Nucleotide excision repair: from molecular mechanisms to patient phenotypes

Apelt, K.

Chapter 4

Exploring the dynamics of high-mobility group proteins at UV-induced DNA lesions

Katja Apelt, Lisa Schubert, Annelotte P. Wondergem, Niels Mailand, Martijn S. Luijsterburg

Abstract

Nucleotide excision repair (NER) eliminates bulky DNA lesions from the genome. The wrapping of genomic DNA around histones forms an obstacle for the binding of DNA repair factors in chromatin. The full repertoire and dynamics of protein factors that interact with UV-irradiated chromatin and that might modulate the binding of DNA repair factors is unknown. To systematically monitor the landscape of proteins interacting with UV-irradiated chromatin in an unbiased manner, we employed chromatin mass spectrometry (CHROMASS) in *Xenopus* egg extract. Using CHROMASS, we identified NER factors, mismatch repair (MMR) factors, and HMGB-type proteins among the most prominent proteins that associate with UV-damaged chromatin. Conversely, HMGN- and HMGA-type proteins were released from *Xenopus* chromatin in response to UV irradiation. Live-cell imaging of GFP-tagged HMG-type proteins in human cells revealed that these responses were evolutionary conserved with all four HMGB-type proteins showing strong recruitment in response to UV-C laser irradiation, and HMGN- and HMGA-type proteins being released from chromatin following UV-C laser irradiation. We established an efficient siRNA-mediated procedure to individually deplete all HMGB-type isoforms to characterize the functional relevance of their recruitment during the response to UV-induced DNA lesions. Our study provides insights into the landscape of chromatin interactions in response to UV irradiation and reveals that HMGB-type proteins are readily recruited to UV-damaged DNA in an evolutionary conserved manner. Future studies will reveal the functional relevance of these UV-induced chromatin interactions.

Introduction

The human genome contains the genetic information that encodes all proteins expressed in human cells. Genome integrity is continuously challenged by DNA damage, which can have a detrimental impact on our cells, including cytotoxicity leading to apoptosis, genomic instability and increased mutagenicity. Genomic DNA lesions are caused by different DNA-damaging agents, such as sunlight, chemotherapeutical agents and endogenous metabolism. Solar ultra-violet (UV) light introduces two different types of genomic DNA lesions that covalently link two adjacent pyrimidine bases together in the same DNA strand: 6-4 photoproducts (6-4PPs) and cyclobutane pyrimidine dimers (CPDs) (Rastogi et al., 2010). The repair of these DNA lesions is mediated by nucleotide excision repair (NER), an evolutionary conserved DNA repair mechanism that involves two distinct sub-pathways. Transcription-coupled repair (TCR) eliminates DNA lesions in transcribed DNA strands, while global genome repair (GGR) recognizes and eliminates DNA lesions throughout the rest of the genome. The main DNA-damage sensor in GGR is the XPC-RAD23A-CEN2 complex (Sugasawa et al., 1998), which indirectly recognizes helix-destabilization lesions, such as 6-4PPs by binding to the undamaged complementary single-stranded DNA opposite to a DNA lesion (Maillard et al., 2007; Min and Pavletich, 2007). Conversely, the XPC complex binds poorly to lesser helix-destabilization lesions, such as CPDs, which requires the aid of the $CRL4^{DDB2}$ E3 ubiquitin ligase complex (Tang et al., 2000). Unlike XPC, DDB2 binds directly to the lesion and flips out the damaged bases, which causes helix destabilization and stimulates the binding of XPC (Scrima et al., 2008). Upon binding, XPC recruits the TFIIH complex, which unwinds the DNA through its helicase activity (Coin et al., 1998) leading to the recruitment of XPA, RPA and the endonucleases ERCC1-XPF and XPG (Li et al., 2015). A single-stranded stretch of 30 nucleotides containing the lesions is excised and the resulting gap is resynthesized by a DNA polymerase.

It is becoming increasingly clear that the presence of UV-induced DNA lesions in chromatin not only triggers the assembly of NER complexes, but also elicits a variety of other cellular responses. For instance, unrepaired DNA lesions that are encountered by replicative DNA polymerases cause replication stalling and trigger the recruitment of translesion synthesis (TLS) polymerases to bypass UV-induced DNA lesions, such as DNA pol η (Novarina et al., 2011). The advantage is that TLS polymerases prevent the toxic consequences of replication stalling, but due to their

lower fidelity these enzymes have a higher probability of incorporating incorrect nucleotides opposite to DNA lesions. Proteins of the mismatch repair (MMR) pathway recognize such TLS-mediated misincorporated nucleotides (Tsaalbi-Shtylik et al., 2015) and excise them. These findings show that multiple protein factors act on UV-damaged chromatin possibly in concert or even in parallel with NER. However, the full repertoire and dynamics of protein factors that interact with UV-irradiated chromatin is unknown.

Previous chromatin-fractionation approaches coupled with mass spectrometry have identified proteins that associate with or dissociate from UV-irradiated chromatin in human cells (Chou et al., 2010; Tresini et al., 2015). One study found that NER and MMR proteins are strongly enriched in chromatin following global irradiation with 50 J/m^2. However, due to the high dose employed, DNA double-strand break and interstrand crosslink repair proteins were found to be enriched as well (Chou et al., 2010), making it difficult to link cellular responses to a specific type of DNA lesion. Another study found that late-stage spliceosomes are displaced from chromatin in response to UV-induced DNA lesions that block transcription (Tresini et al., 2015). However, the number of proteins found enriched in chromatin was rather limited and did not include many NER proteins.

We developed a method called chromatin mass spectrometry (CHROMASS) to study protein recruitment dynamics in response to DNA damage in cell-free extracts of *Xenopus laevis* in an unbiased manner. CHROMASS was used to provide a global overview of the protein dynamics in response to interstrand crosslinks (ICLs) and revealed a new pathway in which the SLF1-SLF2/RAD18 complex recruits the SMC5/6 cohesion complex to ICLs (Raschle et al., 2015). Key advantages of CHROMASS are that the method is highly specific for one type of DNA lesion and that it enables us to study protein recruitment in the absence of transcription, considering that *Xenopus laevis* egg extracts are devoid of transcriptional activity. Here, we applied CHROMASS to systematically monitor the landscape of proteins interacting with UV-irradiated chromatin in an unbiased manner. In addition to NER and MMR proteins, our analysis revealed HMGB-type proteins among the most prominent proteins that associate with UV-damaged chromatin, while HMGN- and HMGA-type proteins were released from Xenopus chromatin in response to UV irradiation. We show that this behavior of HMG protein is evolutionary conserved in human cells by live-cell imaging and we establish procedures to individually deplete all human HMGB-type isoforms to characterize the functional relevance of their recruitment during the response to UV-induced DNA lesions in future studies.

Results

CHROMASS monitors the landscape of proteins that interacts with UV-irradiated chromatin in *Xenopus* egg extracts

To identify proteins that either associate or dissociate with chromatin in response to UV irradiation, we used a method that combined the *Xenopus laevis* egg extract system for cell-free DNA repair with quantitative mass spectrometry called CHROMASS (Raschle et al., 2015). To this end, undamaged or UV-irradiated (2 kJ/m2) *Xenopus* sperm chromatin was injected into egg extracts and allowed to replicate for 45 min. Subsequently, we isolated chromatin by sedimentation through a sucrose cushion and analyzed chromatin-bound proteins by mass spectrometry (Figure 1A). Our analysis revealed around 500 proteins that became significantly enriched on UV-irradiated chromatin relative to undamaged control chromatin (Figure 1B). Among the most highly enriched proteins were factors involved in nucleotide excision repair (NER; such as six of the ten TFIIH subunits, XPA, XPG, ERCC1-XPF and CUL4A/RBX1) and mismatch repair (MMR; such as MSH6, MLH1, MSH2, MSH6, PMS2, PSM1 and EXO1). These findings are largely in agreement with an earlier chromatin localization screen after UV in human cells (Chou et al., 2010). Additionally, we detected strong enrichment of all members of the *Xenopus laevis* HMGB family (HMGB1-3), while HMGA-type and HMGN-type proteins were among the most significantly displaced proteins from UV-irradiated chromatin (Figure 1B). Our CHROMASS approach reveals the landscape of chromatin interactions in response to UV irradiation in a cell-free system.

Generation of human cells expressing GFP-tagged HMG proteins

Our CHROMASS screen revealed the dynamics of HMG-type proteins on UV-damaged chromatin in a cell-free *Xenopus laevis* system. We next sought to establish whether this behavior of HMG proteins is conserved in human cells. Members of each of the HMG families contain distinct DNA-binding domains, including HMG box domains (HMGB), AT hooks (HMGA) or nucleosome-binding domains (NBD; HMGN) (Figure 2A). The HMGB family shows strong evolutionary conservation. For example, human and *Xenopus leavis* HMGB1 share around 90% sequence homology (Figure S1A). There are four HMGB-type proteins (HMGB1-HMGB4) in human cells that are encoded by different genes (Figure S1B). Additionally, each of the four HMGB loci encodes various splice variants (Figure S2A-D). All HMGB proteins contain two DNA-binding domains, called box A and

Figure 1

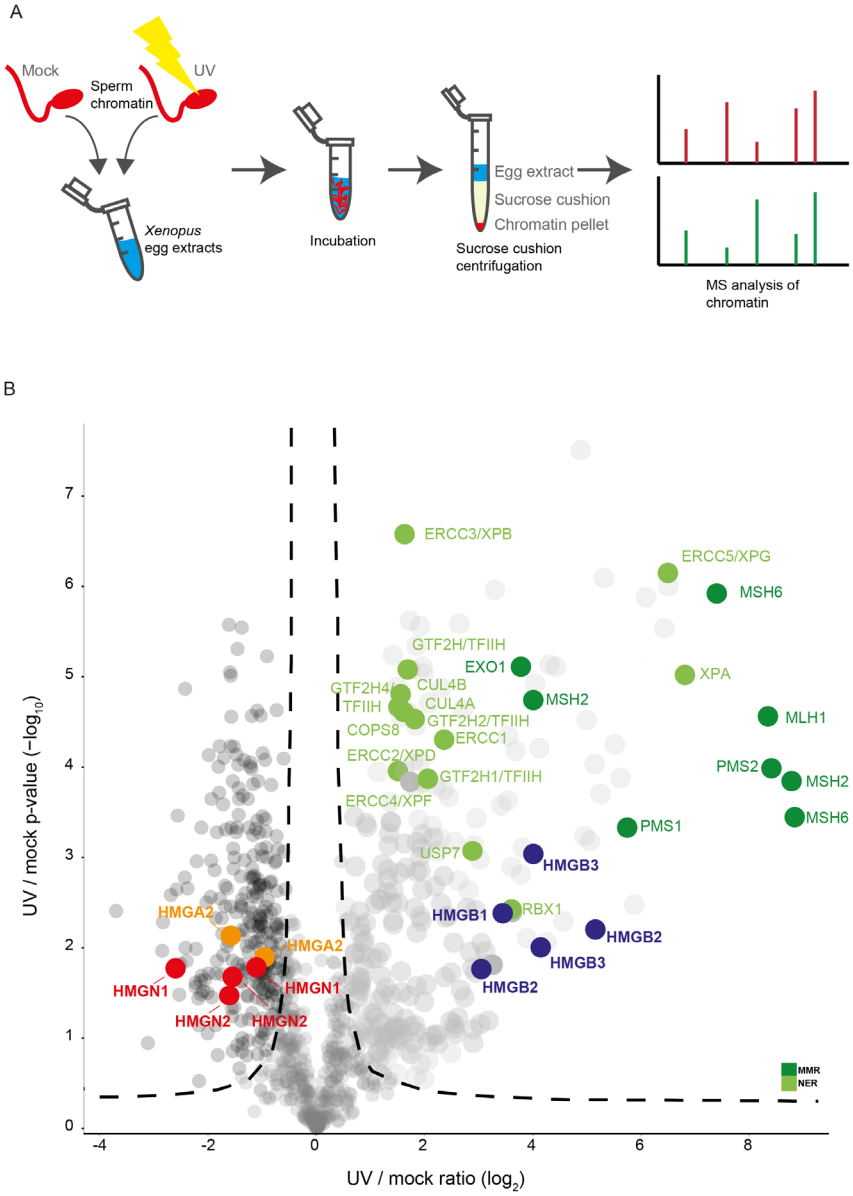

Figure 1. (A) Workflow of CHROMASS showing the preparation of the sperm chromatin for the Mass Spectrometry analysis. Sperm chromatin was UV-C irradiated (2 kJ/m^2) or mock treated and incubated for 30 min. Next the chromatin was isolated by sucrose cushion centrifugation and analyzed by label-free mass spectrometry. (B) CHROMASS results represented in volcano plots depicting chromatin binders and chromatin releasers after UV-C irradiation. The dashed lines indicate an enrichment of one-fold on the x axis ($\log_2$ of 1)

box B, and an acidic C-terminal tail. HMGB1, HMGB2 and HMGB3 have around 77% sequence homology, while HMGB4 is more divergent (36% homology) and lacks the acidic C-terminal tail (Figure S1B). In addition to the full-length gene products, the *HMGB1* and the *HMGB2* genes each produce a shorter isoform lacking the C-terminal tail. Three *HMGB3* isoforms are produced with slightly shorter C-termini, while only a single HMGB4 isoform is known (Figure S2A-D).

For each of the four HMGB-type proteins, as well as for HMGA1 and HMGN1, we generated C-terminal GFP-tagged fusion proteins by inserting cDNAs encoding HMG proteins into pCDNA vectors that are compatible with the Flp-In T-REx system for inducible protein expression. In addition, we generated HMGA1 and HMGB1 which was tagged at their N-termini with GFP. We subsequently generated polyclonal stable cell-lines for all eight HMG proteins by integrating the corresponding HMG-GFP plasmids into human osteosarcoma cells (U2OS-FRT) equipped with the Flp-In/T-REx system. Western blot analysis confirmed the expression of all C-terminal GFP-tagged HMG proteins at the expected molecular height following induction of protein expression by doxycycline (Figure 2B).

HMGA1 and HMGN1 dissociate from UV-damaged chromatin in human cells

Having established cell-lines stably expressing GFP-tagged HMG proteins, we next monitored their behavior in response to UV-C irradiation, first focusing on HMGA and HMGN. In undamaged cells, GFP-tagged HMGA1 and HMGN1 were strictly nuclear without nucleolar accumulation and displayed a non-uniform pattern typically observed for chromatin-bound proteins (Figure 2C, 2E). To study the behavior of HMG proteins in response to UV-induced DNA lesions, we employed irradiation of cells with a pulsed 266 nm UV-C laser on a live-cell imaging set-up in which all glass optics were replaced by quartz optics to allow full UV-C transmission (Apelt et al., 2020; Dinant et al., 2007). Immediately after UV-C irradiation with 30% laser power, we observed a reduction in HMGN1-GFP levels at the site of irradiation, suggesting HMGN1 is displaced or dissociates from UV-damaged chromatin (Figure 2C, 2D), similar to its *Xenopus* counterpart (Figure 1B). Likewise, HMGA1-GFP was rapidly displaced in response to laser-inflicted UV-C irradiation (Figure 2E, 2F). Considering that HMGA1 fusion proteins with GFP on either side showed the same behavior, suggests that the orientation of the GFP-tag does not account for, nor interferes with this behavior. Strikingly, the levels of HMGA1 and HMGN1 at damaged sites gradually recovered within 15 min, suggesting this is a biologically relevant phenomenon that is not caused by bleaching.

Figure 2

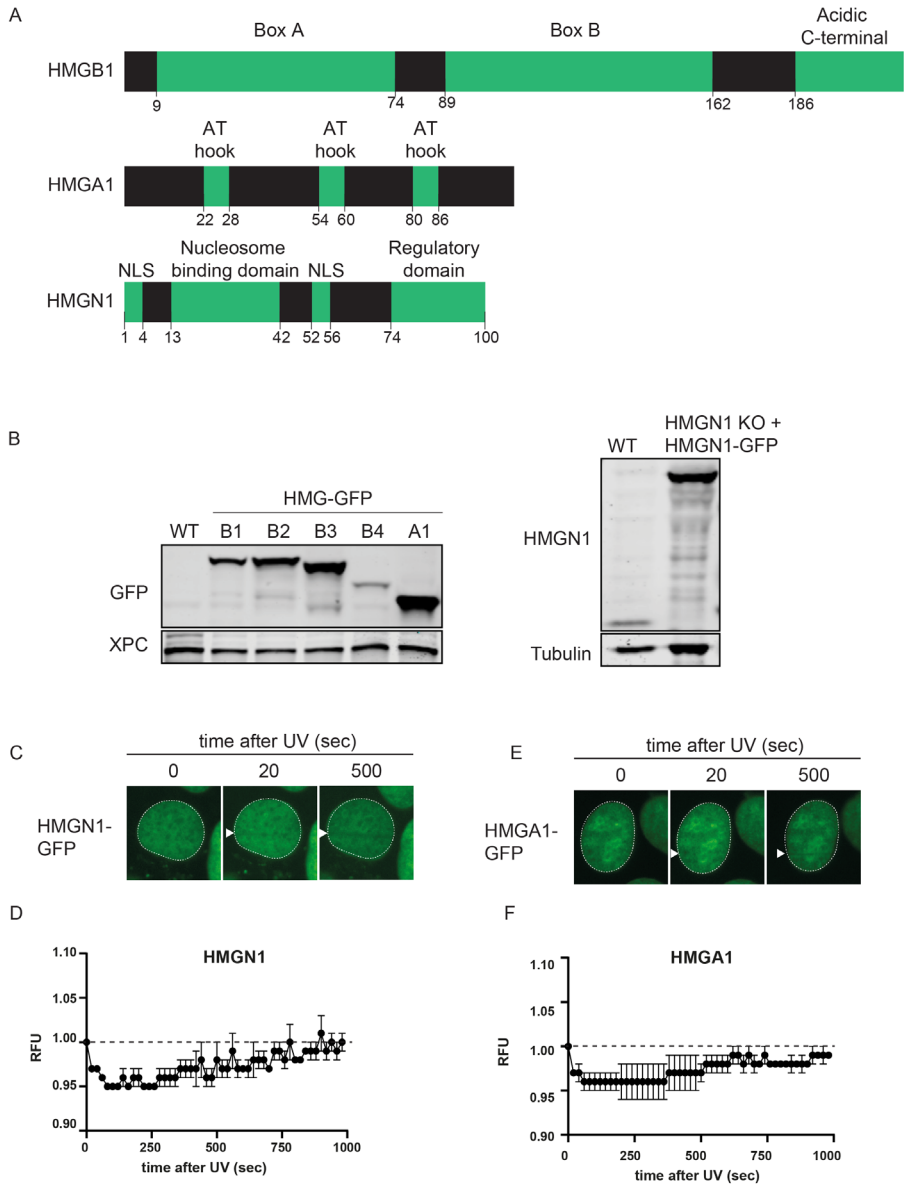

Figure 2. (A) Scheme representation of the different domains of HMGB1, HMGA1 and HMGN1. (B) Western blot of U2OS cells expressing a GFP-tagged HMG protein as indicated. (C) Microscopic images (D) and quantification of the recruitment of HMGN1-GFP to sites of local UV-C irradiation. (E) Microscopic images (F) and quantification of the recruitment of HMGA1-GFP to sites of local UV-C irradiation.

Figure 3

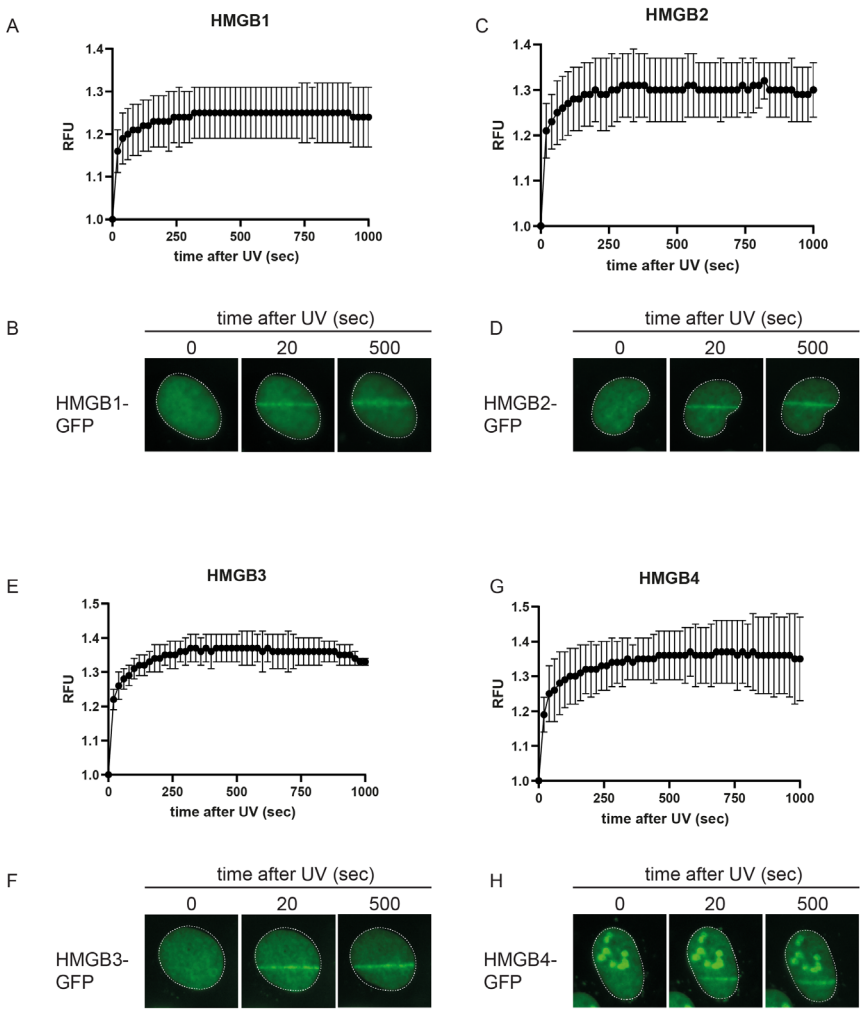

Figure 3. (A) Microscopic images (B) and quantification of the recruitment HMGB1-GFP, (C, D) HMGB2-GFP, (E, F) HMGB3-GFP and (G, H) HMGB4-GFP.

In conclusion, the dissociation of HMGA1 and HMGN1 from UV-damaged chromatin is evolutionary conserved between *Xenopus* and humans.

All four HMGB-type proteins are recruited to UV-damaged chromatin in human cells

Having found that the behavior of HMGA1 and HMGN1 proteins is conserved between human and frog, we next monitored how HMGB-type proteins respond to UV irradiation. In undamaged cells, GFP-tagged versions of HMGB1, HMGB2, and HMGB3 were localized in the nucleoplasm, while HMGB4-GFP was additionally enriched in nucleoli, which are sites where ribosomal RNAs are produced (Figure 3B-H) (Melese and Xue, 1995). Following micro-irradiation with a UV-C laser, we observed rapid and robust recruitment of all four HMGB-type proteins to UV-damaged chromatin, which remained visible for up to 15 min (Figure 3A-F). These findings show that, like in *Xenopus* egg extracts, all HMGB-type proteins associate with UV-damaged chromatin, revealing a remarkably conservation of the behavior of all classes of HMG proteins between frog and humans.

A method to deplete HMGB-type proteins by RNA interference

Following the observation that all four human HMGB-type proteins are recruited to UV-induced DNA lesions, we decided to establish a method to deplete all HMGB proteins as a means to study whether these proteins have a role in the cellular response to DNA damage triggered by UV irradiation. To this end, we designed sRNAs that target each HMGB variant including its splice variants if possible. The siRNAs targeting *HMGB1* and *HMGB3* recognize all known splice variants of these genes. The *HMGB2* siRNA targets the long, but not the short splice variant, while the *HMGB4* siRNA targets the single known isoform (Figure S3-5, S6A). We initially tested the ability of the siRNAs to knock-down ectopically expressed HMGB-GFP proteins. To this end, U2OS cells were transfected twice with siRNAs targeting each HMGB variant. After the second siRNA transfection, we induced the expression of the corresponding HMGB-GFP protein by incubating with doxycycline. Western blot analysis revealed a near-complete knock-down of GFP-tagged HMGB1, HMGB3 and HMGB4 compared to control siRNA transfection, while the levels of HMGB2-GFP were strongly reduced, although residual expression remained (Figure 4A). The expression of HMGB-GFP variants is driven by a strong viral promoter and we currently do not know how the expression level compares to endogenous expression of these HMG-type proteins. To expand this analysis, we transfected parental U2OS cells not expressing HMGB-GFP variants twice with siRNAs, targeting either *HMGB1* or *HMGB2*. Western

Figure 4

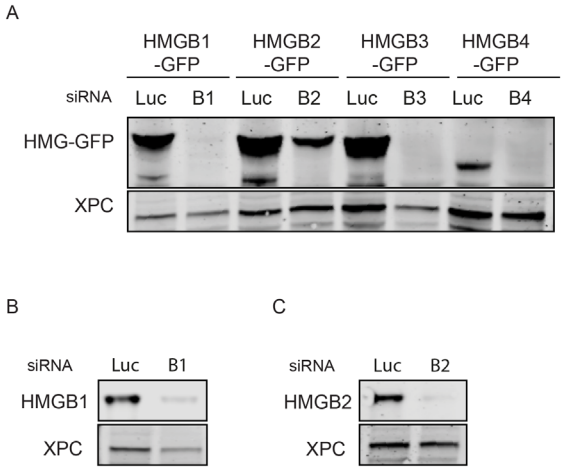

Figure 4. (A) Western blot of U2OS cells expressing a GPF-tagged HMGB protein transfected with siRNAs. (B) Western blot of U2OS cells transfected with siRNA against HMGB1 or (C) HMGB2.

blot analysis using specific antibodies revealed a strong depletion of endogenous HMGB1 or HMGB2 proteins (Figure 4B, C). In conclusion, we have established a method to individually deplete HMGB-type proteins in human cells, which can be used to assess how the loss of HMGB-type proteins, alone or in combination, affects the cellular response to UV-induced DNA lesions.

Discussion

To systematically monitor the landscape of proteins interacting with UV-irradiated chromatin in an unbiased manner, we employed chromatin mass spectrometry (CHROMASS) in *Xenopus* egg extract. We identified nucleotide excision repair (NER), mismatch repair (MMR), and HMGB-type proteins among the most prominent factors that associate with UV-damaged chromatin, while HMGA- and HMGN-type proteins were released from chromatin in response to UV irradiation. Live-cell imaging of GFP-tagged HMGB, HMGA and HMGN-type proteins confirms that this behavior is evolutionary conserved from *Xenopus* to human cells. Future studies will be needed to reveal the potential functional relevance of these UV-induced chromatin interactions.

Comparison with earlier chromatin localization screens

Our CHROMASS analysis revealed that NER and MMR proteins became strongly enriched in *Xenopus* chromatin within 45 min after UV irradiation. This is in agreement with an earlier chromatin localization screen from the Elledge lab in human cells irradiated with 50 J/m^2 and analyzed at 90 min (Chou et al., 2010). The association of HMGB-type proteins was not detected in this screen, which could be due to the analysis at a later time-point. We detected the recruitment of all four HMGB isoforms up to 15 min after UV-C micro-irradiation. A second chromatin localization screen from the Marteijn lab was analyzed at 1 hour after UV irradiation (20 J/m^2) in non-dividing human cells (Tresini et al., 2015). This screen revealed that spliceosome proteins are released from chromatin after UV irradiation, but the number of proteins found to be enriched after UV was rather limited and did not include the known NER proteins. Several HMG proteins were detected in this screen (HMGN1, HMGN2, HMGN3, HMGA1, HMGA2, HMGB1, HMGB2), but none were specifically enriched or depleted compared to unirradiated control cells. This is either due to a difference in sensitivity of the different screens, or it could be explained by the fact that our *Xenopus* extracts are fully proficient in replication but devoid of transcriptional activity, while the screen from the Marteijn lab used transcriptionally active cells that do not replicate (Tresini et al., 2015). Regardless, our validation experiments in human cells show that HMGB-type proteins accumulate, while HMGA1 and HMGN1 are released from chromatin after UV irradiation in human cells.

Release of HMGA1 and HMGN1 from chromatin after UV irradiation

Our experiments show that HMGA1 and HMGN1 are temporarily released from chromatin within the first 15 min after UV irradiation. This release coincides with the rapid accumulation of HMGB-type proteins to chromatin. One intriguing possibility is that the HMGB-type proteins displace HMGA1/HMGN1 from chromatin. This could easily be tested by monitoring the behavior of GFP-tagged HMGA1/HMGN1 in cells depleted for HMGB-type proteins. Earlier experiments revealed that linked histone H1 as well as core histones H2A and H4 are released from sites of local UV-induced DNA damage (Luijsterburg et al., 2012). The release of core histones H4 (Luijsterburg et al., 2012) and H3 (Adam et al., 2016) was found to be dependent on the NER factor DDB2. It is possible that the release of HMGA1/N1 is caused by the DDB2-dependent destabilization and unfolding of chromatin. It will therefore be interesting to test whether the release of HMGA1/HMGN1 is dependent on DDB2 as well. Earlier work revealed that linker H1 and HMGN may have overlapping binding sites on chromatin (Catez et al., 2002), potentially explaining why both chromosomal proteins show the same behavior and are released from chromatin after UV. However, HMGA and HMGB proteins were similarly found to compete with H1 in a synergistic fashion, suggesting that the three classes of HMG proteins function cooperatively to weaken the binding of H1 to chromatin (Catez et al., 2002). Despite this cooperation among HMG proteins, we find that HMGB-type proteins are recruited to UV-damaged chromatin at times when HMGA1, HMGN1 and linker H1 histones are released. The heterochromatin protein 1 (HP1ß) was found to be released from chromatin in response to DNA double-strand breaks in a manner that involved the phosphorylation of a residue essential for its chromatin localization (Ayoub et al., 2008). It is possible that the release of HMGA1/HMGN1 may be a regulated event that similarly requires a post-translational modification of the HMG proteins themselves. The precise physiological relevance of the release of HMGA1 and HMGN1 remains to be established.

The recruitment of HMGB-type proteins to chromatin after UV irradiation

Our CHROMASS screen and validation experiments in human cells reveal that all HMGB-type proteins associate with UV-damaged chromatin. This behaviour could reflect the direct binding of HMGB proteins to UV-induced photolesions, or could be due to the recruitment of HMGB proteins by, for instance, NER or MMR proteins. Earlier biochemical experiments have shown that HMGB1 binds to single-stranded DNA (Bidney and Reeck, 1978) and supercoiled DNA (Hamada and Bustin, 1985), but also to various

types of DNA lesions, including UV-induced photolesions (Pasheva et al., 1998), cisplatin-induced adducts (Lanuszewska and Widlak, 2000) and psoralen-induced DNA lesions (Reddy et al., 2005). This biochemical property of HMGB1 may very well explain our observations that *Xenopus* laevis HMGB1 associates with UV-damaged chromatin in cell extracts and that human HMGB1 accumulates at UV-induced lesions *in vivo*. Our findings also extend this behavior to human HMGB2, HMGB3 and HMGB4. Although likely, it would be important to confirm the binding of these HMGB-type proteins to UV-induced photolesions *in vitro* using recombinant proteins.

HMGB1 contains two DNA-binding domains, box A and box B, in which three α-helices are arranged in an L-shaped configuration (Figure S6B) (Reeves and Adair, 2005). Both DNA-binding domains are highly conserved (>90%) throughout evolution (Figure S1A). Two specific residues in each of the DNA-binding domain are inserted into the DNA minor groove (box A: A17 and F38, box B: F103 and I122 (Thomas and Travers, 2001). Interestingly, the DNA-binding domains of HMGB2 and HMGB3 share a high degree of sequence homology with HMGB1 (75%; Figure S1B), while HMGB4 is more divergent and only shares 35% similarity with HMGB1 (Figure S6C). Of the four residues important for binding to DNA, only the F103 residue is shared between HMGB4 and HMGB1. It would be interesting to mutate this residue and monitor its impact on the recruitment of all HMGB-type proteins to UV-induced DNA lesions.

Another distinctive feature of HMGB4 is the absence of the acidic C-terminal tail, which is present in the other three HMGB proteins (Figure S1B). The acidic tail functions as a negative intramolecular regulator of HMGB1-DNA interactions by binding to the DNA-binding surface of both basic HMG boxes A and B (Stott et al., 2010). It has been suggested that the known post-translational modification of the K2 residue by acetylation may disrupt the intramolecular association of the acidic tail and stimulates the binding of HMGB1 to DNA (Stott et al., 2010). Whether such a modification of HMGB-type proteins regulates their recruitment to UV-induced DNA lesions remains to be determined.

Do HMGB-type proteins affect the response to UV-induced DNA damage?

The precise impact and potential physiological relevance of the observed recruitment of all four human HMGB proteins to UV-induced photolesions remains to be determined. It can be envisioned that the recruitment of these proteins, if they bind directly to UV-induced photolesions, might actually inhibit the binding of DNA damage-recognition proteins that initiate NER,

such as XPC-RAD23 or DDB2. *In vitro* experiments using a reconstituted NER system in the absence or presence of recombinant HMGB1 might indeed support such a scenario (Malina et al., 2002; Mitkova et al., 2005). Such an inhibitory impact of HMGB proteins may be a protective mechanism to inhibit DNA repair and drive cells into apoptosis when the DNA damage load is too high.

On the other hand, HMGB proteins may stimulate DNA repair by NER through an unknown mechanism. It is known that HMGB1 can bind to nucleosomes (Catez et al., 2002) and is able to bend DNA (Stros and Muselikova, 2000). This characteristic of HMGB1 may contribute to somehow increase the local accessibility of chromatin for DNA repair proteins, potentially stimulating histone acetylation or aiding chromatin remodeling complexes. Consistent with this scenario, mouse cells lacking HMGB1 are sensitive to UV-C irradiation (Lange et al., 2008). Moreover, HMGB1 was reported to interact with NER proteins, including XPC-RAD23, XPA and RPA (Lange et al., 2009), which could support a role in NER. Although intriguing, these results will need to be confirmed in human cells, considering that we recently reported that a role of murine HMGN1 in transcription-coupled DNA repair is not conserved in human cells (Apelt et al., 2020). To confirm a potential role of human HMGB1 in NER and to address whether the other HMGB-type proteins have a similar role, we developed an siRNA-based procedure to deplete all HMG proteins individually or in combination. It will be interesting to determine whether depletion of HMGB-type proteins renders cells sensitive to UV-C irradiation or results in a defect in unscheduled DNA synthesis following UV irradiation. Since the four HMGB proteins may act in a functionally redundant manner, it will be of interest to attempt to knock-down all four HMGB proteins simultaneously.

Earlier work has shown that murine HMGB has a role in various DNA repair pathways, including DNA double-strand break repair and mismatch repair (Stros et al., 2000; Yuan et al., 2004). Considering that MMR proteins were highly enriched in UV-damaged chromatin in our CHROMASS screen, it cannot be fully excluded that the recruitment following UV irradiation is linked to this pathway. MMR proteins were indeed found to excise misincorporated nucleotides opposite to UV-induced photolesions by TLS polymerases during replication (Tsaalbi-Shtylik et al., 2015). However, if this were the case, HMGB-type proteins would be expected to accumulate only in those cells that are undergoing DNA replication, which is not what we observed. To fully exclude this possibility, the recruitment of HMGB-type proteins should be monitored in non-dividing cells. Finally, it would be interesting to map the interactome of HMGB-type proteins following UV irradiation to identify potential

interactors that could provide clues to their function during the response to UV-induced DNA lesions. The future studies outlined here will reveal the functional relevance of the UV-induced chromatin interaction of HMGB-type proteins.

Figure S1

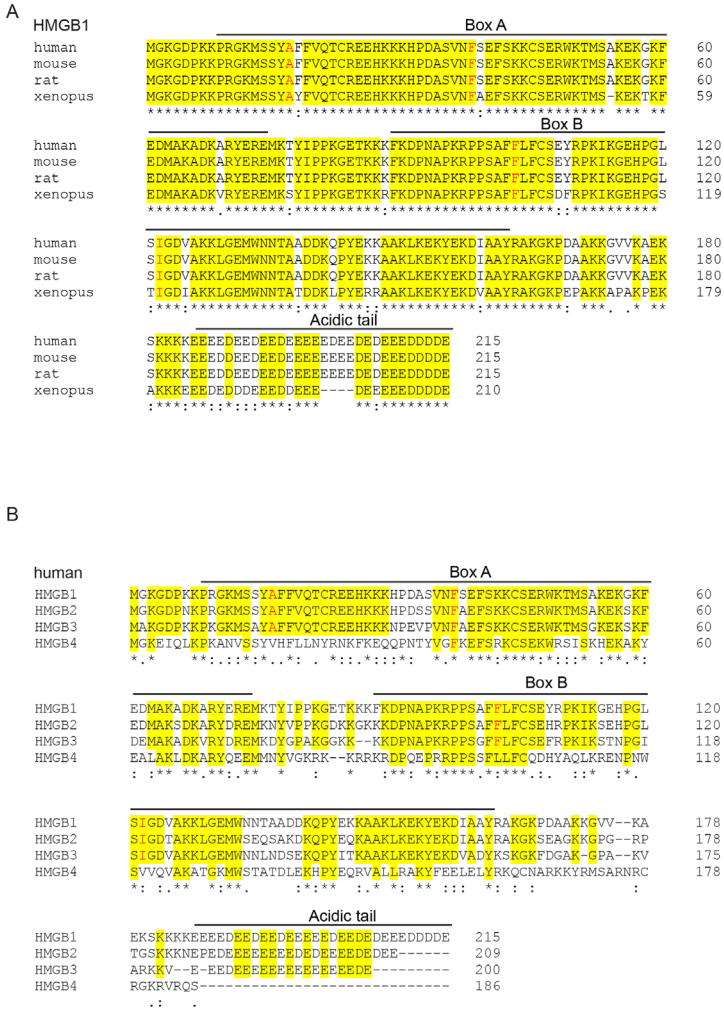

Figure S1. (A) Sequence alignment of human, mouse, rat and *Xenopus laevis* HMGB1 proteins. (B) Sequence alignment of human HMGB1, HMGB2, HMGB3 and HMGB4 protein. The four amino acids that are inserted into the minor groove of the DNA are highlighted in red.

Figure S2

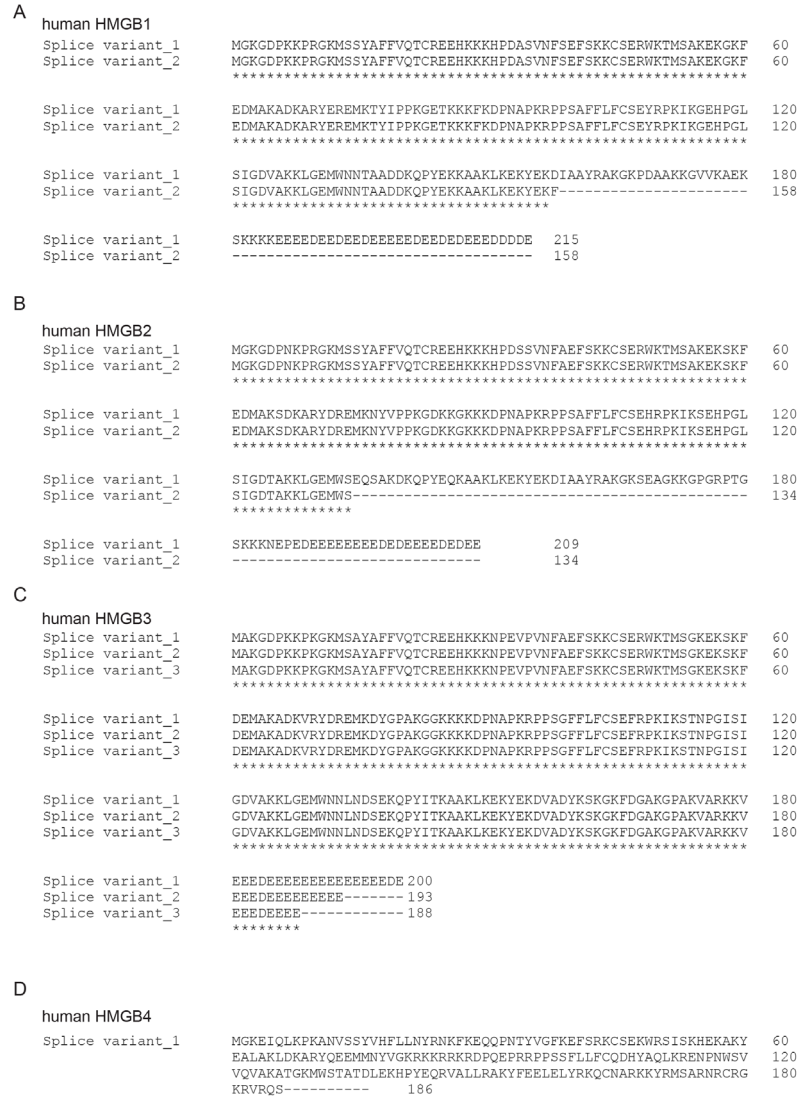

Figure S2. (A) Alignment of the splice variants of human HMGB1, (B) HMGB2, (C) HMGB3 and (D) HMGB4 protein. The four amino acids that are inserted into the minor groove of the DNA are highlighted in red.

Figure S3


```
human HMGB1

Splice variant_1      ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG   60
Splice variant_2      ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG   60
                      ************************************************************

Splice variant_1      CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG   120
Splice variant_2      CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG   120
                      ************************************************************

Splice variant_1      TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT   180
Splice variant_2      TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT   180
                      ************************************************************

Splice variant_1      GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT   240
Splice variant_2      GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT   240
                      ************************************************************

Splice variant_1      CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG   300
Splice variant_2      CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG   300
                      ************************************************************

Splice variant_1      GCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG   360
Splice variant_2      GCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG   360
                      ************************************************************
                                                 siRNA
Splice variant_1      TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC   420
Splice variant_2      TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC   420
                      ************************************************************

Splice variant_1      AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATATGAAAAGGATATTGCT   480
Splice variant_2      AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATATGAAAAGTTCTAG---   477
                      ********************************************************

Splice variant_1      GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA   540
Splice variant_2      ------------------------------------------------------------   477

Splice variant_1      AGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAG   600
Splice variant_2      ------------------------------------------------------------   477

Splice variant_1      GAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAATAA       648
Splice variant_2      ---------------------------------------------------       477
```

Figure S3. Alignment of cDNA of the splice variants of human HMGB1 and in yellow the sequence of the siRNA is indicated.

Figure S4

Figure S4. Alignment of cDNA of the splice variants of human HMGB2 and in yellow the sequence of the siRNA is indicated.

Figure S5

Figure S5. Alignment of cDNA of the splice variants of human HMGB3 and in yellow the sequence of the siRNA is indicated.

Figure S6

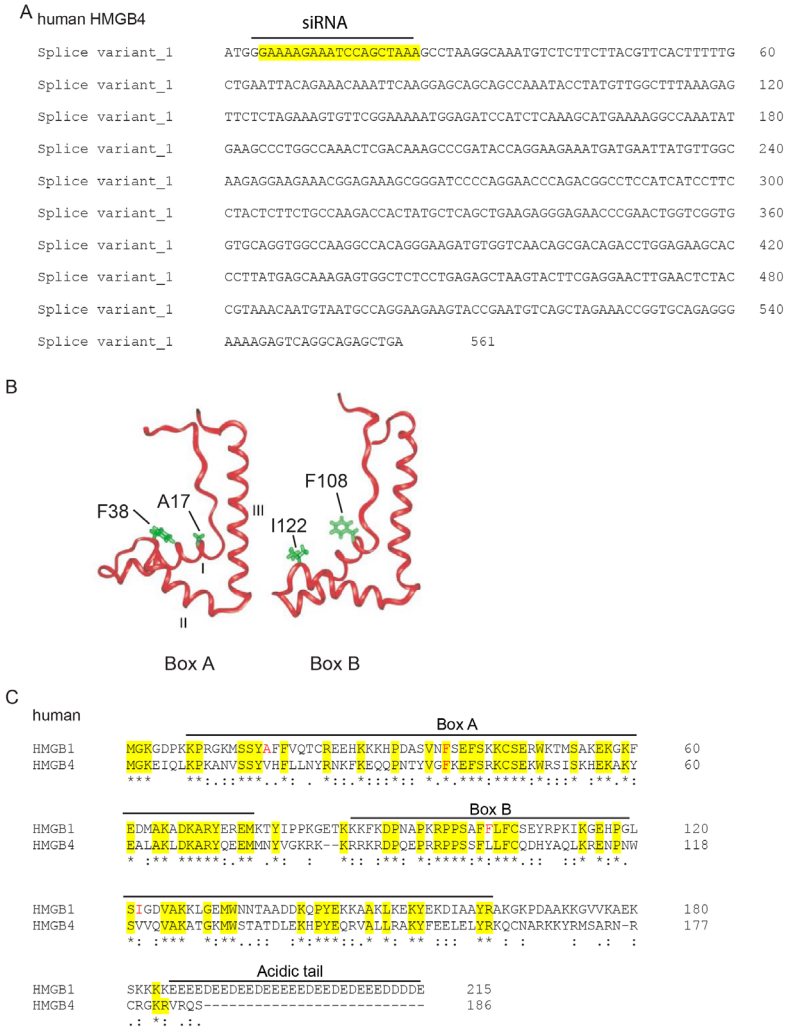

Figure S6. (A) Alignment of cDNA of the splice variants of human HMGB3 and in yellow the sequence of the siRNA is indicated. (B) Representation of the structure of the box A and box B of HMGB1 (Thomas and Travers, 2001). The two amino acids of each box which are inserted into the minor groove of the DNA are indicated in green. (C) Sequence alignment of human HMGB1 and HMGB4 protein. The four amino acids that are inserted into the minor groove of the DNA are highlighted in red.

Material and methods

Cell lines. All cell lines were cultured at 37°C in an atmosphere of 5% CO_2 in DMEM (Thermo Fisher Scientific) supplemented with penicillin/streptomycin (Sigma) and 10% fetal bovine serum (FBS; Bodinco BV).

Generation of knockout cell lines. To generate an HMGN1 knockout, U2OS(FRT) cells were co-transfected with pLV-U6g-PPB encoding a guide RNA from the LUMC / Sigma-Aldrich sgRNA library targeting a specific gene together with an expression vector encoding Cas9-2A-GFP (pX458; Addgene #48138) using lipofectamine 2000 (Invitrogen). Transfected U2OS(FRT) cells were selected on puromycin (1 µg/ml) for three days, plated at low density after which individual clones were isolated. Knockout clones were verified by western blot analysis.

Generation of Flip-In cell lines. To generate stable Flip-In cells, U2OS(FRT) cells were transfected with 5 µg of a pcDNA5/FRT/TO-Puro plasmid which contained the sequence of GFP and the cDNA of a protein of interest. Next the cells were incubated with 1 µg/mL puromycin for a couple of days and the survival cells were pooled. The expression of the GFP-tagged protein was induced by adding 2 µg/ml doxycycline for 24 h. Western blot. Cells were spun down, washed with PBS, and boiled for 10 min in Laemmli buffer (40 mM Tris pH 6.8, 3.35% SDS, 16.5% glycerol, 0.0005% Bromophenol Blue and 0.05 M DTT). Proteins were separated on 4-12% Criterion XT Bis-Tris gels (Bio-Rad, #3450124) in NuPAGE MOPS running buffer (NP0001-02 Thermo Fisher Scientific), and blotted onto PVDF membranes (IPFL00010, EMD Millipore). The membrane was blocked with blocking buffer (Rockland, MB-070-003) for 1 h at RT. The membrane was then probed with antibodies as indicated. An Odyssey CLx system (LI-COR Biosciences) was used for detection.

CHROMASS. CHROMASS experiments were performed as previously described. Sperm chromatin was isolated and mock-treated or irradiated with 2 kJ/m^2 UV-C. Sperm chromatin was subsequently incubated in non-licensing extracts at a final concentration of 16 ng/ml. After 45 min the reaction was stopped with 60 ml of ELB buffer supplemented with 0.2% Triton-X. The chromatin was spun down and the chromatin pellet was resuspended in 50 ml denaturation buffer (8 M urea, 100 mM Tris-HCl, pH 8) and transferred to a new low-binding tube. Then the cysteines were reduced (1 mM DTT for 15 min at RT) and alkylated (0.55 M chloroacetamide

for 40 min at RT protected from light). Hereafter the peptides were digested with 0.5 mg LysC for 2.5 h at RT, followed by an incubation overnight at 30°C with 0.5 mg trypsin. Before purifying the peptides by stage tipping (C18 material), the peptides were first acidified with 10% trifluoroacetic acid (pH < 4) and then neutralized with 400 mM NaCl. For the stage tipping, stage tips were first activated in 100% methanol, then equilibrated in 80% acetonitrile/ 10% formic acid, and finally washed twice in 0.1% formic acid. Then the samples were loaded on the stage tips and washed twice with 50 ml 0.1% formic acid. StageTip elution was performed with 80 ml of 25% acetonitrile in 0.1% formic acid, eluted samples were dried to completion in a SpeedVac at 60°C, dissolved in 10 ml 0.1% formic acid, and stored at -20°C until MS analysis.

UV-C laser micro irradiation. Cells were grown on 18-mm quartz coverslips and placed in a Chamlide CMB magnetic chamber in which growth medium was replaced by CO_2-independent Leibovitz's L15 medium (Thermo Fisher). UV-C laser tracks were made using a diode-pumped solid-state 266 nm Yttrium Aluminum Garnet laser (Average power 5 mW, repetition rate up to 10 kHz, pulse length 1 ns). The laser is integrated in a UGA-42-Caliburn/2L Spot Illumination system (Rapp OptoElectronic). Micro-irradiation was combined with live-cell imaging in an environmental chamber set to 37°C on an all-quartz widefield fluorescence Zeiss Axio Observer 7 microscope, using a 100x (1.2 NA) ultrafluar glycerol-immersion objective (UV-C). The laser system is coupled to the microscope via a triggerbox and a neutral density (ND-1) filter blocks 90% of the laser light. An HXP 120 V metal-halide lamp was used for excitation. Images were acquired in Zeiss ZEN and quantified in Image J.

Knockdown experiments. For the knockdown 75.000 cells were plated in a 12-well. The next day, the cells were transfected with 40 nM siRNA duplexes for 6 h, followed by a second transfection overnight using Lipofectamine RNAiMAX (Invitrogen). Before the analysis, the cells are washed and incubated for at least 24 h in DMEM (10% FBS; 1% PS) with doxycycline.

References

Adam, S., J. Dabin, O. Chevallier, O. Leroy, C. Baldeyron, A. Corpet, P. Lomonte, O. Renaud, G. Almouzni, and S.E. Polo. 2016. Real-Time Tracking of Parental Histones Reveals Their Contribution to Chromatin Integrity Following DNA Damage. Mol Cell. 64:65-78.

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

Aydin, O.Z., J.A. Marteijn, C. Ribeiro-Silva, A. Rodriguez Lopez, N. Wijgers, G. Smeenk, H. van Attikum, R.A. Poot, W. Vermeulen, and H. Lans. 2014. Human ISWI complexes are targeted by SMARCA5 ATPase and SLIDE domains to help resolve lesion-stalled transcription. Nucleic Acids Res. 42:8473-8485.

Birger, Y., K.L. West, Y.V. Postnikov, J.H. Lim, T. Furusawa, J.P. Wagner, C.S. Laufer, K.H. Kraemer, and M. Bustin. 2003. Chromosomal protein HMGN1 enhances the rate of DNA repair in chromatin. EMBO J. 22:1665-1675.

Brueckner, F., U. Hennecke, T. Carell, and P. Cramer. 2007. CPD damage recognition by transcribing RNA polymerase II. Science. 315:859-862.

Bustin, M. 2001. Chromatin unfolding and activation by HMGN(*) chromosomal proteins. Trends Biochem Sci. 26:431-437.

Caron, P., T. Pankotai, W.W. Wiegant, M.A.X. Tollenaere, A. Furst, C. Bonhomme, A. Helfricht, A. de Groot, A. Pastink, A.C.O. Vertegaal, M.S. Luijsterburg, E. Soutoglou, and H. van Attikum. 2019. WWP2 ubiquitylates RNA polymerase II for DNA-PK-dependent transcription arrest and repair at DNA breaks. Genes & development. 33:684-704.

Catez, F., D.T. Brown, T. Misteli, and M. Bustin. 2002. Competition between histone H1 and HMGN proteins for chromatin binding sites. EMBO Rep. 3:760-766.

Catez, F., H. Yang, K.J. Tracey, R. Reeves, T. Misteli, and M. Bustin. 2004. Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. Mol Cell Biol. 24:4321-4328.

Cho, I., P.F. Tsai, R.J. Lake, A. Basheer, and H.Y. Fan. 2013. ATP-dependent chromatin remodeling by Cockayne syndrome protein B and NAP1-like histone chaperones is required for efficient transcription-coupled DNA repair. PLoS Genet. 9:e1003407.

Citterio, E., V. Van Den Boom, G. Schnitzler, R. Kanaar, E. Bonte, R.E. Kingston, J.H. Hoeijmakers, and W. Vermeulen. 2000. ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor. Mol Cell Biol. 20:7643-7653.

Damsma, G.E., A. Alt, F. Brueckner, T. Carell, and P. Cramer. 2007. Mechanism of transcriptional stalling at cisplatin-damaged DNA. Nat Struct Mol Biol. 14:1127-1133.

de Laat, W.L., N.G. Jaspers, and J.H. Hoeijmakers. 1999. Molecular mechanism of nucleotide excision repair. Genes & development. 13:768-785.

Deng, T., Z.I. Zhu, S. Zhang, Y. Postnikov, D. Huang, M. Horsch, T. Furusawa, J. Beckers, J. Rozman, M. Klingenspor, O. Amarie, J. Graw, B. Rathkolb, E. Wolf, T. Adler, D.H. Busch, V. Gailus-Durner, H.Fuchs, M. Hrabe de Angelis, A. van der Velde, L. Tessarollo, I. Ovcherenko, D. Landsman, and M. Bustin. 2015. Functional compensation among HMGN variants modulates the DNase I hypersensitive sites at enhancers. Genome research. 25:1295-1308.

Dinant, C., G. Ampatziadis-Michailidis, H. Lans, M. Tresini, A. Lagarou, M. Grosbart, A.F. Theil, W.A. van Cappellen, H. Kimura, J. Bartek, M. Fousteri, A.B. Houtsmuller, W. Vermeulen, and J.A. Marteijn. 2013. Enhanced chromatin dynamics by FACT promotes transcriptional restart after UV-induced DNA damage. Mol Cell. 51:469-479.

Dinant, C., M. de Jager, J. Essers, W.A. van Cappellen, R. Kanaar, A.B. Houtsmuller, and W. Vermeulen. 2007. Activation of multiple DNA repair pathways by sub-nuclear damage induction methods. J Cell Sci. 120:2731-2740.

Ding, H.F., M. Bustin, and U. Hansen. 1997. Alleviation of histone H1-mediated transcriptional repression and chromatin compaction by the acidic activation region in chromosomal protein HMG-14. Mol Cell Biol. 17:5843-5855.

Fei, J., and J. Chen. 2012. KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). J Biol Chem.

287:35118-35126.
Fitch, M.E., I.V. Cross, S.J. Turner, S. Adimoolam, C.X. Lin, K.G. Williams, and J.M. Ford. 2003. The DDB2 nucleotide excision repair gene product p48 enhances global genomic repair in p53 deficient human fibroblasts. DNA Repair (Amst). 2:819-826.
Fitch, M.E., S. Nakajima, A. Yasui, and J.M. Ford. 2003. In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. J Biol Chem. 278:46906-46910.
Gerlitz, G. 2010. HMGNs, DNA repair and cancer. Biochim Biophys Acta. 1799:80-85.
Gonzalez-Romero, R., J.M. Eirin-Lopez, and J. Ausio. 2015. Evolution of high mobility group nucleosome-binding proteins and its implications for vertebrate chromatin specialization. Mol Biol Evol. 32:121-131.
Hanawalt, P.C., and G. Spivak. 2008. Transcription-coupled DNA repair: two decades of progress and surprises. Nat Rev Mol Cell Biol. 9:958-970.
Jaspers, N.G., A. Raams, M.J. Kelner, J.M. Ng, Y.M. Yamashita, S. Takeda, T.C. McMorris, and J.H. Hoeijmakers. 2002. 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). 1:1027-1038.
Jiang, Y., X. Wang, S. Bao, R. Guo, D.G. Johnson, X. Shen, and L. Li. 2010. INO80 chromatin remodeling complex promotes the removal of UV lesions by the nucleotide excision repair pathway. Proc Natl Acad Sci U S A. 107:17274-17279.
Kim, Y.C., G. Gerlitz, T. Furusawa, F. Catez, A. Nussenzweig, K.S. Oh, K.H. Kraemer, Y. Shiloh, and M. Bustin. 2009. Activation of ATM depends on chromatin interactions occurring before induction of DNA damage. Nat Cell Biol. 11:92-96.
Luijsterburg, M.S., M. Lindh, K. Acs, M.G. Vrouwe, A. Pines, H. van Attikum, L.H. Mullenders, and N.P. Dantuma. 2012. DDB2 promotes chromatin decondensation at UV-induced DNA damage. J Cell Biol. 197:267-281.
Luijsterburg, M.S., and H. van Attikum. 2011. Chromatin and the DNA damage response: the cancer connection. Mol Oncol. 5:349-367.
Marteijn, J.A., H. Lans, W. Vermeulen, and J.H. Hoeijmakers. 2014. Understanding nucleotide excision repair and its roles in cancer and ageing. Nat Rev Mol Cell Biol. 15:465-481.
Masaoka, A., N.R. Gassman, P.S. Kedar, R. Prasad, E.W. Hou, J.K. Horton, M. Bustin, and S.H. Wilson. 2012. HMGN1 protein regulates poly(ADP-ribose) polymerase-1 (PARP-1) self-PARylation in mouse fibroblasts. J Biol Chem. 287:27648-27658.
Mellon, I., G. Spivak, and P.C. Hanawalt. 1987. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. Cell. 51:241-249.
Moser, J., M. Volker, H. Kool, S. Alekseev, H. Vrieling, A. Yasui, A.A. van Zeeland, and L.H. Mullenders. 2005. The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions. DNA Repair (Amst). 4:571-582.
Nakazawa, Y., S. Yamashita, A.R. Lehmann, and T. Ogi. 2010. A semi-automated non-radioactive system for measuring recovery of RNA synthesis and unscheduled DNA synthesis using ethynyluracil derivatives. DNA Repair (Amst). 9:506-516.
Oksenych, V., A. Zhovmer, S. Ziani, P.O. Mari, J. Eberova, T. Nardo, M. Stefanini, G. Giglia-Mari, J.M. Egly, and F. Coin. 2013. Histone methyltransferase DOT1L drives recovery of gene expression after a genotoxic attack. PLoS Genet. 9:e1003611.
Panier, S., Y. Ichijima, A. Fradet-Turcotte, C.C. Leung, L. Kaustov, C.H. Arrowsmith, and D. Durocher. 2012. Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. Mol Cell. 47:383-395.
Pines, A., M.G. Vrouwe, J.A. Marteijn, D. Typas, M.S. Luijsterburg, M. Cansoy, P. Hensbergen, A. Deelder, A. de Groot, S. Matsumoto, K. Sugasawa, N. Thoma, W. Vermeulen, H. Vrieling, and L. Mullenders. 2012. PARP1 promotes nucleotide excision repair through DDB2 stabilization and recruitment of ALC1. J Cell Biol. 199:235-249.
Postnikov, Y., and M. Bustin. 2010. Regulation of chromatin structure and function by HMGN proteins. Biochim Biophys Acta. 1799:62-68.
Rapic-Otrin, V., V. Navazza, T. Nardo, E. Botta, M. McLenigan, D.C. Bisi, A.S. Levine, and M. Stefanini. 2003. True XP group E patients have a defective UV-damaged DNA binding protein complex and mutations in DDB2 which reveal the functional domains of its p48

product. Human molecular genetics. 12:1507-1522.
Ruthemann, P., C. Balbo Pogliano, T. Codilupi, Z. Garajova, and H. Naegeli. 2017. Chromatin remodeler CHD1 promotes XPC-to-TFIIH handover of nucleosomal UV lesions in nucleotide excision repair. EMBO J. 36:3372-3386.
Scrima, A., R. Konickova, B.K. Czyzewski, Y. Kawasaki, P.D. Jeffrey, R. Groisman, Y. Nakatani, S. Iwai, N.P. Pavletich, and N.H. Thoma. 2008. Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. Cell. 135:1213-1223.
Sellou, H., T. Lebeaupin, C. Chapuis, R. Smith, A. Hegele, H.R. Singh, M. Kozlowski, S. Bultmann, A.G. Ladurner, G. Timinszky, and S. Huet. 2016. The poly(ADP-ribose)-dependent chromatin remodeler Alc1 induces local chromatin relaxation upon DNA damage. Mol Biol Cell. 27:3791-3799.
Sugasawa, K., J.M. Ng, C. Masutani, S. Iwai, P.J. van der Spek, A.P. Eker, F. Hanaoka, D. Bootsma, and J.H. Hoeijmakers. 1998. Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Mol Cell. 2:223-232.
Tan, T., and G. Chu. 2002. p53 Binds and activates the xeroderma pigmentosum DDB2 gene in humans but not mice. Mol Cell Biol. 22:3247-3254.
Tantin, D., A. Kansal, and M. Carey. 1997. Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. Mol Cell Biol. 17:6803-6814.
van der Horst, G.T., L. Meira, T.G. Gorgels, J. de Wit, S. Velasco-Miguel, J.A. Richardson, Y. Kamp, M.P. Vreeswijk, B. Smit, D. Bootsma, J.H. Hoeijmakers, and E.C. Friedberg. 2002. UVB radiation-induced cancer predisposition in Cockayne syndrome group A (Csa) mutant mice. DNA Repair (Amst). 1:143-157.
van der Horst, G.T., H. van Steeg, R.J. Berg, A.J. van Gool, J. de Wit, G. Weeda, H. Morreau, R.B. Beems, C.F. van Kreijl, F.R. de Gruijl, D. Bootsma, and J.H. Hoeijmakers. 1997. Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. Cell. 89:425-435.
Wienholz, F., D. Zhou, Y. Turkyilmaz, P. Schwertman, M. Tresini, A. Pines, M. van Toorn, K. Bezstarosti, J.A.A. Demmers, and J.A. Marteijn. 2019. FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. Nucleic Acids Res. 47:4011-4025.
Wilson, B.T., Z. Stark, R.E. Sutton, S. Danda, A.V. Ekbote, S.M. Elsayed, L. Gibson, J.A. Goodship, A.P. Jackson, W.T. Keng, M.D. King, E. McCann, T. Motojima, J.E. Murray, T. Omata, D. Pilz, K. Pope, K. Sugita, S.M. White, and I.J. Wilson. 2016. The Cockayne Syndrome Natural History (CoSyNH) study: clinical findings in 102 individuals and recommendations for care. Genetics in medicine : official journal of the American College of Medical Genetics. 18:483-493.