTTGGC	Genotype front of KI	ELOF1
fw3	GGTTGACGGCAATTTTCGATG	Genotype back of KI	ELOF1
rv3	GACCCCTGGAATGTCTCTGG	Genotype back of KI	ELOF1
rv4	CGGCTGTGACAGCCAGGACC	Genotype complete KI	ELOF1
fw5	CACCTGCAGGAAGCTTCTGC	Genotype front of KI (front, in CSB)	CSB

rv5	CAATCCAAGTATTTTCTCCTTAGC	Genotype CSB KI (reverse, in CSB)	CSB
fw6	CACCACAGAACACGATGACC	Genotype CSB KI (front, in CSB)	CSB
rv6	TCCATGTGCACCTTGAACCG	Genotype CSB KI (front, in HR template)	CSB
fw7	CATCCGGAGCTTGCAGGATCG	Genotype CSB KI (back, in HR template)	CSB
rv7	TCTCCTTTAGCTAGCATTATTA	Genotype CSB KI (back, in CSB)	CSB
fw8	ACGCGGATTTGCGCTCCAAC	Genotype UVSSA KI (back, in HR template)	UVSSA
rv8	TTCTGCGAGGCCAGACCCAT	Genotype UVSSA KI (reverse, in UVSSA)	UVSSA
fw9	ATCCTGCTCCCCGGAATGCC	Genotype UVSSA KI (front, in UVSSA)	UVSSA
rv9	CCACCGCTTGATTTTGGCAGG	Genotype UVSSA KI (front, in HR template)	UVSSA
PCR1 fw	ACACTCTTTCCTACACGACGC TCTTCCGATCTNNNNNNGGCTT TATATATCTTGTGGAAAGGACG	Barcoding PCR	Screen PCR1
PCR1 rv	GTGACTGGAGTTCAGACGTGTG CTCTCCGATCTACTGACGGGC ACCGGAGCCAATTCC	Barcoding PCR	Screen PCR1
PCR2 fw	AATGATACGCGACACCGAGA TCTACACTCTTTCCTACACGA CGCTCTCCGATCT	Adaptor PCR	Screen PCR2
PCR2 rv	CAAGCAGAAGACGGCATAACGAG ATTGGTCAGTGAAGTGGAGTTCA GACGTGTGCTCTCCGATCT	Adaptor PCR	Screen PCR2
Barcode 1	ACATCG	Wt 0J, barcode 1	Screen PCR1
Barcode 2	TCAAGT	Wt 0J, barcode 2	Screen PCR1
Barcode 3	TGGTCA	Wt 6.8J, barcode 3	Screen PCR1
Barcode 4	AAGCTA	Wt 6.8J, barcode 5	Screen PCR1

ELOF1 complemented cell lines were generated by lentiviral transduction in ELOF1 ^{-/-} cells. Therefore, full-length expression constructs with ELOF1-Flag-GFP, or Wt or mutated ELOF1-Flag were synthesized (Genscript) and inserted in a pLenti-CMV-puro-DEST plasmid⁶⁵. After transduction, cells were selected with 1 µg/ml puromycin.

HCT116 osTIR1 knock-in (KI) cells⁶⁶ were generated by transiently transfecting cells with an sgRNA-containing pLentiCRISPR.v2 plasmid (sgRNA sequences in table 1, see below)

targeting the stop codon of *ELOF1*, *CSB* or *UVSSA* and co-transfecting a homology-directed repair template, which included an Auxin-inducible Degron, fluorescent mScarletI-tag, HA-tag, hygromycin resistance cassette and homology arms (140 bp for *ELOF1*, 200 bp for *CSB* and *UVSSA*, sequence upon request)⁶⁷. Subsequently, cells were seeded in a low density to allow expansion and were kept in presence of 100 µg/ml hygromycin for two weeks to select for successful recombination. Single-cell clones were genotyped and homozygous KI clones were selected for further analysis. A GFP-RPB1 KI was generated in HCT116 Wt or *ELOF1*-KI cells as previously described by Steurer *et al.*⁶⁸. MRC-5 GFP-RPB1 KI cells⁶⁸ expressing CPD-PL-mCherry were generated as described previously⁶⁹.

Genotyping PCR was performed on genomic DNA (isolated using a PureLink™ Genomic DNA Mini Kit according to manufacturer's protocol) with Phusion (NEB) or taq (Invitrogen) polymerases according to manufacturer's protocol. Primer sequences can be found in table 1, see below. If necessary for assessing genomic alterations, PCR fragments were sequenced with forward primers and indels were analyzed using TIDE analysis⁷⁰.

siRNA transfections were performed 2 or 3 days before each experiment using Lipofectamine RNAiMax (Invitrogen) according to manufacturer's protocol. siRNAs were purchased from Dharmacon: siELOF1 #1: 5'-CCGUGGCCUAGAGGAAUUUU-3', siELOF1 #2: 5'-GAAAUCCUGUGAUGUGAAUUU-3', siCSB: 5'-GCAUGUGUCUACGAGAUUUU-3', siXPF: M-019946-00, siSPT4: L-012602-00-0005, siSPT5: L-016234-00-0005, siCSA: L-011008-00-0005. Knock-down efficiency was determined by immunoblot or RT-qPCR.

For UV-C irradiation, cells were washed with PBS, and placed under a 254 nm germicidal UV-C lamp (Philips). Duration of irradiation was controlled with an air-pressured shutter connected to a timer and cells were irradiated with doses as indicated. Cells were treated with VCP inhibitor (Seleck Chemicals, 5 µM) directly after UV irradiation or pre-treated 1 hour before irradiation with proteasome inhibitor MG132 (Enzo, 50 µM) or NEDD8 E1 Activating Enzyme Inhibitor (NAEi) MLN4924 (R&D systems, 10 µM) where indicated. Cells were treated for 1 hour with the following chemicals: Actinomycin D (Sigma, 1 µg/ml), Flavopiridol (Sigma, 1 µM), THZ1 (Xcessbio, 2 µM), Mitomycin C (Sigma, 10 µg/ml unless indicated differently), or potassium bromate (KBrO₃, Sigma). Cells were exposed continuously to camptothecin or treated for 24 hours with cisplatin, illudin S, or hydroxyurea (all Sigma). Final concentrations of all inhibitors were diluted in culture media and cells were washed once with PBS before putting fresh media after removing damaging agent when necessary. For ionizing radiation, plates were irradiated using an RS320 X-ray cabinet (X-Strahl). For photoreactivation, cells were washed with PBS and covered with a thin layer of HBSS (ThermoFisher) before exposing them to white-light tubes (General Electric Lighting Polylux LX F36W/840) for 10 minutes at 37 °C⁶⁹. Mock-treated samples were covered with tinfoil during photo-reactivation.

GeCKO v2 lentiviral library production and transduction

We used the lentiCRISPRv2 human library designed by Shalem *et al.*⁷¹ and obtained from Addgene. The sgRNA library was synthesized using array synthesis as previously described⁷¹ and cloned as a pool into the lentiCRISPR transfer plasmid for virus production.

To produce the pooled lentiviral library, twelve T-225 flasks of HEK293T cells were seeded at ~40% confluency the day before transfection. Per flask 10 µg of pVSVg, and 15 µg of psPAX2 (Addgene) packaging plasmids and 20 µg of lentiCRISPR plasmid library were transfected using Lipofectamine 2000 and Plus reagent (Life Technologies), according to manufacturer's instructions. After 6 hours the medium was changed, and after 60 hours the medium was collected and centrifuged at 3,000 rpm for 10 minutes at 4 °C to pellet cell debris. The supernatant was filtered through a 0.45 µm low protein-binding membrane (Millipore Steriflip HV/PVDF). To achieve a 300 times concentration of the GeCKO pooled library, the virus was ultracentrifuged (Sorvall) at 24,000 rpm for 2 hours at 4 °C and then resuspended overnight at 4 °C in D10 supplemented with 1% BSA. Aliquots were stored at –80°C.

Per condition 20 million MRC-5 cells were transduced at 75% confluency in 145 cm² dishes with concentrated lentivirus diluted in 18 ml of culture medium supplemented with 12 µg/mL polybrene (Sigma). The virus titer was determined to achieve a multiplicity of infection of <0.25. The next day, cells were re-seeded at 25% confluency in culture medium containing 2 µg/ml Puromycin. Cells were expanded for 1 week in puromycin-containing medium. Culture medium was refreshed every other day.

Genome-wide CRISPR screen

For UV irradiation or mock treatment 30 million transduced and puromycin-selected cells were seeded per condition at 40% confluency in 145 cm² dishes (2.25 million cells per dish) in medium without puromycin. The next day (day 0) dishes were mock-treated or irradiated with 6.8 J/m² UV-C. Control cells (mock-treated) and UV irradiated cells were washed with PBS and (mock) irradiated every day for 10 consecutive days. The culture medium was refreshed after each irradiation. Mock-treated cells were reseeded to 40% confluency when they reached a confluency > 90%. After the last irradiation cells were given 24 hours to recover and gDNA was isolated using the Blood & Cell Culture DNA Midi Kit (Qiagen) according to the manufacturers protocol (DNA content of MRC-5 cells was estimated at 10 pg per cell, genomic DNA of max 15 million cells was loaded per column). The screen was performed in duplicate.

PCR and next-generation sequencing

Per condition, sgRNA sequences of at least 300 µg of DNA (of ~30 million cells) were amplified by PCR (PCR1) using barcoded forward primers to be able to deconvolute multiplexed samples after next-generation sequencing (primers and barcodes are listed in table 1, see below). PCR1 was performed on 3 µg of gDNA in a total volume of 50 µl per reaction. Each PCR1 reaction contained 1 U of Phusion Hot Start II Polymerase (Thermo Fisher Scientific), 1x reaction buffer, 200 nm of each dNTP, 0.5 µM of both forward and reverse primer, and 3% DMSO. The following PCR program was used: initial denaturation for 3 minutes at 98°C; 35 cycles of denaturation for 1 sec at 98°C, primer annealing for 30 sec at 60°C, extension for 30 sec at 72°C, and final extension of 10 minutes at 72°C. Individual PCR reaction products were pooled per condition and 2 µl of pooled PCR product was used for a second PCR (PCR2) using primers containing adapters for next-generation sequencing (table 1, see below). The same PCR program was used as for PCR1, except that only 15 cycles were applied. 30 µl of PCR2 product was

cleaned up to remove primer pairs using the NucleoSpin Gel & PCR clean up kit (Bioké). Equal DNA content between conditions was checked by gel electrophoresis and samples were equimolarly pooled and subjected to Illumina next-generation sequencing as described before²⁴. Mapped read-counts were subsequently used as input for the Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout (MAGeCK) analysis software package, using version 0.5. For each condition, two biological replicates were performed. All conditions were sequenced simultaneously. To determine which genes showed a significant negative selection after 10 days of UV treatment, the sequencing data were analyzed with the MAGeCK tool²⁵. Gene ontology (GO) term enrichment analysis was performed using the g:Profiler website. Genes with a FDR<0.1 were analyzed and the top 10 biological processes affected by UV were identified.

Survival assays

For clonogenic survival assay, 200-300 cells were seeded per well in triplicate in a 6-well plate. The following day, cells were treated with different DNA damaging agents. Following treatment, colonies were grown for 7 to 10 days after which they were fixed and stained using Coomassie blue (50% methanol, 7% acetic acid and 0.1% coomassie blue (all Sigma)). To assess the growth speed of siRNA-transfected cells, 10,000 (HCT116) or 20,000 (ELOF1 -/-A) cells were seeded in a 6-well plate and grown for 10 days after transfection. Colony numbers were counted using GelCount (Oxford Optronix Ltd.). Relative colony number was plotted of at least 2 independent experiments, each performed in triplicate. Levels were normalized to mock-treated, set to 100 and plotted with SEM. Statistics was performed using independent T-test.

For AlamarBlue survival assay, siRNA-transfected cells were seeded to confluency in presence of 0.5% serum in triplicate in 96-well plates to arrest cells in G₀, and UV-irradiated after 30 hours. 72 hours after UV irradiation, AlamarBlue® (Invitrogen) was added for 4 hours and fluorescence was measured at 570 nm using a SpectraMax iD3 reader. Data were background corrected and normalized to mock-treated conditions.

RNA isolation, cDNA synthesis and RT-qPCR

To determine ELOF1 expression levels, RNA was isolated using the RNeasy mini kit (Qiagen) and cDNA was synthesized using the SuperScript™ II Reverse Transcriptase (Invitrogen), both according to the manufacturer's protocol. The generated cDNA was amplified using 1x taqman assay (ELOF1: Hs00361088_g1, GAPDH: 4333764T, both Thermofisher) and 1x taqman gene expression master mix (Thermofisher) by activating UNG for 2 minutes at 50°C, activating the polymerase for 10 minute at 95°C, followed by 40 cycles of 15 seconds of denaturing at 95°C and 1 minute of annealing and extending at 60°C in a CFX96 Touch Real-Time PCR Detection System. mRNA expression levels were normalized to GAPDH using the $2^{-\Delta\Delta Ct}$ method⁷².

Cell lysis and immunoblotting

Cells were directly lysed in SDS Page loading buffer (0.125M Tris pH 6.8, 2% SDS, 0.005% bromophenol blue, 21% glycerol, 4% β-mercaptoethanol) or, for assessing the chromatin fraction, one confluent 9.6 cm² dish was lysed for 30 minutes at 4°C in buffer containing 30 mM HEPES pH 7.5, 130 mM NaCl, 1 mM MgCl₂, 0.5% Triton X-100, cComplete™

EDTA-free protease inhibitors (Roche), Phosphatase inhibitor cocktail 2 (Sigma), N-ethylmaleimide (Sigma), and 50 μ M MG132i. Chromatin was pelleted at 15,000 for 10 minutes at 4°C and washed once. Finally, the chromatin was digested for 30 minutes at 4°C in presence of 50 U of benzonase (Millipore) before adding SDS Page loading buffer and incubating 5 minutes at 95°C. Chromatin fractions or cell lysates were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad). Proteins were transferred onto PVDF membranes (0.45 μ m, Merck Millipore) at 4°C, either 1.5h at 90V with 1x transfer buffer (25mM TRIS, 190mM Glycine, 10% methanol) or overnight at 25V in 2x transfer buffer (50mM TRIS, 380mM Glycine). Membranes were blocked with 5% BSA (Sigma) in PBS-tween (0.05%) and probed with primary antibodies (Table 2, see below). Subsequently, membranes were extensively washed with PBS-tween and incubated with secondary antibodies coupled to IRDyes (LI-COR, table 3, see below) to visualize proteins using an Odyssey CLx infrared scanner (LI-COR).

Table 2. primary antibodies.

Antibody	Host	Source	Dilutions	
			WB	IF
53BP1	Rb	Santa Cruz, sc-22760	N.A.	1/1000
BrdU (CldU)	Rat	Abcam, ab6326	N.A.	1/500
BrdU (IdU)	Ms	BD Biosciences, B44, 347580	N.A.	1/100
CSA/ERCC8	Ms	Santa Cruz, sc376981	1/250	N.A.
CSB/ERCC6	G	Santa Cruz, sc10459	1/500	N.A.
CSB/ERCC6	Rb	Antibodies-online, ABIN2855858	1/1000	N.A.
GFP	Ms	Roche, 14314500	1/1000	N.A.
GFP	Rb	Abcam, Ab290	1/1000	N.A.
HA	R	Roche, 11867423001	1/1000	N.A.
Lamin B1	Rb	Abcam,	1/1000	N.A.
p62/GTF2H1	Ms	Sigma Aldrich, WH0002965M1	1/1000	N.A.
RPB1 (Pol II)	Rb	Cell signalling, D8L4Y	1/1000	N.A.
RPB3	Rb	Abcam, ab138436	1/1000	N.A.
RPB9	Rb	Abcam, ab192407	1/500	N.A.
Ser2	R	Chromotek, 3E10	1/1000	N.A.
SPT4/SUPT4H1	Rb	Cell signalling, D3P2W	1/1000	N.A.
SPT5/SUPT5H	Rb	Bethyl, A300-869A	1/500	N.A.
SSRP1	Ms	Biologend, 609701	1/1000	N.A.
Tubulin	Ms	Sigma Aldrich, B512	1/5000	N.A.
XPD	Ms	Abcam, ab54676	1/1000	N.A.
XPF	Ms	Santa Cruz, sc-136153	1/500	N.A.

Table 3. secondary antibodies.

Antibody	Host	Source	Dilutions	
			WB	IRDye
Rabbit	goat	Sigma, sab4600215	1/10000	770
Rabbit	goat	Sigma, sab4600200	1/10000	680
Mouse	goat	Sigma, sab4600199	1/10000	680
Mouse	goat	Sigma, sab4600214	1/10000	770
Goat	donkey	Sigma, sab4600375	1/10000	770
Rat	goat	Sigma, sab4600479	1/10000	770
Antibody	Host	Source	IF	AlexaFluor
Rabbit	donkey	Invitrogen, A21207	1/1000	594
Mouse	goat	Invitrogen, A11001	1/300	488
Rat	donkey	Jackson Immuno-Research lab., 712-166-153	1/150	Cy3

Fluorescence Recovery After Photobleaching (FRAP)

For FRAP, a Leica TCS SP5 microscope (LAS AF software, Leica) equipped with a HCX PL APO CS 63x 1.40 NA oil immersion lens (ELOF1, RPB1, CSB) or Leica TCS SP8 microscope (LAS AF software, Leica) equipped with a HC PL APO CS2 63x 1.40 NA oil immersion lens (UVSSA) was used. Cells were maintained at 37°C and at 5% CO₂ during imaging. A narrow strip of 512 x 32 pixels (for ELOF1 and RPB1) or 512x16 (for CSB and UVSSA) spanning the nucleus was imaged every 400 ms (200 ms for UVSSA during pre-bleach) at 400 Hz using a 488 nm laser (RPB1) or 561 nm laser (ELOF1, CSB, UVSSA). 25 (RPB1), 40 (ELOF1), or 5 (CSB, UVSSA) frames were measured to reach steady state levels before photobleaching (1 frame 100% laser power for RPB1 and ELOF1, 2 frames for CSB and UVSSA). After photobleaching, the recovery of fluorescence was measured with 600 (ELOF1 and RPB1), 40 (CSB) or 20 (UVSSA) frames until steady-state was reached. Fluorescence intensity was measured inside and outside of the nucleus and recovery was determined by correcting for background signal and normalizing the values to the average pre-bleach fluorescence intensities. Relative fluorescence intensity levels were calculated using the pre-bleach intensities corrected for background. Immobile fractions (F_{imm}) were calculated using the individual and average (indicated by <brackets>) fluorescence intensities after bleaching (I_{bleach}) and fluorescence intensities after recovery from the bleaching ($I_{recovery}$):

$$F_{imm} = 1 - (I_{recovery,UV} - \langle I_{bleach} \rangle) / (\langle I_{recovery,unc} \rangle - \langle I_{bleach} \rangle)$$

Experimental FRAP curves of Pol II were simulated using Monte-Carlo-based computational modeling as described previously⁶⁸ to determine the residence time of elongating Pol II and the fraction size of promoter-bound and elongating Pol II.

Native immunoprecipitation (IP)

Cells were mock-treated or irradiated with 16 J/m² UV-C 1 hour prior to cell harvest. Cell pellets were prepared from 3 confluent 145 cm² dishes per condition for IP followed by immunoblot or 8 confluent 145 cm² dishes per condition for mass spectrometry. Cells were collected by trypsinization and pelleted in cold PBS using centrifugation for 5 minutes at 1500 rpm. After one wash with cold PBS, cell pellets were stored at -80°C until immunoprecipitation.

For immunoprecipitation, pellets were thawed on ice and lysed for 20 minutes at 4°C in HEPES buffer containing 30 mM HEPES pH 7.6, 1 mM MgCl₂, 150 mM NaCl, 0.5% NP-40, and 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche). Chromatin was pelleted by spinning 5 minutes at 10,000 g at 4°C and subsequently incubated for 1 hour at 4°C in HEPES buffer containing 500 units of Benzonase (Millipore) and 2 µg Pol II antibody (ab5095, abcam) or IgG (sc2027, SantaCruz) to digest the chromatin. After 1 hour, the NaCl was increased to 300 mM to inactivate benzonase and antibody-binding was continued for another 30 minutes. The undigested fraction was pelleted at 13,200 rpm for 10 minutes at 4°C and the soluble, antibody-bound fraction was immunoprecipitated for 90 minutes at 4°C using 25 µL slurry salmon sperm protein A agarose beads (Millipore). Unbound proteins were removed by washing the beads 5 times in wash buffer (30 mM HEPES pH 7.6, 150 mM NaCl, 1mM EDTA, 0.5% NP-40, and 0.2x cOmplete™ EDTA-free Protease Inhibitor Cocktail). Bound proteins were eluted in SDS page loading buffer and separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad). Samples were processed for immunoblotting or fixed and stained for mass spectrometry using Imperial protein stain (Pierce) according to manufacturer's protocol.

For ELOF1 IP, the same protocol was followed but instead of adding antibody during chromatin digestion, precipitation was performed using RFP-Trap® agarose beads (Chromotek) and binding control agarose beads (Chromotek).

Cross-linked immunoprecipitation

Cells were mock-treated or irradiated with 16 J/m² UV-C one hour prior to cell harvest. Cell pellets were prepared from 8 confluent 145 cm² dishes per condition for mass spectrometry. MRC-5 GFP-RPB1 KI cells were used for Pol II IP (Flag-beads) and HCT116 ELOF1-KI cells were used for ELOF1 IP (HA-beads).

Crosslinked IP was performed as described previously⁷³ with modifications as indicated. Cells were cross-linked with 1% paraformaldehyde (PFA) in serum-free DMEM for 7 minutes with constant shaking before quenching the reaction for 5 minutes with glycine (final concentration of 0.125 M). Cells were collected by scraping in PBS with 10% glycerol and 1 mM PMSF and pelleted for 15 minutes at maximum speed at 4°C. Consequently, chromatin was purified by washing the cell pellets for 30 minutes at 4°C in buffer 1 (50 mM HEPES, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100, 0.5% NP-40, 10% glycerol), pelleting the cells 10 minutes at 1300 rpm, washing the pellet twice with buffer 2 (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) and finally pelleting the chromatin, all at 4°C. Chromatin was sonicated in RIPA buffer (10 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% SDS, 1% Sodium Deoxycholate and 0.5 mM EGTA) using

the Bioruptor Sonicator (Diagenode) with 14 cycles of 15s on/15s off using the highest amplitude. Extracted chromatin was collected by spinning 15 minutes at maximum speed and pre-cleared for 30 minutes with Protein G agarose beads (Pierce) at 4°C. IP was performed by incubating 4 hours at 4°C with Flag M2 agarose beads (Sigma). Finally, aspecific interactors were removed by washing five times with RIPA buffer and proteins were eluted and crosslinking was reversed by incubating 30 minutes at 95°C in SDS Page loading buffer. Samples were separated on 4-15% Mini-PROTEAN TGX™ Precast Protein Gels (BioRad) and fixed and stained using imperial protein stain in preparation of mass spectrometry. To all buffers, 1 mM PMSF, 0.5 mM Na₂VO₄, 5 mM NaF, 5 mM NaPPi, 10 mM β-glycerol and cOmplete™ EDTA-free Protease Inhibitor Cocktail were added.

For ELOF1 IP, the same protocol was followed with minor alterations. Cells were crosslinked in 1 mM dithiobis(succinimidyl propionate) (DSP) in PBS for 30 minutes and quenched by adding Tris pH 7.5 to a final concentration of 25 mM for 10 minutes. IP was performed using HA-agarose beads (Sigma) and beads were incubated for 5 minutes at 95°C to elute and reverse cross-linked immunocomplexes.

Mass spectrometry

SDS-PAGE gel lanes were cut into slices and subjected to in-gel reduction with dithiothreitol (Sigma, D8255), alkylation with iodoacetamide (Sigma, I6125) and digestion with trypsin (sequencing grade; Promega) as previously described⁵⁷. Nanoflow liquid chromatography tandem mass spectrometry (nLC-MS/MS) was performed on an EASY-nLC 1200 coupled to a Lumos Tribid Orbitrap mass spectrometer (ThermoFisher Scientific) operating in positive mode. Peptide mixtures were trapped on a 2 cm x 100 μm Pepmap C18 column (Thermo Fisher 164564) and then separated on an in-house packed 50 cm x 75 μm capillary column with 1.9 μm Reprosil-Pur C18 beads (Dr. Maisch) at a flowrate of 250 nL/min, using a linear gradient of 0–32% acetonitrile (in 0.1% formic acid) during 90 min. The eluate was directly sprayed into the electrospray ionization (ESI) source of the mass spectrometer. Spectra were acquired in continuum mode; fragmentation of the peptides was performed in data-dependent mode by HCD. Mass spectrometry data were analyzed using the MaxQuant software (version 1.6.3.3). The false discovery rate (FDR) of both PSM and protein was set to 0.01 and the minimum ratio count was set to 1. The Andromeda search engine was used to search the MS/MS spectra against the UniProt database (taxonomy: Homo sapiens, release June 2017), concatenated with the reversed versions of all sequences. A maximum of two missed cleavages was allowed. In case the identified peptides of two proteins were the same or the identified peptides of one protein included all peptides of another protein, these proteins were combined by MaxQuant and reported as one protein group. Before further analysis, known contaminants and reverse hits were removed. Gene ontology (GO) term enrichment analysis was performed using the g:Profiler website. Genes with an average SILAC ratio of >2.5 were analyzed and the top 10 biological processes affected by UV were identified.

DRB/TT_{chem}-seq method

The DRB/TT_{chem}-seq was carried out as described in Gregersen *et al.*⁴⁰ in two biological replicates. Briefly, 8×10^6 cells were incubated in 100 μM DRB (Sigma-Aldrich) for 3.5 hours. The cells were then washed twice in PBS and fresh, DRB-free medium was

added to restart transcription. The RNA was labelled *in vivo* with 1 mM 4SU (Glentham Life Sciences) for 10 minutes prior to the addition of TRIzol (Thermo Fisher Scientific), which was used to stop the reaction at the desired time point. Following extraction, 100 µg of RNA was spiked-in with 1 µg 4-thiouracile labelled *S. cerevisiae* RNA (strain BY4741, MATa, his3D1, leu2D0, met15D0, ura3D0), and then fragmented with NaOH and biotinylated with MTSEA biotin-XXlinker (Biotium). The biotinylated RNA was then purified using µMACS Streptavidine MicroBeads (Miltenyi Biotec) and used for library preparation. The libraries were amplified using the KAPA RNA HyperPrep kit (Roche) with modifications as described in¹⁹. The fragmentation step was omitted and the RNA, resuspended in FPE Buffer, was denatured at 65°C for 5 min. Two SPRI bead purifications were carried out, with a bead-to-sample volume ratio of 0.95x and 1x, respectively. The libraries were then sequenced with single end 75bp reads on the HiSeq4000, with ~50,000,000 reads per sample.

Computational Analysis

DRB/TT_{chem}-seq data were processed using previously published protocol⁴⁰. Briefly, reads were aligned to human GRCh38 Ensembl 86. Read depth coverage was normalized to account for differences between samples using a scale factor derived from a yeast spike-in aligned and counted against *Saccharomyces cerevisiae* R64-1-1 Ensembl 86.⁷⁴ Biological replicate alignments were combined for the purpose of visualization and wave-peak analysis in order to increase read-depth coverage.

A set of non-overlapping protein-coding genes (200kb+) were selected for wave-peak analysis. A meta-gene profile was calculated by taking a trimmed mean of each basepairs coverage in the region -2kb:+200kb around the TSS. This was further smoothened using a spline. Wave peaks were called at the maximum points on the spline, with the stipulation that the peak must advance with time before being subjected to manual review. Elongation rates (kb/min) were calculated by fitting a linear model to the wave peak positions as a function of time.

EU incorporation

Cells were grown on coverslips and transcription levels were measured by pulse labeling with 5'ethynyl uridine (EU, Jena Bioscience) in Ham's F10 medium supplemented with 10% dialyzed FCS and 20 mM HEPES buffer (both Gibco). Cells were labeled for 30 minutes using 400 µM EU (MRC-5 cells) or for 1 hour with 200 µM EU (HCT116 cells) before fixation with 3.7% formaldehyde (FA, Sigma) in PBS for 15 minutes at room temperature (RT). After permeabilisation with 0.1% Triton X-100 in PBS for 10 minutes and blocking in 1.5% BSA in PBS for 10 minutes, Click-it chemistry-based azide coupling was performed by incubation for 1 hour with 60 µM Atto594 Azide (Attotec, Germany) in 50 mM Tris buffer (pH 8) with 4 mM CuSO₄ (Sigma), and 10 mM freshly prepared ascorbic acid (Sigma). DAPI (Brunschwig Chemie) was added to visualize the nuclei. Coverslips were washed with 0.1% Triton in PBS and PBS only and mounted with Aqua-Poly/Mount (Polysciences). Cells were imaged with a Zeiss LSM 700 Axio Imager Z2 upright microscope equipped with a 40x Plan-apochromat 1.3 NA oil immersion lens or 63x Plan-apochromat 1.4 NA oil immersion lens (Carl Zeiss Micro Imaging Inc.). Integrated density of the EU signal in the nuclei was quantified using ImageJ. Therefore, the surface of each nucleus was

determined based on the DAPI signal and mean fluorescence intensity was determined, corrected for the background signal. With these values, the integrated density was calculated, and plotted as single cell point with the average and SEM.

For assessing recovery of transcription after UV, cells were mock-treated or irradiated with 8 J/m² UV-C 2 or 18 hours before EU incorporation. For recovery after mitomycin C, cells were mock-treated or incubated for 2 hours with 10 µg/ml Mitomycin C followed by a recovery period of 2 or 22 hours in normal medium. Integrated density was normalized to mock-treated.

TC-NER-specific UDS

Amplified UDS was performed as described previously⁴³. Briefly, siRNA transfected primary XP186LV (XP-C patient cells) were serum-deprived for at least 24 hours in Ham's F10 (Lonza) containing 0.5% FCS and antibiotics to arrest cells in G₀. Cells were irradiated using 8 J/m² UV and labelled for 7 hours with 20 µM 5-Ethynyl-2'-deoxyuridin (EdU) and 1 µM Floxouridine (Sigma). Subsequently, a 15-minute chase was performed with normal medium (0.5% FCS) supplemented with 10 µM thymidine (Sigma) to remove unincorporated EdU and cells were fixed and permeabilized with 3.7% FA and 0.5% Triton X-100 for 15 minutes. After permeabilizing the cells for 20 minutes with 0.5% Triton in PBS and washing with 3% BSA in PBS, endogenous peroxidase activity was quenched using 2% hydrogen peroxide (Sigma) for 15 minutes and incubated with PBS+ (0.5% BSA + 0.15 % glycine). Click-it chemistry was performed using the Click-it reaction cocktail containing Azide-PEG3-Biotin Conjugate (20 µM, Jena Bioscience), 1× Click-it reaction buffer (ThermoFisher Scientific), copper(III) sulfate (0.1 M) and 10× reaction buffer additive (ThermoFisher Scientific) for 1 hour and washed with PBS. To amplify the signal, coverslips were incubated for 1 hour using HRP-streptavidin conjugate (500 µg/ml), followed by PBS washes and a 10-minute incubation with Alexa-Fluor 488 labeled tyramide (100x stock, Thermofisher Scientific). Coverslips were washed with PBS and PBS+ and the nuclei were stained with DAPI in 0.1% triton. DAPI was washed away with 0.1% triton and slides were mounted using Aqua-Poly/Mount.

Unscheduled DNA Synthesis (UDS)

Cells were grown to confluency on coverslips and serum-deprived (0.5%) for 2 days to arrest cells in G₀. Cells were irradiated with 16 J/m² and labeled with 20 µM EdU (Invitrogen) in Ham's F10 supplemented with 10% dialyzed FCS and 20 mM HEPES buffer (both Gibco) for 3 hours before fixation for 15 minutes (3.7% FA and 0.5% triton X-100). Background signal was blocked by washing twice with 3% BSA in PBS for 10 minutes and nuclei were permeabilized for 20 minutes using 0.5% triton in PBS. EdU incorporation was visualized using Click-it chemistry, imaged and analyzed as described in the section *EU incorporation* with the adjustment that click-it reaction was performed for 30 minutes.

Yeast strains

Yeast deletion strains used in this study are derivatives of the wild type strain BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) and Y452 (*MATα, ura3-52, his3-1, leu2-3, leu2-112, cir^o*). The gene deletions were made by transformation of yeast cells with PCR products bracketing selection markers⁷⁵ or following published methods⁷⁶.

Yeast UV sensitivity assay

Yeast cells were grown in YPD medium to mid-log phase. For spotting assay, cells were serially 10-fold diluted in fresh YPD medium and spotted on YPD plates. After exposure to different doses of UV-C light (254 nm), plates were incubated at 30°C in the dark and images were taken after 3-5 days of incubation. For quantitative UV survival assay, diluted yeast cells were plated on YPD plates and exposed to the indicated UV doses. The number of colonies on each plate was counted after incubating for 3 days at 30°C in the dark. The survival graph depicts the mean and SEM of three independent experiments.

CPD-seq library preparation and sequencing

CPD-seq analysis of repair in Wt and *elf1Δ* mutant strains was performed as previously described⁴⁹. Briefly, yeast cells were grown to mid-log phase, pelleted, re-suspended in dH₂O, and irradiated with 125 J/m² UV-C light (254 nm). After UV treatment, cells were incubated in the dark in pre-warmed, fresh YPD medium for repair. Cells were collected before UV irradiation (No UV), immediately after UV (0 hours), and following a 2-hour repair incubation. The cells were pelleted and stored at -80°C until genomic DNA isolation.

Genomic DNA extraction, CPD-seq library preparation and quality control, sequencing with an Ion Proton sequencer, and data processing were performed as previously described⁴⁹. The resulting sequencing reads were aligned to the yeast genome (saccer3) using Bowtie 2⁷⁷. Only CPD-seq reads associated with lesions at dipyrimidine sequences (i.e., TT, TC, CT, CC) were retained for further analysis.

Bin analysis for CPD repair along the transcribed strand (TS) and non-transcribed strand (NTS) of ~4500 yeast genes was performed as previously described⁷⁸, using transcription start site (TSS) and polyadenylation site (PAS, also referred to as transcription termination site, TTS) coordinates from Park *et al.*⁷⁹. A similar gene bin analysis was displayed for each yeast gene using the Java Treeview program^{80,81}. Genes were sorted by transcription rate⁸². Single nucleotide resolution repair analysis adjacent to the TSS was performed as previously described^{78,83}. Nucleosome dyad coverage from MNase-seq experiments were obtained from Weiner *et al.*,⁸⁴ as reference. CPD-seq data for *elf1Δ* and Wt yeast was normalized using the fraction of CPDs remaining determined for bulk genomic DNA by T4 endonuclease V digestion and alkaline gel electrophoresis (see below).

Analysis of bulk CPD repair in UV irradiated yeast

Alkaline gel electrophoresis to assay global DNA repair of bulk DNA was conducted as previously described⁸⁵. Yeast cell cultures were grown to mid-log phase in YPD media. Yeast cell cultures were briefly centrifuged to pellet, resuspended in dH₂O, and exposed to 100 J/m² UV-C light or left unirradiated for the “No UV” sample. Following irradiation, yeast cells were resuspended in YPD and incubated at 30°C. Aliquots were taken at each repair time point, briefly centrifuging to discard media supernatant prior to storing yeast cells at -80°C. Genomic DNA was isolated by bead beating the yeast cell pellets in 250 μL lysis buffer (2% Triton-X 100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl pH 8, 1 mM Na₂EDTA) and 300 μL Phenol-Chloroform-Isoamyl alcohol (25:24:1). 300 μL TE pH 8 was added to each tube, briefly vortexing to mix. Samples were centrifuged and the DNA-containing

aqueous layer was transferred to a fresh tube for ethanol precipitation. DNA pellets were resuspended in TE pH 8 containing 0.2 mg/mL RNase A, incubating at 37°C for 15 minutes prior to enzymatic digestion. Equal amounts of DNA were then treated with T4 endonuclease V (T4 PDG; NEB) and resolved by electrophoresis on a 1.2% alkaline agarose gel. Following neutralization and staining with SYBR Gold (Invitrogen), alkaline gels were imaged using the Typhoon FLA 7000 (GE Healthcare) and analyzed using ImageQuant TL 8.2 (GE Healthcare). The number of CPD lesions per kb was estimated using the ensemble average pixel density of each lane, corrected by the no enzyme control lane. Percent repair was calculated by normalizing the number of CPDs per kb to the no repair time point. Graphs represent the mean and SEM of at least 3 independent experiments.

Repair analysis of UV induced CPDs in *RPB2* locus

Yeast cells were grown in synthetic dextrose (SD) medium at 30°C to late log phase ($A_{600} \approx 1.0$), irradiated with 120 J/m² of UV-C and incubated in YPD medium at 30°C in the dark. At different times of the repair incubation, aliquots were removed, and the genomic DNA was isolated. To map the induction and repair of UV-induced CPDs at the nucleotide resolution in a specific gene, libraries of DNA fragments adjoining the lesions were created by using the LAF-Seq (Lesion-Adjoining Fragment Sequencing) strategy⁸⁶ with some modifications. Briefly, the isolated genomic DNA was restricted with HincII and NruI to release a 553 bp *RPB2* gene fragment (168 bp upstream and 385 bp downstream of the transcription start site) and incised at the CPDs with T4 endonuclease V and treated *E. coli* endonuclease IV (New England Biolabs). The 3' ends of the restricted and CPD-incised DNA fragments were ligated to Illumina sequencing adapters by using CircLigase (Lucigen). After PCR amplification, the libraries were sequenced by using an Illumina HiSeq platform.

The sequencing reads were aligned to the *RPB2* gene by using Bowtie 2⁷⁷. The numbers of reads from the UV-irradiated samples were normalized to those from the control (unirradiated) samples. Reads corresponding to CPDs at individual sites along the *RPB2* gene fragment were counted after subtraction of the background counts (in the unirradiated samples) by using codes in R. To more directly 'visualize' the CPD induction and repair profiles, images with band intensities corresponding to counts of aligned sequencing reads were created by using codes in R and MATLAB.

C. elegans strains and UV sensitivity assays

C. elegans strains were cultured according to standard methods and outcrossed against Bristol N2, which was used as wild type. Mutant alleles were *xpc-1(tm3886)*, *csb-1(ok2335)*, and *elof-1(emc203)*. The loss of function *elof-1(emc203)* (Suppl. Fig. 6I) mutant strain was generated by injection of Cas9 protein together with tracrRNA and two crRNAs targeting *elof-1* (CAGTTGAATTGGGTGTCGAG and AGACGTCGATTGGCTCGGAG; Integrated DNA Technologies). Deletion animals were selected by genotyping PCR and sequencing. UV survival experiments were performed as described previously⁵². Animals were irradiated at the indicated dose using two Philips TL-12 (40W) tubes emitting UV-B light. Briefly, 'germ cell and embryo UV survival' was determined by allowing UV-irradiated staged young adults to lay eggs on plates for 3 hours. To calculate the survival

percentage, the total number of hatched and unhatched eggs was counted after 24 hours. For the 'L1 larvae UV survival', staged L1 larvae were UV irradiated and grown for 48 hours. Survival percentage was calculated by counting surviving animals that developed beyond the L2 stage and arrested animals as L1/L2 larvae.

Metaphase spreads and chromosomal aberrations

Metaphase spreads were carried out as described previously⁸⁷. Briefly, cells were irradiated with 4 J/m² or mock-treated 48 or 72 hours before preparing metaphase spreads (final confluence of 50-80%). Cells were arrested at metaphase by incubating with colcemid (*N*-methyl-*N*-deacetyl-colchicine, Roche, 10295892001) for the last 14 hours before harvesting the cells. Collected cells were treated with hypotonic solution (KCl 0.075 M) for 30 minutes at 37 °C and fixed with methanol:acetic acid 3:1. Telomere-FISH was further carried out to study chromosomal aberrations. Metaphases were hybridized with telomere-repeat specific peptide nucleic acid (PNA) probes (Applied Biosystems) as described to label telomeres⁸⁸. A minimum 60 metaphase images were obtained using Carl Zeiss Axio Imager D2 microscope using 63x Plan Apo 1.4 NA oil immersion objective and analyzed with ImageJ software for chromosomal aberrations.

DNA fiber analysis

DNA fiber analysis was carried out as described previously^{87,89}. Briefly, cells were sequentially pulse-labeled with 30 μM CldU (c6891, Sigma-Aldrich) and 250 μM IdU (I0050000, European Pharmacopoeia) for 15 min. For assessing fork progression after DNA damage, cells were irradiated with 4 J/m² UV and incubated for 2 hours before pulse-labeling. After labeling, cells were collected and resuspended in PBS at 2.5 × 10⁵ cells per ml. The labeled cells were mixed 1:1 with unlabeled cells, and 2.5 μl of cells was added to 7.5 μl of lysis buffer (200 mM Tris-HCl, pH 7.5, 50 mM EDTA, and 0.5% (w/v) SDS) on a glass slide. After 8 min, the slides were tilted at 15–45°, and the resulting DNA spreads were air dried, fixed in 3:1 methanol/acetic acid overnight at 4 °C. The fibers were denatured with 2.5 M HCl for 1 hour, washed with PBS and blocked with 0.2% Tween-20 in 1% BSA/PBS for 40 min. The newly replicated CldU and IdU tracks were incubated (for 2.5 hours in the dark, at RT with anti-BrdU antibodies recognizing CldU and IdU (Table 2, see below), followed by a 1-hour incubation with secondary antibodies at RT in the dark: anti-mouse Alexa Fluor 488 and anti-rat Cy3 (Table 3, see below). Fibers were visualized and imaged by Carl Zeiss Axio Imager D2 microscope using 63X Plan Apo 1.4 NA oil immersion objective. Data analysis was carried out with ImageJ software. A one-way ANOVA was applied for statistical analysis using the GraphPad Prism Software.

Immunofluorescence

Immunofluorescence was carried out as described previously⁹⁰. Cells were grown on 24-mm glass coverslips and mock-treated or irradiated with 8 J/m² 48, 24 or 6 hours prior to fixation for 15 minutes in PBS with 3.7% FA. Subsequently, cells were permeabilized with 0.1% Triton X-100 in PBS and washed with PBS+ (0.15% BSA and 0.15% glycine in PBS). Cells were incubated for 2 hours at RT with rabbit anti-53BP1 antibody (table 2, see below) in PBS+. Thereafter, cells were washed with PBS+, 0.1% Triton and PBS+ before incubating 2 hours at RT with donkey anti-rabbit Alexa Fluor 594 conjugated antibody (table 3, see below) and DAPI. After washes with PBS+ and 0.1% Triton,

coverslips were mounted with Aqua-Poly/Mount. Images were acquired with a Zeiss LSM700 Axio Imager Z2 upright microscope equipped with a 63x Plan-apochromat 1.4 NA oil immersion lens (Carl Zeiss Micro Imaging Inc.). Number of foci per nucleus was counted by using ImageJ.

References

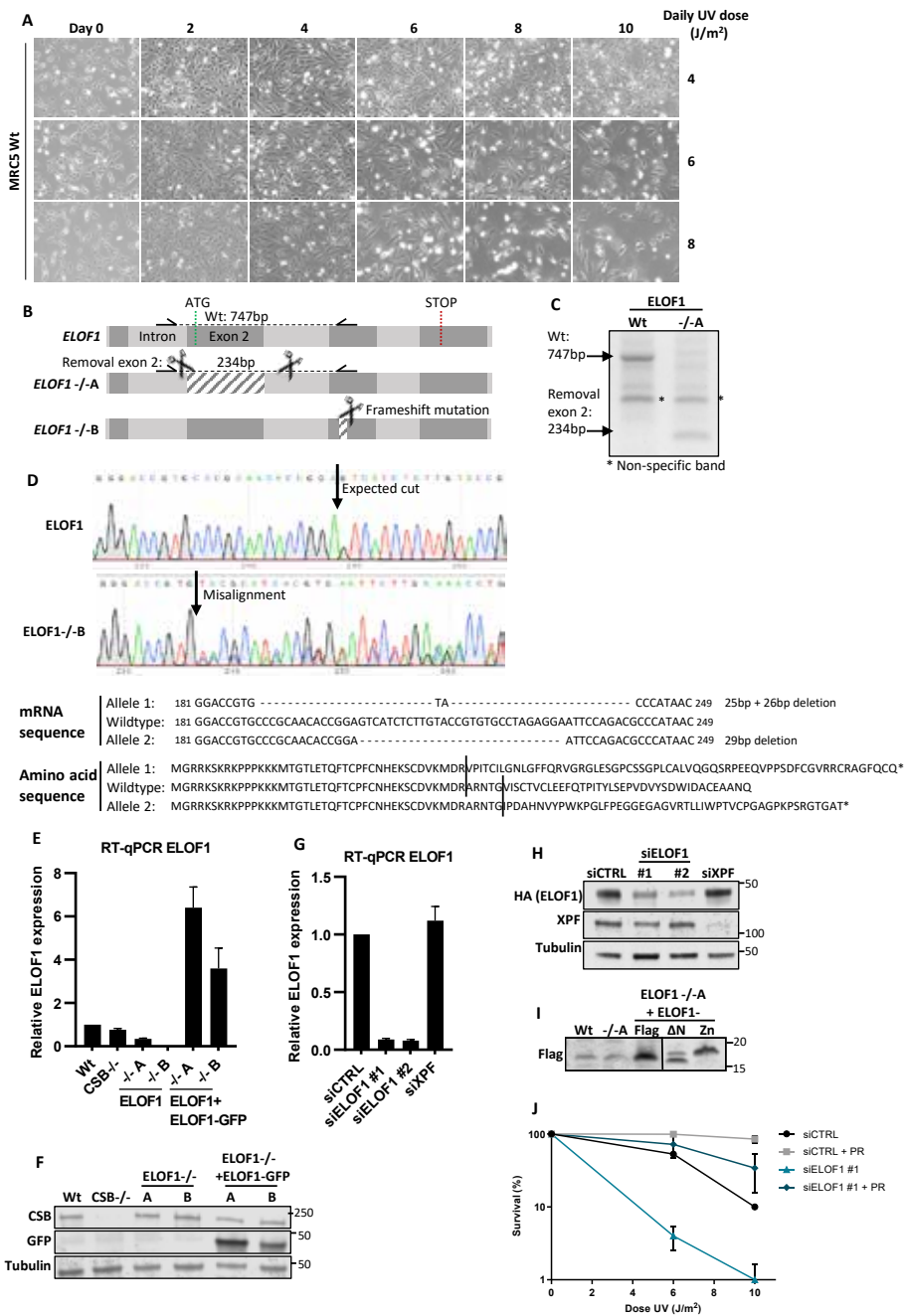
1. Lans, H., Hoeijmakers, J.H.J., Vermeulen, W. & Marteijn, J.A. The DNA damage response to transcription stress. *Nat Rev Mol Cell Biol* 20, 766-784 (2019).
2. Gregersen, L.H. & Svejstrup, J.Q. The Cellular Response to Transcription-Blocking DNA Damage. *Trends Biochem Sci* 43, 327-341 (2018).
3. Doetsch, P.W. Translesion synthesis by RNA polymerases: occurrence and biological implications for transcriptional mutagenesis. *Mutat Res* 510, 131-40 (2002).
4. Marietta, C. & Brooks, P.J. Transcriptional bypass of bulky DNA lesions causes new mutant RNA transcripts in human cells. *EMBO Rep* 8, 388-93 (2007).
5. Hoeijmakers, J.H. DNA damage, aging, and cancer. *N Engl J Med* 361, 1475-85 (2009).
6. Helmrigh, A., Ballarino, M., Nudler, E. & Tora, L. Transcription-replication encounters, consequences and genomic instability. *Nat Struct Mol Biol* 20, 412-8 (2013).
7. Ljungman, M. & Zhang, F. Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis. *Oncogene* 13, 823-31 (1996).
8. Gomez-Gonzalez, B. & Aguilera, A. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes Dev* 33, 1008-1026 (2019).
9. Crossley, M.P., Bocek, M. & Cimprich, K.A. R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell* 73, 398-411 (2019).
10. Gaillard, H. & Aguilera, A. Transcription as a Threat to Genome Integrity. *Annu Rev Biochem* 85, 291-317 (2016).
11. Hanawalt, P.C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol* 9, 958-70 (2008).
12. Laugel, V. Cockayne syndrome: the expanding clinical and mutational spectrum. *Mech Ageing Dev* 134, 161-70 (2013).
13. Xu, J. et al. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* 551, 653-657 (2017).
14. Saijo, M. et al. Functional TFIIH is required for UV-induced translocation of CSA to the nuclear matrix. *Mol Cell Biol* 27, 2538-47 (2007).
15. Kamiuchi, S. et al. Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proc Natl Acad Sci U S A* 99, 201-6 (2002).
16. Sin, Y., Tanaka, K. & Saijo, M. The C-terminal Region and SUMOylation of Cockayne Syndrome Group B Protein Play Critical Roles in Transcription-coupled Nucleotide Excision Repair. *J Biol Chem* 291, 1387-97 (2016).
17. Groisman, R. et al. CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes Dev* 20, 1429-34 (2006).
18. Nakazawa, Y. et al. Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair. *Cell* 180, 1228-1244 e24 (2020).
19. Tufegdžić Vidaković, A. et al. Regulation of the RNAPII Pool Is Integral to the DNA Damage Response. *Cell* 180, 1245-1261 e21 (2020).
20. Okuda, M., Nakazawa, Y., Guo, C., Ogi, T. & Nishimura, Y. Common TFIIH recruitment mechanism in global genome and transcription-coupled repair subpathways. *Nucleic Acids Res* 45, 13043-13055 (2017).
21. Schärer, O.D. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol* 5, a012609 (2013).

22. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* 15, 465-81 (2014).
23. Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11, 783-784 (2014).
24. Evers, B. et al. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* 34, 631-3 (2016).
25. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15, 554 (2014).
26. Yang, W. & Gao, Y. Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu Rev Biochem* 87, 239-261 (2018).
27. Daniels, J.P., Kelly, S., Wickstead, B. & Gull, K. Identification of a crenarchaeal orthologue of Elf1: implications for chromatin and transcription in Archaea. *Biol Direct* 4, 24 (2009).
28. Prather, D., Krogan, N.J., Emili, A., Greenblatt, J.F. & Winston, F. Identification and characterization of Elf1, a conserved transcription elongation factor in *Saccharomyces cerevisiae*. *Mol Cell Biol* 25, 10122-35 (2005).
29. Joo, Y.J., Ficarro, S.B., Chun, Y., Marto, J.A. & Buratowski, S. In vitro analysis of RNA polymerase II elongation complex dynamics. *Genes Dev* 33, 578-589 (2019).
30. Ehara, H. et al. Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* 357, 921-924 (2017).
31. Mayer, A. et al. Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* 17, 1272-8 (2010).
32. Ehara, H. et al. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* 363, 744-747 (2019).
33. Steurer, B. et al. Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. *Proc Natl Acad Sci U S A* 115, E4368-E4376 (2018).
34. Nilson, K.A. et al. THZ1 Reveals Roles for Cdk7 in Co-transcriptional Capping and Pausing. *Mol Cell* 59, 576-87 (2015).
35. Chao, S.H. et al. Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. *J Biol Chem* 275, 28345-8 (2000).
36. Sobell, H.M. Actinomycin and DNA transcription. *Proc Natl Acad Sci U S A* 82, 5328-31 (1985).
37. Van Oss, S.B., Cucinotta, C.E. & Arndt, K.M. Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. *Trends Biochem Sci* 42, 788-798 (2017).
38. Chen, F.X., Smith, E.R. & Shilatifard, A. Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* 19, 464-478 (2018).
39. van den Boom, V. et al. DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J Cell Biol* 166, 27-36 (2004).
40. Gregersen, L.H., Mitter, R. & Svejstrup, J.Q. Using TTchem-seq for profiling nascent transcription and measuring transcript elongation. *Nat Protoc* 15, 604-627 (2020).
41. Nudler, E. RNA polymerase backtracking in gene regulation and genome instability. *Cell* 149, 1438-45 (2012).
42. Jia, N. et al. A rapid, comprehensive system for assaying DNA repair activity and cytotoxic effects of DNA-damaging reagents. *Nat Protoc* 10, 12-24 (2015).
43. Wienholz, F., Vermeulen, W. & Marteijn, J.A. Amplification of unscheduled DNA synthesis signal enables fluorescence-based single cell quantification of transcription-coupled nucleotide excision repair. *Nucleic Acids Res* 45, e68 (2017).

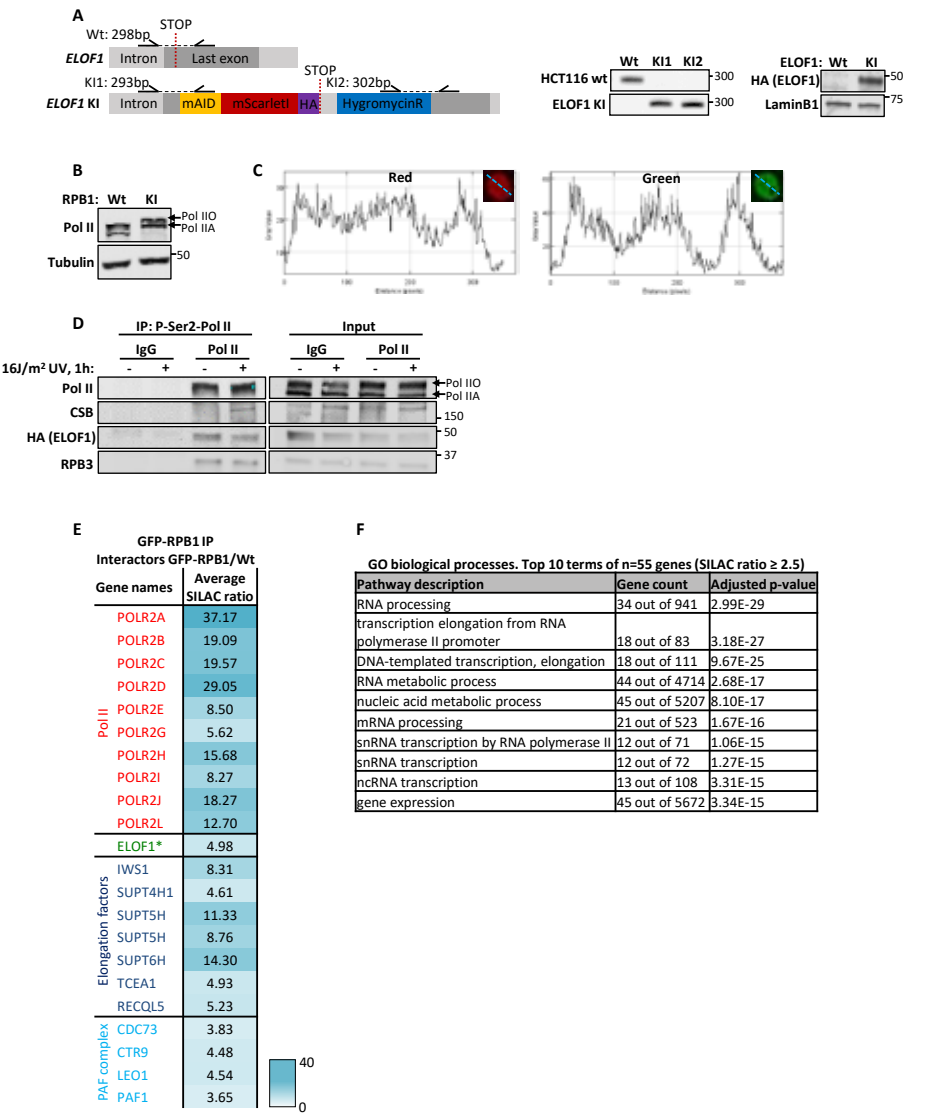
44. Jaspers, N.G. et al. Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. *DNA Repair (Amst)* 1, 1027-38 (2002).
45. Slyskova, J. et al. Base and nucleotide excision repair facilitate resolution of platinum drugs-induced transcription blockage. *Nucleic Acids Res* 46, 9537-9549 (2018).
46. Veloso, A. et al. Genome-wide transcriptional effects of the anti-cancer agent camptothecin. *PLoS One* 8, e78190 (2013).
47. Brooks, P.J. et al. The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *J Biol Chem* 275, 22355-62 (2000).
48. van Gool, A.J. et al. RAD26, the functional *S. cerevisiae* homolog of the Cockayne syndrome B gene ERCC6. *EMBO J* 13, 5361-9 (1994).
49. Mao, P., Smerdon, M.J., Roberts, S.A. & Wyrick, J.J. Chromosomal landscape of UV damage formation and repair at single-nucleotide resolution. *Proc Natl Acad Sci U S A* 113, 9057-62 (2016).
50. Tijsterman, M., Verhage, R.A., van de Putte, P., Tasseront-de Jong, J.G. & Brouwer, J. Transitions in the coupling of transcription and nucleotide excision repair within RNA polymerase II-transcribed genes of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 94, 8027-32 (1997).
51. Li, S. & Smerdon, M.J. Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO J* 21, 5921-9 (2002).
52. Lans, H. et al. Involvement of global genome repair, transcription coupled repair, and chromatin remodeling in UV DNA damage response changes during development. *PLoS Genet* 6, e1000941 (2010).
53. Geverts, B., van Royen, M.E. & Houtsmuller, A.B. Analysis of biomolecular dynamics by FRAP and computer simulation. *Methods Mol Biol* 1251, 109-33 (2015).
54. Williamson, L. et al. UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. *Cell* 168, 843-855 e13 (2017).
55. Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. *Science* 315, 859-62 (2007).
56. van der Weegen, Y. et al. The sequential and cooperative action of CSB, CSA and UVSSA targets the TFIIH complex to DNA damage-stalled RNA polymerase II. *bioRxiv*, 707216 (2019).
57. Schwertman, P. et al. UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat Genet* 44, 598-602 (2012).
58. Zhang, X. et al. Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nat Genet* 44, 593-7 (2012).
59. He, J., Zhu, Q., Wani, G., Sharma, N. & Wani, A.A. Valosin-containing Protein (VCP)/p97 Segregase Mediates Proteolytic Processing of Cockayne Syndrome Group B (CSB) in Damaged Chromatin. *J Biol Chem* 291, 7396-408 (2016).
60. Lukas, C. et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat Cell Biol* 13, 243-53 (2011).
61. Tellier, A.P., Archambault, D., Tremblay, K.D. & Mager, J. The elongation factor Elof1 is required for mammalian gastrulation. *PLoS One* 14, e0219410 (2019).
62. Fei, J. & Chen, J. KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). *J Biol Chem* 287, 35118-26 (2012).
63. Wienholz, F. et al. FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res* (2019).

64. Poli, J. et al. Mec1, INO80, and the PAF1 complex cooperate to limit transcription replication conflicts through RNAPII removal during replication stress. *Genes Dev* 30, 337-54 (2016).
65. Campeau, E. et al. A Versatile Viral System for Expression and Depletion of Proteins in Mammalian Cells. *PLOS ONE* 4, e6529 (2009).
66. Yesbolatova, A., Natsume, T., Hayashi, K.-i. & Kanemaki, M.T. Generation of conditional auxin-inducible degron (AID) cells and tight control of degron-fused proteins using the degradation inhibitor auxinole. *Methods* 164-165, 73-80 (2019).
67. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M.T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* 15, 210-218 (2016).
68. Steurer, B. et al. Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. *Proceedings of the National Academy of Sciences* 115, E4368-E4376 (2018).
69. Steurer, B. et al. Fluorescently-labelled CPD and 6-4PP photolyases: new tools for live-cell DNA damage quantification and laser-assisted repair. *Nucleic Acids Research* 47, 3536-3549 (2019).
70. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research* 42, e168-e168 (2014).
71. Shalem, O. et al. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* 343, 84-87 (2014).
72. Livak, K.J. & Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods* 25, 402-408 (2001).
73. Wienholz, F. et al. FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res* 47, 4011-4025 (2019).
74. Tufegdžić Vidaković, A. et al. Regulation of the RNAPII Pool Is Integral to the DNA Damage Response. *Cell* 180, 1245-1261.e21 (2020).
75. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44, W160-5 (2016).
76. Gardner, J.M. & Jaspersen, S.L. Manipulating the yeast genome: deletion, mutation, and tagging by PCR. *Methods Mol Biol* 1205, 45-78 (2014).
77. Brachmann, C.B. et al. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14, 115-32 (1998).
78. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-9 (2012).
79. Mao, P., Smerdon, M.J., Roberts, S.A. & Wyrick, J.J. Asymmetric repair of UV damage in nucleosomes imposes a DNA strand polarity on somatic mutations in skin cancer. *Genome Res* 30, 12-21 (2020).
80. Park, D., Morris, A.R., Battenhouse, A. & Iyer, V.R. Simultaneous mapping of transcript ends at single-nucleotide resolution and identification of widespread promoter-associated non-coding RNA governed by TATA elements. *Nucleic acids research* 42, 3736-3749 (2014).
81. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences* 95, 14863 (1998).
82. Saldanha, A.J. Java Treeview--extensible visualization of microarray data. *Bioinformatics* 20, 3246-8 (2004).
83. Holstege, F.C. et al. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717-28 (1998).
84. Mao, P. et al. Genome-wide maps of alkylation damage, repair, and mutagenesis in yeast reveal mechanisms of mutational heterogeneity. *Genome Res* 27, 1674-1684 (2017).
85. Weiner, A. et al. High-Resolution Chromatin Dynamics during a Yeast Stress Response. *Molecular Cell* 58, 371-386 (2015).

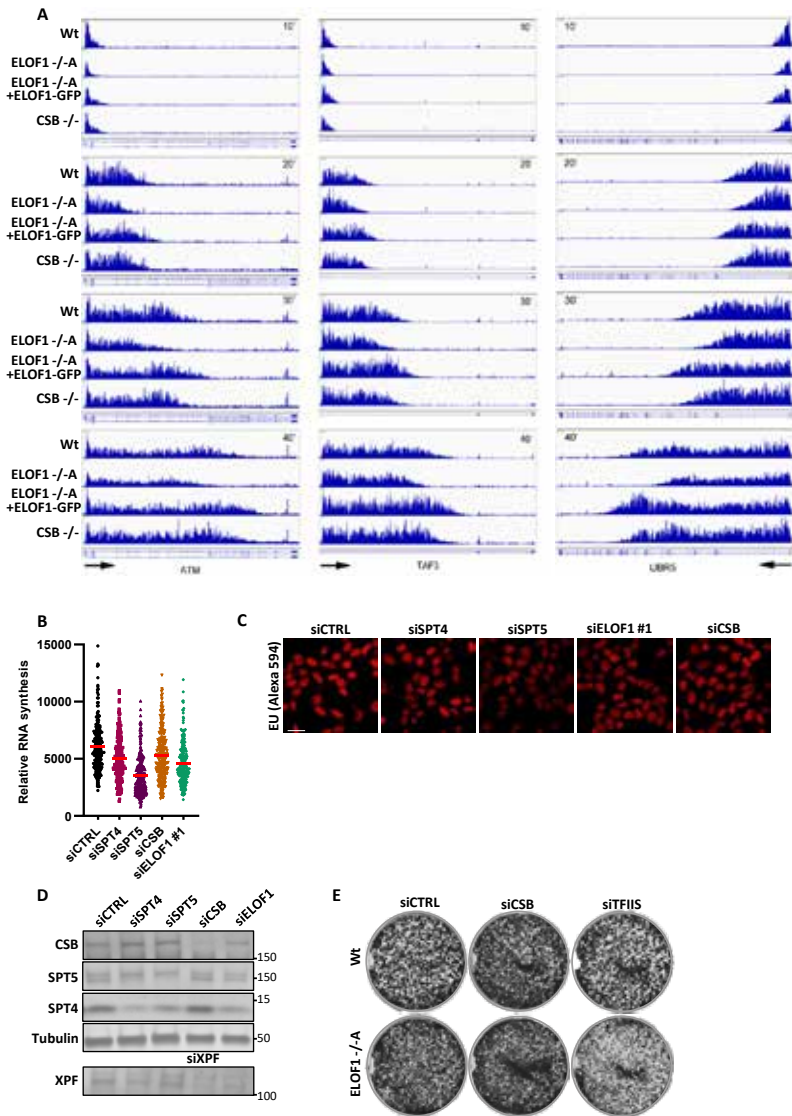
86. Hodges, A.J., Plummer, D.A. & Wyrick, J.J. NuA4 acetyltransferase is required for efficient nucleotide excision repair in yeast. *DNA Repair (Amst)* 73, 91-98 (2019).
87. Li, M., Ko, T. & Li, S. High-resolution Digital Mapping of N-Methylpurines in Human Cells Reveals Modulation of Their Induction and Repair by Nearest-neighbor Nucleotides. *J Biol Chem* 290, 23148-61 (2015).
88. Mukherjee, C. et al. RIF1 promotes replication fork protection and efficient restart to maintain genome stability. *Nature Communications* 10, 3287 (2019).
89. Callen, E. et al. ATM prevents the persistence and propagation of chromosome breaks in lymphocytes. *Cell* 130, 63-75 (2007).
90. Cornacchia, D. et al. Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *Embo J* 31, 3678-90 (2012).
91. van Cuijk, L. et al. SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair. *Nat Commun* 6, 7499 (2015).



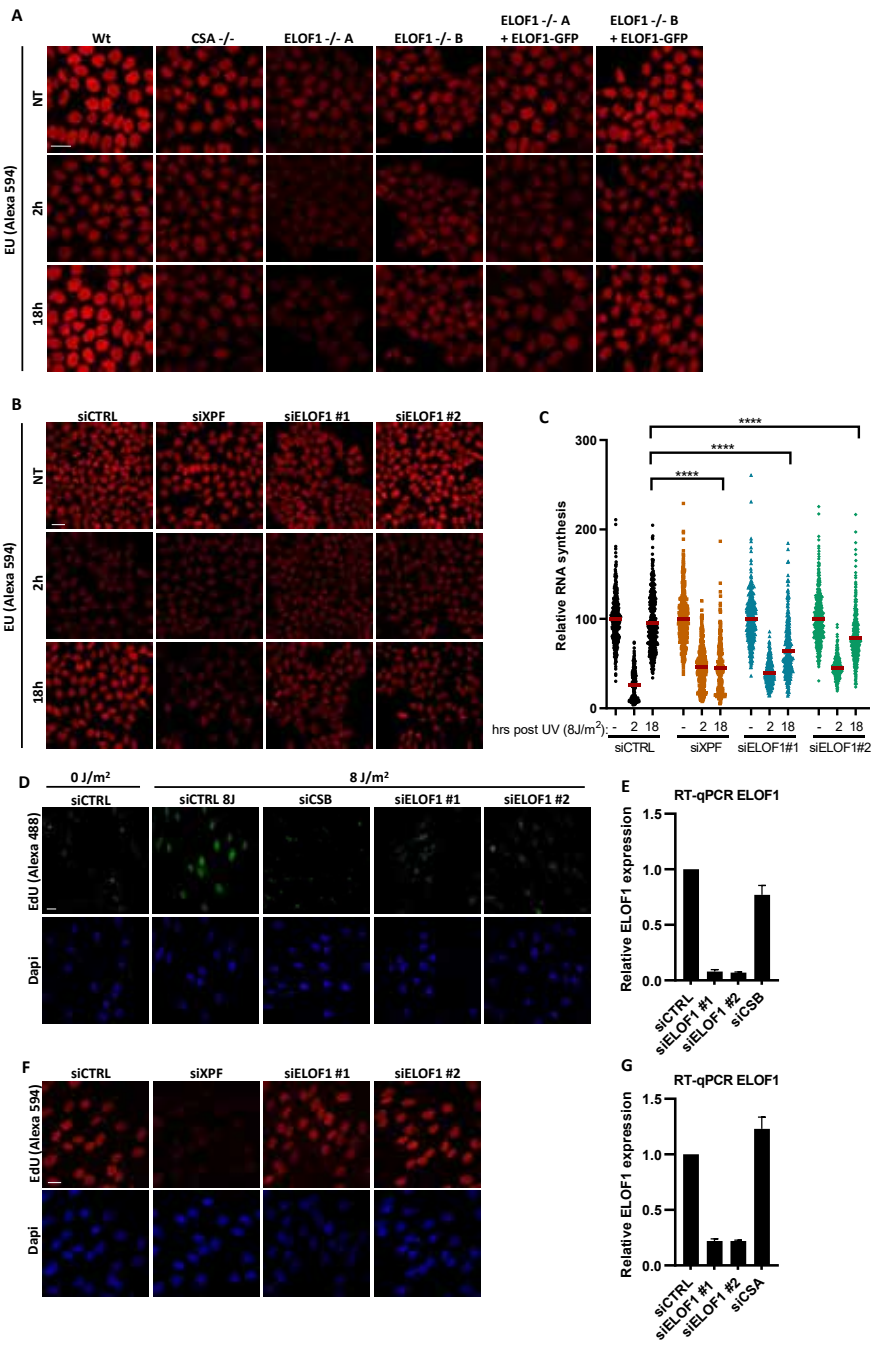
◀ **Supplemental figure 1.** (A) Brightfield images of MRC-5 cells irradiated with indicated doses of UV-C for 10 consecutive days. Images were taken every other day. (B) Schematic of the genomic *ELOF1* locus. Scissors indicate target regions of the sgRNAs used to generate *ELOF1* KO (-/-) cells, half arrows indicate primers used for genotyping as shown in (C). (C+D) Genotyping of *ELOF1* KO (-/-) cells, both originating from a single cell clone. (C) Genotyping PCR of loss of exon 2 in *ELOF1* -/-A cells. (D) Top panel: Sequencing results showing frameshift mutations in the targeted genomic locus of *ELOF1* -/-B. Bottom panel: Amino acid sequence of ELOF1 in *ELOF1* -/-B cells. (E) Relative ELOF1 levels in indicated HCT116 Wt and *ELOF1* KO (-/-) cells, with ELOF1 re-expression where indicated, as determined by RT-qPCR. Relative ELOF1 mRNA expression was normalized to GAPDH signal and levels in Wt cells were set to 1. Error bars indicate SEM. (F) Immunoblot of indicated HCT116 cell lines showing CSB or ELOF1-GFP expression. Tubulin was used as loading control. (G) Relative ELOF1 levels in HCT116 cells transfected with indicated siRNAs as determined by RT-qPCR. Relative ELOF1 expression was normalized to GAPDH signal and siCTRL levels were set to 1. Error bars indicate SEM. (H) Immunoblot showing endogenous ELOF1 and XPF levels in *ELOF1-mScarlet1-HA* KI cells (suppl. fig. 2A) transfected with indicated siRNAs. Tubulin was used as loading control. (I) Immunoblot showing expression of Flag-tagged Wt or indicated ELOF1 mutants in HCT116 *ELOF1* -/-A cells. (J) Relative colony survival of CPD photolyase cells transfected with indicated siRNAs. PR indicates CPD removal by photoreactivation. Plotted curves represent averages of 2 independent experiments \pm SEM.



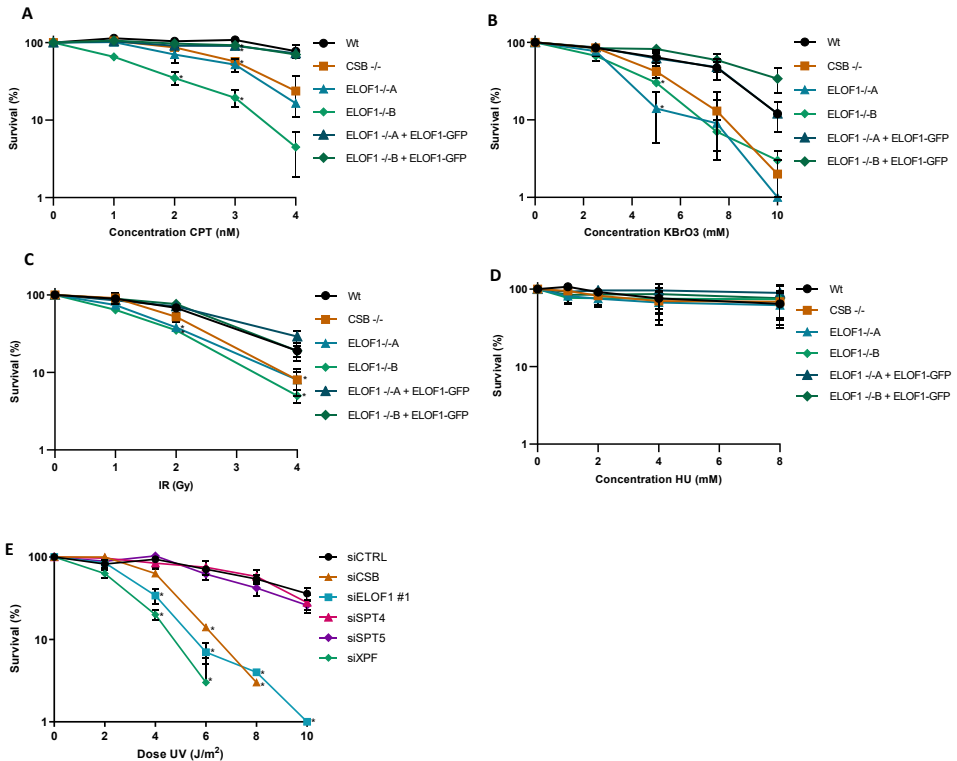
Supplemental figure 2. (A) Left panel: Schematic of the genomic locus of *ELOF1* for generating *ELOF1-mScarlet1*-HA KI cell line. Half arrows indicate primer locations. Middle and right panel: Genotyping PCR and immunoblot for *ELOF1*-KI cell line. LaminB1 was used as loading control. **(B)** Immunoblot of HCT116 *GFP-RPB1* KI. Tubulin was used as loading control. **(C)** Histograms showing intensities of GFP and mScarlet1 measured over the indicated dotted line in HCT116 double KI cells. **(D)** Native immunoprecipitation of P-Ser2-modified Pol II in HCT116 cells followed by immunoblotting for indicated proteins. Cells were harvested 1 hour after mock treated or irradiation with 16 J/m² UV-C. IgG was used as binding control. **(E)** Interaction heat map based on the SILAC ratios of MRC-5 *GFP-RPB1*-interacting proteins as determined by quantitative interaction proteomics. Average SILAC ratios of duplicate experiments are plotted and represent *RPB1*-interactors relative to empty beads. SILAC ratio >1 indicates increase in interaction. * indicates proteins quantified in one experiment. **(F)** Top 10 enriched GO terms (biological processes) identified using g:Profiler of 55 proteins identified as *ELOF1* interactor with an average SILAC ratio of 2.5 or higher.



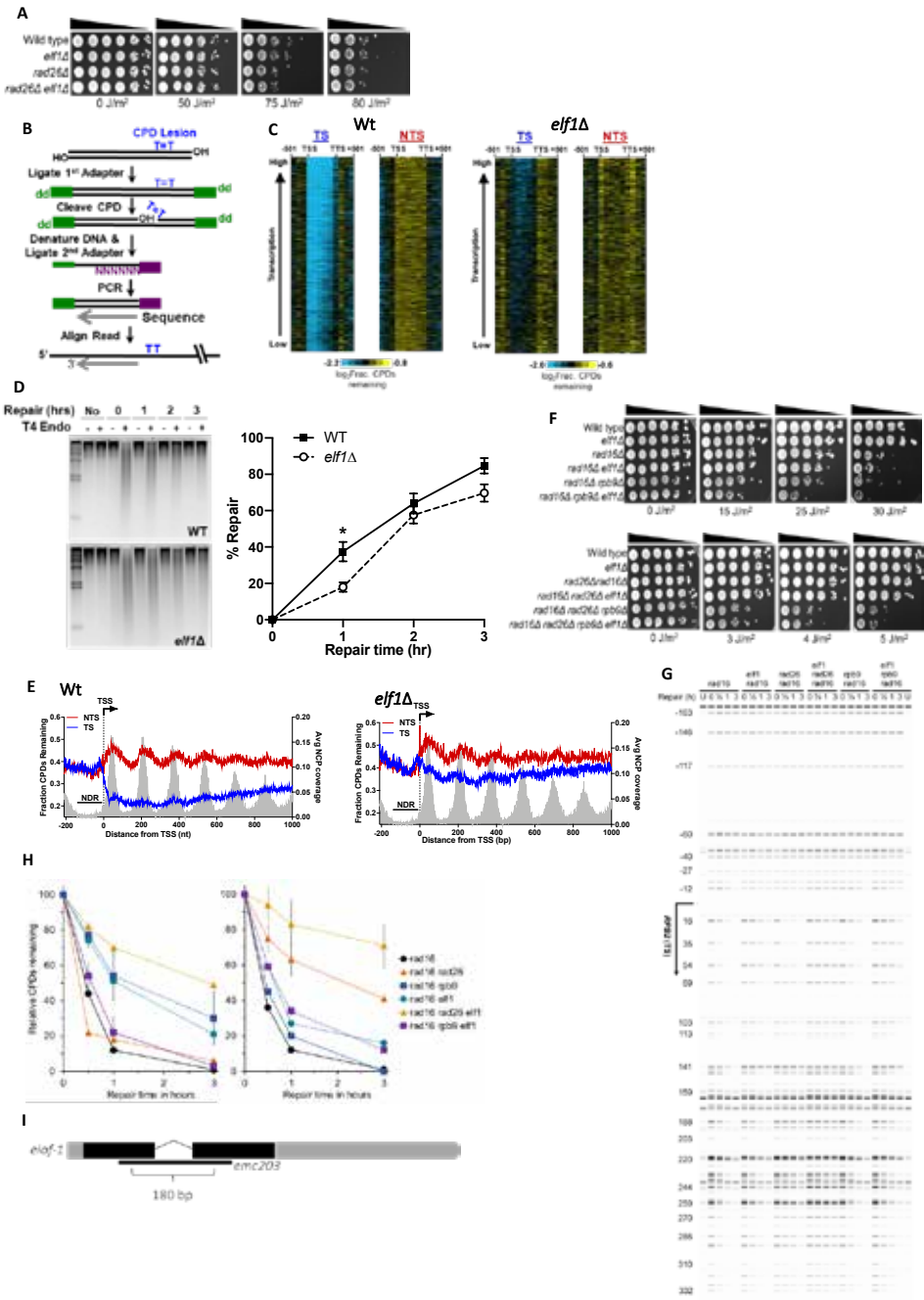
Supplemental figure 3. (A) Browser tracks from DRB/TT_{chem}-seq experiment at *ATM*, *TAF3* and *UBR5*. Results are shown 10, 20, 30 or 40 minutes after DRB release. (B) Transcription levels as determined by relative EU incorporation in HCT116 cells transfected with indicated siRNAs. Red lines indicate average integrated density \pm SEM. $n \geq 200$ cells from two independent experiments. (C) Representative images of EU incorporation in HCT116 cells transfected with indicated siRNAs. Scale bar: 20 μ m. (D) Immunoblot for indicated proteins in HCT116 cells transfected with indicated siRNAs. Tubulin was used as loading control. (E) Images of HCT116 Wt and ELOF1 -/-A cells transfected with indicated siRNAs, stained with coomassie blue 10 days after transfection.



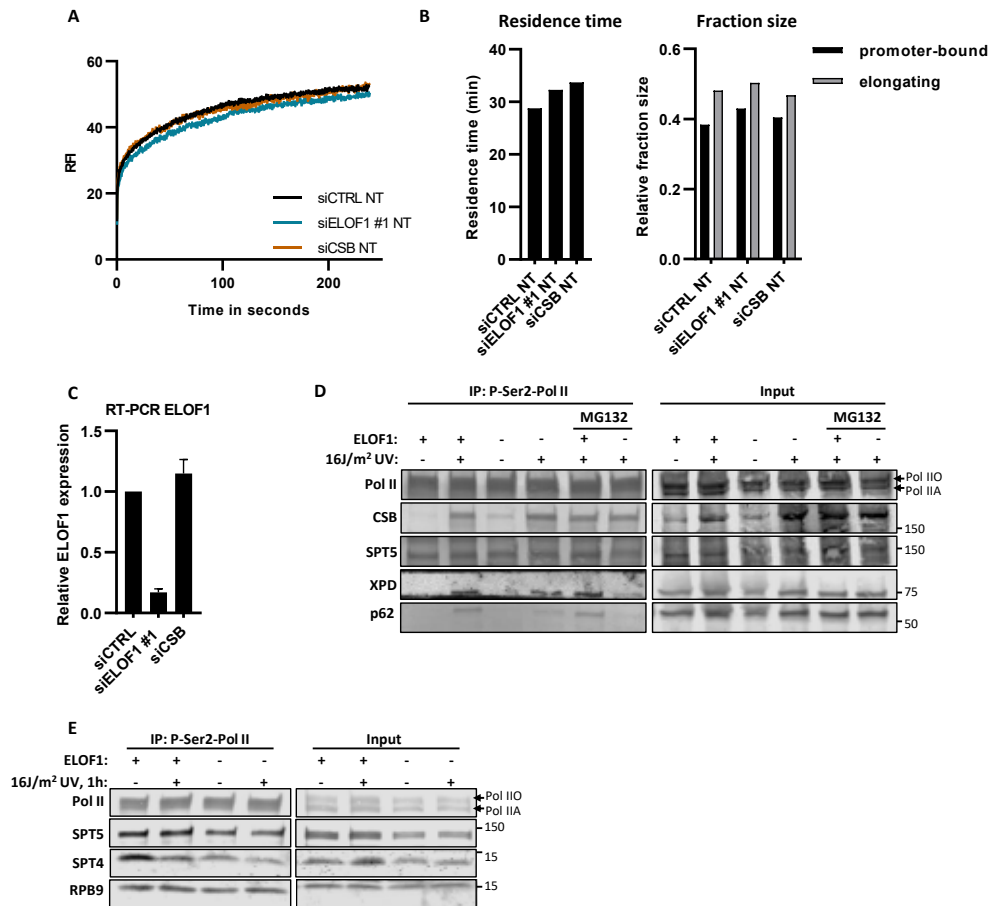
Supplemental figure 4. (A+B) Representative immunofluorescence images of EU incorporation in **(A)** indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, or **(B)** HCT116 cells transfected with indicated siRNAs, 2 or 18 hours after 8 J/m² UV-C or mock treatment (NT). Scale bar: 20 μ m. **(C)** Transcription restart after UV damage as determined by relative EU incorporation in HCT116 cells transfected with indicated siRNAs, 2 or 18 hours after 8 J/m² UV-C or mock treatment (NT). Relative integrated density of UV-irradiated samples are normalized to mock-treated and set to 100%. Red lines indicate average integrated density \pm SEM. $n \geq 300$ cells from three independent experiments. **(D)** Representative immunofluorescence images of amplified EdU signal in XP186LV fibroblasts (XP-C) transfected with indicated siRNAs, 7 hours after exposure to 8 J/m² UV-C. Scalebar: 20 μ m. **(E)** Relative ELOF1 mRNA levels in XP186LV fibroblasts (XP-C) following transfection with indicated siRNAs as determined by RT-qPCR. ELOF1 expression was normalized to GAPDH expression and siCTRL levels were set to 1. Error bars indicate SEM. **(F)** Representative fluorescence images of EdU incorporation 3 hours after irradiation with 16 J/m² UV-C in C5RO (hTert) cells transfected with indicated siRNAs. Scale bar: 20 μ m. **(G)** Relative ELOF1 mRNA levels in C5RO (hTert) cells following transfection with indicated siRNAs as determined by RT-qPCR. ELOF1 expression was normalized to GAPDH expression and siCTRL levels were set to 1. Error bars indicate SEM. **** $p \leq 0.0001$.



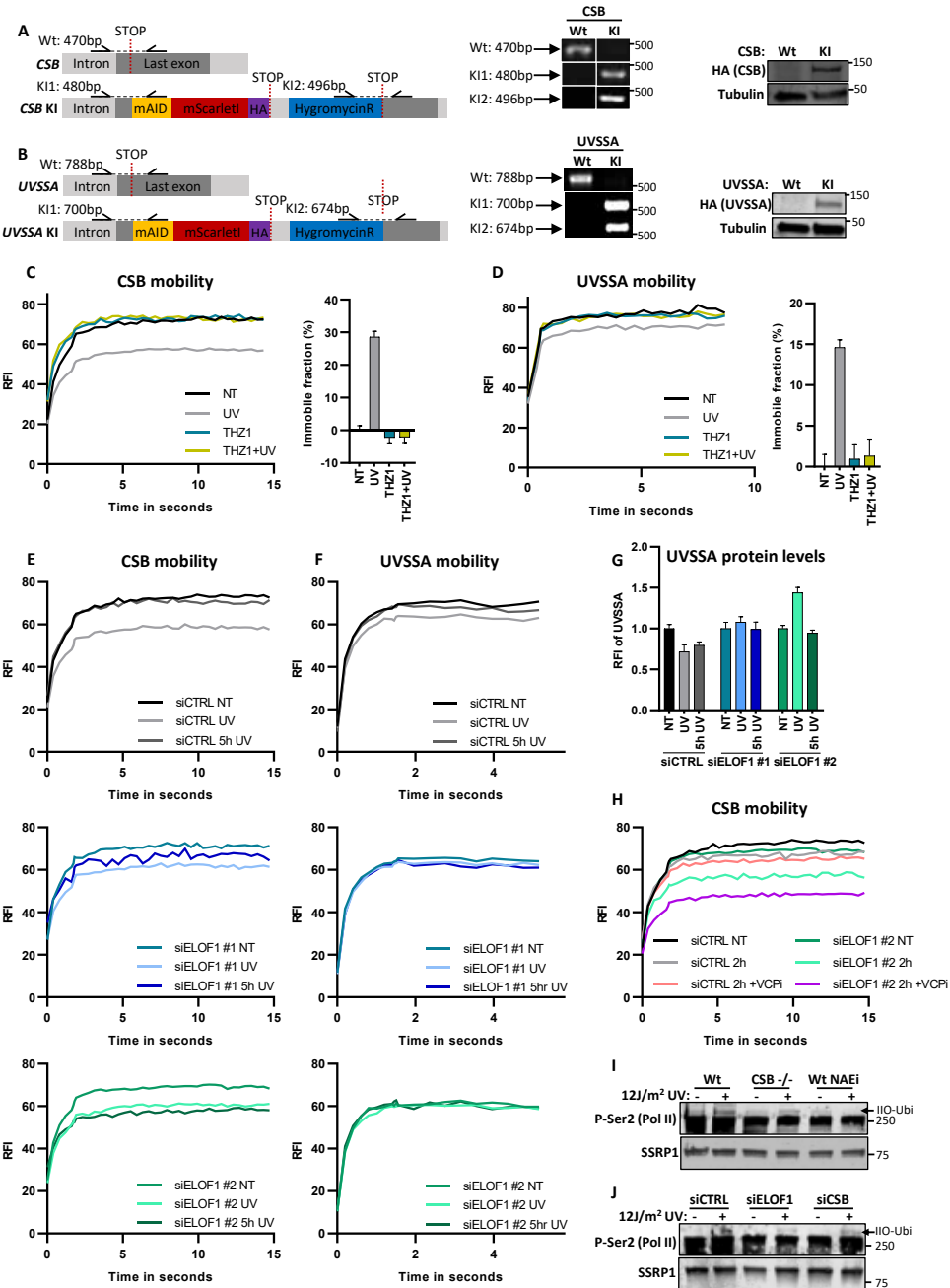
Supplemental figure 5. (A-D) Relative colony survival of indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, continuously exposed to indicated concentrations of **(A)** camptothecin (CPT) or **(B)** potassium bromate (KBrO₃), or irradiated with indicated doses of **(C)** ionizing radiation (IR), or exposed **(D)** to indicated concentrations of hydroxyurea (HU). Plotted curves represent averages of at least two independent experiments \pm SEM. **(E)** Relative colony survival of HCT116 cells transfected with indicated siRNAs following exposure to indicated doses of UV-C. Plotted curves represent averages of three independent experiments \pm SEM. * $P \leq 0.05$



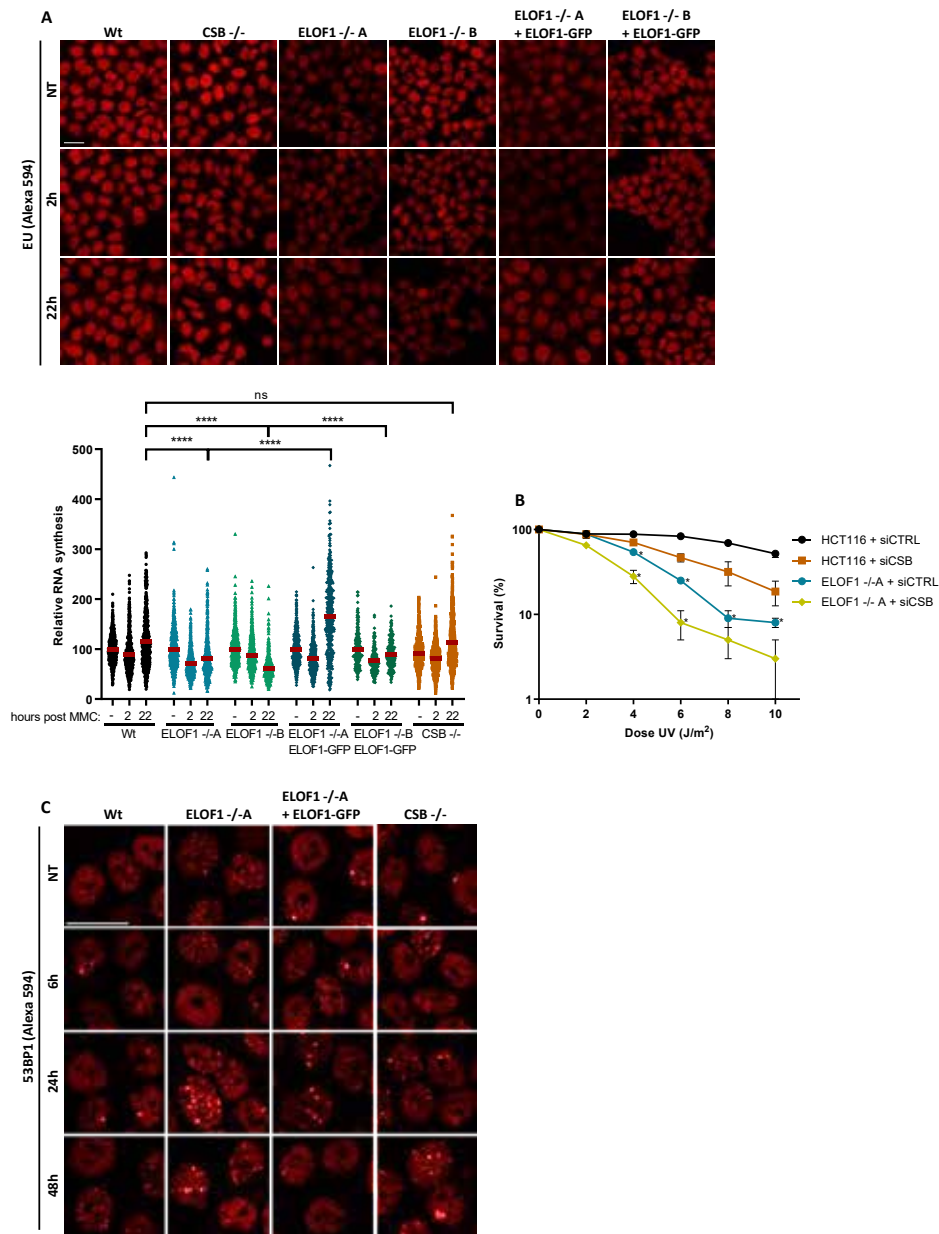
◀ **Supplemental figure 6.** (A) Indicated mutant yeast strains were serially 10-fold diluted, spotted, and exposed to indicated UV-C doses. (B) Schematic showing the CPD-seq method. Isolated DNA is sonicated and adaptors are ligated. Subsequently, CPDs are cleaved by T4 endonuclease V and APE1 nuclease to generate 3' ends. Following denaturing of the DNA, the ends are ligated to a second adaptor that allows sequencing of CPDs. (C) Gene plot analysis of CPD-seq data following 2-hour repair for ~4500 yeast genes, ordered by transcription frequency. Plots depict fraction of unrepaired CPDs following 2-hour repair relative to no repair for both the transcribed strand (TS) and non-transcribed strand (NTS) for gene coding regions, regions upstream of the transcription start site (TSS), and downstream of the transcription termination site (TTS). Each row represents approximately 10 genes to display the plot in a compact manner. (D) Analysis of bulk repair of UV-induced CPD lesions in Wt and *elf1Δ* mutant yeast. The repair of CPD lesions at various time points was measured by T4 endonuclease V digestion and alkaline gel electrophoresis of genomic DNA isolated from UV-irradiated yeast (100 J/m² UV-C light). A representative gel is shown on the left. The right panel depicts the quantification of CPD repair at each time point from at least three independent experiments ±SEM. **P*≤0.05. (E) Single nucleotide resolution analysis of CPD-seq data downstream of the TTS of ~4500 yeast genes. Plots depict fraction of unrepaired CPDs following 2-hour repair relative to no repair for both TS and NTS. Nucleosome positioning data is shown for reference. (F) Controls for UV spotting assays shown in Fig. 3I. (G) Image showing repair of CPDs in the TS of the *RPB2* gene for indicated yeast strains. The image was generated by converting counts of sequencing reads aligned to the sites of the RPB2 fragment into bands. 'U' indicates samples from unirradiated cells. Nucleotide positions relative to the TSS (+1) of the *RPB2* gene are indicated on the left. (H) Left panel: Relative percentage of CPDs remaining in the short region (within 54 bp) immediately downstream of the transcription start site of the *RPB2* gene. Right panel: Relative percentage of CPDs remaining in the more downstream region (from 69 to 353 bp) of the *RPB2* gene. Error bars (S.D.) are shown only for most relevant strains for clarity. (I) Schematic representation of the *C. elegans eloF-1* genomic organization, depicting the 180 bp *emc203* deletion allele generated with CRISPR/Cas9. Shaded boxes represent exons with coding sequences shown in black.



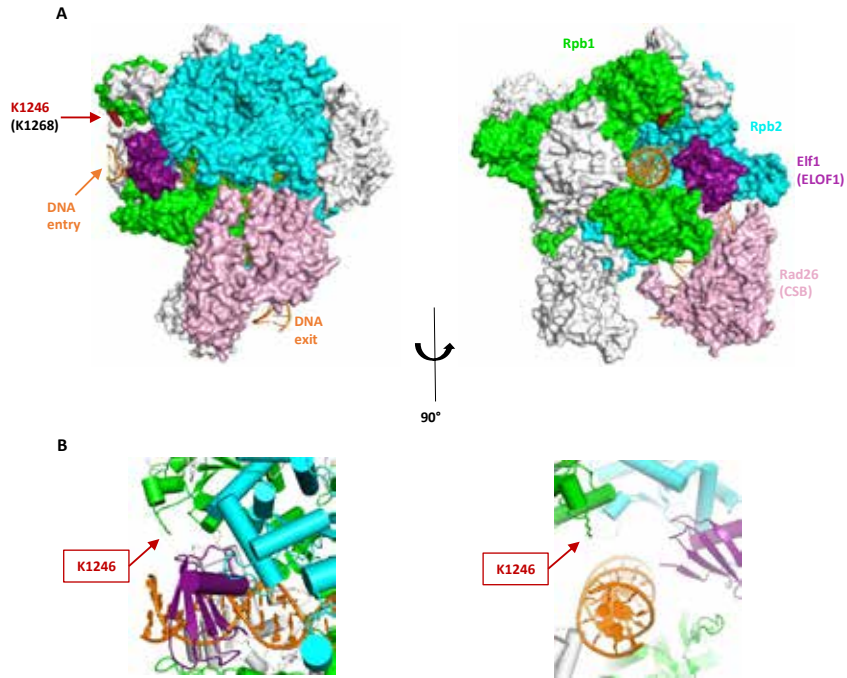
◀ **Supplemental figure 7. (A)** FRAP analysis of GFP-RPB1 mobility after depletion of indicated factors. Mock-treated curves corresponding to figure 4A. $n \geq 14$ cells. **(B)** Left panel: Residence time of elongating Pol II or right panel: relative fraction size of promoter-bound or elongating Pol II as determined by Monte-Carlo-based modeling of RPB1 mobility as shown in (A). **(C)** Relative ELOF1 mRNA levels in *GFP-RPB1* KI cells transfected with indicated siRNAs as determined by RT-qPCR. ELOF1 expression was normalized to GAPDH signal and levels of control cells were set to 1. Error bars indicate SEM. **(D)** Native immunoprecipitation of Pol II in Wt and ELOF $-/-$ A cells followed by immunoblotting for indicated proteins. Cells were harvested 1 hour after mock treatment or irradiation with 16 J/m^2 UV-C. MG132: treatment with $50 \mu\text{M}$ proteasome inhibitor MG132, 1 hour before UV irradiation. **(E)** Native immunoprecipitation of Pol II in Wt and ELOF $-/-$ A cells followed by immunoblotting for indicated proteins. Cells were harvested 1 hour after mock treatment or irradiation with 16 J/m^2 UV-C.



◀ **Supplemental figure 8.** (A) Left panel: Schematic of the genomic locus of *CSB* and used strategy for generating the homozygous *CSB-mScarletl-HA* KI cell line. Half arrows indicate primer locations. Middle and right panel: Genotyping PCR and immunoblot for *CSB*-KI cell line. (B) Left panel: Schematic of the genomic locus of *UVSSA* and used strategy for generating the homozygous *UVSSA- mScarletl-HA* KI cell line. Half arrows indicate primer locations. Middle and right panel: Genotyping PCR and immunoblot for *UVSSA*-KI cell line. (C) Left panel: CSB mobility was determined by FRAP analysis of *CSB-mScarletl* after the indicated treatments. THZ1: 1 hour treatment (2 μ M) before UV-C irradiation (4 J/m²) or mock treatment. Right panel: Relative immobile fraction of CSB as determined by FRAP analysis. Plotted values represent mean \pm SEM and are normalized to mock treated. $n \geq 15$ cells. (D) Same as C but for *UVSSA-mScarletl*. $n \geq 10$ cells. (E+F) FRAP analyses of *CSB-mScarletl* (E) or *UVSSA-mScarletl* (F) mobility after transfection with indicated siRNAs in individual graphs. Cells were mock treated (NT) or analyzed directly (UV) or 5 hours (5hr UV) after irradiation with 4 J/m² UV-C. (G) Relative fluorescence intensity of *UVSSA* in *UVSSA*-KI cells transfected with indicated siRNAs as determined by live-cell imaging. Plotted values represent mean \pm SEM. $n \geq 16$ cells. (H) FRAP analysis of CSB in *CSB*-KI cells transfected with indicated siRNAs 2 hours after UV. VCPi: VCP inhibitor (5 μ M) was directly added after UV-C (4 J/m²). (I) Immunoblot of chromatin fraction of indicated cell lines 1 hour after 12 J/m² UV-C or mock treatment. NAEi = 1 hour treatment with NEDDylation inhibitor (10 μ M). SSRP1 is shown as loading control. (J) Immunoblot of chromatin fraction of HCT116 cells transfected with indicated siRNAs 1 hour after 12 J/m² UV-C or mock treatment. SSRP1 is shown as loading control.



◀ **Supplemental figure 9. (A)** Top panel: Representative immunofluorescence images of EU incorporation in indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, 2 or 22 hours after a 2-hour exposure to 10 $\mu\text{g}/\text{ml}$ mitomycin C or mock treatment. Scale bar: 20 μm . Bottom panel: Transcription restart after mitomycin C as determined by relative EU incorporation in the indicated HCT116 cells. Mitomycin C-treated samples are normalized to mock treated levels and set to 100%. Red lines indicate average integrated density \pm SEM. $n \geq 300$ cells from four independent experiments. **(B)** Relative colony survival of indicated cell lines with siRNA transfection following exposure to indicated doses of UV-C. Plotted curves represent averages of three independent experiments \pm SEM. **(C)** Representative immunofluorescence images of 53BP1 foci in indicated HCT116 Wt and KO (-/-) cells, with ELOF1 re-expression where indicated, 6, 24 or 48 hours after exposure to 8 J/m^2 UV-C or mock treatment. Scale bar: 20 μm . * $p \leq 0.05$, *** $p \leq 0.0001$.



Supplemental figure 10. (A) *S.cerevisiae* Pol II (5vvr.pdb) with Rpb1 in green, Rpb2 in cyan, DNA in orange and Rad26 (CSB) in pink. The *P.pastoris* Pol II in complex with elongation factors (5xog.pdb) was superimposed onto this structure (Rpb1 subunits aligned onto each other), and all subunits except Elf1 (ELOF1; purple) were omitted for clarity. Conserved lysine K1246 (K1268 in mammalian Pol II) is indicated in dark red. **(B)** Close up of Elf1 (ELOF1) binding region.

Supplemental table 1. Table showing negatively regulated genes from the CRISPR/cas9 screen resulting from MaGeCK analysis of the change in abundance of sgRNAs in UV-treated over mock-treated samples. Experiment was performed in duplicate.

<https://1drv.ms/x/s!AvEm1Jm7uzqFnnZDvTykqJWZQwla?e=p6Yi3E>

Supplemental table 2. Table with SILAC ratios and peptide numbers as determined using quantitative interaction proteomics. Each tab represents a different experiment as indicated.

<https://1drv.ms/x/s!AvEm1Jm7uzqFnnX9wYHxxC9rEOfV?e=h18Zoc>



4

DNA damage-induced replication stress results in PA200-proteasome mediated degradation of acetylated histones

I.K. Mandemaker^{1,#}, M.E. Geijer¹, I. Kik¹, K. Bezstarosti²,
E. Rijkers², A. Raams¹, R.C Janssens¹, H. Lans¹,
J.H. Hoeijmakers¹ J.A. Demmers², W. Vermeulen¹ and
J.A. Marteijn¹

¹ Department of Molecular Genetics, Onco Institute, Erasmus University Medical Center, Rotterdam, The Netherlands

² Proteomics Center, Erasmus University Medical Center, Rotterdam, The Netherlands

Current address: Biomedical Center Munich, Faculty of Medicine, Ludwig-Maximilians-Universität München, Germany

Published in EMBO reports, 2018; 19 (10)

Abstract

Histone acetylation influences protein interactions and chromatin accessibility and plays an important role in the regulation of transcription, replication and DNA repair. Conversely, DNA damage affects these crucial cellular processes and induces changes in histone acetylation. However, a comprehensive overview of the effects of DNA damage on the histone acetylation landscape is currently lacking. To quantify changes in histone acetylation we developed an unbiased quantitative mass spectrometry analysis on affinity purified acetylated histone peptides, generated by differential parallel proteolysis. We identify a large number of histone acetylation sites and observe an overall reduction of acetylated histone residues in response to DNA damage, indicative of a histone-wide loss of acetyl-modifications. This decrease is mainly caused by DNA damage-induced replication stress coupled to specific proteasome-dependent loss of acetylated histones. Strikingly, this degradation of acetylated histones is independent of ubiquitylation but requires the PA200-proteasome activator, a complex that specifically targets acetylated histones for degradation. The uncovered replication stress-induced degradation of acetylated histones represents an important chromatin-modifying response to cope with replication stress.

Introduction

DNA transacting processes such as transcription, replication and DNA repair take place in the context of chromatin. Chromatin is a highly organized structure in which DNA is wound around histone octamers to form nucleosomes. The histone octamer consists of a histone H3-H4 heterotetramer flanked by two histone H2A-H2B heterodimers. The linker histone H1 binds linker DNA entering and exiting nucleosomes, thereby regulating chromatin compaction¹. Histones are a target for many post-translational modifications (PTMs), like methylation, phosphorylation, ubiquitylation and acetylation, that are predominantly located at histone tails protruding from the nucleosome². Histone PTMs can directly influence the strength of the histone interactions with each other and with the DNA². Furthermore, several PTMs provide a docking site for specific readers, like chromatin remodeling complexes containing bromodomains that have affinity for acetylated histones³. Together this interplay of histone PTMs and chromatin remodeling proteins controls the accessibility of the chromatin thereby playing an important role in transcription, replication and DNA repair.

For example, histone PTMs play a crucial role during DNA repair and DNA damage signaling^{2,4,5}. Histone acetylation was one of the first histone PTMs, shown to be involved in DNA repair. More than 3 decades ago it was found that after UV irradiation histones undergo a wave of rapid hyperacetylation followed by a hypoacetylation phase⁶. UV light causes helix-distorting lesions such as 6-4 photoproducts (6-4PP) and cyclobutane pyrimidine dimers (CPD), which block transcription and interfere with replication. These lesions can be removed by a specific DNA repair mechanism called nucleotide excision repair (NER)⁷. The fact that hyperacetylated nucleosomes both increased chromatin accessibility *in vivo* and stimulated repair efficiency⁸, led to the formulation of the access-repair-restore concept⁹. This model proposed that chromatin is remodeled by ATP-dependent chromatin remodelers, histone chaperones and modifying enzymes to provide access of repair proteins to damaged sites. After repair, the chromatin conformation is restored to pre-damage conditions to preserve epigenetic information and inhibit DNA damage signaling^{9,10}. Over the years, several histone acetyltransferases (HATs) have been implicated in the response to UV damage¹¹⁻¹⁴. For instance, p300 interacts with PCNA and is associated with newly synthesized DNA after UV irradiation¹⁴. UV-induced H3K9/K14 acetylation by GCN5 increases nucleosome accessibility at the repressed MFA2 locus in yeast^{11,15} through binding of the RSC remodeling complex, which stimulated CPD repair¹⁶. In human cells, GCN5 is necessary for efficient recruitment of NER factors and repair¹⁷. In addition to H3 acetylation, also histone H4 is rapidly acetylated after UV by ING2 leading to the recruitment of XPA to the lesion¹³. Notably, besides acetylation also deacetylation plays an important role during the UV-DDR. For instance, histone deacetylase enzymes (HDACs) 1 and 2 are recruited to damaged sites by the DNA damage recognition proteins, DDB1 and DDB2, resulting in H3K56 deacetylation¹⁸.

Although these studies underscore the crucial interplay of histone acetylation and DNA repair, thus far a comprehensive overview of UV-induced histone acetylation and deacetylation events during repair, but also caused by UV-induced replication or transcription blocks is missing. Changes in histone modifications are often studied using

modification-specific antibodies^{15,19}. However, these techniques rely on the specificity and availability of antibodies and unknown modification sites can therefore not be identified. Especially the acetylation status of histone H2A and H2B after UV irradiation remains largely unclear. In this study we used a quantitative mass spectrometry approach to identify histone acetylation changes in response to UV irradiation in an unbiased manner. Surprisingly, we found that UV damage induces a histone-wide reduction in acetylation levels. This loss of acetylated histones was not dependent on active transcription or NER. Instead, we show that it is the result of replication stress-induced histone degradation by a specific type of proteasome, containing the PA200 subunit, which recognizes acetylated proteins.

Results

Isolation of acetylated histone peptides

To identify the effects of UV-induced DNA damage on histone acetylation in an unbiased and quantitative manner we used a stable isotope labeling by amino acids in cell culture (SILAC) MS based approach. Cells labelled with light-isotope-containing amino acids (KOR0) were mock treated and cells labelled with heavy-isotope-containing amino acids (K6R10) were UV irradiated (16 J/m²) 1 hour before harvesting. In order to increase detection and quantification of histone acetylation sites, including less abundant ones, we established an isolation procedure to enrich for acetylated histone peptides. To this end we combined histone acid extraction^{20,21} with recently developed acetyl-lysine immunoprecipitation (Ac-IP) procedures^{22,23} (Fig 1A). Histones were isolated by a two-step histone acid extraction protocol: extracting first linker histone H1 and the high mobility group proteins, followed by isolation of the core histones. Specificity of the histone extraction was confirmed by comparing the pellet fraction containing the precipitated non-acid-soluble proteins with the histone H1 and the core histone fractions using Coomassie staining (Fig 1B). Protein bands of the expected sizes of H1 (21 kDa) and the core histones (10-15 kDa) could be detected in the designated lanes. The presence of histone H1.2 in the histone H1 fraction and histone H2B in the core histone fraction was confirmed by western blot analysis. Furthermore, the induction of γ H2AX following UV irradiation^{24,25} indicates that histone PTMs are preserved during the acid extraction procedure (Fig 1C). Prior to digestion the isolated H1 and core histone fractions were pooled. To obtain peptide sizes that are compatible with MS analysis to ensure a high coverage of all histone proteins, we split our sample in fractions, each digested with a different protease; trypsin, pepsin or GluC (Fig 1A). The trypsin fraction was further split in four and digested for different durations. After pooling these differential digested fractions, this approach led to the detection of many different unique and overlapping peptides, covering 74-98% of the core histone sequences and 27-62% of the different histone H1 variants (Fig 1D and Dataset EV1).

Acetylated peptides were isolated by Ac-IP and measured on LC-MS/MS. While several acetylated histone peptides could already be identified without the specific isolation procedure (Dataset EV1), the Ac-IP resulted in a 4-fold increase in the number of identified unique acetylated histone peptides. We identified 301 different acetylated histone peptides in total, 75% of which carry more than one acetyl-group (Dataset

EV2). Of note, when we performed additional MaxQuant analysis for the presence of phosphorylation, methylation and diGLY modifications, we found that about 19% of the peptides carried additional histone PTMs (Fig EV1). This suggests that the affinity of the antibody used for the acetylation enrichments is not negatively affected by the presence of additional PTMs on the same peptide.

Within this set of peptides, 40 unique histone acetylation sites were identified, including many previously described sites, like histone H3 acetylation on lysine 9, 14, 18, 23, 27 and 64^{26,27}. Interestingly, we found H2B to have most unique acetylation sites (Fig 1E).

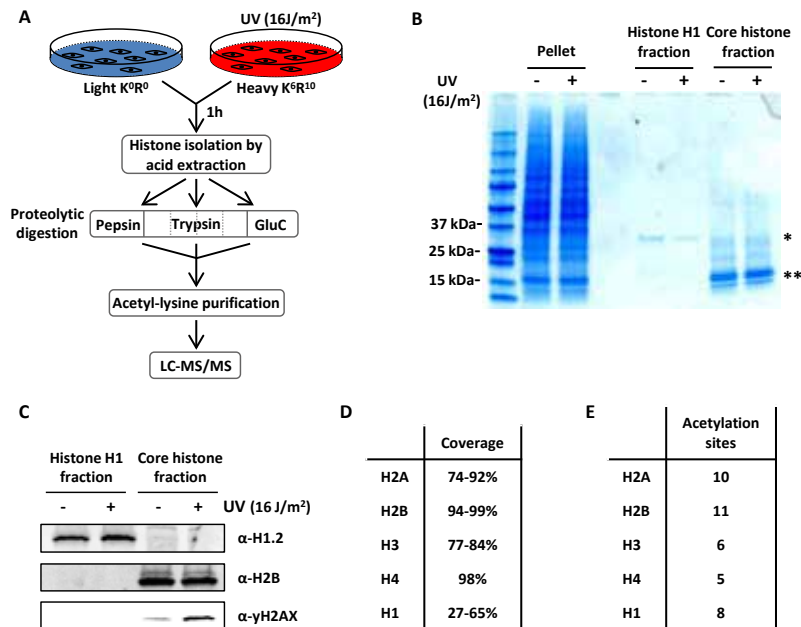


Figure 1. Isolation of acetylated histone peptides. (A) Schematic overview of the experimental set up to isolate and identify acetylated histone peptides following UV irradiation. SILAC labeled cells are mock or UV (16 J/m²) treated one hour before harvesting. Heavy and light labelled cells are mixed in a 1:1 ratio followed by acid extraction to isolate histones. Histones are separated in 3 fractions each digested with either pepsin, trypsin or GluC. The limited digestion with trypsin is performed for 2, 10, 30 and 120 min. The different fractions of digested peptides are combined and followed by acetyl-lysine immunopurification. The acetylated-lysine enriched histone peptides are analyzed by LC-MS/MS. (B) Histones were isolated by acid extraction from HeLa cells one hour after UV (16 J/m²) or mock treatment and were loaded on a 4-15% gradient SDS-PAGE gel and stained with Coomassie Brilliant Blue. The different fractions originate from the same amount of cells and contain the non-acid-soluble proteins (pellet), histone H1 and high mobility group proteins (H1 fraction) or the core histones (core histone fraction). * indicates 21 kDa protein band, most likely representing histone H1. ** indicates several protein bands ranging between 10-15 kDa most likely representing core histones. (C) Western blot of isolated histone fractions from HeLa cells one hour after UV (16 J/m²) or mock treatment. Western blots were stained with α-histone H1.2 (top panel), α-histone H2B (middle panel) and α-γH2AX (bottom panel). (D) Table listing the coverage of the histone sequences using our peptide digestion procedure by LC-MS/MS analysis. (E) Table listing the number of acetylation sites on core histones identified by MS after acetylated peptide enrichment.

As expected, acetylated sites of core histones were mainly found in the tails²⁸ whereas lysine 64 of histone H3 was the only acetylated residue found within the globular domains (Dataset EV3). In contrast to the core histones, the majority of acetyl modified lysines of the linker histone H1 were located in the globular domain²⁹.

Histone acetylation levels are decreased after UV-irradiation

After validating our approach to identify histone acetylation sites efficiently, we analyzed differences in the extent of histone acetylation one hour after UV-induced DNA damage. To visualize UV-induced changes in the acetylation status, all unique acetyl-modified histone peptides were plotted against their SILAC based UV/mock \log_2 ratio (Fig 2A). Surprisingly, the vast majority of the acetylated histone peptides had a negative normalized SILAC ratio following UV treatment (61% had a UV/mock \log_2 ratio < -0.5), indicative of an overall reduction in acetylation level of histones one hour after UV-induced DNA damage (Fig 2B and Dataset EV2). This is in contrast to previously observed UV-induced acetylation at chromatin locations nearby DNA lesions, like H3K9/14Ac, which was observed using site specific antibodies combined with a ChIP procedure^{11,15}. Our mass spectrometry approach, which quantifies all histones in the cell, showed a loss of histone H3 peptides that were acetylated at lysine 9 or 14 (Dataset EV3). In addition to these 2 acetylation sites on histone H3, we found several other H3 acetylation sites to be reduced after UV damage. All core histones displayed lower acetylation levels for most quantified acetyl modifications after UV-induced DNA damage, however histone H2B and H2A were affected the most (Fig 2B, Fig EV2A and Dataset EV2&3). Most acetylated peptides we identified, showed multiple acetyl-modifications on different residues, which were all reduced following DNA damage. Together this data indicates a non-site specific, general loss of acetylated histones. This loss can be either caused by deacetylation of histones, reduced activity of histone acetylases or by degradation of acetylated histones. Of note, only for histone H2A variant H2AZ sites with increased acetylation after UV irradiation (UV/mock \log_2 ratio > 0.5) were observed, while its unmodified peptides are not changed (Dataset EV2&3).

To confirm this striking histone-wide reduction of acetylation detected by MS, we quantified overall histone acetylation levels on western blot using the α -acetyl-lysine antibody. Western blotting of sonicated whole cell extract (WCE), obtained by lysing HeLa cells directly in Laemmli buffer, showed that the vast majority of the signal from the α -acetyl-lysine antibody is confined in two bands around 10-15 kDa (Fig 2C). This α -acetyl-lysine signal fully overlapped with bands obtained by staining against the different core histones, the lower band mainly representing acetylated H4, while the upper band represent acetylated histones H2A, H2B and H3 (Fig EV2B). A similar α -acetyl-lysine signal was obtained from an acid extracted histone fraction (Fig 2C and Fig EV2C). This strongly suggests that these low molecular weight acetyl-lysine signals represent acetylated histones and that the majority of lysine modifications with acetyl within the cell take place at the highly expressed and heavily modified core histones. This was further corroborated by the reduction in acetylation signal following incubation with HAT inhibitors and the increase in signal after HDACs were inhibited (Fig 2C). Together this shows that the overall histone acetylation status can be assessed in a quantitative manner by western blot using α -acetyl-lysine staining.

In line with our MS results, we observed a UV-induced decrease in acetylated histones 1 hour after UV (Fig 2D-E). Interestingly, the histone acetylation levels decreased even further over time. Quantification of the histone acetylation levels normalized to either histone H4 or tubulin showed a 30-50% reduction of acetylated histones 4

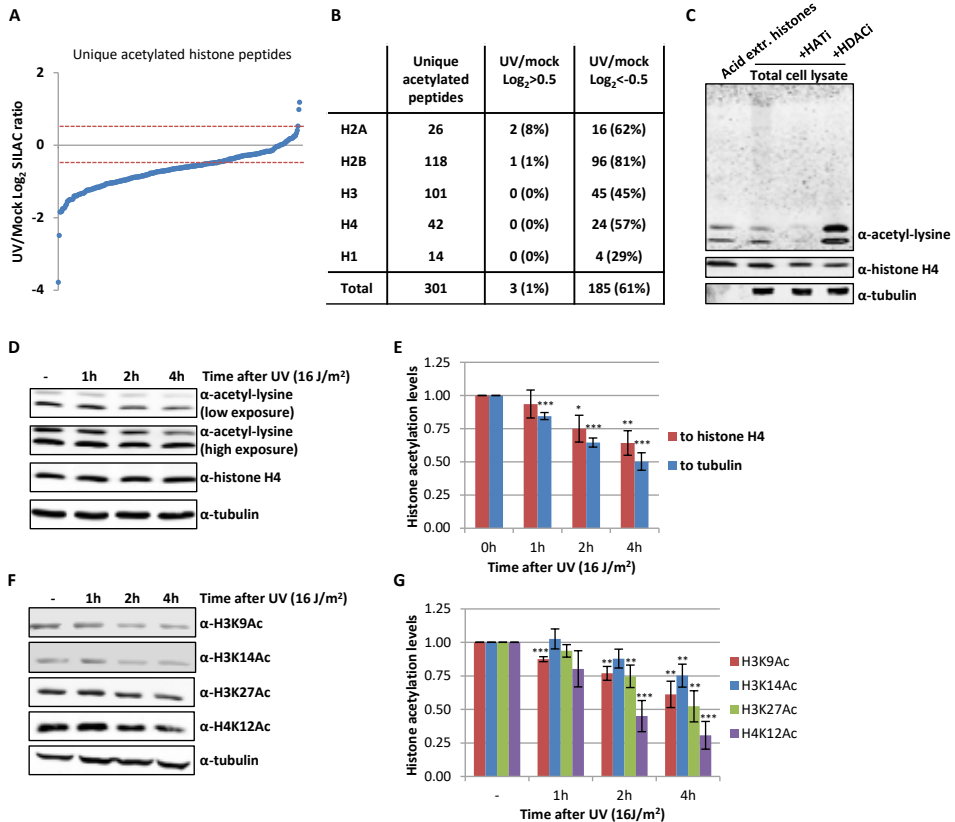


Figure 2. UV-induced decrease in acetylated histones. (A) Identified acetylated histone peptides plotted against their log₂ SILAC ratio (1h after 16 J/m² UV/mock), ranked by SILAC ratio. (B) Table listing the number of identified peptides per histone using MS following enrichment of acetylated peptides, and the number of peptides identified that are decreased (UV/mock log₂<-0.5) or increased (UV/mock log₂>0.5) one hour after UV irradiation (16 J/m²). (C) Western blot of acid extracted histones and WCE from HeLa cells, treated with HATi (CTK7A, 100 μM and CPTH2, 50 μM) or HDACi (TSA, 1 μM) for 4h or mock treated as indicated. (D) A representative western blot of histone acetylation levels of HeLa cells lysed at indicated time points after UV-irradiation (16 J/m²). Blots were stained with α-acetyl-lysine (top panels, high and low exposure), α-histone H4 (middle panel) and α-tubulin (bottom panel). (E) Quantification of histone acetylation levels, normalized against either the histone H4 levels (red bars) or tubulin levels (blue bars). Error bars represent SEM. N=5 independent experiments. Significant differences between UV treated and mock treated conditions were calculated using a T-test and are indicated with * (p<0.1), ** (p<0.05) and *** (p<0.01). (F) Representative western blots of HeLa WCE from different time points after UV-C irradiation (16 J/m²) using the indicated histone modification specific antibodies. Histone H4 and tubulin were used as loading controls. (G) Quantification of specific histone acetylation marks normalized to tubulin levels. Significant differences were calculated using a T-test and are indicated with ** (p<0.05) and *** (p<0.01). N=3, error bars represent SEM.

hours following UV exposure (Fig 2D-E). Our MS results were further confirmed using western blot in combination with antibodies recognizing specific histone acetylation marks (H3K9, H3K14, H3K27 and H4K12) that were identified to be decreased in our MS experiments (Fig 2F-G). In addition, using another acetyl-lysine antibody²², we confirmed the UV-induced reduction in overall histone acetylation levels by both MS (Dataset EV4) and western blotting experiments (Fig EV2D-E), excluding the possibility that our results were due to preferred recognition motifs or other biases of the used antibodies.

Recovery of acetylation levels 16 hours after UV is dependent on NER

To study the temporal behavior of histone acetylation levels after UV irradiation in more detail, western blot experiments were performed with WCE obtained at later time points after UV irradiation. While the acetylation signal was still reduced 8 h after UV, it recovered to levels similar as mock treated cells 16 h after irradiation (Fig 3A (left panel) and B), a time point when most DNA repair is finished^{30,31}. This suggests that repair of UV-induced damage by NER might be necessary for the recovery of histone acetylation levels. Indeed, the recovery of the histone acetylation signal is abolished in cell lines deficient for NER proteins XPC or XPA, indicating that repair of UV-induced DNA damage is crucial for recovery of histone acetylation levels and that the persistent presence of DNA damage prevents this (Fig 3). Interestingly, in NER-deficient cells a similar loss of acetylated histones was observed compared to NER-proficient cells in the first 8 h after UV, showing that the UV-induced decrease in acetylated histones is not dependent on

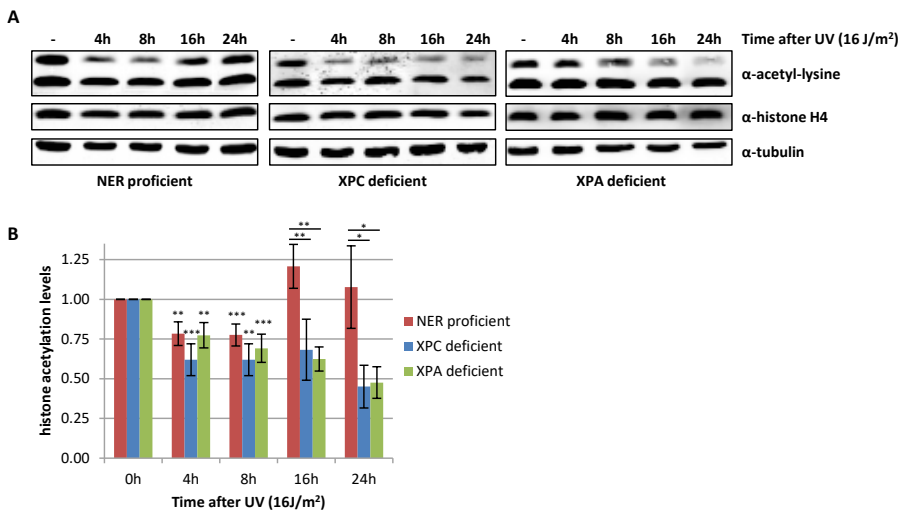


Figure 3. The recovery of histone acetylation levels at later time points after UV is dependent on NER. (A) Representative western blots of WCE from NER-proficient HeLa cells and NER-deficient XP4PA (XP-C) and XP2OS (XP-A) cells obtained at indicated time points after UV irradiation (16 J/m²) and stained with the indicated antibodies. (B) Quantification of histone acetylation levels of NER-proficient cells (HeLa) and NER-deficient (XP-A and XP-C) patient cells at the indicated time points after UV-irradiation (16 J/m²). Histone acetylation levels are normalized against histone H4 levels. Average of at least 5 independent experiments and error bars represent SEM. Significant differences (T-test) are indicated with * (p<0.1), ** (p<0.05) and *** (p<0.01).

DNA damage recognition or on repair by NER (Fig 3). Together these data suggest that the trigger for the observed decrease in histone acetylation levels and occurs upstream or in parallel of the damage recognition step of NER.

Transcription does not influence the decrease in acetylated histones after UV-irradiation

Persistent UV-induced DNA lesions severely impede transcription and replication^{32,33}. As active transcription is highly associated with increased levels of histone acetylation² we tested whether transcription inhibition, in the absence of DNA damage, affects the acetylation status of histones. Transcription was impeded by THZ1 and flavopiridol, which both inhibit transcription preceding productive elongation, and α -amanitin that blocks elongating RNAPII³⁴. In contrast to UV, the histone acetylation levels remain rather stable after transcription inhibition, indicating that transcription inhibition is not the main cause of the UV-induced loss of acetylated histones (Fig 4A-B). However, even though chemical transcription inhibition did not result in a decrease in acetylated histones, we could not exclude that active transcription or lesion stalled RNAPII might initiate this UV-induced process. To test this, cells were pre-treated with the transcription inhibitors THZ1 or flavopiridol to deplete cells of elongating RNAPII before UV irradiation. This, however, did not affect the decrease in histone acetylation levels after UV irradiation (Fig 4C-D), indicating that the UV-induced loss of acetylated histones is a process independent of transcription.

Decrease in histone acetylation levels is the result of UV-induced replication stress

In addition to transcription inhibition, UV induced DNA lesions also cause replication stress, by slowing down or stalling replication forks. To test whether the loss of acetylated histones is caused by UV-induced replication stress, we also induced replication stress by exposing cells to a combination of hydroxyurea (HU) and arabinofuranosyl cytidine (AraC) or by mitomycin C (MMC). These experiments showed that induction of replication stress results in a similar loss of acetylated histones as after UV irradiation (Fig 5A-B). This indicates that replication stress itself can indeed induce a general decrease in histone acetylation levels and suggests that the loss of acetylated histones following DNA damage might be the direct consequence of UV-induced replication stress. To test whether the loss of acetylated histones following UV exposure could be attributed to replication stress, we studied UV-induced effects on histone acetylation in non-replicating cells. To this end, HeLa cells were blocked in S-phase using the Cdc7/CDK9 inhibitor PHA 767491 hydrochloride (Fig EV3A). These cells displayed lower histone acetylation levels under unperturbed conditions (Fig 5C and Fig EV3B), which suggests that the high histone acetylation levels in cycling cells are mainly induced by replication related events. However, no additional UV-induced decrease was observed in these cells (Fig 5C-D). Similar results were obtained in contact inhibited non-replicating VH10 cells, which showed less reduction in histone acetylation levels compared to cycling VH10 cells following UV-exposure (Fig EV3C-E). Together these results indicate that the loss of acetylated histones after UV irradiation is mainly the result of replication stress.

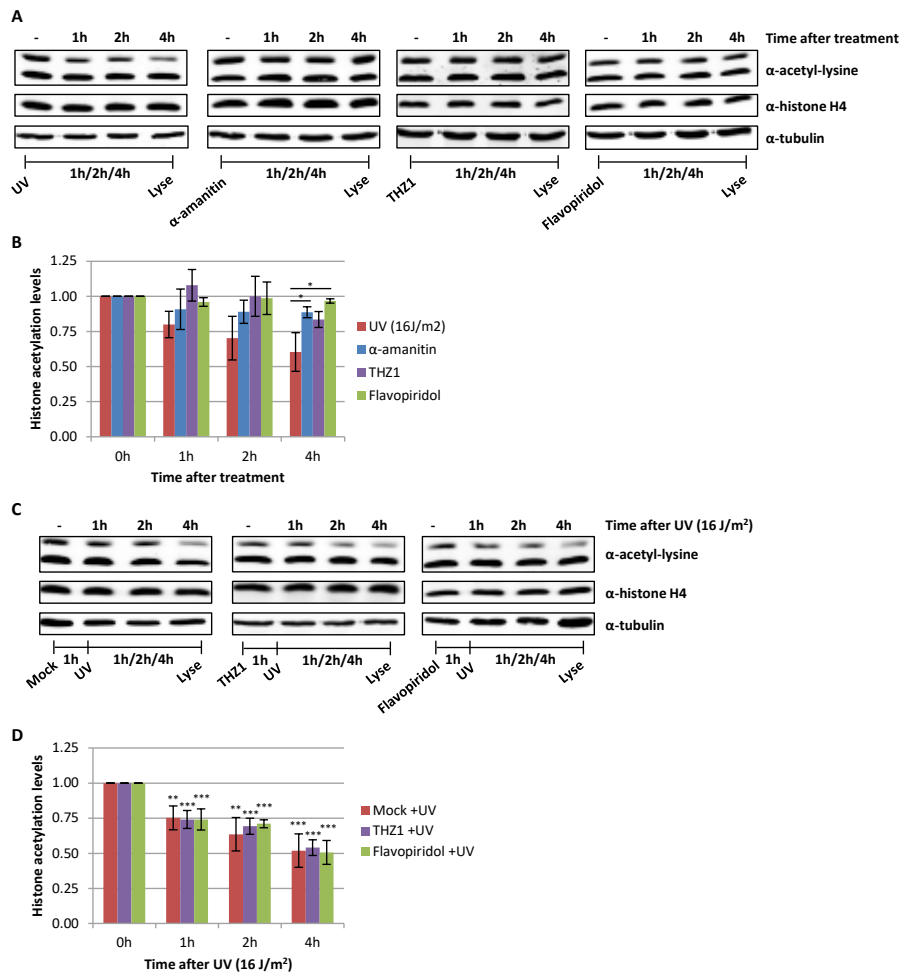


Figure 4. Loss of acetylated histones is independent of transcription inhibition. (A) Representative western blots of histone acetylation levels of HeLa cells lysed at indicated time points after UV irradiation (16 J/m², left panel), α-amanitin (second panel, 100 μg/ml), THZ1 (third panel, 1 μM) or flavopiridol (right panel, 1 μM) treatment. Blots were stained with α-acetyl-lysine (top panel), α-histone H4 (middle panel) and α-tubulin (bottom panel). **(B)** Quantified histone acetylation levels, normalized against histone H4 levels for quantification. Average of at least 3 experiments. Error bars represent SEM. Significant differences between UV-irradiated and inhibitor treated conditions are indicated with * ($p < 0.1$). **(C)** Representative western blots of histone acetylation levels of HeLa cells pre-treated with transcription inhibitors (mock (left), THZ1 (middle, 1 μM) or flavopiridol (right, 1 μM) an hour before UV irradiation and lysed at the indicated time points after UV (16 J/m²). Blots were stained with the indicated antibodies. **(D)** Quantification of α-acetyl-lysine signal, normalized against histone H4 levels. Average of 5 experiments. Error bars represent SEM. Significant differences, calculated by T-test, between UV treated and mock treated conditions are indicated with * ($p < 0.1$), ** ($p < 0.05$) and *** ($p < 0.01$).

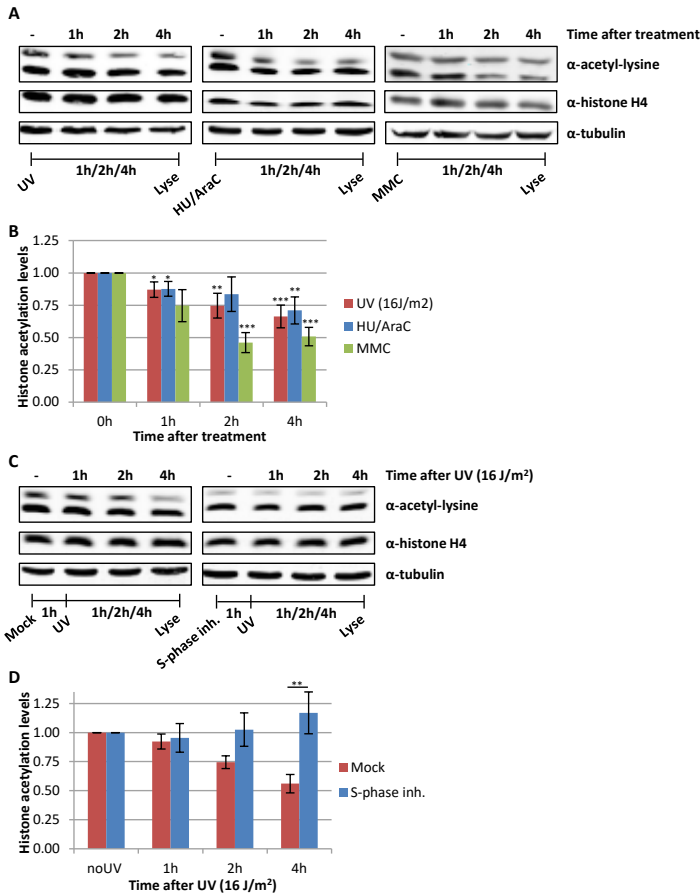


Figure 5. Replication stress induces loss of acetylated histones. (A) Representative western blots of histone acetylation levels of HeLa cells obtained at the indicated time points after UV irradiation (16 J/m²), HU/AraC treatment (100 mM/10 µM) or MMC treatment (10 µg/ml). Blots were stained with α-acetyl-lysine (top panel), α-histone H4 (middle panel) and α-tubulin (bottom panel). **(B)** Quantification of average histone acetylation signal, normalized to loading control. N≥3. Error bars represent SEM. Significant differences, calculated with T-test, between treated and mock conditions are indicated with * (p<0.1), ** (p<0.05) and *** (p<0.01). **(C)** Representative Western blots, stained with the indicated antibodies of HeLa cells pre-treated o/n with an S-phase inhibitor (PHA 767491 hydrochloride, 10 µM) and lysed at indicated time points after UV (16 J/m²). **(D)** Quantification of the α-acetyl-lysine signals, normalized against histone H4 levels, N= 5, error bars represent SEM. Significant differences, calculated with T-test, between UV-irradiated and inhibitor treated conditions are indicated with ** (p<0.05).

Although ATR and ATM signaling pathways play an important role following DNA damage and replication stress^{35,36}, the replication stress-induced loss of acetylated histones was not dependent on the activity of these kinases as histone acetylation was still reduced in the presence of both ATM and ATR inhibitors (Fig EV3F-G).

Chromatin bound acetylated histones are degraded

Besides the reduction in histone acetylation levels, we observed a concomitant subtle decrease in the histone H4 levels after UV irradiation, which was more pronounced at later time points (Fig 6A-B), however to a lesser extent than the loss of acetylated histones (Fig 6C). This suggests that the observed UV-induced decrease in histone acetylation levels could be caused by degradation of a specific subset of acetylated histones. Accordingly, we observed that the decrease in histone acetylation after UV was larger when normalized to tubulin compared to normalization to H4 levels (Fig 2E). To test this hypothesis, cells were pre-treated with the proteasome inhibitor MG132 one hour before UV irradiation. Proteasome inhibition completely rescued the UV-induced loss of both acetylated histones and histone H4 (Fig 6A-C). Of note, this loss of H4 is, like the loss of histone acetylation, dependent on replication (Fig EV3I).

During replication, the levels of soluble histones are increased. In addition, it has previously been shown that newly synthesized histone H3/H4 dimers are acetylated to stimulate incorporation into chromatin³⁷⁻³⁹. Therefore, we hypothesized that this pool of newly synthesized, yet unincorporated acetylated histones might be degraded upon replication stress, thereby preventing an excess of non-incorporated histones. To test this, we performed cell fractionation experiments to separate newly synthesized, free histones from chromatin bound histones. Compared to chromatin-bound histones, the levels of soluble histones is, as expected, very low and often hardly detectable on blots (Fig 6D). Most importantly, the acetylation signal is almost completely restricted to the chromatin fraction. This suggests that the observed 30-50% decrease in acetylation levels can only be explained by degradation of chromatin-bound histones. Indeed, after UV irradiation the acetylation levels in the chromatin fraction are decreasing over time, to a similar extent as was observed in whole cell extract (Fig 6D-E). This replication stress-induced decrease in acetylation levels of chromatin-bound histones was confirmed by additional stainings using acetylation site specific histone antibodies (Fig 6D, F-G).

Acetylated histones are degraded by PA200-proteasome complexes

Interestingly, inhibition of the ubiquitin-activating enzyme (E1) using PYR-41, did not rescue the decrease in acetylated histones (Fig 7A-B). This indicates, in contrast to most proteasome dependent protein degradation, that the proteasomal degradation of acetylated histones is independent of protein ubiquitylation. It has been shown that acetylated histones can be recognized and degraded by a specific proteasome complex independent of protein ubiquitylation⁴⁰. This specific proteasome complex consists of the 20S core complex and the nuclear proteasome activator PA200⁴⁰. In contrast to the 26S proteasome, in which the 19S regulatory cap recognizes ubiquitylated proteins, the PA200-proteasome complex does not recognize poly-ubiquitin chains but binds to acetylated proteins and targets them for degradation⁴⁰. Therefore, we tested the involvement of PA200 in the UV-induced decrease in histone H4 and histone

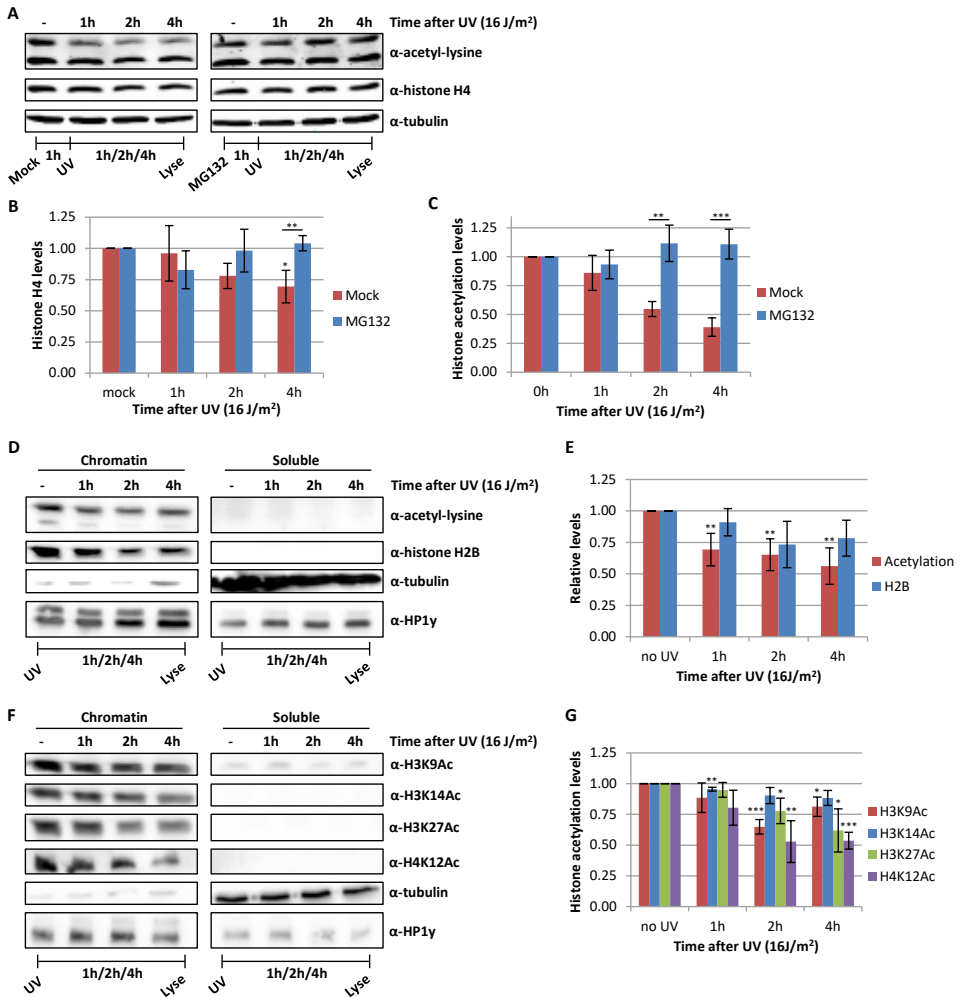


Figure 6. Proteasomal degradation of acetylated histones following replication stress. (A) Western blots of HeLa cells treated with proteasome inhibitor MG132 one hour before UV irradiation and lysed at different time points after UV (16 J/m^2). Blots were stained with α -acetyl-lysine (top panel), α -histone H4 (middle panel) and α -tubulin (bottom panel). Quantifications of **(B)** histone H4 levels and **(C)** acetylation levels, normalized against tubulin levels. Average of 6 experiments and error bars represent SEM. Representative blots are shown in panel A. **(D)** Fractionation experiments comparing histone and acetylation levels in chromatin and soluble fractions at the indicated time points after UV irradiation (16 J/m^2). HP1y was used as loading control for chromatin fraction, tubulin for soluble fraction. **(E)** Quantification of histone acetylation and H2B levels in chromatin fraction. Non-irradiated samples were set as 1. $N \geq 4$, error bars indicate SEM. **(F)** Representative western blots of fractionated cell extracts of HeLa cells at indicated time after UV irradiation (16 J/m^2) using histone modification specific antibodies. Loading controls for H3K9 and H3K27 are depicted in figure F. **(G)** Quantification of the signal of the indicated histone modification specific antibodies stainings in the chromatin fraction. $N \geq 4$ and error bars represent SEM. Significant differences were calculated with T-test and are indicated with * ($p < 0.1$), ** ($p < 0.05$) and *** ($p < 0.01$).

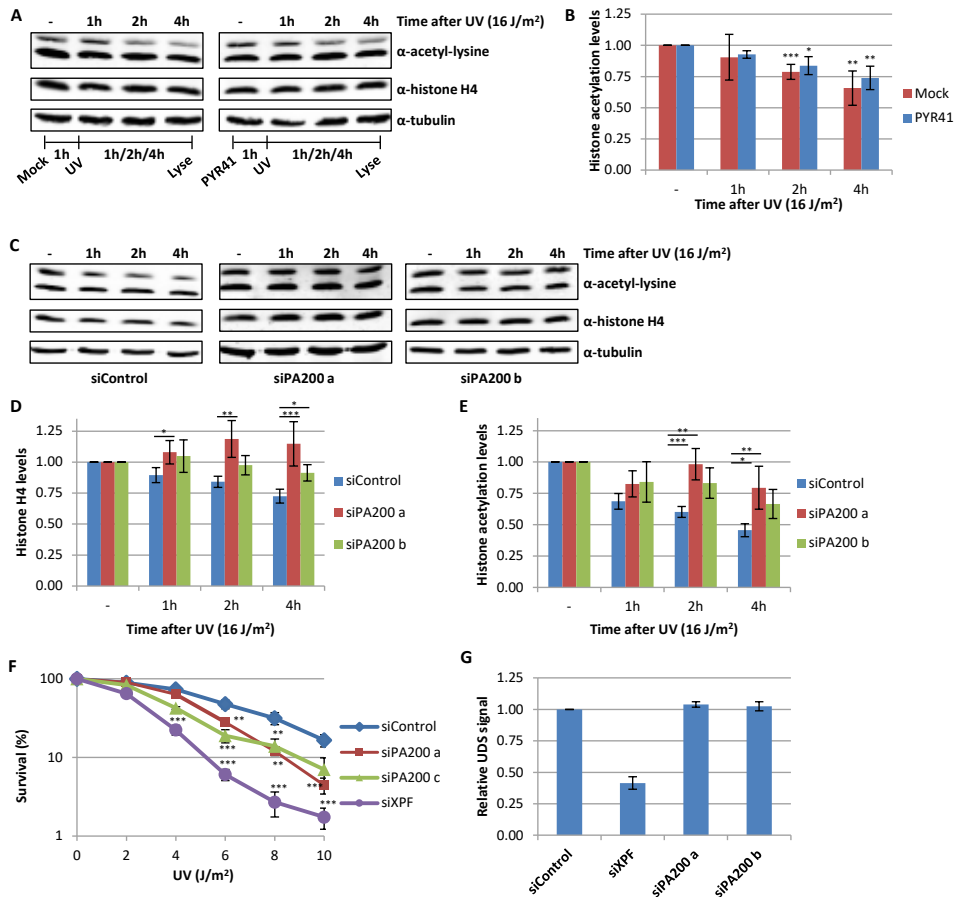


Figure 7. PA200-proteasome degrades acetylated histones. (A) Western blots of HeLa cells pre-treated with the E1 enzyme inhibitor PYR-41 (10 μ M) one hour before UV irradiation and lysed at different time points after UV (16 J/m²). Blots were stained with the indicated antibodies and representative blots are shown. (B) Quantification of histone acetylation levels, normalized against histone H4 levels, average of 5 experiments. Error bars represent SEM. (C) Representative western blots of siControl or siPA200 (2 independent siRNAs, a and b) transfected cells, lysed at indicated time points after UV-irradiation (16 J/m²). (D) Quantification of histone H4 and (E) histone acetylation levels, normalized against tubulin and non-treated levels are set as 1. The average of at least 9 independent experiments and SEM is shown. Significant differences were calculated by T-test and are indicated with * ($p < 0.1$), ** ($p < 0.05$) and *** ($p < 0.01$). (F) Clonogenic survival of HeLa cells treated with different doses of UV-C. Relative survival (in percentage) is plotted against UV-C dose. $N \geq 2$, error bars represent SEM. Significant differences, calculated by T-test are indicated with * ($p < 0.1$), ** ($p < 0.05$) and *** ($p < 0.01$). (G) Unscheduled DNA synthesis in non-cycling C5RO cells measured by the incorporation of EdU (20 μ M) after UV irradiation (16 J/m²) visualized by ATTO594 labelling using Click-chemistry. Levels in siControl cells were set as 1. $N = 2$, >200 cells analyzed, error bars represent SEM.

acetylation levels. siRNA mediated knockdown of PA200 (Fig EV3A) inhibited the UV-induced proteasomal degradation of acetylated histones (Fig 7C-E). Together these results show that following replication stress, for example induced by UV-induced DNA damage, acetylated histones are degraded by the PA200-20S proteasome complex in an ubiquitin independent manner. To address the biological relevance of the degradation of acetylated histones we performed cellular survival assays following UV-induced DNA damage. Interestingly, PA200 depleted cells were more sensitive to UV-induced DNA damage compared to control transfected cells (Fig 7F, Fig EV3C), suggesting that the degradation of acetylated histones play a crucial role during the cellular responses to replication stress. Of note, this UV-sensitivity following PA200 depletion is not caused by an effect on NER-mediated repair rate, since knockdown of PA200 did not affect NER efficiency as shown by unscheduled DNA synthesis (UDS) experiments, which quantifies the gap filling DNA synthesis step NER (Fig 7G). Together these data indicate that the replication stress-induced degradation of acetylated histones is important for cells to cope with replication stress.

Discussion

To study the effects of UV-induced DNA damage on histone acetylation in an unbiased manner, we established a protocol to efficiently isolate acetyl-modified histone peptides. This stepwise purification approach, consisting of an acid extraction of histones, combined with proteolytic digestion with different proteases in parallel and followed by immuno-purification of acetylated peptides and analysis by mass spectrometry, resulted in the identification of 40 histone acetylation sites originating from over 300 unique acetylated peptides. Using this procedure, we uncovered a striking UV-induced reduction of the vast majority of detected acetylated residues originating from all histones, except for variant H2AZ. Western blot experiments confirmed this UV-induced overall loss of acetylated histones and showed a loss of overall histone acetylation levels up to 40% 4 hours after UV exposure. Thus far most research on histone acetylation was focused on H3 and H4, for example on the role of H4K16Ac in chromatin compaction⁴¹ and H3K9/14Ac in transcription⁴² and during the UV-DDR¹⁵⁻¹⁷. Our data shows that acetylation levels of H2A and H2B are more reduced in response to UV than H3 and H4 acetylation levels (Fig 2B, Fig EV2A and Dataset EV2&3), suggesting an important role for the regulation of H2A and H2B acetylation levels in the UV-DDR or in the extensive chromatin remodeling processes observed following DNA damage^{43,44}. In striking contrast to the here described global decrease of histone acetylation, multiple other studies have observed an increase in acetylation in response to UV irradiation^{11-13,15-17,45}. For example, the UV-induced histone acetylation on H3K9/14 was shown to stimulate efficient NER¹⁷ by opening the chromatin structure as a result of recruitment of the chromatin remodeling complex RSC¹⁶. This UV-induced acetylation is mediated by histone acetyltransferases GCN5 and p300, which are both found at UV lesions and interact with the UV-DDB complex^{11,12,15,17,46,47}. Our MS results show a reduction in H3K9 and K14 acetylation an hour after UV damage. In addition, western blot analysis using antibodies specifically recognizing H3K9 and K14 acetylation shows a further decrease in the levels of these acetylation marks at later time points after damage (Fig 2F-G). This apparent discrepancy may be explained by the fact that the observed UV-induced increase of these marks may specifically happen in close vicinity of UV-induced DNA

lesions¹⁷. It is thus likely that this local increase of acetylated histones is masked by the overall replication-stress-derived loss of acetylated Histones. In addition, in our MS dataset we do not find any peptides that only carry H3K9Ac or only H3K14Ac, but all peptides containing these modifications carry at least one other additional acetylation mark. Additionally, this discrepancy could also be due to differences in cell cycle or caused by the fact that following replication stress a specific subset of hyperacetylated histones is degraded, while histones in the vicinity of UV-induced lesions carry a single H3K9 or H3K14 acetylation mark.

We propose that our observed overall loss of histone acetylation is a separate event, that acts independent of the reported histone acetylation events during NER^{11-13,15-17,45}. First, the large 40% decrease in total histone acetylation levels occurs on all core histones and includes multiple lysines, suggesting a chromatin-wide effect in contrast to the reported histone acetylation events that are often lysine specific and are mainly locally observed centered around the DNA lesion^{11,15,17}. Furthermore, NER-stimulating acetylation was proposed to be facilitated by the recruitment of HATs to the site of DNA damage^{14,17,46-48}, while the here described replication stress-associated loss of acetylated histones is due to proteasomal degradation rather than caused by changes in HATs or HDACs activity. In addition, the degradation of acetylated histones is only observed in cycling cells (Fig. 5C-D), while NER and the previously described histone acetylation events are expected to take place throughout the cell cycle. Finally, we found that the loss of acetylated histones after UV irradiation is completely independent on NER; it is still observed in NER deficient cells (Fig 3) and the degradation of acetylated histones does not affect NER efficiency (Fig 7G). These observations suggest that the overall decrease in histone acetylation in response to UV is not directly associated with DNA repair of UV lesions, but more likely a consequence of UV-induced replication stress. Indeed, induction of replication stress by other mechanism, treatment with MMC or HU/AraC, resulted in a similar loss of acetylated histones (Fig 5A-B). In line with this, it was previously observed, using acetyl-histone specific antibodies, that H3K9Ac and H3K56Ac are reduced after hydroxyurea treatment¹⁹. The fact that mainly an overall loss of histone acetylation was observed, despite the existence of independent and opposing UV-induced histone acetylation influencing events, indicates that in dividing cells the replication stress-induced loss of acetylated histones is much more abundant than the DNA damage-localized, site-specific acetylation that regulates NER. To specifically investigate acetylation events that are induced by UV-lesions or associated with NER, it is thus important to execute experiments in non-replicating cells.

Of note, unlike the loss of acetylated histones, the recovery of histone acetylation levels after UV-induced replication stress is dependent on NER. Since NER is the mammalian DNA repair system capable of removing UV-induced lesions⁷, its absence will prohibit repair of replication stress inducing UV-lesions. The recovery of histone acetylation levels is likely the result of synthesis of new histones required for the resumption of replication, which are subsequently acetylated, to replace the acetylated histones that were degraded in response to DNA damage.

The decrease in histone acetylation levels following replication stress is proteasome dependent and is likely explained by the degradation of a specific subset of histones, i.e. acetylated histones, as the histone acetylation levels are more reduced than the histone levels in response to UV irradiation (Fig 6A-C). Surprisingly, the observed histone degradation is not dependent on ubiquitylation, but instead it depends on the PA200-proteasome that specifically targets acetylated proteins for degradation (Fig 7A-E)⁴⁰. It is not known whether or how PA200 can discriminate between differentially acetylated histones. For example, it is not expected that histones in transcriptionally active regions, which also carry acetylation marks⁴⁹, are constantly targeted by the PA200-proteasome. We observed a reduction in almost all acetylated peptides, many of which carried multiple acetyl groups (Dataset EV2). This might suggest that PA200 may not recognize specific histone marks, but rather targets hyperacetylated histones for degradation. Future studies should uncover if and how PA200 can discriminate between different subsets of acetylated histones, or that for example the activity of the PA200 complex itself is regulated to control histone degradation following replication stress. It is also not clear yet whether incorporated histones can be degraded by the PA200 proteasome or whether histones need to be evicted by a specific histone chaperone from chromatin prior to degradation.

An important question that remains to be answered is why histones are degraded in response to replication stress. In response to DNA damage, chromatin is remodeled to a more accessible conformation to enable efficient repair and this might be facilitated by the degradation of histones. In line with this, histone degradation in response to zeocin treatment was recently observed in yeast. This degradation was shown to lead to enhanced chromatin dynamics and recombination rates⁵⁰. We have shown that PA200 dependent histone degradation does not affect nucleotide excision repair, however it might be necessary to efficiently overcome stalled replication forks. In accordance with this hypothesis, PA200 was found to localize to chromatin in response to DNA damage⁵¹, which suggests that it might degrade nucleosomal histones. Previously, it was shown that an excess of histones interferes with homologous recombination in yeast⁵². It is thus plausible that the observed histone degradation in response to replication stress is required to prevent an excess of histones to resolve stalled replication forks efficiently. Degradation of chromatin-bound histones around stalled forks might make space for the proteins involved in resolving stalled forks for example by facilitating reversal of the replication fork followed by template switching to allow resumption of replication⁵³. Additionally, this mechanism might be involved in the repair of DNA breaks, caused by the collapse of stalled replication forks. As degradation of acetylated histones is much more abundant as the UV-induced histone acetylation, this suggest that chromatin remodeling by histone degradation around stalled replication forks possibly involves longer stretches of chromatin than remodeling around UV-induced lesions that are resolved by NER. However, it is currently not known whether histones are indeed hyperacetylated near stalled replication forks. Interestingly, in yeast all core histones, except H2AZ, are degraded in response to DNA damage⁵⁰. The only two sites that in our study were identified to be more acetylated in response to UV irradiation are both located on this histone variant (Dataset EV3), which was previously described to be important for UV survival and CPD repair in yeast¹⁶. This might imply that also in

human cells H2AZ might not be degraded in response to replication stress. However, the function of this UV-induced H2AZ acetylation¹⁶ and why the PA200-proteasome would not recognize this specific histone variant remains unclear.

In summary, our data indicates that in response to replication stress acetylated histones are specifically degraded by the PA200-proteasome in an ubiquitin independent manner. This degradation most likely represents an important chromatin-modifying mechanism for cells to cope with stalled replication forks.

Materials and methods

Cell culture and treatments

HeLa and VH10, SV40 or hTert immortalized, cells were cultured in DMEM/F10 (Lonza) supplemented with 10% fetal calf serum (FCS) and 1% penicillin-streptomycin (PS, P0781 Sigma) at 37°C in a humidified incubator containing 5% CO₂. For stable isotope labeling by amino acids in cell culture (SILAC), HeLa cells were cultured for at least 10 cell doublings in lysine, arginine and L-glutamine (PAA) deficient DMEM with 10% dialyzed FCS (Invitrogen), 1% PS and 1% ultraglutamine (200mM Lonza), supplemented with either 73 µg/mL light [¹²C₆]lysine and 42 µg/mL light [¹²C₆, ¹⁴N₄]arginine (Sigma) or similar concentrations of heavy [¹³C₆]lysine and heavy [¹³C₆, ¹⁵N₄]arginine (Cambridge Isotope Laboratories). If not stated otherwise in the figure legends, cells were at 70-80% confluency when treated. For UV treatments cells were washed with phosphate-buffered saline (PBS) and irradiated with 16 J/m² (254nm, Philips TUV lamp). To induce replication stress a combination of hydroxyurea (100 mM) and AraC (10µM) or mitomycin C (10 µg/ml) was used. Transcription was inhibited using; α-amanitin (0.1mg/ml), THZ1 (1 µM) and flavopiridol (1 µM). To inhibit cell cycle PHA 767491 hydrochloride (10 µM) was added o/n to the culture medium. MG132 (50 10 µM) and PYR41 (10 µM) were added to inhibit the proteasome or ubiquitylation respectively. DNA damage signaling was inhibited by the addition of ATM (Ku55933, 10 µM) and ATR (VE-821, 10 µM) inhibitors. Sequences of siRNAs used: siControl UGGUUACAUGUCGACUAA, siXPF (Dharmacon smartpool), siPA200 a GAAAAGAGAUGCAAAGUUA, siPA200 b AGAAAUAAGGCUCAGCAAA and siPA200 c GCUUCAACUUAGUAAAGAA.

Histone isolation and protein digestion

Cells were washed twice with cold PBS, scraped and mixed in a 1:1 ratio based on cell pellet size (UV:mock). The cell pellet was resuspended in 1 volume 5% perchloric acid. After incubating for 10 min, the solution was centrifuged for 10 min at 13,000 rpm. The supernatant containing histone H1 was collected and the procedure was repeated twice. The pellet was then resuspended in 2 volumes 0.4N hydrochloric acid (HCl) and incubated for 15 min before spinning down (10 min 13.000 rpm). This procedure was repeated twice and all supernatants containing the core histones were collected. The histone containing supernatants were precipitated by the addition of trichloric acid to 25% and incubated on ice for 30 min and centrifuged for 20 min at 13,000 rpm. Pellets were washed in acetone with 0.006%HCl and subsequently in acetone and finally the histone pellets were dissolved in 50 mM ammonium bicarbonate. Digestions were performed with proteomics-grade trypsin (Roche, 1:100, 30°C, 2/10/30/120 min) (50% of

sample), GluC (100x, o/n, room temperature (RT)) (25% of sample) or on an immobilized pepsin (Pierce) column (in 0.5% trifluoroacetic acid (TFA)) (25% of sample). Tryptic and GluC digestion was stopped by acidification to 0.5% TFA. The digested peptides were mixed and subsequently purified using 200 mg tC18 SEP-PAK SPE cartridges (Waters) and eluted with 40% acetonitrile (ACN) containing 0.1% TFA. The peptides were then lyophilized for 48 h (Scanvac CoolSafe 110-4, Scala Scientific).

Ac-K peptide enrichment

Lyophilized peptides were dissolved in 1.4 mL of IAP buffer (PTMscan, cell signaling) and incubated with anti-Ac-K antibody beads (PTMscan, cell signaling) for 2 h at 4°C on a rotating unit. Beads were washed three times in IAP buffer followed by two washes in H₂O and subsequently the peptides were eluted using 0.1% of TFA in H₂O. The enrichments with the anti-Ac-K antibody from Immunechem were performed as described previously²². In short, peptides were dissolved in 1.4 ml IP buffer (50 mM MOPS pH7.2, 10 mM sodium phosphate, 50 mM NaCl) and incubated o/n with 62.5 µg anti-Ac-K antibody conjugated to 12 µl protein A beads (GE Healthcare). Beads were washed 3 times in IP buffer and two times with water before peptide elution with 0.15% TFA. Eluted peptides were purified with C18 stage tips (Millipore).

Mass spectrometry

Digested histone peptides, not enriched for acetyl-lysines, were analyzed with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) or a quadrupole Orbitrap (Q-Exactive, Thermo Fisher Scientific) and samples enriched for acetylated peptides were analyzed with the Orbitrap Fusion Tribrid mass spectrometer according to protocols below.

Mass spectra were acquired on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC 1000 system (Thermo Fisher Scientific). Peptide samples were loaded onto a ReproSil C18 reversed phase column (20 cm x 75 µm ID) and eluted with a gradient of 5-80% (acetonitrile containing 0.1% formic acid) over 90 min at 300 nl/min. For all experiments, the instrument was operated in data-dependent acquisition (DDA) mode. MS1 spectra were collected at a resolution of 120,000, with an automated gain control (AGC) target of 2E5 and a max injection time of 50 ms in the scan range from 375-1500 m/z. The most intense ions were selected for MS/MS, top speed method with a 3 second cycle time. The normalized collision energy was optimized at 30% for HCD. Precursors were filtered according to charge state (2-7z), and monoisotopic peak assignment. Previously interrogated precursors were dynamically excluded for 70 s. Peptide precursors were isolated with a quadrupole mass filter set to a width of 0.7 Th.

For ETD and ETHcd fragmentation experiments a decision tree was applied in the Orbitrap Tune software for instrument control. For precursor ions with charges 2-3 ETHcd was used for fragmentation, while for precursors with charges 4-8 ETD was used. Calibrated charge dependent ETD parameters were applied and fragment ions were measured in the ion trap. Alternatively, peptides were analyzed on a quadrupole Orbitrap (Q-Exactive, Thermo Fisher Scientific) mass spectrometer equipped with an

EASY-nLC 1000 (Thermo Fisher Scientific). Peptide samples were loaded onto a ReproSil C18 reversed phase column (20 cm x 75 μ m ID) and eluted with a gradient of 5-80% (acetonitrile containing 0.1% formic acid) over 70 min at 300 nl/min. Fragmentation of the peptides was performed in DDA mode. MS1 spectra were collected at a resolution of 70,000, with an automated gain control (AGC) target of 1E6 and a max injection time of 50 ms. The 10 most intense ions were selected for MS/MS. Precursors were filtered according to charge state (2-7z) and monoisotopic peak assignment. Previously interrogated precursors were dynamically excluded for 30 sec and peptide precursors were isolated with a quadrupole mass filter set to a width of 2.0 Th.

Peptide identification

Raw data files were analyzed using MaxQuant software (version 1.5.2.8). MS/MS spectra were searched against a histone database, containing all uniprot entries of which the protein name contains the string "histone" and organism "*Homo sapiens*", using the Andromeda search engine. The protease specificity was set to nonspecific cleavage. Cysteine carbamidomethylation was included as a fixed modification, whereas methionine oxidation, N-terminal protein acetylation and lysine acetylation were set as variable modifications. To check for the presence of additional histone modifications, as shown in Figure EV1, dedicated MaxQuant searches were performed in which also ubiquitylation, phosphorylation, methylation, dimethylation and trimethylation were set as variable modifications. A false discovery rate of 0.05 for peptides and a minimum peptide length of 7 were set. Before further data analysis, known contaminants and reverse hits were removed from the modification-specific peptides list.

Cell fractionations

To separate chromatin-bound proteins from soluble proteins, cells (80% confluent, 3 cm² dishes) were washed 2x with PBS and scraped from the plates. Cells were lysed for 1h in buffer A (30 mM HEPES pH7.5, 250 mM NaCl, 1 mM MgCl₂, 0.5% Triton X-100, protease inhibitor) on ice and centrifuged for 15 min at full speed at 4°C. Supernatants containing soluble proteins were taken and pellets, containing chromatin, were washed 1x with 1ml buffer A. Both the soluble and chromatin fractions 2x SDS sample buffer was added and samples were boiled for 5 min at 95°C prior to loading on 14% SDS-PAGE gels.

Western blotting

Cell lysates were made in 2x Laemmli buffer. Lysates were boiled and sonicated with a Diagenode Bioruptor (30 sec on; 30 sec off for 10min) to shear the DNA. Lysates were separated on a 6 or 14% SDS-PAGE acrylamide gel and transferred to a PVDF membrane (0.45 μ m, Merck Millipore Ltd). Membranes were blocked with 5% milk in PBS for 1 hour at RT and incubated with primary antibody for 2 hours at RT or overnight at 4°C. Secondary Alexa Fluor 795 donkey anti-mouse antibodies and Alexa Fluor 680 donkey anti-rabbit antibodies (Sigma) were used to visualize the proteins using an infrared imaging system (Odyssey; LI-COR Biosciences). Primary antibodies were: rabbit- α -histone H2B (1:2000, Cat # ab64039 abcam), mouse- α -phospho-H2A.X (SER139) (1:1000, Cat # 05-636, Millipore), rabbit- α -histone H1.2 (1:1000, Abcam), rabbit- α -histone H2A (1:200, ab13923, Abcam), goat- α -histone H2B (1:100, Cat #

SC-8650, Santacruz), goat- α -histone H3 (1:1000, Cat # CS-8654, Santacruz), mouse- α -histone H4 (1:1000, Cat # ab31830, Abcam), rabbit- α -Acetyl-lysine (1:2000, Ac-K-100, Cat # 9814S, PTMscan), rabbit- α -Acetyl-lysine (1:1000, Cat # ICP0380, Immunechem), rabbit- α -Acetyl histone H3 Lys9 (1:1000, Cat # 06-942, Millipore), rabbit- α -Acetyl histone H3 Lys14 (1:2000, Cat # 07-353, Millipore), rabbit- α -Acetyl histone H3 Lys27 (1:2000, Cat # ab4729, Abcam), rabbit- α -Acetyl histone H4 Lys12 (1:2000, Cat # ab46983, Abcam), mouse-anti-HP1 γ (1:1000, Cat # ab56978, Abcam) and mouse- α -tubulin (1:5000, Cat # T5168, Sigma Aldrich).

For western blots the rabbit- α -Acetyl-lysine from PTMscan was used, except when stated otherwise. Quantifications were performed using the Odyssey software (LI-COR Biosciences). The intensity of the acetyl-lysine signal was normalized against the histone H4 or tubulin signals and the mock treated time point was set as 1. Each experiment was performed at least three times and mean values and standard error of the means (SEM) are shown. A two-tailed t-test was performed and P-values <0.1 (*), <0.05 (**) and <0.005 (***) are depicted.

EdU incorporation

Cells were incubated with 20 μ M 5-ethynyl-2'-deoxyuridine (EdU, Invitrogen) and 1 μ M 5-Fluoro-2'-deoxyuridine for 2 hours. Cells were fixed in 3.6% formaldehyde in PBS and permeabilized in 0.5% Triton-X in PBS. EdU was visualized with a click-it reaction using Alexa fluorophore 488 nm according to manufactures protocol (Invitrogen). For Unscheduled DNA synthesis experiments, cells were labeled with 20 μ M EdU in medium containing 20 mM Hepes buffer (Lonza) directly after UV irradiation (16J/m²). After 3h of labeling a chase with 10 μ M thymidine was performed after which the cells were fixed in 3.7% formaldehyde and 0.5% Triton in PBS. Cells were shortly washed twice in PBS and blocked 2x in 3% BSA in PBS for 10 min. Permeabilization was performed in 0.5% Triton in PBS for 20 min. Click-it reaction was performed using a 594 nm azide (Atto Tec) for 30 min. Images were obtained using a LSM700 microscope (Carl Zeiss Microimaging Inc.) and analyzed using ImageJ software⁵⁴.

Clonogenic survival assays

HeLa cells were seeded in 6-well plates (400 cells/well) a day before treatment with a single doses of UV-C. Each experiment was performed in triplicate. After 6 days the colonies were fixed and stained with 50% methanol, 43% H₂O, 7% acetic acid and 0.1% Brilliant blue R (Sigma). Number of colonies was counted using a GelCountTM (Oxford Optronix, version 1.1.2.0). The survival was plotted as the relative amount of colonies after treatment compared to the non-treated samples.

Immunofluorescence

Cells were grown on coverslips until 70-80% confluency, washed in PBS and fixed in 2% paraformaldehyde in PBS containing 0.1% triton-X. Cells were permeabilized for 20 min in 0.5% Triton-X in PBS and washed in PBS containing 0.5% bovine serum albumin (BSA) and 0.15% glycine prior to primary antibody mouse- α -phospho-H2A.X (SER139) (1:1000, Cat # 05-636, Millipore) staining for 1 h at RT. Coverslips were washed three times short and twice for 10 min in 0.1% triton-X in PBS and once in PBS with BSA and glycine and

subsequently stained with secondary antibodies labeled with alexa fluorochromes 555 (Invitrogen) and DAPI (0.1 µg/ml) for 1 h at RT. Images were obtained using a Leica DM4000B microscope.

Data availability

The mass spectrometry data from this publication have been deposited to the PeptideAtlas database (<http://www.peptideatlas.org>) and assigned the identifier PASS01209 (<http://www.peptideatlas.org/PASS/PASS01209>).

Acknowledgements

This work is dedicated to the memory of Iris Kik. This work was supported by the Dutch Organization for Scientific Research (NWO) TOP Grants of Earth and Life Sciences and ZonMW (854.11.002 and 912.12.132), European Research Council Advanced Grant (340988-ERC-ID), Dutch Organization for Scientific Research Earth and Life Sciences VIDI grant (846.13.004) and Cancer Genomics Netherlands. This work is part of the OncoCode Institute which is partly financed by the Dutch Cancer Society.

Author contribution

I.M. performed the majority of the experiments. M.G., I.K and R.J. performed western blots. A.R. performed UDS experiments. K.B., E.R. and J.D. performed MS experiments and data analysis. I.M., J.H., H.L., W.V. and J.A.M. designed the experiments. I.M. and J.M. analyzed the data and wrote the manuscript. All authors reviewed the manuscript.

Conflict of interest

The authors declare they have no conflict of interest.

References

1. Zhu, P. & Li, G. Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr Opin Struct Biol* **36**, 106-15 (2016).
2. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).
3. Langst, G. & Manelyte, L. Chromatin Remodelers: From Function to Dysfunction. *Genes (Basel)* **6**, 299-324 (2015).
4. Escargueil, A.E., Soares, D.G., Salvador, M., Larsen, A.K. & Henriques, J.A. What histone code for DNA repair? *Mutat Res* **658**, 259-70 (2008).
5. van Attikum, H. & Gasser, S.M. Crosstalk between histone modifications during the DNA damage response. *Trends Cell Biol* **19**, 207-17 (2009).
6. Ramanathan, B. & Smerdon, M.J. Changes in nuclear protein acetylation in u.v.-damaged human cells. *Carcinogenesis* **7**, 1087-94 (1986).
7. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* **15**, 465-81 (2014).
8. Ramanathan, B. & Smerdon, M.J. Enhanced DNA repair synthesis in hyperacetylated nucleosomes. *J Biol Chem* **264**, 11026-34 (1989).
9. Smerdon, M.J. DNA repair and the role of chromatin structure. *Curr Opin Cell Biol* **3**, 422-8 (1991).
10. Green, C.M. & Almouzni, G. When repair meets chromatin. First in series on chromatin dynamics. *EMBO Rep* **3**, 28-33 (2002).
11. Teng, Y., Yu, Y. & Waters, R. The *Saccharomyces cerevisiae* histone acetyltransferase Gcn5 has a role in the photoreactivation and nucleotide excision repair of UV-induced cyclobutane pyrimidine dimers in the MFA2 gene. *J Mol Biol* **316**, 489-99 (2002).
12. Kim, M.K., Shin, J.M., Eun, H.C. & Chung, J.H. The role of p300 histone acetyltransferase in UV-induced histone modifications and MMP-1 gene transcription. *PLoS One* **4**, e4864 (2009).
13. Wang, J., Chin, M.Y. & Li, G. The novel tumor suppressor p33ING2 enhances nucleotide excision repair via inducement of histone H4 acetylation and chromatin relaxation. *Cancer Res* **66**, 1906-11 (2006).
14. Hasan, S., Hassa, P.O., Imhof, R. & Hottiger, M.O. Transcription coactivator p300 binds PCNA and may have a role in DNA repair synthesis. *Nature* **410**, 387-91 (2001).
15. Yu, Y., Teng, Y., Liu, H., Reed, S.H. & Waters, R. UV irradiation stimulates histone acetylation and chromatin remodeling at a repressed yeast locus. *Proc Natl Acad Sci U S A* **102**, 8650-5 (2005).
16. Duan, M.R. & Smerdon, M.J. Histone H3 lysine 14 (H3K14) acetylation facilitates DNA repair in a positioned nucleosome by stabilizing the binding of the chromatin Remodeler RSC (Remodels Structure of Chromatin). *J Biol Chem* **289**, 8353-63 (2014).
17. Guo, R., Chen, J., Mitchell, D.L. & Johnson, D.G. GCN5 and E2F1 stimulate nucleotide excision repair by promoting H3K9 acetylation at sites of damage. *Nucleic Acids Res* **39**, 1390-7 (2011).
18. Zhu, Q. *et al.* Damaged DNA-binding protein down-regulates epigenetic mark H3K56Ac through histone deacetylase 1 and 2. *Mutat Res* **776**, 16-23 (2015).
19. Tjeertes, J.V., Miller, K.M. & Jackson, S.P. Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J* **28**, 1878-89 (2009).
20. Bergink, S. *et al.* DNA damage triggers nucleotide excision repair-dependent monoubiquitylation of histone H2A. *Genes Dev* **20**, 1343-52 (2006).
21. Citterio, E. *et al.* ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor. *Molecular and Cellular Biology* **20**, 7643-7653 (2000).
22. Elia, A.E. *et al.* Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. *Mol Cell* **59**, 867-81 (2015).

23. Svinkina, T. *et al.* Deep, Quantitative Coverage of the Lysine Acetylome Using Novel Anti-acetyl-lysine Antibodies and an Optimized Proteomic Workflow. *Mol Cell Proteomics* **14**, 2429-40 (2015).
24. 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* **186**, 835-47 (2009).
25. Hanasoge, S. & Ljungman, M. H2AX phosphorylation after UV irradiation is triggered by DNA repair intermediates and is mediated by the ATR kinase. *Carcinogenesis* **28**, 2298-304 (2007).
26. Turner, B.M. Cellular memory and the histone code. *Cell* **111**, 285-91 (2002).
27. Kallin E, Z.Y. Chromatin remodeling. in *Encyclopedia of biological chemistry*. (ed. Lennarz WJ, L.M.) 456-462 (Academic Press, London, 2004).
28. Sadakierska-Chudy, A. & Filip, M. A comprehensive view of the epigenetic landscape. Part II: Histone post-translational modification, nucleosome level, and chromatin regulation by ncRNAs. *Neurotox Res* **27**, 172-97 (2015).
29. Wisniewski, J.R., Zougman, A., Kruger, S. & Mann, M. Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Mol Cell Proteomics* **6**, 72-87 (2007).
30. van Cuijk, L. *et al.* SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair. *Nat Commun* **6**, 7499 (2015).
31. Choi, J.H. *et al.* Highly specific and sensitive method for measuring nucleotide excision repair kinetics of ultraviolet photoproducts in human cells. *Nucleic Acids Res* **42**, e29 (2014).
32. Orren, D.K., Petersen, L.N. & Bohr, V.A. Persistent DNA damage inhibits S-phase and G2 progression, and results in apoptosis. *Mol Biol Cell* **8**, 1129-42 (1997).
33. Steurer, B. & Marteijn, J.A. Traveling Rocky Roads: The Consequences of Transcription-Blocking DNA Lesions on RNA Polymerase II. *J Mol Biol* (2016).
34. Bensaude, O. Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* **2**, 103-108 (2011).
35. Bensimon, A., Aebersold, R. & Shiloh, Y. Beyond ATM: the protein kinase landscape of the DNA damage response. *FEBS Lett* **585**, 1625-39 (2011).
36. Nam, E.A. & Cortez, D. ATR signalling: more than meeting at the fork. *Biochem J* **436**, 527-36 (2011).
37. Sobel, R.E., Cook, R.G., Perry, C.A., Annunziato, A.T. & Allis, C.D. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A* **92**, 1237-41 (1995).
38. Li, Q. *et al.* Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* **134**, 244-55 (2008).
39. Verreault, A. De novo nucleosome assembly: new pieces in an old puzzle. *Genes Dev* **14**, 1430-8 (2000).
40. Qian, M.X. *et al.* Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* **153**, 1012-24 (2013).
41. Shogren-Knaak, M.I., H.; Sun, J.M.; Pazin, M.J.; Davie, J.R.; Peterson, C.L. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844-847 (2006).
42. Agalioti, T.C., G.; Thanos, D. Deciphering the transcriptional histone acetylation code for a human gene. *Cell* **111**, 381-92 (2002).
43. Lans, H., Marteijn, J.A. & Vermeulen, W. ATP-dependent chromatin remodeling in the DNA-damage response. *Epigenetics & Chromatin* **5**(2012).
44. Soria, G., Polo, S.E. & Almouzni, G. Prime, repair, restore: the active role of chromatin in the DNA damage response. *Molecular Cell* **46**, 722-34 (2012).

45. Brand, M. *et al.* UV-damaged DNA-binding protein in the TFIIIC complex links DNA damage recognition to nucleosome acetylation. *EMBO J* **20**, 3187-96 (2001).
46. Rapic-Otrin, V., McLenigan, M.P., Bisi, D.C., Gonzalez, M. & Levine, A.S. Sequential binding of UV DNA damage binding factor and degradation of the p48 subunit as early events after UV irradiation. *Nucleic Acids Res* **30**, 2588-98 (2002).
47. Datta, A. *et al.* The p48 subunit of the damaged-DNA binding protein DDB associates with the CBP/p300 family of histone acetyltransferase. *Mutat Res* **486**, 89-97 (2001).
48. Yang, X. *et al.* Histone acetyltransferase 1 promotes homologous recombination in DNA repair by facilitating histone turnover. *Journal of Biological Chemistry* **288**, 18271-82 (2013).
49. Kurdistan, S.K., Tavazoie, S. & Grunstein, M. Mapping global histone acetylation patterns to gene expression. *Cell* **117**, 721-33 (2004).
50. Hauer, M.H. *et al.* Histone degradation in response to DNA damage enhances chromatin dynamics and recombination rates. *Nat Struct Mol Biol* **24**, 99-107 (2017).
51. Blickwedehl, J. *et al.* Role for proteasome activator PA200 and postglutamyl proteasome activity in genomic stability. *Proc Natl Acad Sci U S A* **105**, 16165-70 (2008).
52. Liang, D., Burkhart, S.L., Singh, R.K., Kabbaj, M.H. & Gunjan, A. Histone dosage regulates DNA damage sensitivity in a checkpoint-independent manner by the homologous recombination pathway. *Nucleic Acids Res* **40**, 9604-20 (2012).
53. Mailand, N., Gibbs-Seymour, I. & Bekker-Jensen, S. Regulation of PCNA-protein interactions for genome stability. *Nat Rev Mol Cell Biol* **14**, 269-82 (2013).
54. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-82 (2012).

Histone peptide	Modifications
(H2B) KGSKKAVTKAQKKDGKKR	4 Acetyl (K) 4 Acetyl (K);Dimethyl (KR)
(H3) QTARKSTGGKAPRKQLATKAAR	Acetyl (K);Methyl (KR) 2 Acetyl (K) 2 Acetyl (K);Phospho (STY) 2 Acetyl (K);Methyl (KR) 2 Acetyl (K);Dimethyl (KR) 3 Acetyl (K) 3 Acetyl (K);Phospho (STY) 3 Acetyl (K);Methyl (KR) 3 Acetyl (K);Dimethyl (KR) 4 Acetyl (K)
(H4) KGLGKGGAKR	3 Acetyl (K) Acetyl (K);GlyGly (K)

Figure EV1. Isolation of acetylated histone peptides. Table listing three representative examples of acetylated histone peptides from H2B, H3 and H4, isolated by acetyl-IP and identified with MS followed by MaxQuant analysis. In addition to acetylation marks, these peptides can be modified with additional histone PTMs.

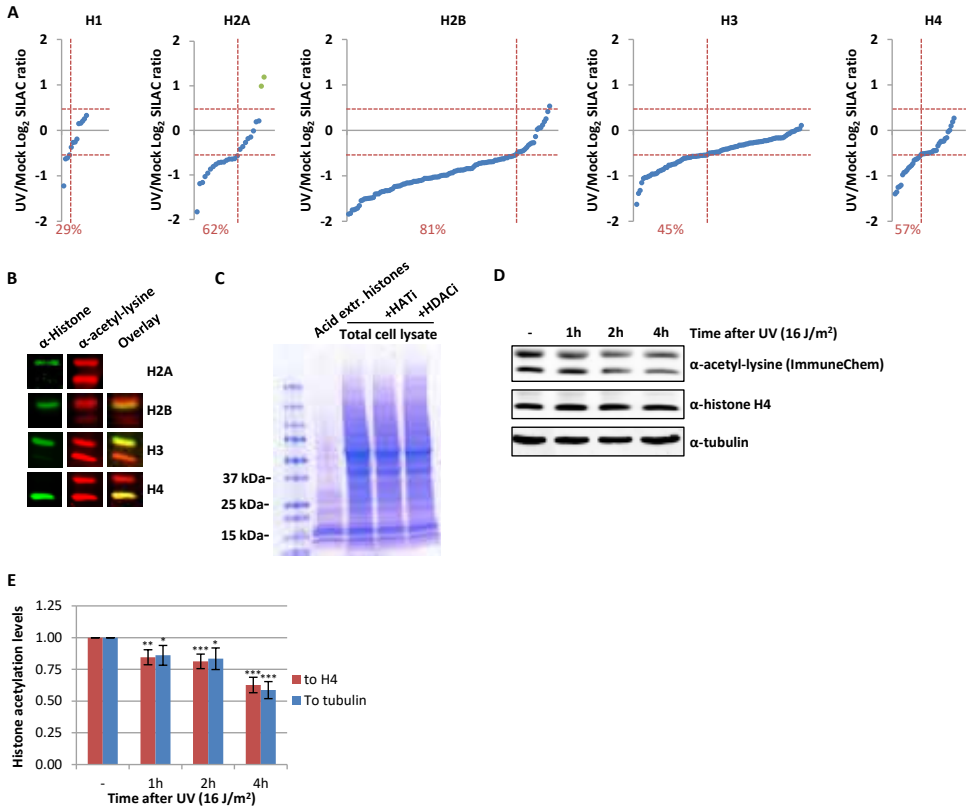


Figure EV2. UV-induced decrease in acetylated histones. (A) Identified acetylated peptides per histone plotted against their log₂ SILAC ratio (1h after 16 J/m² UV/mock), ranked by SILAC ratio. Green dots represent the identified peptides from histone variant H2AZ. The percentage of peptides with a SILAC ratio < -0.5 are indicated in red. (B) Western blots of WCE from HeLa cells stained with α-histone H2A, α-histone H2B, α-histone H3 α-histone H4 (depicted in green) and α-acetyl-lysine (depicted in red) indicating that the strong acetyl-lysine signal of the low molecular weight bands originates from acetylated core histones. H2A and acetyl-lysine antibodies were raised in the same species and therefore did not allow an overlay. (C) Coomassie staining of extracted histones and WCE from the same number of HeLa cells, treated with HATi (CTK7A, 100 μM and CPTH2, 50 μM) or HDACi (TSA, 1 μM) or mock treated. (D) Western blots of HeLa cells lysed at different time points after irradiation with UV (16 J/m²). Blots were stained with α-acetyl-lysine (Immunechem, top panel), α-histone H4 (middle panel) and α-tubulin (bottom panel). (E) Quantification of the α-acetyl-lysine signal normalized against either the histone H4 levels (red bars) or tubulin levels (blue bars). Error bars represent SEM. N= 7 and SEM is shown. Significant differences between UV treated and mock treated conditions are calculated with T-test and indicated with * (p<0.1), ** (p<0.05) and *** (p<0.01).

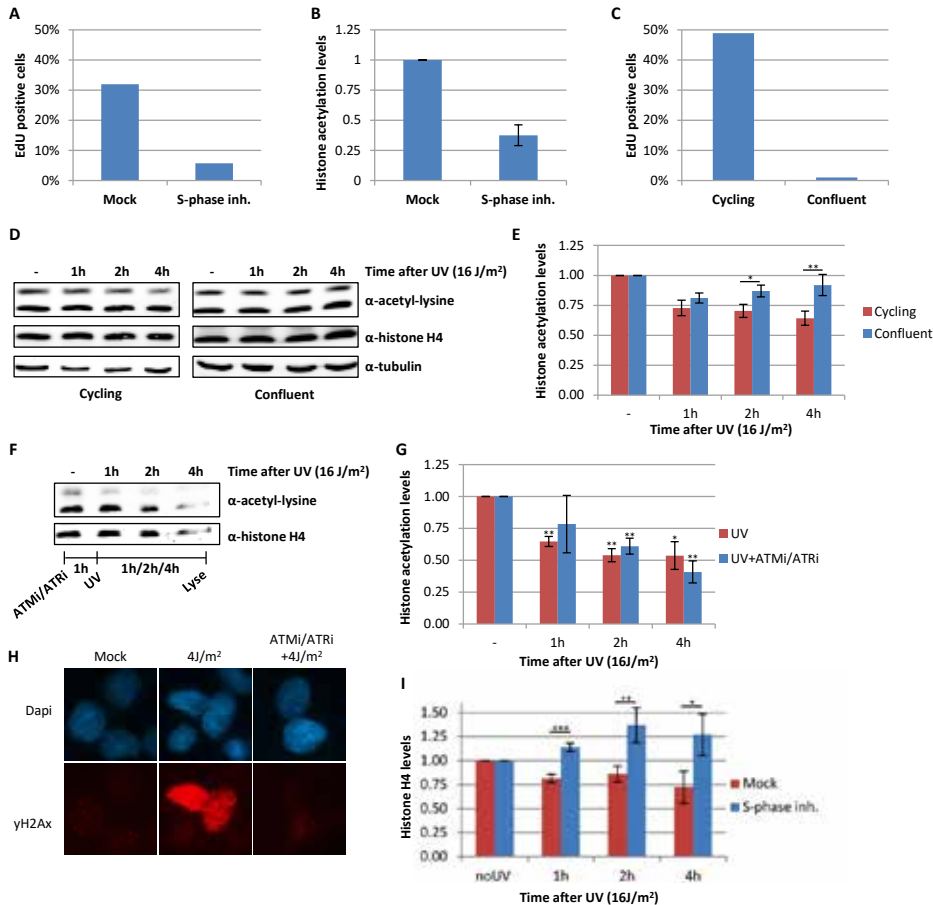


Figure EV3. Loss of acetylated histones is caused by replication stress. (A) Relative amount of HeLa cells positively stained for EdU incorporation (2h, 20 μM), quantified from immunofluorescence images (>120 cells analyzed). (B) Quantification of histone acetylation levels, normalized against histone H4 levels of HeLa cells that are cultured with S-phase inhibitor (10 μM, o/n) or mock treated. Error bars represent SEM. (C) Relative amount of EdU positive, cycling or contact inhibited (confluent) VH10 cells, quantified from immunofluorescence images (>100 cells analyzed). (D) Western blots of WCE from confluent (right panel) or cycling (left panel) VH10 cells lysed at different time points after UV (16J/m²). Blots were stained with α-acetyl-lysine (top panel), α-histone H4 (middle panel) and α-tubulin (bottom panel). (E) Quantification of histone acetylation levels in confluent or cycling VH10 cells at different time points in which the histone acetylation levels are normalized against histone H4 levels after UV (16J/m²). N= 3, error bars represent SEM. Significance of differences between cycling and confluent cells are indicated with * (p<0.1) and ** (p<0.05). (F) Western blot of WCE, made at different time points after UV irradiation (16J/m²) from HeLa cells pre-treated for 1 hr with a combination of ATM (Ku55933, 10 μM) and ATR (VE-821, 10 μM) inhibitors. (G) Quantification of histone acetylation levels in cells treated with ATM and ATR inhibitors an hour before UV irradiation (16J/m²). Bar graph represent mean, dots represent individual data points. N=2 (H) Representative immunofluorescence images of yH2AX staining following, mock treatment, UV-induced DNA damage (4 J/m²) with or without ATM and ATR inhibitors as indicated. Scale bar 10 μm. (I) Quantification of histone H4 levels in HeLa cells pre-treated o/n with an S-phase inhibitor (PHA 767491 hydrochloride, 10 μM) and lysed at indicated time points after UV (16 J/m²). N= 5, error bars represent SEM, data is normalized against tubulin levels. Significant differences, calculated with T-test, between UV-irradiated and inhibitor treated conditions are indicated with * (P<0.1), ** (p<0.05) and *** (p<0.01).

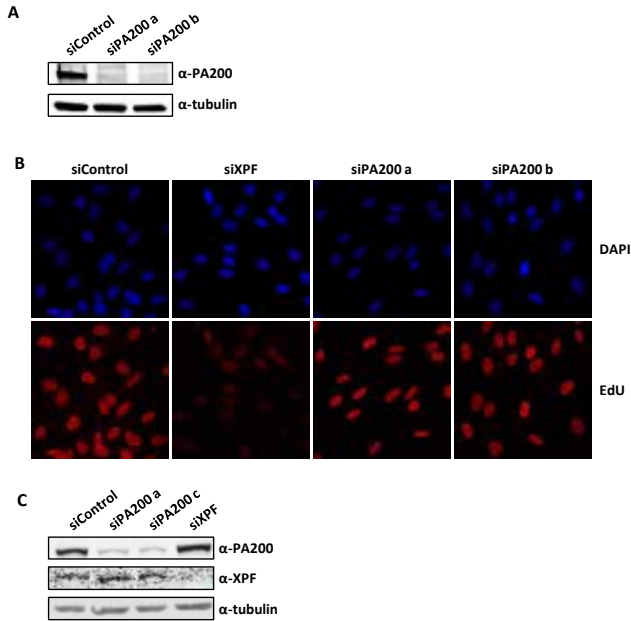


Figure EV4. Histone degradation does not affect repair of UV-lesions. (A) Western blot showing the knock down efficiency of the siRNAs targeting PA200. **(B)** Representative images of non-cycling C5RO cells treated with EdU (20 μ M) for 3h after UV irradiation (16J/m²) visualized by ATTO594 labelling using Click-chemistry. Scale bar 50 μ m. **(C)** Western blot showing the knock down efficiency of the siRNAs targeting PA200 and XPF.



5

GENERAL DISCUSSION

Faithful transcription is essential to maintain cellular homeostasis and is therefore tightly regulated at the different steps of the transcription cycle. However, transcription is daily threatened by endogenous and exogenous DNA damaging agents. These agents may induce transcription blocking lesions (TBLs) that impair progression of RNA polymerase II (Pol II). Impaired transcription can lead to a loss of *de novo* transcription, mutagenic transcripts, formation of R-loops¹ and transcription-replication conflicts. If TBLs are not repaired correctly, Pol II stalling at TBLs can lead to cellular dysfunction and premature aging^{2,3}. Moreover, transcription-replication conflicts can induce genome instability, eventually contributing to the development of cancer^{4,5}. Fortunately, all organisms are equipped with the dedicated TBL-removing DNA repair mechanism transcription-coupled nucleotide excision repair (TC-NER) to prevent these deleterious effects⁶. In addition to the direct transcription blockage induced by TBLs, Pol II stalling induces a genome-wide regulation of transcription, likely to prevent further catastrophe, to preserve genome stability and regulate repair. These direct and indirect consequences of TBLs, also called the transcription stress response, are described in chapter 1⁷. In this thesis, we focused on dissecting the DNA damage-induced transcription stress response by identifying new factors in this response using Pol II interaction proteomics (Chapter 2) and a genome-wide CRISPR/Cas9 screen (Chapter 3).

Over the years, several factors have been identified to be involved in directly regulating Pol II in response to DNA damage. However, likely not all factors involved have yet been identified. Identifying additional factors that are involved may provide a better insight in the mechanisms underlying the cellular response to DNA damage-induced transcription stress. One of the key steps remaining to be elucidated is how repair factors could gain access to lesions underneath TBL-stalled Pol II. It was proposed that Pol II could be removed through reverse translocation⁸ or could physically be removed⁹. Several factors have been suggested to be involved in this step. For example, TFIIH might be involved in reverse translocating Pol II^{7,10} and TFIIIS was suggested to stimulate the intrinsic exonuclease activity of Pol II to realign the extruding mRNA and facilitate transcription restart^{11,12}. However, the exact role and function of these proteins is not clear yet and other factors might be involved in this processing step. In chapter 2, we used a SILAC-based interaction proteomics approach to isolate Pol II and identify interactors involved in regulating Pol II.

Although Pol II is a tightly regulated protein, surprisingly, only few labs have studied the Pol II complex composition using quantitative mass spectrometry¹³⁻¹⁵ and even fewer have studied its composition in response to UV-induced DNA damage¹⁶. This already suggests that studying the Pol II complex composition is a challenging enterprise as described in chapter 2. This difficulty is likely caused by the tight binding of Pol II to the chromatin, making it difficult to extract this complex together with its interactors. Treating the chromatin with stringent fragmentation methods resulted in the desired release of Pol II from chromatin, but unfortunately also in the loss of Pol II or its interactors by degradation or in release of the interactors from the Pol II complex. Nevertheless, we were able to successfully extract and immunoprecipitate the Pol II complex and we identified several basal transcription factors as Pol II interactors. However, identifying new UV-induced interactors remained challenging. This could partially be explained by

proteins that are part of the elongation complex in undamaged conditions, which are also involved in regulating Pol II function in response to DNA damage. An example of such an interactor is ELOF1, as described in chapter 3 and discussed below. Moreover, the pool of Pol II in cells is rather large and only a small fraction of Pol II will be bound to lesions at a given moment. Possibly, this lesion-bound fraction will be too small to identify DNA damage-induced Pol II interactors using mass spectrometry. Increasing the UV dose to generate more TBLs and thereby presumably more lesion-bound Pol II might not be efficient, as with higher UV doses Pol II might be degraded¹⁷. Alternatively, accumulating lesion-stalled Pol II complexes by preventing progression of the repair reaction might be successful. However, as discussed in Chapter 2, accumulating lesion-stalled Pol II using NER-deficient XPA KO cells did not enhance the number of identified interactors. Since XPA functions rather downstream in NER where it helps TFIIH to verify the lesion and is a central organizer of the NER incision complex^{6,18,19}, its absence might arrest the repair reaction at a step where Pol II has already been removed from the lesion. Therefore, it may be better to use knock-out cells of upstream factors that specifically function in TC-NER to accumulate lesion-stalled Pol II and consequently immunoprecipitate these Pol II complexes. In this case, it might be helpful to additionally inactivate GG-NER to further enhance the number of lesion-stalled complexes by knocking out XPC. Indeed, using ELOF1 KO cells showed that we could accumulate CSA and CSB at TBLs since they remain immobilized and bound to Pol II five hours after damage induction (Chapter 3). ELOF1 functions at the beginning of the repair reaction and therefore might sufficiently accumulate lesion-stalled Pol II to identify new UV-specific interactors. Alternatively, we could make use of a CSA or CSB KO cell line, which are even further upstream or make use of a UVSSA ubiquitin dead mutant. In this case, we would inhibit the currently known last step of the TC-NER pathway, before TFIIH is recruited and the combined NER pathway is initiated.

Thus far, we mainly focused on isolating new factors involved in regulating Pol II in response to DNA damage and the DNA repair reaction itself. However, an equally important step of the transcription stress response is the resumption of transcription following successful removal of the TBL. Although several factors have been described to be involved in restarting transcription²⁰, the exact mechanism and factors involved have yet to be determined. We can use our SILAC-based mass spectrometry approach to study Pol II-interacting proteins at later time points after UV, particularly at a time point that transcription is being resumed. As shown in Chapter 3, CSA and CSB binding to Pol II and their immobilization is resolved five hours after 4 J/m² UV-C irradiation. This suggests that from this point onwards, on average, transcription could be resumed. The moment transcription is resumed will depend on the provided damage load of UV since higher damage load may take longer to be repaired and resume transcription.

An alternative approach for the identification of DNA damage response factors is using genetic screens like genome-wide pooled CRISPR/Cas9 libraries²¹. CRISPR/Cas9 targets specific genomic loci and could thereby introduce stable loss-of-function mutations in protein coding genes that facilitate higher screening sensitivity than shRNA libraries²¹⁻²³. CRISPR/Cas9 screening is unbiased and uses the selective pressure of a treatment to

identify genes that are important for cell survival²¹. The development of this new, sequence-specific technique enables the identification of previously unidentified genes in the DNA damage response upon exposure to many genotoxic agents as can be seen in the paper by Olivieri *et al.*²⁴. In chapter 3, we described a CRISPR/Cas9 screen using UV irradiation as a selective pressure for identifying new genes involved in regulating the transcription stress response.. Olivieri *et al.* used Illudin S to identify new genes involved in the DNA damage response. Illudin S is a natural occurring drug isolated from mushrooms that induces lesions that are specifically recognized and repaired by the TC-NER pathway but not by global genome (GG-)NER²⁵. Therefore, it can be used to more specifically identify new TC-NER factors than using UV irradiation and might therefore be a more interesting genotoxic agent to identify transcription stress response proteins. More importantly, it can be used to identify Pol II-interacting proteins that are already binding irrespective of DNA damage but which interaction it changed upon DNA damage, for example the interaction of Pol II and ELOF1. This underscores the importance of using complementary types of screens to identify novel DNA damage response genes, as exemplified by the different hits identified in our CRISPR/Cas9 and UV-induced Pol II interaction proteomics screens described in this thesis.

To even more specifically identify factors involved in the DNA damage induced transcription stress pathway, it might be interesting to make use of the CRISPR/Cas9 screen in combination with labeling nascent RNA using 5-Ethynyl-uridine (EU) to identify restart factors. Transcription restart can be visualized using click-it chemistry²⁶ and FACS sorting can be used to select cells based on their transcription levels. sgRNA-transduced cells that cannot restart transcription as characterized by low EU incorporation, are interesting candidates which might be involved in transcription restart. In a similar way we can also apply this method to identify genes that are important for TC-NER by labeling with 5-ethynyl-2'-deoxyuridine (EdU)²⁷. EdU is used to label DNA damage-induced DNA replication in non-cycling cells and was previously used in XPC knock out cells to determine the unscheduled DNA synthesis specifically during TC-NER as a measure for successful repair²⁸. By combining such an assay with a CRISPR/Cas9 library and label them with EdU in response to UV irradiation, we expect a low EdU level in TC-NER-deficient cells. FACS sorting the cells based on the EdU levels will identify genes involved in the repair of TBLs.

Further studying the function of ELOF1 showed that it is essential for functional TC-NER and that, in its absence, Pol II cannot be ubiquitinated at Lysine 1268 (K1268) in response to UV and UVSSA was not recruited to the TBL (chapter 3). This suggests that ELOF1 is involved in ubiquitinating Pol II, which may occur through two different mechanisms. First, ELOF1 might facilitate reorganization of the lesion-stalled Pol II complex including the TC-NER components. In that case, ELOF1 could be responsible for repositioning the CRL4^{CSA} E3 ubiquitin ligase complex in such a way that it could ubiquitinate Pol II. Second, ELOF1 could be recruiting a yet unknown E3 ubiquitin ligase that specifically ubiquitinates Pol II at K1268. Even though, the CRL4^{CSA} complex has a flexible region with a big reach, this seems to be the less likely option since ELOF1 binds the DNA entry tunnel of Pol II while CSB and thus CSA bind the DNA extruding from Pol II^{29,30} and are less likely to directly interact with ELOF1. However, the exact mechanism how ELOF1

facilitates ubiquitylation of Pol II and consequently the recruitment of UVSSA has yet to be determined.

Impaired recruitment of UVSSA to Pol II in the absence of ELOF1 also resulted in the impaired recruitment of TFIIF, since UVSSA recruits TFIIF via a direct interaction with p62³¹. The interaction of UVSSA to Pol II is stabilized in response to DNA damage and this facilitates ubiquitylation of UVSSA³². The ubiquitylation of UVSSA induces recruitment of TFIIF to Pol II³². Since TFIIF was suggested to be involved in backtracking of Pol II due to its translocase activity (Chapter 1)⁷, this could imply that in the absence of ELOF1, Pol II cannot be reverse translocated. This could lead to more persistent Pol II stalling that could be further enhanced by the constant forward translocase activity of CSB, that pushes Pol II onto the lesion (Chapter 1)⁷. This persistent stalling could lead to induction of transcription-replication conflicts⁴, which is indeed observed in ELOF1 KO cells (Chapter 3). Surprisingly, in CSB KO cells, although TC-NER is similarly affected as in ELOF1 KO cells, transcription-replication conflicts do not occur to the same extent. This might be explained by the absence of CSB, that does not keep on pushing Pol II on the TBLs and that therefore Pol II might be released from the lesion by advancing replication forks. However, in the absence of CSA there is also additive sensitivity for ELOF1 to UV, this suggests that the absence of CSB's forward translocating activity is not the only reason that there are less transcription-replication conflicts induced in CSA- or CSB-deficient cells. The absence of these conflicts in CSA- and CSB-deficient cells suggests that Pol II is stalled differently on a lesion in the absence of ELOF1 or suggests that ELOF1 has a more active role in preventing these conflicts. The latter seems more likely as in siRNA-depleted cells, Pol II stalling is similar without CSB and ELOF1. It is interesting to speculate that ELOF1 might be involved in preventing transcription-replication conflicts by actively regulating release of Pol II from the chromatin or that in the absence of ELOF1 Pol II is differently stalled on a lesion as ELOF1 acts near the DNA entry tunnel of Pol II. Finally, as ELOF1 KO and CSB KO stall Pol II to the same extent but only ELOF1 KO leads to transcription-replication conflicts, comparing the UV-induced Pol II interactors in these cells to might reveal proteins that are involved in resolving transcription-replication conflicts.

In addition to its uncovered role in TC-NER, we showed that ELOF1 also is a transcription elongation factor. In the absence of ELOF1, the transcription elongation rate is clearly reduced, possibly due to increased stalling of Pol II at endogenous lesions. This can be seen using SILAC-based interaction proteomics in ELOF1 KO cells where in unperturbed conditions, CSA and CSB already bind more to Pol II than in Wt cells. Interestingly, using SILAC-based interaction proteomics (Chapter 2), we identified RECQL5, a candidate that specifically binds Pol II in response to UV but has the opposite effect on transcription in unperturbed conditions compared to ELOF1. RECQL5, which has previously been identified as a direct interactor of Pol II¹³, is important for inhibiting transcription elongation and thereby allows dispersed transcription^{33,34}. In the absence of RECQL5, transcription is much faster and more frequently stalled, resulting in genome instability³⁴. Interestingly, both the depletion of ELOF1 and RECQL5 lead to genome instability without DNA damage infliction, however, the mechanism is different as depletion of ELOF1 inhibits transcription elongation whereas depletion of RECQL5 increases transcription

elongation. This suggests that the transcription elongation rate should be in perfect balance to maintain faithful transcription, prevent increased Pol II stalling and eventually prevent genome instability.

DNA transacting processes like replication, transcription and DNA repair have to function in the dynamic chromatin environment. Therefore, the chromatin organization needs to be actively remodeled to provide repair factors access to the lesion³⁵, to restore chromatin after repair and to facilitate transcription restart^{20,36,37}. Chromatin reorganization is for example regulated by remodeling of the nucleosomes, eviction or exchange of histones, and histone modifications³⁶. Histone modifications can be responsible for changing nucleosome-DNA contacts, correct nucleosome assembly and for example for degradation of histones. The latter was shown for acetylation^{38,39} and is in line with our results shown in Chapter 4 where we illustrate the ubiquitylation-independent PA200 proteasome-mediated degradation of acetylated histones upon DNA damage⁴⁰. Such degradation of histones promotes decompaction and induces mobility of the chromatin, which has been shown to facilitate DNA damage repair and overcome replication fork stalling⁴¹. Previously, in yeast, acetylation on newly synthesized H3K56 has been shown to be important for correct nucleosome assembly during incorporation of the histones into newly replicated DNA during S-phase and the absence of this acetylation inhibits DNA replication³⁸. Moreover, H3K56 acetylation and H3K9 acetylation were also reduced upon DNA damage in eukaryotic cells⁴².

We propose that acetylated histones are being degraded in response to replication stress to inhibit replication or to stimulate repair of stalled replication forks. Of note, this degradation was independent of ongoing transcription and is therefore likely not induced by lesion-stalled Pol II or the induction of TC-NER. However, opening of the chromatin may stimulate TC-NER or result in more transcription in the presence of replication stress and thereby might result in more transcription-replication conflicts. Therefore, it would be interesting to determine if inhibiting the PA200 proteasome upon DNA damage, that interferes with transcription and replication, increases transcription-replication conflicts.

In this thesis, we focused on the UV-induced transcription stress response. Although UV irradiation is a widely studied method to induce TBLs, it will be very interesting to use other types of DNA damaging agents. Especially since different types of TBLs are expected to induce a different transcriptional response or trigger different factors involved in regulating the repair and cellular responses to these TBLs. For example, oxidative damage induces mainly lesion bypass rather than long-term stalling^{30,43,44} and therefore might use and recruit different proteins than used during repair of UV-induced lesions. Moreover, the various consequences for Pol II when confronted with structurally different TBLs may trigger differential interactors. For instance, platinum agents such as cisplatin induce dGpG crosslinks that inhibit transcription and are repaired by TC-NER^{18,45,46} but the stalling of Pol II is mechanistically different as upon UV. In case of an intrastrand crosslink, this is caused by bending of the DNA⁴⁷. These crosslinks are likely not entering the active site of Pol II but Pol II is stalled in front of the lesion⁴⁸. In contrast, UV-induced lesions such as CPDs can enter the catalytic site of Pol II and

thereby trigger its stalling and this will result in Pol II covering the lesion⁴⁹⁻⁵¹. 8-oxo-G lesions are caused by endogenous reactive oxygen species, which are mainly repaired by base excision repair (BER) and Pol II is proposed to stall at the base excision repair intermediate instead of directly at the lesion⁵²⁻⁵⁴. DNA protein-crosslinks can be caused by aldehydes as well as topoisomerase-inhibiting drugs⁶ and prevent progression of Pol II⁵⁵⁻⁵⁷. Having different mechanisms to stall Pol II on a lesion could indicate that the mechanism of initiating repair is also different and therefore might require alternative factors to induce repair, regulate transcription effects *in trans*, or stimulate transcription restart. Therefore, it will be interesting to compare if there are different factors involved in the response to different types of TBLs. First, we can identify the interactors binding to Pol II in response to these different types of lesions by using SILAC-based mass spectrometry and compare them to UV-induced interactors. Pol II-interacting proteins that are specifically recruited to a certain type of lesion are interesting for follow-up studies. Second, we could use a CRISPR/Cas9 screen to identify DNA damage response genes in response to different lesions. TBLs that require different genes to survive are particularly interesting for follow-up studies to further understand the mechanism of the DNA damage-induced transcription stress response. Moreover, when candidates are identified using multiple DNA damaging agents, they are more likely to be a trustworthy candidate to be involved in an overall mechanism. For example, using CRISPR/cas9 screens following different DNA damaging agents, successfully identified most of the TC-NER factors²⁴.

In this thesis, we aimed to identify new regulators involved in the DNA damage-induced transcription stress response to improve our comprehension of the molecular mechanism of this highly orchestrated response to transcription-blocking DNA damage. Understanding the transcription stress response will provide more insight in DNA damage-induced ageing and the induction of genome instability as a consequence of transcription-replication conflicts. Identifying ELOF1 as a new DNA damage response gene and further unraveling its function helped to reveal its mechanism in the transcription stress response and solved a small piece of the transcription stress response puzzle. More research is needed to provide a complete picture of the factors involved in regulating the cellular response to transcription stress and dissect the complete molecular mechanism.

References

1. Crossley, M.P., Bocek, M. & Cimprich, K.A. R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell* **73**, 398-411 (2019).
2. Gómez-González, B. & Aguilera, A. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes & development* **33**, 1008-1026 (2019).
3. López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194-1217 (2013).
4. Gaillard, H. & Aguilera, A. Transcription as a Threat to Genome Integrity. *Annu Rev Biochem* **85**, 291-317 (2016).
5. Aguilera, A. & Garcia-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell* **46**, 115-24 (2012).
6. Lans, H., Hoeijmakers, J.H.J., Vermeulen, W. & Marteijn, J.A. The DNA damage response to transcription stress. *Nat Rev Mol Cell Biol* **20**, 766-784 (2019).
7. Geijer, M.E. & Marteijn, J.A. What happens at the lesion does not stay at the lesion: Transcription-coupled nucleotide excision repair and the effects of DNA damage on transcription in cis and trans. *DNA Repair (Amst)* (2018).
8. Wang, W., Xu, J., Chong, J. & Wang, D. Structural basis of DNA lesion recognition for eukaryotic transcription-coupled nucleotide excision repair. *DNA repair* **71**, 43-55 (2018).
9. Chiou, Y.Y., Hu, J., Sancar, A. & Selby, C.P. RNA polymerase II is released from the DNA template during transcription-coupled repair in mammalian cells. *J Biol Chem* **293**, 2476-2486 (2018).
10. Li, C.-L. *et al.* Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Molecular Cell* **59**, 1025-1034 (2015).
11. Sigurdsson, S., Dirac-Svejstrup, A.B. & Svejstrup, J.Q. Evidence that Transcript Cleavage Is Essential for RNA Polymerase II Transcription and Cell Viability. *Molecular Cell* **38**, 202-210 (2010).
12. Kettenberger, H., Armache, K.J. & Cramer, P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**, 347-57 (2003).
13. Aygün, O., Svejstrup, J. & Liu, Y. A RECQ5–RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proceedings of the National Academy of Sciences* **105**, 8580 (2008).
14. Harlen, Kevin M. *et al.* Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Reports* **15**, 2147-2158 (2016).
15. Lynch, C.J. *et al.* The RNA Polymerase II Factor RPAP1 Is Critical for Mediator-Driven Transcription and Cell Identity. *Cell reports* **22**, 396-410 (2018).
16. Boeing, S. *et al.* Multiomic Analysis of the UV-Induced DNA Damage Response. *Cell Reports* **15**, 1597-1610 (2016).
17. Wilson, M.D., Harreman, M. & Svejstrup, J.Q. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim Biophys Acta* **1829**, 151-7 (2013).
18. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* **15**, 465-81 (2014).
19. Kokic, G. *et al.* Structural basis of TFIIH activation for nucleotide excision repair. *Nature Communications* **10**, 2885 (2019).
20. Mandemaker, I.K., Vermeulen, W. & Marteijn, J.A. Gearing up chromatin: A role for chromatin remodeling during the transcriptional restart upon DNA damage. *Nucleus* **5**, 203-10 (2014).
21. Shalem, O., Sanjana, N.E. & Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nature Reviews Genetics* **16**, 299-311 (2015).

22. Evers, B. *et al.* CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* **34**, 631-3 (2016).
23. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343**, 84-87 (2014).
24. Olivieri, M. *et al.* A Genetic Map of the Response to DNA Damage in Human Cells. *Cell* **182**, 481-496. e21 (2020).
25. Jaspers, N.G. *et al.* Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. *DNA Repair (Amst)* **1**, 1027-38 (2002).
26. Jao, C.Y. & Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proceedings of the National Academy of Sciences* **105**, 15779 (2008).
27. Salic, A. & Mitchison, T.J. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proceedings of the National Academy of Sciences* **105**, 2415 (2008).
28. Wienholz, F., Vermeulen, W. & Marteijn, J.A. Amplification of unscheduled DNA synthesis signal enables fluorescence-based single cell quantification of transcription-coupled nucleotide excision repair. *Nucleic Acids Res* **45**, e68 (2017).
29. Ehara, H. *et al.* Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363**, 744-747 (2019).
30. Xu, J. *et al.* Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653-657 (2017).
31. Okuda, M., Nakazawa, Y., Guo, C., Ogi, T. & Nishimura, Y. Common TFIIH recruitment mechanism in global genome and transcription-coupled repair subpathways. *Nucleic Acids Res* **45**, 13043-13055 (2017).
32. Nakazawa, Y. *et al.* Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair. *Cell* **180**, 1228-1244 e24 (2020).
33. Aygün, O. *et al.* Direct inhibition of RNA polymerase II transcription by RECQL5. *J Biol Chem* **284**, 23197-203 (2009).
34. Saponaro, M. *et al.* RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell* **157**, 1037-49 (2014).
35. Lans, H., Marteijn, J.A. & Vermeulen, W. ATP-dependent chromatin remodeling in the DNA-damage response. *Epigenetics Chromatin* **5**, 4 (2012).
36. Polo, S.E. Reshaping Chromatin after DNA Damage: The Choreography of Histone Proteins. *Journal of Molecular Biology* **427**, 626-636 (2015).
37. Wienholz, F. *et al.* FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res* **47**, 4011-4025 (2019).
38. Li, Q. *et al.* Acetylation of Histone H3 Lysine 56 Regulates Replication-Coupled Nucleosome Assembly. *Cell* **134**, 244-255 (2008).
39. Qian, M.-X. *et al.* Acetylation-Mediated Proteasomal Degradation of Core Histones during DNA Repair and Spermatogenesis. *Cell* **153**, 1012-1024 (2013).
40. Mandemaker, I.K. *et al.* DNA damage-induced replication stress results in PA200-proteasome-mediated degradation of acetylated histones. *EMBO Rep* **19**(2018).
41. Hauer, M.H. *et al.* Histone degradation in response to DNA damage enhances chromatin dynamics and recombination rates. *Nat Struct Mol Biol* **24**, 99-107 (2017).
42. Tjeertes, J.V., Miller, K.M. & Jackson, S.P. Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *The EMBO Journal* **28**, 1878-1889 (2009).
43. Charlet-Berguerand, N. *et al.* RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *Embo J* **25**, 5481-91 (2006).

44. Damsma, G.E. & Cramer, P. Molecular basis of transcriptional mutagenesis at 8-oxoguanine. *J Biol Chem* **284**, 31658-63 (2009).
45. Enoiu, M., Jiricny, J. & Scharer, O.D. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res* **40**, 8953-64 (2012).
46. Slyskova, J. *et al.* Base and nucleotide excision repair facilitate resolution of platinum drugs-induced transcription blockage. *Nucleic Acids Res* **46**, 9537-9549 (2018).
47. Johnstone, T.C., Park, G.Y. & Lippard, S.J. Understanding and improving platinum anticancer drugs--phenanthriplatin. *Anticancer Res* **34**, 471-6 (2014).
48. Damsma, G.E., Alt, A., Brueckner, F., Carell, T. & Cramer, P. Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* **14**, 1127-33 (2007).
49. Tornaletti, S. Transcription arrest at DNA damage sites. *Mutat Res* **577**, 131-45 (2005).
50. Tornaletti, S., Reines, D. & Hanawalt, P.C. Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. *J Biol Chem* **274**, 24124-30 (1999).
51. Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. *Science* **315**, 859-62 (2007).
52. Kathe, S.D., Shen, G.P. & Wallace, S.S. Single-stranded breaks in DNA but not oxidative DNA base damages block transcriptional elongation by RNA polymerase II in HeLa cell nuclear extracts. *J Biol Chem* **279**, 18511-20 (2004).
53. 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* **39**, 5926-34 (2011).
54. Menoni, H. *et al.* The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage. *Nucleic Acids Res* **46**, 7747-7756 (2018).
55. Fielden, J., Ruggiano, A., Popović, M. & Ramadan, K. DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans. *DNA Repair (Amst)* **71**, 198-204 (2018).
56. Solier, S. *et al.* Transcription poisoning by Topoisomerase I is controlled by gene length, splice sites, and miR-142-3p. *Cancer Res* **73**, 4830-9 (2013).
57. Veloso, A. *et al.* Genome-wide transcriptional effects of the anti-cancer agent camptothecin. *PLoS One* **8**, e78190 (2013).



APPENDIX

Summary

Samenvatting

Curriculum Vitae

List of publications

PhD portfolio

Acknowledgements

Summary

The genetic code for all organisms is stored in the nucleotide sequence of their DNA. Protecting the genetic code is essential to faithfully transcribe the genetic information of our DNA into RNA. Progression of RNA polymerase II (Pol II), the protein involved in RNA transcription, is regularly impeded by DNA damage. Such transcription-blocking lesions (TBLs) can be induced by endogenously produced reactive oxygen species formed as side products from metabolic processes, or by exogenous sources such as UV irradiation and chemicals. Elongating Pol II that encounters a TBL can cause reduced transcription fidelity that eventually can result in mutant RNA transcripts or can even completely block the synthesis of new mRNAs. This transcriptional burdens can lead to reduced cell function and eventually induce DNA damage-induced aging. Moreover, if stalling of Pol II is persistent, this can induce the formation of R-loops and induce transcription-replication conflicts that can lead to genome instability and eventually can induce cancer. Collectively, these transcriptional impediments and consequences are referred to as the DNA damage-induced transcription stress response.

To prevent these detrimental effects of prolonged stalling, cells have a dedicated repair mechanism called transcription-coupled nucleotide excision repair (TC-NER) that repairs TBLs and is initiated upon stalling of Pol II at a lesion. Lesion-stalled Pol II results in the recruitment of TC-NER factors that help to recognize and verify the lesion and excise the damaged part of the DNA. Eventually, new DNA is synthesized and the DNA is ligated after which transcription can be resumed. During the repair reaction, Pol II needs to be displaced to allow repair factors to access the lesion. Moreover, in addition to the direct consequences of stochastic Pol II stalling, transcription is also highly regulated at a genome-wide level, as part of the DNA damage-induced transcription stress response. These include complex activation and subsequent inhibition of transcription initiation, a release of promoter-paused Pol II into the gene body and changes in alternative splicing. These events are all important for cells to cope with TBLs and contribute to recognizing the lesions, initiating repair and maintaining transcription fidelity. In **Chapter 1**, the different direct and indirect consequence of TBLs on Pol II and genome-wide transcription regulation have been reviewed. In addition, we described the factors that have thus far been identified in regulating this cellular response to DNA damage-induced transcription stress including the TC-NER reaction.

Although several factors have been identified to be involved in regulating the cellular response to DNA damage-induced transcription stress, the exact mechanisms of this pathway have yet to be determined. To identify new proteins involved in regulating Pol II in response to UV-induced DNA damage, we used quantitative interaction proteomics. In **Chapter 2**, we described different approaches to extract and immunoprecipitate Pol II and its interactors. First, we used a native extraction protocol to isolate both the initiating and elongation fraction of Pol II. Subsequently, we used these fractions to successfully immunoprecipitate Pol II and used SILAC-based interaction proteomics to identify its basal and UV-induced interactors. We identified several Pol II complex components and regulatory factors but, surprisingly, only a few known and no new UV-specific interactors. Therefore, we switched to a protocol that includes chemical crosslinking to stabilize loosely associated and transiently interacting factors. We assessed both PFA and DSP as crosslinking agents

to identify transient interactors. Both agents successfully identified Pol II interactors but again only a few UV-specific interactors, besides the known UV-induced interactor CSB, were identified. Even inhibiting degradation of proteins using VCP segregase inhibition, which extracts ubiquitylated Pol II from the chromatin to allow proteasomal degradation, or accumulating NER-intermediates, using NER-deficient XPA knock out cells, did not improve identification of UV-induced interactors of Pol II and left just a few interesting new candidates. Alternatively, we used two independent proximity labeling methods to covalently biotin-tag proteins that are in close proximity of Pol II. Both approaches identified basal complex interactors but again few new UV-induced interactors were identified. Finally, we used a native fractionation method to isolate Pol II and focused on immunoprecipitating elongating Pol II to enrich for lesion-stalled complexes. This procedure successfully identified multiple of the known UV-induced Pol II interactors but also identified only a few additional interactors. Finally, we combined the mass spectrometry results of the different extraction methods and discuss the putative implications of UV-specific Pol II interactors that were identified in four or more experiments.

To further identify new factors involved in regulating the cellular response to transcription stress, we performed a genome-wide CRISPR/Cas9 loss-of-function screen to identify genes involved in the UV-induced DNA damage response, **Chapter 3**. In this screen, we identified many known TC-NER factors such as *CSB*, *CSA* and *UVSSA* and additionally identified Elongation Factor 1 Homolog (*ELOF1*) as a new UV-sensitive gene. During basal conditions, *ELOF1* is an integral part of the transcription elongation complex. In the absence of *ELOF1*, transcription resumption in response to UV was severely impaired due to impaired TC-NER. We also showed that this repair function was highly conserved between species. Further studying the function of *ELOF1* in TC-NER revealed that without *ELOF1*, *UVSSA* and thus the downstream repair factors, like the TFIIH complex, could not be recruited. This is likely a result of the absence of Pol II ubiquitylation, which suggests that *ELOF1* is important for facilitating this ubiquitylation. Interestingly, we showed that there is an additive sensitivity when both *ELOF1* and *CSB* were knocked out, suggesting an additional role for *ELOF1* outside of TC-NER. We finally showed that this supplementary function was linked to preventing transcription-replication conflicts.

In **Chapter 4**, we described another regulatory mechanism implicated in the UV-induced DNA damage response. Here, using a SILAC-based mass spectrometry approach to specifically study changes in histone PTMs, we show that there is a genome-wide loss of acetylated histones. This loss of acetylation is independent of transcription but can be induced by different types of DNA damage that induce replication-stress. We showed that the histone-wide deacetylation is a result of degradation of acetylated histones by the PA200 proteasome via a ubiquitin-independent mechanism. This uncovered replication stress-induced degradation of acetylated histones represents an important chromatin-modifying response to cope with replication stress.

Finally, in **Chapter 5**, we summarized the main findings of the in this thesis described experimental work and discussed their implications. We propose future experiments that can be helpful to expand our knowledge on the DNA damage-induced transcription stress response.

Samenvatting

De genetische code voor alle organisme ligt opgeslagen in de nucleotide volgorde van hun DNA en vormt de basis voor alle eiwitten. Het beschermen van deze code is essentieel om de genetische informatie van ons DNA foutloos over te schrijven in RNA, dit proces wordt transcriptie genoemd. RNA polymerase II (Pol II) is het eiwitcomplex dat verantwoordelijk is voor het overschrijven van de genetische informatie van het DNA in RNA. Echter, de voortgang van Pol II wordt regelmatig verstoord door schades in het DNA. Deze zogenoemde transcriptie-blokkerende laesies (TBLs) kunnen veroorzaakt worden door schadelijke stoffen die gevormd worden in ons lichaam als bijproduct van ons metabolisme of door stoffen van buiten ons lichaam zoals chemicaliën of UV-straling in zonlicht. Wanneer Pol II tijdens het aflezen van het DNA een TBL tegen komt kan dit de betrouwbaarheid van het overschrijven van het DNA verminderen of zelfs helemaal blokkeren. Dit kan uiteindelijk leiden tot het foutief kopiëren van het DNA of de volledige afwezigheid van nieuwe RNA-kopieën. Deze problemen met het overschrijven van het DNA kunnen leiden tot een verminderde cel functie en uiteindelijk tot versnelde veroudering. Als Pol II langdurig geblokkeerd blijft, kan dit leiden tot botsingen met de eiwitten die het DNA kopiëren, het replicatie mechanisme. Deze zogenoemde transcriptie-replicatie conflicten (TRCs) kunnen instabiliteit van het genoom veroorzaken wat uiteindelijk kan bijdragen aan het ontstaan van kanker. De cel heeft verschillende systemen ontwikkeld om met DNA-schade om te kunnen gaan die het overschrijven van ons DNA belemmeren. Deze verschillende mechanismen worden gezamenlijk de DNA-schade geïnduceerde transcriptie stress reactie genoemd.

Om de schadelijke effecten van Pol II wat vastgelopen is op DNA-schade te voorkomen, hebben cellen een herstel mechanisme genaamd transcriptie-gekoppeld nucleotide excisie herstel (TC-NER). TC-NER herstelt TBLs en dit proces wordt gestart zodra Pol II vastloopt op een schade. Vastgelopen Pol II wordt herkend door de TC-NER factoren CSA en CSB, waarna UVSSA en het eiwitcomplex TFIIH worden gerekruteerd welke helpen met het in gang zetten van de reparatie reactie. Vervolgens wordt het beschadigde deel van het DNA eruit geknipt en vervangen door nieuw DNA zodat het overschrijven van dit stukje DNA kan worden hervat. Daarnaast wordt transcriptie op vele verschillende manieren gereguleerd na DNA schade. Dit zorgt ervoor dat de TBLs worden herkend, opgeruimd, en transcriptie correct kan plaatsvinden wat nodig is voor het overleven van de cel. In **Hoofdstuk 1** worden de gevolgen van TBLs op Pol II en transcriptie regulatie samengevat.

Ondanks de identificatie van verschillende factoren die betrokken zijn bij het reguleren van de cellulaire reactie op DNA-schade geïnduceerde transcriptie stress is het exacte mechanisme nog niet duidelijk. Om nieuwe eiwitten te identificeren die betrokken zijn bij deze cellulaire reactie, hebben we gebruik gemaakt van methodes om op een kwantitatieve manier eiwit interacties van Pol II te bestuderen. In **Hoofdstuk 2** beschrijven we verschillende manieren om Pol II te isoleren uit cellen. Als eerste gebruiken we een extractie protocol dat gebruikt maakt van milde condities om zowel initiërend als elongerend Pol II te isoleren. Vervolgens kunnen we deze fracties gebruiken om Pol II te isoleren en daarmee kwantitatieve massa spectrometrie uit te voeren om zo basale en UV geïnduceerde Pol II interactie partners te identificeren. Met behulp van deze

methode hebben we verschillende componenten en regulerende factoren van het Pol II complex geïdentificeerd, maar helaas geen nieuwe UV geïnduceerde interactie partners. Daarom zijn we overgestapt naar een protocol wat gebruik maakt van een chemische reactie (crosslinken) om kort gebonden of zwakke eiwit interacties te stabiliseren en deze zo detecteerbaar te maken. We hebben twee verschillende chemische stoffen onderzocht om de eiwit interacties te crosslinken, PFA en DSP. Beide stoffen konden succesvol gebruikt worden om Pol II interactie partners te identificeren maar konden, naast de al bekende TC-NER factor CSB, weinig UV geïnduceerde interactie partners identificeren. Het remmen van afbraak van eiwitten of het halverwege blokkeren van de TC-NER reactie, door cellen zonder XPA te gebruiken, leidden niet tot een verrijking van de UV geïnduceerde interactie partners van Pol II. Als alternatief hebben we gebruik gemaakt van twee verschillende technieken die eiwitten in dichte nabijheid een biotine-label geven, een zogenoemde nabijheidsmarkering. Beide methodes konden Pol II interactie partners identificeren maar weinig nog onbekende UV geïnduceerde factoren. Uiteindelijk hebben we gebruik gemaakt van een ander milde extractie techniek die de cellen eerst in fracties opdeelt. Daarnaast hebben we ons specifiek gericht op de elongerende Pol II complexen om op de manier te verrijken voor schade-gebonden Pol II, aangezien dit de complexen zijn die kunnen vastlopen op de DNA-schades. Deze methode was van alle geteste methodes het meest succesvol in het identificeren van UV verrijkte eiwitten, desondanks resulteerde ook deze methode maar in de identificatie van weinig nieuwe factoren. Tot slot hebben we de resultaten van alle hierboven genoemde experimenten gecombineerd en op deze manier gekeken naar potentieel interessante eiwitten die meerdere keren in verschillende experimenten werden geïdentificeerd.

Een andere manier om te kijken naar nieuwe factoren die betrokken zijn bij de cellulaire reactie op transcriptiestress is door gebruik te maken van een genoom-brede CRISPR/Cas9 verlies-van-functie analyse om genen te ontdekken die betrokken zijn bij de DNA-schade reactie, zoals beschreven in **Hoofdstuk 3**. Het feit dat deze analyse veel reeds bekende TC-NER factoren identificeerde als UV-gevoelige genen, zoals CSB, CSA en UVSSA, liet zien dat deze methode goed werkte. Verder werd in deze screen ook een eiwit geïdentificeerd waarvan nog geen rol in TC-NER bekend was: Elongatie Factor 1 Homoloog (ELOF1). We ontdekten dat ELOF1 bindt aan Pol II en een belangrijk onderdeel van het transcriptie elongatie complex is in condities zonder DNA-schade. Verder hebben we gevonden dat ELOF1 een belangrijke rol speelt in de TC-NER reactie en dat het zorgt dat het overschrijven van het DNA na UV weer goed hervat kan worden. We laten ook zien dat deze herstel functie van ELOF1 sterk geconserveerd is tussen verschillende soorten. We hebben aangetoond dat ELOF1 er in de TC-NER-reactie voor zorgt dat UVSSA en het TFIIH-complex correct naar het op schade vastgelopen Pol II worden gerekruteerd. Dit komt waarschijnlijk doordat Pol II niet wordt geubiquityleerd zonder ELOF1, wat suggereert dat ELOF1 belangrijk is voor het verzorgen van deze ubiquitine modificatie. Interessant genoeg hebben we ook laten zien dat ELOF1 ook een rol heeft in de DNA-schade geïnduceerd transcriptie stress reactie, buiten de TC-NER-reactie. Deze tweede rol voor ELOF1 is gelinkt aan het voorkomen van TRCs.

In **Hoofdstuk 4** beschrijven we een ander mechanisme dat belangrijk is in de UV geïnduceerde DNA-schade reactie waarbij we histonen bestuderen. Histonen zijn eiwitten

waar het DNA omheen gewikkeld zit om zo onze genetische code in de cel compact op te slaan en te organiseren. In dit hoofdstuk maken we gebruik van massa spectrometrie om veranderingen in modificaties van histonen te kunnen bestuderen en laten we zien dat er een genoom-breed verlies is van geacetyleerde histonen. Het verlies van deze acetylactie is onafhankelijk van transcriptie maar kan worden geïnduceerd door DNA-schades die leiden tot replicatie stress. We hebben laten zien dat het verminderen van de acetylactie een gevolg is van de afbraak van histonen met een acetylactie modificatie. De histonen worden afgebroken door het PA200 proteosoom, onafhankelijk van ubiquitine. Dit leidt tot een DNA schade geïnduceerde afbraak van geacetyleerde histonen, een proces wat waarschijnlijk een belangrijke rol speelt bij het veranderen van chromatine, zodat cellen beter om kunnen gaan met replicatie stress veroorzakende DNA schade.

Ten slotte vatten we in **Hoofdstuk 5** de belangrijkste bevindingen samen van het in dit proefschrift beschreven experimentele werk. Daarnaast bespreken we de conclusies hiervan en bespreken we tot welke nieuwe inzichten de conclusies hebben geleid. Verder stellen we toekomstige experimenten voor die van belang kunnen zijn om onze kennis van de DNA-schade geïnduceerde transcriptie stress reactie verder uit te breiden.

Curriculum vitae

Name: Marit Esmée Geijer
Date of birth: October 19, 1991
Place of birth: Leiden
E-mail: marit_geijer@live.nl

Work experience

- 2020 – present **Scientist** at Toxys, Leiden.
- 2015 – 2020 **PhD student** at the department of Molecular Genetics, Erasmus Medical Center, Rotterdam.
- 2015 **Research internship** at the department of Internal Medicine, Erasmus Medical Center, Rotterdam.
- 2015 **Communication internship** at the department of Science Communication and Society, Leiden University.
- 2013 – 2014 **Research internship** at the former department of Toxicogenetics, Leiden University Medical Center.
- 2013 **Research internship** at the department of Endocrinology, Leiden University Medical Center.

Education

- 2013 – 2015 **Master's degree, Biomedical Sciences**, Leiden University.
- 2010 – 2013 **Bachelor's degree, Biomedical Sciences**, Leiden University.
- 2004 – 2010 **Pre-university education (VWO)**, 't Vlietland College, Leiden.

List of publications

Elongation factor ELOF1 drives transcription-coupled repair and prevents genome instability.

Marit E Geijer, Di Zhou, Kathiresan Selvam, Barbara Steurer, Bastiaan Evers, Chirantani Mukherjee, Simona Cugusi, Marvin van Toorn, Melanie van der Woude, Wenzhi Gon5, Roel Janssens, Anja Raams, Joyce HG Lebbink, Bart Geverts, Dalton A Plummer, Karel Bezstarosti, Arjan F Theil, Richard Mitter, Adriaan B Houtsmuller, Wim Vermeulen, Jeroen AA Demmers, Shisheng Li, Hannes Lans, René Bernards, Jesper Q Svejstrup, Arnab Ray Chaudhuri, John J Wyrick, Jurgen A Marteijn.

Revisions Nature Cell Biology.

What happens at the lesion does not stay at the lesion: Transcription-coupled nucleotide excision repair and the effects of DNA damage on transcription in cis and trans.

Geijer ME, Marteijn JA. DNA Repair (Amst). 2018. PMID: 30195642.

DNA damage-induced replication stress results in PA200-proteasome-mediated degradation of acetylated histones. Mandemaker IK, Geijer ME, Kik I, Bezstarosti K, Rijkers E, Raams A, et al. EMBO Rep. 2018;19(10). PMID: 30104204.

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, et al. Proceedings of the National Academy of Sciences. 2018;115(19):E4368-E76. PMID: 29632207.

PhD portfolio

Name: Marit Geijer
PhD period: October 2015 – July 2020
Promoter: Wim Vermeulen
Co-promoter: Jurgen Marteijn
Erasmus MC department: Molecular Genetics
Research school: Biomedical Sciences, MGC

General courses

Safely working in the laboratory	2015
Special topic course on chromatin (Chromatin day)	2015
Special topic course on signalling (Signalling day)	2016
Genetics course	2016
Research integrity	2017
The ins and outs of CRISPR-Cas	2017
Biostatistical methods I: basic principles part A	2018
Biomedical English Writing and Communication	2019

Specific courses

Microscopy introductory courses	2015
OIC Course Functional Imaging and Super Resolution	2015
MGC Technology Facilities :Genomics&Transcriptomics	2016
MGC Technology Facilities: Proteomics	2016
MGC Genome Maintenance and Cancer	2016
Microscopic Image Analysis: From Theory to Practice	2017

Teaching assistance and project supervision

Supervising BSc Nanobiology Genetics Exercises	2016-2018
Supervisor MolMed Master student	2017
Supervisor Nanobiology Master student	2018
Supervisor LST Master student	2019

Meetings, Workshops and Conferences

Vermeulen Group discussions	2015-2020
Molecular Genetics Workdiscussions	2015-2020
Vermeulen Literature discussions	2015-2020
MGC DNA Repair meeting Leiden/Rotterdam	2015-2020
Annual MGC PhD workshop	2016-2018
Annual MGC Symposium	2016-2018
Cancer Genomics Meetings, Amsterdam	2016, 2018
Responses to DNA damage: from molecule to disease, Egmond aan Zee, The Netherlands	2016
Oncode Launch, Amsterdam	2018
Oncode Kick off meeting: Genome Instability	2018
EMBO Workshop: RNA and genome maintenance, Mainz, Germany	2018
Oncode Meeting	2019
Life Congress	2019

Acknowledgements

Eindelijk is het dan zover: mijn proefschrift is klaar! Het waren 5 mooie, gezellige en vooral drukke jaren maar wat ben ik trots dat ik dit volbracht heb. Zoals zoveel dingen heb ik dit natuurlijk niet alleen gedaan. Ik wil graag iedereen bedanken die mij hebben geholpen maar een aantal mensen in het bijzonder.

Beste promotor, beste **Wim**, bedankt voor je hulp en advies tijdens mijn promotietraject en het afronden van het schrijven van mijn boekje. Werken in jouw lab heb ik altijd als heel fijn ervaren, bedankt voor alle wijze woorden tijdens alle werkbesprekingen.

Beste co-promotor, beste **Jurgen**, bedankt dat ik onderdeel uit mocht maken van jouw groep! Ik heb met veel plezier met je samen gewerkt en waardeer de discussies die we gevoerd hebben over mijn werk en jouw steun tijdens mijn gehele PhD. Sorry voor het gebrek aan koffies tijdens onze vele (te laat begonnen) werkbesprekingen en bedankt voor je (slechte) grapjes. Ik waardeer de fijne en soepele samenwerking die we over de afgelopen jaren hebben gehad. Veel succes met je verdere carrière en hopelijk zien we elkaar snel weer tijdens je oratie ;).

Off course my paranymphs, Karen and Barbara. I am very happy to be supported by two such strong women on this important day in my life.

Lieve **Karen**, bedankt voor jouw steun de afgelopen jaren. Ik denk dat we over werkelijk alles kunnen praten, klagen over het OV, kletsen over tuinieren, maar ook onderwerpen waar we het moeilijker mee hadden. Bedankt dat ik altijd bij je terecht kon voor hulp, advies of soms wat bemoedigende woorden. Ik ben heel blij dat jij mijn paranymf wilt zijn!

Dear **Barbara**, as you already wrote it in your acknowledgements, working together was fun, efficient and uncomplicated. I much appreciate the fun we had/have both in and outside the lab. Going through the same process really made it easier to talk and complaint about our PhDs. Thank you for always being there to chat about anything. I am very proud that I have supported you during your defense and I am very happy that you will do the same for me!

Lieve **Angela**, vanaf mijn eerste dag konden we het samen goed vinden. Jouw hulp in mijn eerste weken heeft mij ontzettend geholpen en je hebt mij veel geleerd. Zelfs nu help je mij nog waar nodig. Ook wij konden over werkelijk alles praten, soms was het een beetje té gezellig in ons kantoor ;). Bedankt voor je hulp, gezelligheid en de treinreizen naar huis!

And of course thanks to all other members of the Vermeulen/Lans/Martijn labs. **Roel**, bedankt voor de fijne samenwerking en jouw hulp bij het afronden van het ELOF1 project! **Hannes** en **Arjan**, bedankt voor jullie gezelligheid, grapjes en chaos op vrijdagmiddagen. **Di**, thank you so much for all your help with the ELOF1 project and sharing you precious cell lines for this. **Alex**, thank you for all your help and advice over the past years. Especially all the times I have asked you for help using the Maxquant software.

Your nice supervision during my Master's internship is one of the reasons I pursued a PhD, thank you! **Cristina**, thank you for your enthusiasm about science and for answering all my questions. My office girls **Carlota**, **Diana** and **Alba** thank you for all the chatting and making the last period of my PhD so much fun! **Marvin**, **Maarten**, **Irene**, **Anja**, **Melanie**, and **Maria**, thank you all for your support, chitchats and discussions. **Nils**, bedankt voor de gezellig lunches en alle hulp als mijn computer weer eens niet deed wat ik graag zou willen.

All former lab members, **Imke** bedankt dat ik onderdeel mocht zijn van jouw PA200 manuscript en je hulp op het lab. **Mariangela**, **Yasemin**, **Masaki**, and **Jana**, thank you for your support. Mijn studenten **Lizzy**, **Jeroen** en **Ilse**, bedankt voor jullie hulp. Many thanks to all the collaborators of my precious ELOF1 manuscript and all staff of the Molecular Genetics department. **Joyce**, bedankt voor alle gezellige treinreizen, dat maakte de Leiden-Rotterdam reisjes stukken leuker!

Ook dank aan de **BW meisjes** voor jullie steun tijdens mijn PhD. Ik ben blij dat ik met jullie altijd over mijn onderzoek (of wat dan ook) kan praten en we onze PhD ervaringen kunnen delen.

Mijn lieve schoonfamilie, **Ton**, **Marjan**, **Jorinde** en **Daan**, bedankt voor jullie steun en interesse in mijn werk! Wat fijn dat jullie altijd je best doen om echt te begrijpen waar ik het soms nou toch over heb en de vraag of ik mijn artikel kan printen zodat jullie er naar kunnen kijken.

Mijn lieve broers en schone zussen, **Robin**, **Frederique**, **Gydo** en **Aimée**, heel erg bedankt voor jullie steun en alle vragen om zo goed mogelijk te begrijpen wat ik nou toch precies doe al was het nog steeds vaak abacadabra. Bedankt dat elke stikstof tank mij nu aan minions doet denken ;).

Lieve **papa**, dit boekje is voor jou en mama. Dankzij jullie kon ik überhaupt gaan studeren en daarna aan dit grote avontuur beginnen. Bedankt!

En natuurlijk, lieve **Niek**. Zonder jou zou ik nooit begonnen zijn aan mijn PhD. Bedankt voor je oneindige steun door de jaren heen en dat duwtje in mijn rug wanneer ik het nodig had. Bedankt voor je begrip voor de lange dagen, het regelen van alles en het geven van rust zodat ik mij geen zorgen hoeft te maken over van alles en nog wat. Maar bovenal bedankt voor je humor, liefde en vertrouwen in dat ik alles kan.