



Universiteit
Leiden

The Netherlands

Unravelling molecular mechanisms in transcription-coupled nucleotide excision repair

Weegen, Y. van der

Citation

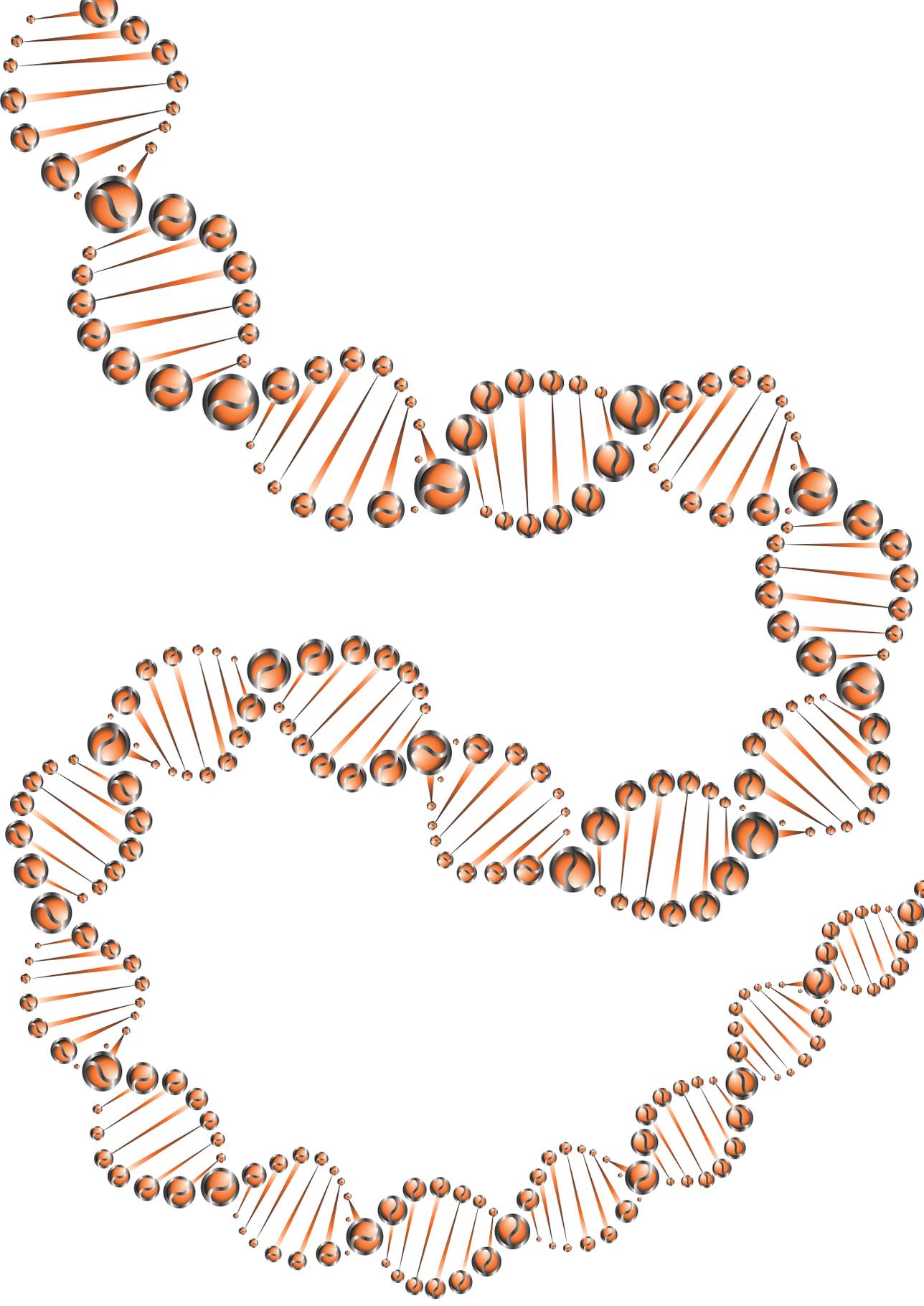
Weegen, Y. van der. (2022, February 17). *Unravelling molecular mechanisms in transcription-coupled nucleotide excision repair*. Retrieved from <https://hdl.handle.net/1887/3275039>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3275039>

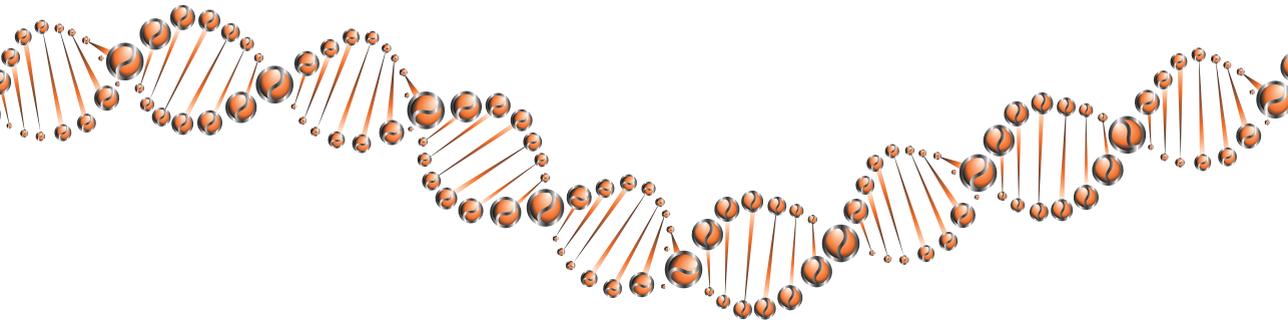
Note: To cite this publication please use the final published version (if applicable).



2

The cooperative action of CSB, CSA and UVSSA target TFIIF to DNA damage-stalled RNA polymerase II

Yana van der Weegen, Hadar Golan Berman, Tycho E.T. Mevissen, Katja Apelt, Román González-Prieto, Joachim Goedhart, Elisheva E. Heilbrun, Alfred C.O. Vertegaal, Diana van den Heuvel, Johannes C. Walter, Sheera Adar, and Martijn S. Lijsterburg



Abstract

The response to DNA damage-stalled RNA polymerase II (RNAPII α) involves the assembly of the transcription-coupled repair (TCR) complex on actively transcribed strands. The function of the TCR proteins CSB, CSA and UVSSA and the manner in which the core DNA repair complex, including transcription factor IIIH (TFIIH), is recruited are largely unknown. Here, we define the assembly mechanism of the TCR complex in human isogenic knockout cells. We show that TCR is initiated by RNAPII α -bound CSB, which recruits CSA through a newly identified CSA-interaction motif (CIM). Once recruited, CSA facilitates the association of UVSSA with stalled RNAPII α . Importantly, we find that UVSSA is the key factor that recruits the TFIIH complex in a manner that is stimulated by CSB and CSA. Together these findings identify a sequential and highly cooperative assembly mechanism of TCR proteins and reveal the mechanism for TFIIH recruitment to DNA damage-stalled RNAPII α to initiate repair.

Introduction

Nucleotide excision repair (NER) is a versatile DNA repair pathway that removes a wide range of helix-distorting DNA lesions from the genome, including ultraviolet (UV) light-induced photolesions. Transcription-coupled repair (TCR) is a specialized NER sub-pathway that specifically removes DNA lesions from actively transcribed DNA strands [2]. It is believed that the TCR pathway is initiated by the stalling of elongating RNA polymerase II (RNAPII α) at DNA lesions, which triggers the recruitment of the core NER machinery to repair these lesions [3]. After lesion recognition, the transcription factor IIIH (TFIIH) complex is recruited to unwind the DNA [4, 5] followed by dual incision and the release of a 22-30 nucleotide-long DNA strand containing the lesion [6, 7]. The generated single-stranded DNA gap is filled by repair synthesis and the nick is sealed [3]. However, the mechanism through which TCR recognizes transcription-blocking lesions and recruits the repair machinery are unknown.

Inherited defects that selectively impair TCR give rise to Cockayne Syndrome (CS) and UV-sensitive syndrome (UV S S). Although cells from both CS and UV S S patients show a defect in TCR [8, 9], the phenotypes associated with these disorders are very different. CS is characterized by severe and progressive neurodegeneration [10, 11], while UV S S is characterized by mild UV sensitivity [12-14]. The majority of CS patients carry mutations in the *CSB* or *CSA* genes [15, 16], while UV S S patients carry mutations in the *UVSSA* gene [17, 18].

The CSB protein contains a central SWI2/SNF2-like DNA-dependent ATPase domain [19], and resides in a complex with RNAPII α [20, 21]. Live-cell imaging suggests that CSB monitors the progression of transcription elongation by continuously probing RNAPII α complexes [22]. It has been suggested that CSB is involved in the repositioning of RNAPII to make the DNA lesion accessible for repair proteins [23]. Although the association of CSB with RNAPII is sufficient to recruit TFIIH in vitro [24], it is unknown whether additional factors are required to trigger the recruitment of the repair machinery in vivo.

Like CSB, the CSA and UVSSA proteins also associate with DNA damage-stalled RNAPII α [17, 18, 25, 26]. The CSA protein contains seven WD40 repeats that form a seven bladed β -propeller [27]. Earlier work has shown that CSA is

incorporated into a DDB1-CUL4-based E3 ubiquitin ligase complex [25, 28] that becomes transiently activated in response to UV irradiation and targets CSB for proteasomal degradation [29]. Furthermore, the CSA complex also targets the UV-induced transcription repressor ATF3 as a means to regulate transcription restart after UV [30]. Current models suggest that CSA is dispensable for the recruitment of the excision repair machinery to stalled RNAPII [31], and that CSA is unlikely to recruit UVSSA to sites of UV-induced DNA damage [32]. Thus, the precise recruitment mechanism and the role of CSA in TCR is currently not clear.

The UVSSA protein contains an N-terminal VHS domain and a C-terminal DUF2043 domain of unknown function. Several studies reported that UVSSA, likely through its binding partner USP7, protects CSB from UV-induced degradation [17, 18, 26, 33]. However, ectopic expression of CSB in UVSSA-deficient cells did not rescue TCR, suggesting that UVSSA has additional functions in this repair mechanism [17]. Moreover, UVSSA was found to associate with RNAPII [18, 26], but whether UVSSA is constitutively bound to RNAPII, or associates with DNA damage-stalled RNAPII through either CSA or CSB is still a topic of debate.

The TFIIH complex consists of seven core subunits, including the XPB and XPD helicases, and three CAK kinase subunits [34]. While the CAK complex is crucial during transcription initiation, it inhibits the XPD helicase activity required for repair [35]. The release of the CAK complex from core TFIIH is triggered by the association of repair factors XPA and XPG, which switches TFIIH from a transcription factor into a repair factor [35, 36].

Despite the knowledge that CSB, CSA, and UVSSA are required for TCR, we still know very little about how the interplay between these proteins targets the core repair machinery, including TFIIH, to DNA damage-stalled RNAPII. In this study, we demonstrate a sequential and highly cooperative assembly of TCR proteins and unveil the mechanism for TFIIH recruitment to DNA damage-stalled RNAPIIo.

Results

Isolation of active TCR complexes under native conditions

Our current understanding of the assembly and functioning of multi-protein complexes that mediate transcription-coupled DNA repair (TCR) is fairly limited. This is largely due to a lack of sensitive methods to isolate active TCR complexes and analyze their composition. To overcome this limitation, we set out to establish a new immunoprecipitation-based method to isolate the elongating form of RNA polymerase II (RNAPIIo) and associated proteins from the chromatin fraction of UV-irradiated cells under native conditions (Fig. 1a). To this end, we employed extensive benzonase treatment to solubilize the chromatin fraction after centrifugation, followed by immunoprecipitation using antibodies that recognize the Ser2-phosphorylated form of RNAPII. This RNAPII modification is absent from transcription start sites (TSS), but increases across gene bodies and is associated with transcription elongation [37]. Immunoprecipitation of RNAPIIo revealed a UV-specific association with the Cockayne syndrome (CS) proteins CSB and CSA, as well as with CUL4, DDB1 and RBX1, which are subunits of a DDB1-CUL4 (CRL4) E3 ubiquitin ligase complex that associates with CSA [25, 28]. Additionally, several subunits of the TFIIH complex (XPD/p80,

XPB/p89, GTFH1/p62, and CDK7) also associated with RNAPII α after UV (Fig. 1b; Supplementary Fig. 1a). Similar results were obtained when using antibodies against Ser5-phosphorylated RNAPII (Supplementary Fig. 1b). Importantly, we did not detect an RNAPII-TFIIH interaction in unirradiated cells, suggesting that our procedure indeed does not capture RNAPII involved in transcription initiation during which it interacts extensively with TFIIH [38]. In line with this, genomic DNA fragments immunoprecipitated together with RNAPII α after UV irradiation were highly enriched for the most abundant UV-induced DNA lesion (cyclobutane pyrimidine dimer; CPD) (Fig. 1c). These findings suggest that a considerable fraction of the RNAPII α molecules we capture under our conditions are stalled at DNA lesions. Although the CS proteins and TFIIH readily assembled with RNAPII α after UV irradiation, downstream repair proteins such as XPA, ERCC1-XPF and XRCC1 could not be detected (Fig. 1b, Supplementary Fig. 1a). It should be noted that we could not detect UVSSA either after pull-down of RNAPII α or in whole cell lysates due to a lack of specific antibodies (Supplementary Fig. 1c). These initial results suggest that CSB, CSA and TFIIH associate with DNA damage-stalled RNAPII, but that the assembly of downstream repair factors may require the removal or backtracking of RNAPII to make the lesion accessible to the repair machinery [23].

CSA is recruited to damage-stalled RNAPII by CSB

To acquire more insights into the initial assembly of TCR factors, we generated *CSB*, *CSA*, and *UVSSA* knockout (KO) cells using CRISPR-Cas9-mediated genome editing in U2OS cells equipped with the Flp-In/T-REx system. The knockout of *CSB*, *CSA*, and *UVSSA* was confirmed by western blot analysis and/or DNA sequencing (Fig. 1d; Supplementary Fig. 2a, b). Clonogenic survival assays revealed that all TCR-KO cells were highly sensitive to transcription-blocking DNA damage induced by Illudin S (Fig. 1e), which is a natural compound from mushroom *O. illudens* causing DNA lesions that are exclusively repaired by TCR [39]. Importantly, complementation of these TCR-KO cells with inducible GFP-tagged versions of *CSB*, *CSA*, and *UVSSA* fully restored their resistance to Illudin S (Fig. 1d, e). We next applied our immunoprecipitation-based method in the different TCR-KO cells to establish how *CSB* and *CSA* recruitment to DNA damage-stalled RNAPII α is regulated. *CSB* associated with RNAPII α in wild-type (WT), *CSA*-KO and *UVSSA*-KO cells specifically after UV irradiation, suggesting that *CSB* is the first of these proteins to associate with DNA damage-stalled RNAPII α (Fig. 1f). The association of *CSA* with stalled RNAPII α was abolished in *CSB*-KO cells, but was not affected in cells lacking *UVSSA* (Fig. 1f). Importantly, re-expressing GFP-tagged *CSB* in the *CSB*-KO cells restored the association between RNAPII α and *CSA* (Fig. 1f, g). Next, we asked whether *CSA* mediates the recruitment of the CRL4 complex, including its E3 ubiquitin ligase partner DDB1 to DNA damage-stalled RNAPII α . As an additional control we also included *XPC*-KO cells, which are deficient in global genome repair (GGR; Supplementary Fig. 2c). Immunoprecipitation of RNAPII α revealed a UV-specific interaction with DDB1 in WT, *XPC*-KO, and *UVSSA*-KO cells (Fig. 1h). However, this interaction was completely abolished in *CSA*-KO and *CSB*-KO cells, showing that *CSA* indeed mediates the recruitment of the CRL4 complex to lesion-stalled RNAPII α (Fig. 1h).

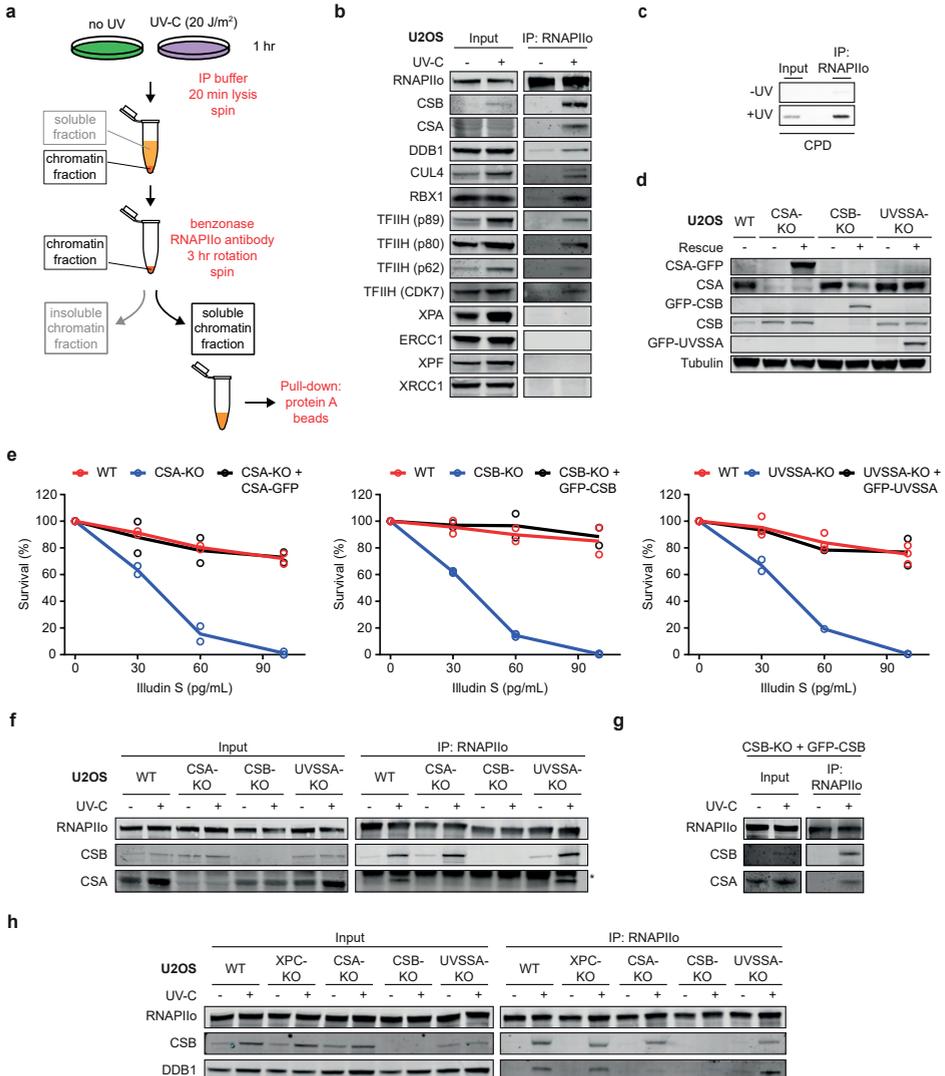


Fig. 1: CSA is recruited to DNA damage-stalled RNAPII α by CSB. (a) Outline of a new IP method to isolate RNAPII α and associated proteins from mock-treated or UV-irradiated (20 J/m²) U2OS (FRT) cells. (b) Endogenous RNAPII Co-IPs on WT cells stained for the indicated TCR proteins. Note that it is not possible to stain for all these proteins on one membrane. This panel is a composite of several representative Co-IPs. See Supplementary Fig. 1a for each individual Co-IP. (c) Endogenous RNAPII Co-IP followed by slot blot analysis of CPDs (d) Western blot analysis of CSB, CSA, and UVSSA knockout cells complemented with inducible GFP-tagged versions of these proteins ($n=2$). See Supplementary Fig. 2a, b for validation of knockouts by sequencing. (e) Clonogenic Illudin S survival of WT, CSB, CSA, and UVSSA knockout and rescue cell lines. Each symbol represents the mean of an independent experiment ($n=2$ for all except for WT in UVSSA-KO figure which is $n=3$) each experiment contains 2 or 3 technical replicates. Endogenous RNAPII Co-IP on (f) WT, CSB, CSA, and UVSSA knockout cells, (g) CSB-KO stably expressing GFP-CSB, and (h) WT, XPC, CSB, CSA, and UVSSA knockout cells. The asterisk in panel e indicates the heavy chain of the RNAPII antibody. At least two independent replicates of each IP experiment were performed obtaining similar results.

Mapping the CSA-interaction motif (CIM) in CSB

In order to gain a better understanding of the CSA recruitment mechanism by CSB, we aimed to identify the region in CSB that is required for the interaction with CSA. To this end, we employed a chromatin-tethering approach making use of the U2OS 2-6-3 cell line harboring an integrated LacO array in the genome [40]. This cell line enables the analysis of protein-protein interactions by tethering proteins of interest fused to the bacterial LacR and fluorescent protein mCherry to a defined chromosomal region [41, 42] (Fig. 2a). Expression of mCherry-LacR fused to full-length CSB (Fig. 2b) resulted in clear localization of the fusion protein to the LacO array and triggered the robust recruitment of CSA-GFP (Fig. 2c). In contrast, expression of LacR alone failed to recruit CSA-GFP to the LacO array (Fig. 2c).

To identify the CSA-interaction domain in CSB, we fused various truncated fragments of CSB to mCherry-LacR and examined their ability to recruit CSA-GFP to the LacO array (Fig. 2b, Supplementary Fig. 3-Fig. 4). Fragments of CSB spanning the N-terminus or the central region containing the conserved ATPase/helicase domain (N, M, and Δ C) were unable to recruit CSA-GFP. Conversely, tethering of a LacR-tagged CSB region spanning the C-terminus (C and Δ N) triggered robust recruitment of CSA-GFP (Fig. 2b-d, Supplementary Fig. 3). These results suggest that the C-terminus of CSB is essential for the interaction with CSA. The C-terminus of CSB contains a ubiquitin-binding domain (UBD; 1400-1428 [43]) and a recently identified winged-helix domain (WHD; 1417-1493) that interacts with RIF1 [44]. Interestingly, we found that the most N-terminal region (1221-1305) of the CSB C-terminus alone, or fragments containing solely the UBD (1400-1493) or WHD (1417-1493) domains do not support CSA recruitment. However, a region just upstream of the UBD (1306-1399) is sufficient to mediate CSA recruitment to the LacO array (Fig. 2b-d, Supplementary Fig. 3). Importantly, we found that tethering full-length CSB lacking this minimal interaction region (Δ 1306-1399) indeed failed to support CSA recruitment (Fig. 2b-d). Further deletion analysis showed that CSB lacking the region just upstream of the UBD (1353-1399) failed to recruit CSA-GFP, whereas CSB lacking the UBD (1400-1428) or amino acids 1306-1352 were fully proficient in interacting with CSA-GFP (Supplementary Fig. 4). Moreover, while CSB Δ 1353-1368 and CSB Δ 1369-1384 were fully proficient in recruiting CSA-GFP to the LacO array, deleting amino acids 1385-1399 abolished the ability of CSB to interact with CSA-GFP (Fig. 2b-d, Supplementary Fig. 4). These findings identify an evolutionary conserved CSA-interaction motif (CIM) in CSB that is located between amino acids 1385-1399 (Fig. 2e; Supplementary Fig. 5).

The C-terminal CIM in CSB recruits CSA to stalled RNAPII

We next set out to address the importance of this new CSB motif under more physiological conditions. To this end, we stably expressed GFP-tagged CSB^{WT} or CSB Δ CIM in CSB-KO cells (Fig. 3a, b). Pull-down of GFP-tagged CSB^{WT} showed a strong UV-induced interaction with CSA, which was virtually absent after pull-down of CSB Δ CIM even though equal amounts of CSB were immunoprecipitated (Fig. 3c). These findings were confirmed by quantitative mass spectrometry (MS) after pull-down of GFP-tagged versions of either CSB^{WT} or CSB Δ CIM (Supplementary Fig. 6a-c). Immunoprecipitation of endogenous RNAPII α in

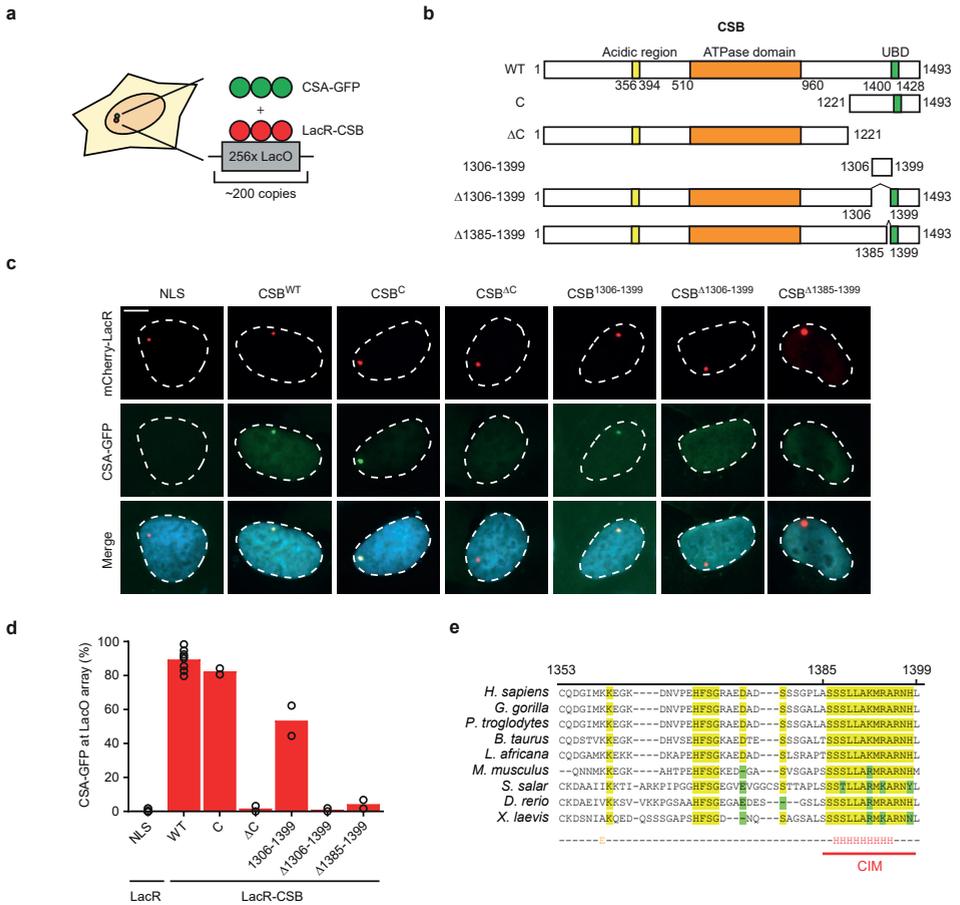


Fig. 2: CSA interacts with the newly identified C-terminal CIM of CSB. (a) Outline of the chromatin-tethering approach in U2OS 2-6-3 cells. (b) A schematic representation of CSB and its deletion mutants. (c) Recruitment of CSA-GFP to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins (scale bar = 5 μ m). See Supplementary Fig. 3-Fig. 4 for a full overview of all tested mutants. (d) Quantification of CSA-GFP and mCherry-LacR-CSB co-localization at the LacO array. Each symbol represents the mean of an independent experiment ($n=2$ for all except LacR-NLS and LacR-CSB^{WT} which is $n=8$, >50 cells collected per experiment). (e) Sequence alignment of CSB orthologues. See Supplementary Fig. 5 for additional alignments.

these cell lines showed that both CSB^{WT} and CSB^{ΔCIM} associated equally with RNAPIIo after UV irradiation. However, CSB^{ΔCIM} failed to recruit CSA to DNA damage-stalled RNAPIIo, while a strong association of CSA was observed in cells expressing CSB^{WT} (Fig. 3d). Importantly, the stable expression of GFP-CSB^{ΔCIM} in CSB-KO cells failed to restore sensitivity to Illudin S, while expression of GFP-CSB^{WT} almost fully rescued this phenotype (Fig. 3e). To determine whether the CIM can mediate a functional interaction between CSB and CSA, we mixed recombinant *Xenopus laevis* CSB^{WT} or CSB^{ΔCIM} with ubiquitin, E1, E2, and the E3 ubiquitin ligase CRL4^{CSA} consisting of *Xenopus laevis* CSA, DDB1, CUL4A, and RBX1 (Supplementary Fig. 6d). While xCRL4^{CSA} promoted the efficient

ubiquitylation of xICSB^{WT} , it did not ubiquitylate $\text{xICSB}^{\Delta\text{CIM}}$ (Fig. 3f). These data suggest that xICSB uses its CIM to interact directly with xICSA . Consistent with this interpretation, immobilized xICSB^{WT} but not $\text{xICSB}^{\Delta\text{CIM}}$ interacted with endogenous xICSA from *Xenopus* egg extract (Fig. 3g). Similar results were observed when xICSB was substituted with hsCSB (Fig. 3f, g). Collectively, these data demonstrate that CSA is recruited to DNA damage-stalled RNAPII α by CSB through direct interactions with the newly identified C-terminal CIM in CSB.

UVSSA is recruited to DNA damage-stalled RNAPII α by CSA

Previous studies have demonstrated that UVSSA associates with RNAPII α , but due to conflicting results, it remains unclear if UVSSA recruitment to RNAPII α is enhanced by UV irradiation and dependent on the CS proteins [18, 26, 32]. Therefore, we monitored GFP-UVSSA recruitment to RNAPII α in *UVSSA*-KO cells complemented with GFP-UVSSA (WT) in which we additionally knocked out either *CSB* or *CSA*. The knockout of *CSB* and *CSA* was verified by western blot analysis, DNA sequencing (Fig. 4a; Supplementary Fig. 2), and Illudin S clonogenic survival assays (Fig. 4b). Immunoprecipitation of endogenous RNAPII α in these cell lines showed that GFP-UVSSA became readily detectable after UV irradiation in WT cells, whereas this interaction was virtually absent in *CSA*-KO and *CSB*-KO cells (Fig. 4c). Thus, GFP-UVSSA is targeted to DNA damage-stalled RNAPII α in a manner that is dependent on the CS proteins [18]. Moreover, pull-down of GFP-UVSSA confirmed a robust UV-induced association with RNAPII α , CSB, and CSA. However, these UV-specific interactions were abolished in *CSB*-KO and *CSA*-KO cells. Interestingly, we detected a weak UV-independent interaction between GFP-UVSSA and CSA, which was enhanced after UV irradiation in a manner that required CSB (Fig. 4d). These findings suggest that the cooperative assembly of the TCR complex is important to mediate efficient targeting of UVSSA to lesion-stalled RNAPII α .

CSB and CSA are required for the recruitment of the TFIID complex

Either CSB, CSA, or UVSSA can each associate with TFIID [24, 27, 45], but which of these proteins is responsible for the recruitment of TFIID to DNA damage-stalled RNAPII α is currently unknown. To directly assess if CSB and CSA are required for the recruitment of TFIID, we monitored TFIID (p62 and p89) recruitment in *UVSSA*-KO complemented with GFP-UVSSA (WT) in which we additionally knocked out either *CSB* or *CSA*. Immunoprecipitation of endogenous RNAPII α revealed a UV-specific interaction with TFIID in WT cells, while these interactions were severely reduced in the *CSB*-KO and *CSA*-KO cells (Fig. 5a). Interestingly, TFIID also failed to associate with RNAPII α in *CSB*-KO cells complemented with GFP- $\text{CSB}^{\Delta\text{CIM}}$ (Supplementary Fig. 6e), consistent with our findings that this mutant is not capable of recruiting CSA (Fig. 3c, g). These initial results suggest that the TFIID complex is recruited in a manner that requires both CS proteins.

UVSSA targets the TFIID complex to lesion-stalled RNAPII α

It has been reported that UVSSA can interact with TFIID [17, 33, 45], but whether this reflects a constitutive interaction or a UV-induced association is unclear. To gain more insight into the nature of this interaction, we immunoprecipitated GFP-UVSSA from the solubilized chromatin fraction of mock-treated and UV-

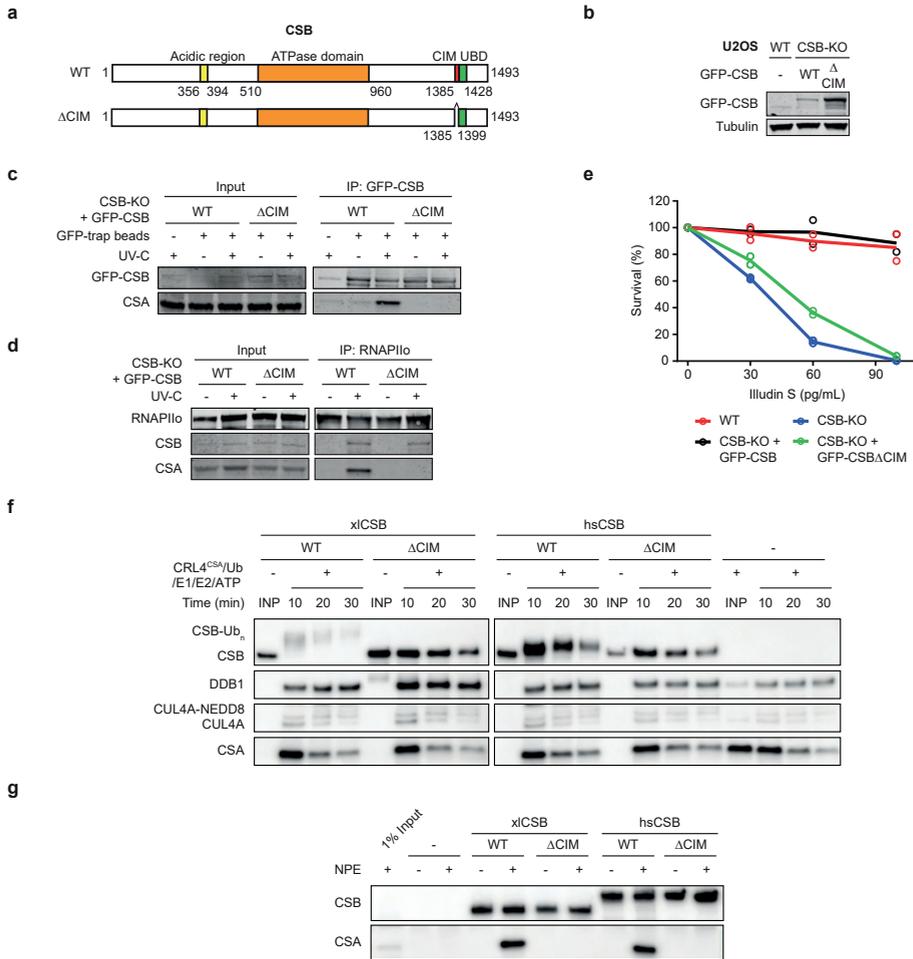


Fig. 3: The CIM of CSB mediates the recruitment of CSA to damage-stalled RNAPII α . (a) A schematic representation of CSB and the CSB Δ CIM mutant. (b) Western blot analysis of U2OS (FRT) and CSB-KO complemented with either GFP-CSB^{WT} or GFP-CSB Δ CIM ($n=2$). (c) Co-IP of GFP-CSB^{WT} and GFP-CSB Δ CIM on the combined soluble and chromatin fraction. (d) Endogenous RNAPII Co-IP in GFP-CSB^{WT} and GFP-CSB Δ CIM cell lines. See also Supplementary Fig. 6e for additional Co-IP data. (e) Clonogenic Illudins S survival of WT and CSB-KO cell lines and the GFP-tagged CSB rescue cell lines. Each symbol represents the mean of an independent experiment ($n=2$) each experiment contains 2 technical replicates. Note that the same survival data for WT, CSB-KO and CSB-KO + GFP-CSB is also shown in Fig. 1e. (f) *In vitro* ubiquitylation of recombinant *Xenopus laevis* (xl) and *Homo sapiens* (hs) CSB variants with recombinant xlCRL4^{CSA}, E1, E2, ubiquitin, and ATP. At indicated times, *in vitro* ubiquitylation reactions were stopped and blotted with anti-FLAG (top three panels) or anti-xlCSA (bottom panel) antibodies. See also Supplementary Fig. 6d. (g) Immobilized recombinant CSB variants were incubated with *Xenopus laevis* nucleoplasmic extract (NPE), recovered, and blotted with anti-FLAG (top panel) or anti-xlCSA (bottom panel) antibody. At least two independent replicates of each IP and *in vitro* ubiquitylation experiment were performed obtaining similar results.

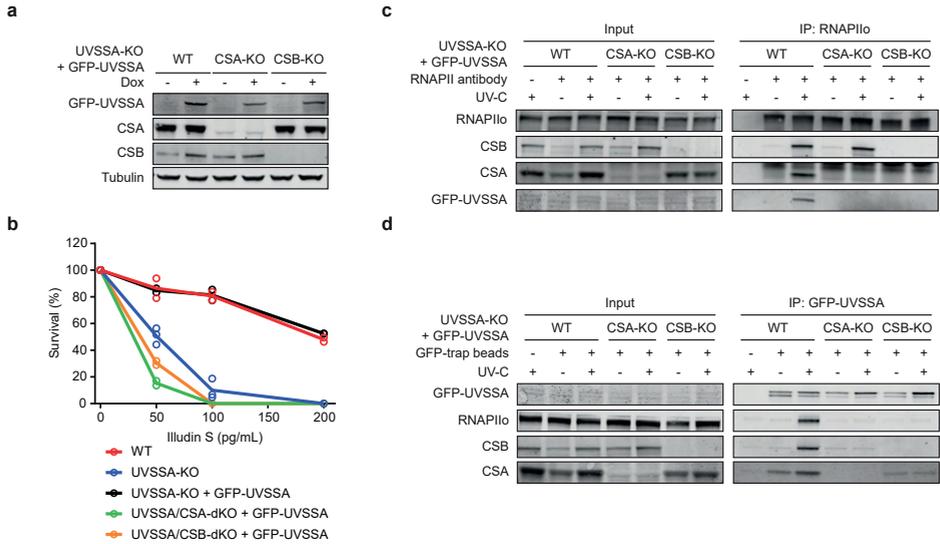


Fig. 4: UVSSA is recruited to DNA damage-stalled RNAPII α by CSA. (a) Western blot analysis of *UVSSA-KO*, *UVSSA/CSA-dKO*, and *UVSSA/CSB-dKO* complemented with GFP-UVSSA ($n=2$). (b) Clonogenic Illudin S survival of WT, *UVSSA-KO*, *UVSSA/CSA-dKO*, and *UVSSA/CSB-dKO* cell lines complemented with GFP-UVSSA. Each symbol represents the mean of an independent experiment ($n=2$ for all except for *UVSSA-KO* which is $n=3$), each of which is based on two technical replicates. (c) Endogenous RNAPII Co-IP on *UVSSA-KO*, *UVSSA/CSA-dKO*, and *UVSSA/CSB-dKO* complemented with GFP-UVSSA. (d) Co-IP of GFP-UVSSA in *UVSSA-KO*, *UVSSA/CSA-dKO*, and *UVSSA/CSB-dKO* cell lines. The asterisk in c indicates the heavy chain of the RNAPII antibody. At least two independent replicates of each IP experiment were performed obtaining similar results.

irradiated cells followed by mass spectrometry (MS). Following UV irradiation, our MS analysis identified 28 UV-specific UVSSA interactors, including CSB, the CSA-interacting protein DDB1, and RNAPII subunits. Additionally, among the most prominent UV-specific interactions were the TFIIH subunits XPB/p89 and XPD/p80 (Fig. 5b). These findings demonstrate that UVSSA interacts in a UV-specific manner with TFIIH.

Immunoprecipitation of GFP-UVSSA indeed confirmed a UV-specific interaction with TFIIH subunits by western blot analysis (Fig. 5c). Strikingly, these interactions were severely reduced in the *CSB-KO* and *CSA-KO* cells, suggesting a cooperative interaction mechanism in which CSB is required to stabilize the interaction between CSA and UVSSA, while CSA is required to stabilize the interaction between UVSSA and TFIIH.

We subsequently asked if UVSSA is also required for TFIIH recruitment. To this end, we employed our immunoprecipitation-based method in *CSB-KO*, *CSA-KO*, and *UVSSA-KO* cells to monitor TFIIH recruitment. In addition, we included *XPA-KO* cells (Supplementary Fig. 2c) as a positive control since XPA recruitment, at least during GGR, occurs downstream of TFIIH [46].

Immunoprecipitation of endogenous RNAPII α in these cell lines revealed a UV-specific interaction with TFIIH in WT and *XPA-KO* cells (Fig. 5d). These findings suggest that XPA recruitment does not only occur downstream of TFIIH

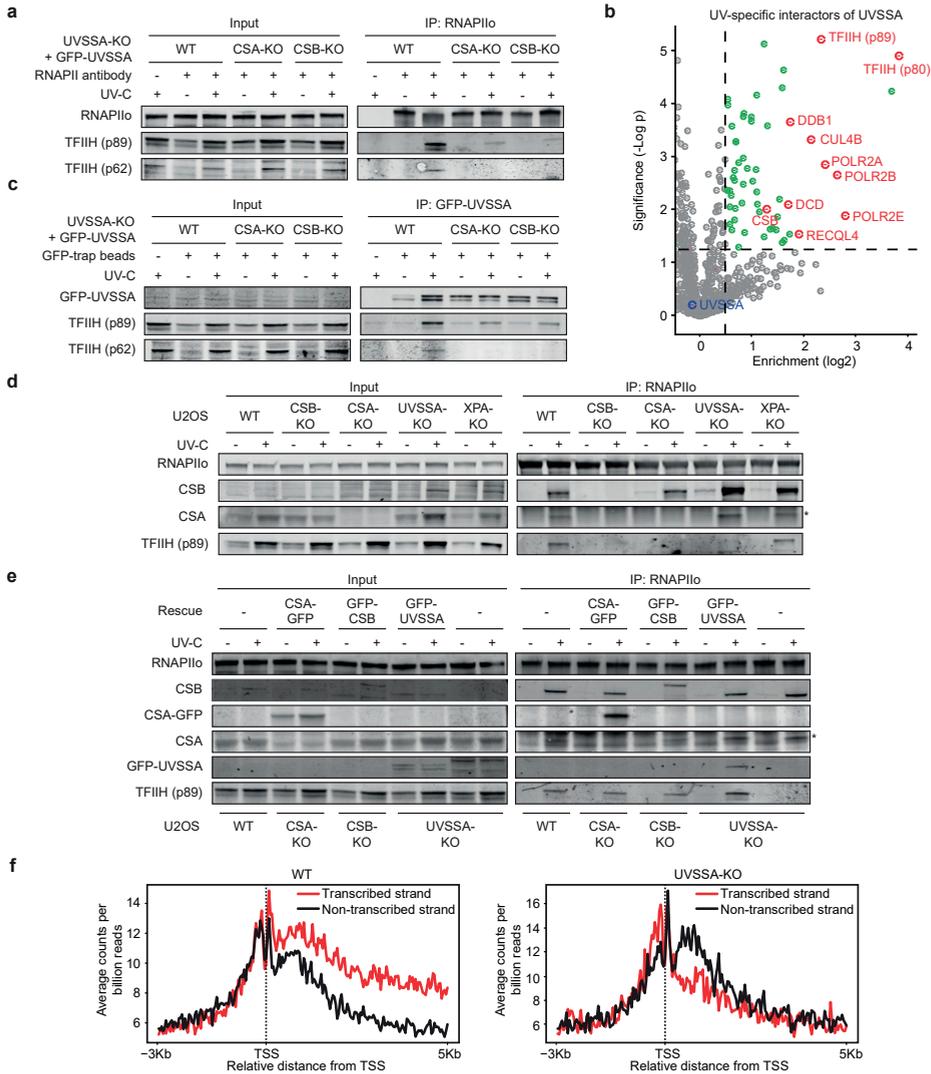
in GGR but also in TCR. Interestingly, similar to *CSB*-KO and *CSA* cells, we found that the UV-induced interaction between RNAPII α and TFIIH was severely reduced in *UVSSA*-KO cells (Fig. 5d). Furthermore, complementation of these TCR-KO cells with inducible GFP-tagged versions of *CSB*, *CSA*, and *UVSSA* fully restored the UV-induced association of TFIIH to RNAPII α (Fig. 5e). These findings demonstrate that *CSB*, *CSA*, and *UVSSA* are equally important for the recruitment of the TFIIH complex to DNA damage-stalled RNAPII α .

Genome-wide XR-seq confirms that *UVSSA* is a core TCR factor

Our findings show that *UVSSA*, just like *CSA* and *CSB*, is required to recruit TFIIH to initiate TCR-mediated repair. To provide further support for a role of *UVSSA* in TCR, we carried out genome-wide XR-sequencing (XR-seq), which enables the generation of genome-wide repair maps by isolating and sequencing the 30-mers that are generated upon dual incision [47, 48]. We generated nucleotide-resolution maps of UV-induced cyclobutane pyrimidine dimer (CPDs) repair in U2OS wild-type cells (Fig. 5f; Supplementary Fig. 7a), which revealed that CPD repair under these conditions is enriched on the transcribed strands within gene bodies consistent with TCR-mediated repair [47]. Importantly, the bias in CPD repair observed in transcribed strands was completely lost in both *CSA*-KO (Supplementary Fig. 7a) and *UVSSA*-KO cells (Fig. 5f). These findings provide direct genome-wide support for an essential role of *UVSSA* in TCR.

UVSSA is the key protein that recruits TFIIH to lesion-stalled RNAPII α

We next asked whether TFIIH is recruited via direct protein-protein contacts with *UVSSA*, or whether *CSB* and *CSA* also contribute to this interaction. To address this, we generated *UVSSA* separation-of-function mutants that are selectively impaired in their interaction with either *CSA* (*UVSSA* $^{\Delta 100-200}$) or the TFIIH complex (*UVSSA* $^{\Delta 400-500}$) [33] (Fig. 6a). These separation-of-function mutants were characterized using our chromatin-tethering approach. mCherry-LacR-*UVSSA*^{WT} clearly localized to the LacO array and triggered the robust recruitment of *CSA*-GFP and endogenous TFIIH (Fig. 6b, c). As expected, mCherry-LacR-*UVSSA* $^{\Delta 100-200}$ was unable to recruit *CSA*-GFP to the LacO array, but triggered robust TFIIH recruitment (Fig. 6b, c). In contrast, mCherry-LacR-*UVSSA* $^{\Delta 400-500}$ was unable to recruit TFIIH to the LacO array, but was proficient in recruiting *CSA*-GFP (Fig. 6b, c). These results confirm that *UVSSA* contains a *CSA*-interacting region (CIR; amino acids 100-200) and a TFIIH-interacting region (TIR; amino acids 400-500). To elucidate the importance of the CIR and TIR in *UVSSA* under more physiological conditions, we stably expressed inducible GFP-*UVSSA*^{WT}, GFP-*UVSSA* $^{\Delta CIR}$, or GFP-*UVSSA* $^{\Delta TIR}$ in *UVSSA*-KO cells (Fig. 6d). Pull-down of GFP-*UVSSA*^{WT} showed a strong UV-induced interaction with RNAPII α , *CSB*, *CSA*, and TFIIH. These interactors were virtually absent after pull-down of GFP-*UVSSA* $^{\Delta CIR}$ (Fig. 6e; Supplementary Fig. 7b, c). The *UVSSA* $^{\Delta CIR}$ mutant was unable to interact with *CSA* and we found that its association with TFIIH was also abolished. This result is consistent with the finding that the *UVSSA*-TFIIH interaction is reduced in *CSA*-KO cells (Fig. 5c) and suggests that *CSA* stabilizes the interaction between *UVSSA* and TFIIH. Pull-down of GFP-*UVSSA* $^{\Delta TIR}$ resulted in a strong UV-induced interaction with RNAPII α , *CSB*, and *CSA*, while its interaction with TFIIH was completely abolished (Fig. 6e; Supplementary Fig.



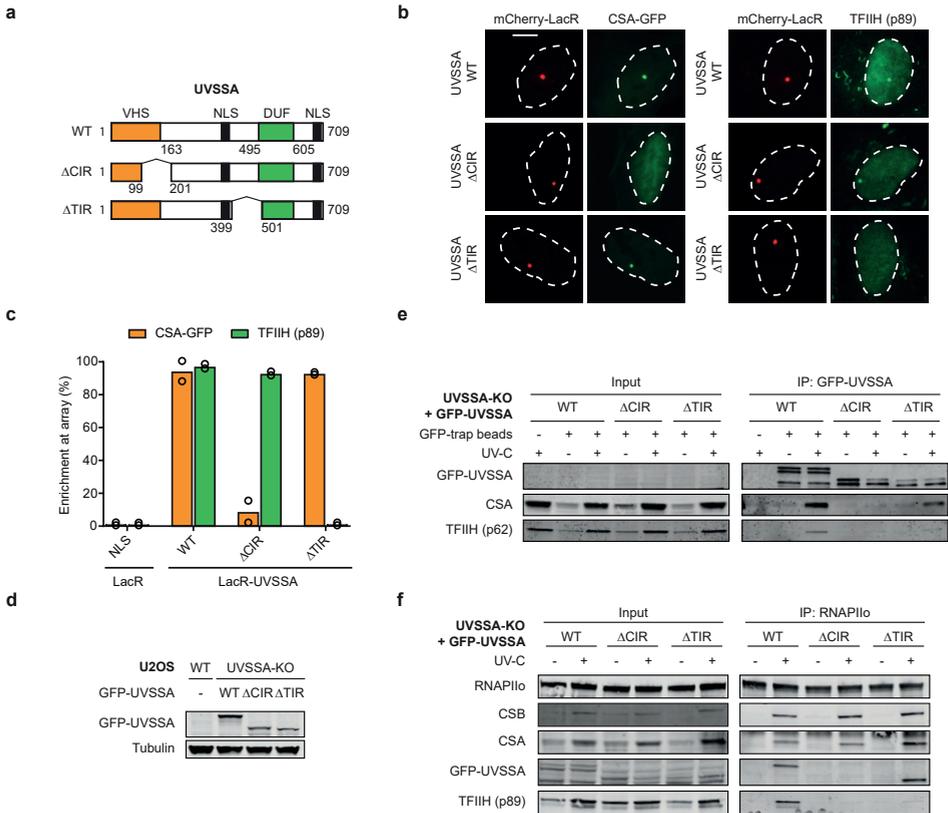


Fig. 6: UVSSA is the key protein that recruits TFIIH. (a) A schematic representation of UVSSA WT and deletion mutants. The CSA-interacting region (CIR) and TFIIH-interacting region (TIR) are indicated. (b) Recruitment of CSA-GFP and TFIIH (p89) to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins (scale bar = 5 μ m). (c) Quantification of CSA-GFP and endogenous TFIIH (p89) co-localization at the LacO array. Each symbol represents the mean of an independent experiment ($n=2$, >50 cells collected per experiment). (d) Western blot analysis of U2OS (FRT) and UVSSA-KO cells complemented with GFP-UVSSAWT, GFP-UVSSA Δ CIR, and GFP-UVSSA Δ TIR ($n=2$). (e) Co-IP of GFP-UVSSAWT, GFP-UVSSA Δ CIR, and GFP-UVSSA Δ TIR. (f) Endogenous RNAPII Co-IP in GFP-UVSSA WT , GFP-UVSSA Δ CIR, and GFP-UVSSA Δ TIR cell lines. At least two independent replicates of each IP experiment were performed obtaining similar results. See also Supplementary Fig. 7b, c for additional Co-IP data.

7b, c). These findings were largely confirmed by quantitative MS after pull-down of GFP-tagged versions of UVSSA. Interestingly, this approach also identified several proteins that interacted with UVSSA WT , but not with the UVSSA mutants (Supplementary Fig. 8a-e).

We next set out to directly assess the ability of these UVSSA mutants to participate in TCR complex assembly. Immunoprecipitation of endogenous RNAPII α showed a UV-specific association of RNAPII α with CSB and CSA in both UVSSA WT and mutant cell lines (Fig. 6f). Endogenous RNAPII α immunoprecipitation revealed a UV-specific interaction with GFP-UVSSA WT and GFP-UVSSA Δ TIR, whereas GFP-UVSSA Δ CIR failed to associate with RNAPII α . The

fact that a mutant of UVSSA that is deficient in its association with CSA fails to be recruited confirms our earlier findings that CSA is essential to recruit UVSSA to DNA damage-stalled RNAPII α (Fig. 4c). Furthermore, the recruitment of TFIIDH (p89) to DNA damage-stalled RNAPII α was abolished in both UVSSA mutant cell lines (Fig. 6f). These experiments strongly suggest that TFIIDH is recruited to DNA damage-stalled RNAPII α via direct protein-protein contacts with UVSSA. Importantly, the stable expression of GFP-UVSSA Δ^{CIR} and GFP-UVSSA Δ^{TIR} in UVSSA-KO cells failed to restore their sensitivity to Illudin S, which was almost fully restored by GFP-UVSSA WT (Fig. 7a). In line with this, genome-wide XR-seq revealed that the stable expression of GFP-UVSSA Δ^{CIR} and GFP-UVSSA Δ^{TIR} in UVSSA-KO cells failed to restore the repair of CPDs in the transcribed strand, which was fully restored by GFP-UVSSA WT (Fig. 7b; Supplementary Fig. 9a). These findings confirm that both the CIR and the TIR of UVSSA have an essential role in TCR.

Altogether, our data reveals a sequential and cooperative assembly mechanism of the human TCR complex, which involves the stepwise assembly of CSB, CSA, and UVSSA to target the TFIIDH complex to DNA damage-stalled RNAPII α to initiate DNA repair (Fig. 7c).

Discussion

Although it has been recognized for some time that CSA, CSB, and UVSSA are required for transcription-coupled repair (TCR), remarkably little is known about how these proteins cooperate to trigger TCR. Our findings suggest a highly cooperative recruitment mechanism that involves the sequential association of CSB, CSA and UVSSA to target the TFIIDH complex to DNA damage-stalled RNAPII α to initiate repair.

We show that both CSB and CSA associate with RNAPII α in a manner that is strongly induced by UV irradiation, and that CSA recruitment is completely dependent on CSB. This is in line with earlier work showing that CSB facilitates the translocation of CSA to the nuclear matrix after UV irradiation [49]. Moreover, we demonstrate that CSA is required for the association of DDB1 with RNAPII α , suggesting that CSA is recruited to DNA damage-stalled RNAPII α as part of a CRL4 CSA complex [25, 28]. Previous findings suggested that CSB dynamically associates with RNAPII α under undamaged conditions and that this interaction is stabilized upon UV irradiation [22, 50]. While our method may not be sensitive enough to capture these transient interactions, our findings do support that the CSB-RNAPII α interaction is stabilized after UV irradiation.

Earlier observations suggested that CSB physically interacts with CSA [27, 29], while other studies failed to detect this association [20, 21]. Our findings fully support a direct UV-induced association between the CS proteins. Importantly, we identified the CSA-interaction motif (CIM) in the C-terminus of CSB that is essential for targeting CSA to stalled RNAPII α . Interestingly, the CIM region in CSB is evolutionary conserved in species that also contain the CSA gene, including mammals, amphibians and fish (Supplementary Fig. 5). In line with this, we demonstrate that both human and *Xenopus leavis* CSB require its CIM to directly interact with CSA *in vitro*. However, the CIM is absent in species without CSA, including yeast, nematodes, but also holometabolous insects, which have lost the

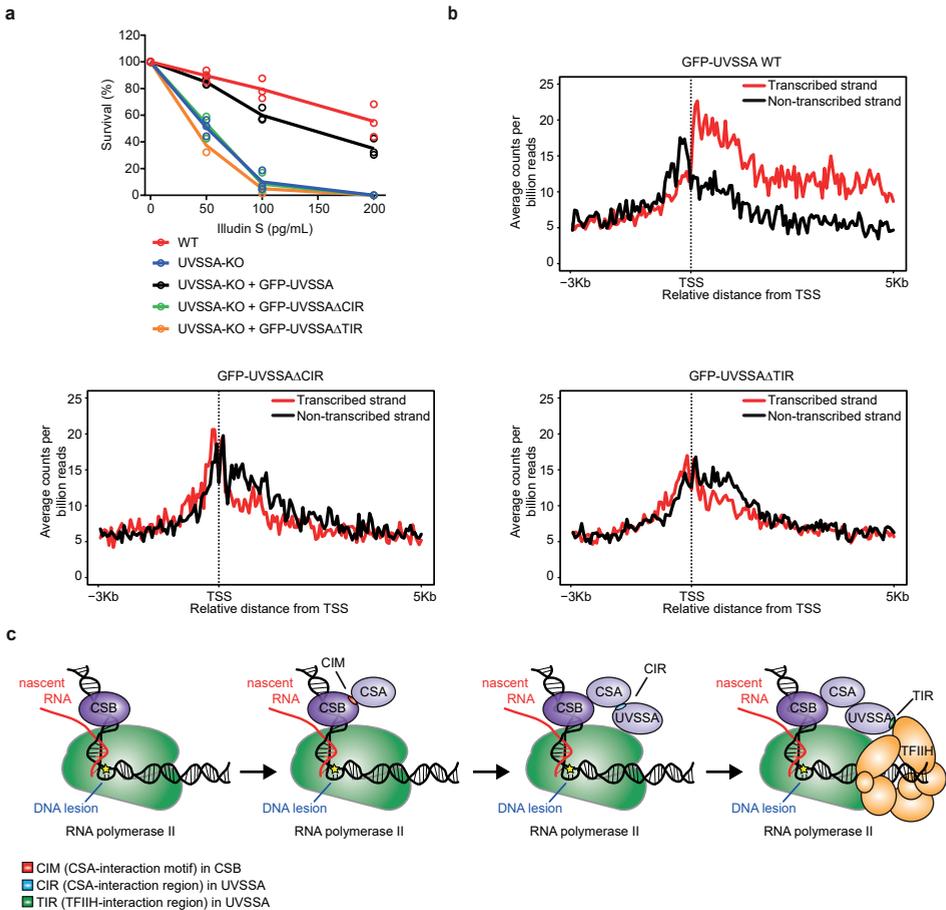


Fig. 7: The CIR and TIR of UVSSA are crucial for CPD repair. (a) Clonogenic Illudin S survival of WT and UVSSA-KO cell lines and the GFP-tagged UVSSA rescue cell lines. Each symbol represents the mean of an independent experiment ($n=3$ for all except for GFP-UVSSA Δ CIR which is $n=2$). Each experiment contains 2 technical replicates. Note that the same survival data for UVSSA-KO is also shown in Fig. 4b. (b) CPD XR-seq repair signal 3 Kb upstream and 5 Kb downstream of the annotated TSS of 16,088 genes in GFP-UVSSA^{WT}, GFP-UVSSA Δ CIR and GFP-UVSSA Δ TIR. Signal is plotted separately for the transcribed (red) and non-transcribed (black) strands. The data represent the average of two independent experiments for GFP-UVSSA^{WT} and one single experiment for GFP-UVSSA Δ CIR and GFP-UVSSA Δ TIR with a bin size of 53 nt. See also Supplementary Fig. 9 for additional XR-seq data. (c) Model of how the assembly of CSB, CSA, and UVSSA targets the TFIIH complex to DNA damage-stalled RNAPII.

CSA gene during the course of evolution (Supplementary Fig. 5).

It is striking that even though CSB contains a CSA-interaction motif (CIM), the association between these proteins is induced by UV irradiation. Previous studies revealed that the association of CSB with lesion-stalled RNAPII triggers a conformational change that repositions the N-terminus, thereby exposing residues in the C-terminus of CSB [50]. It is conceivable that this conformational change exposes the CIM to facilitate efficient CSA recruitment. Interestingly,

while the CIM is located right next to the ubiquitin-binding domain (UBD) in CSB [43], we find that CSB Δ UBD is fully functional in interacting with CSA. However, it is possible that the CIM and the UBD collaborate, as a tandem protein-interaction module [51], to enable optimal CSA recruitment. In this scenario, CSA would have protein-protein interactions with the CIM, which would be stabilized by the binding of the UBD to auto-ubiquitylated CSA [28]. The recently identified UVSSA protein can be isolated as part of a chromatin-bound stalled RNAPII α complex. Our current findings shed light on its recruitment mechanism by demonstrating that the association of UVSSA with RNAPII α is strongly induced by UV irradiation and fully dependent on both CSA and CSB. Moreover, knockout of *UVSSA* did not affect CSA or CSB recruitment to DNA damage-stalled RNAPII α , suggesting that UVSSA is the last of these proteins to be recruited. Consistent with a reported association between CSA and UVSSA [33], we find that CSA targets UVSSA to DNA damage-stalled RNAPII α by interacting with a region in the N-terminal VHS domain (CIR; amino acids 100-200) of UVSSA. Intriguingly, the robust UV-induced association between CSA and UVSSA is stabilized by CSB, suggesting a cooperative assembly mechanism of the TCR complex.

In contrast to our observation that the CS proteins are required for the recruitment of UVSSA to DNA damage-stalled RNAPII α , live-cell imaging experiments showed that UVSSA is recruited to sites of UV-C-induced laser damage independently of the CS proteins [26, 32]. There could be several reasons for these seemingly conflicting results. Firstly, the methodology is very different. We isolate RNAPII α -associated TCR proteins from the chromatin-bound fraction after UV, while live-cell imaging studies monitor the recruitment of GFP-tagged TCR proteins to local UV-C laser damage. Therefore, it is possible that the observed recruitment of CSB and UVSSA could, in part, be triggered by something other than lesion-stalled RNAPII α . In line with this hypothesis, GFP-CSA could not be detected at sites of local UV-C laser damage [32], even though CSA is essential for TCR and showed a robust association with stalled RNAPII α under our conditions. Secondly, the time-frame during which UVSSA association is measured is different. While we isolate RNAPII α -associated UVSSA one hour after UV irradiation, the recruitment studies visualized UVSSA binding in the first 40 seconds after UV-C laser irradiation. It cannot be excluded that UVSSA transiently associates with UV-damaged chromatin independently of the CS proteins, but that the stable association with stalled RNAPII α during productive TCR is fully dependent on CSA and CSB. In line with this, we find that mutants of TCR proteins that display a clear assembly defect under our conditions also show a strong sensitivity to Illudin S reflecting impaired TCR. In conclusion, our findings favor a model in which UVSSA is recruited by CSA and argues for a cooperative assembly mechanism in which CSB stabilizes the association between CSA and UVSSA to ensure efficient targeting to stalled RNAPII α .

A major unresolved question is how the core NER machinery, likely starting with the TFIIH complex, is recruited to DNA damage-stalled RNAPII α to initiate repair. Biochemical *in vitro* experiments have shown that the association of CSB with RNAPII is sufficient to recruit TFIIH [24]. In addition, CSA was shown to associate with the p44 subunit of TFIIH [27], while UVSSA can interact with the p62 subunit of TFIIH [45]. In agreement, we found that GFP-UVSSA associates with several subunits of the TFIIH complex in a UV-specific manner *in vivo*.

Furthermore, our data reveals that CSB, CSA, and UVSSA are equally important for the recruitment of TFIIH to DNA damage-stalled RNAPII *in vivo*. Indeed, similar to previous results with CSB-deficient cells [47, 52], our high-resolution repair maps fully support a crucial role of both CSA and UVSSA in the TCR-mediated clearing of UV-induced lesions on a genome-wide level. Importantly, we found that UVSSA contains a TFIIH-interacting region (TIR; amino acids 400-500), which is crucial for the association of TFIIH with lesion-stalled RNAPII. Consistently, it has been shown that the PH domain of p62 (1-108) associates with a small fragment in UVSSA (400-419) *in vitro* and that mutations within this region causes a defect in recovery of RNA synthesis *in vivo* [45]. Moreover, we found that the UVSSA Δ CIR mutant was not only unable to associate with CSA, but also with the TFIIH complex. Our findings favour a model in which CSA not only recruits UVSSA to stalled RNAPII but also stabilizes the direct interaction between UVSSA and TFIIH, resulting in the recruitment of TFIIH to stalled RNAPII. In this regard, it would be interesting to examine if this interaction between UVSSA and the p62 subunit of TFIIH is the sole mechanism through which TFIIH is recruited to DNA damage-stalled RNAPII *in vivo*, or whether other subunits and regions also contribute.

Here we show that UVSSA is essential to bridge the TFIIH complex to CSB/CSA-bound RNAPII to initiate TCR. Importantly, these findings also suggest that neurodegeneration seen in Cockayne syndrome (CS) is not caused by the inability to remove transcription-blocking DNA lesions, since neurodegeneration is not a feature in UV-sensitive syndrome (UV^SS). Previous findings revealed that CS fibroblasts fail to ubiquitylate and subsequently degrade DNA damage-stalled RNAPII [17], while UV^SS fibroblasts displayed even faster degradation of RNAPII after UV, possibly due to a failure to deubiquitylate RNAPII by the UVSSA binding partner USP7 [17, 18, 26, 33]. These findings suggest that features of CS are caused by toxicity associated with prolonged RNAPII stalling at lesions rather than the inability to remove transcription-blocking DNA lesions [53]. It remains unresolved how RNAPII is repositioned while TCR is in progress to enable access to the DNA lesion. We speculate that the TCR complex that we capture contains an inactive TFIIH complex, which is bound to the CAK complex. Indeed, we detect the association of CAK subunit CDK7 with RNAPII after UV. The association of repair factors XPA and XPG activates TFIIH [35, 36], which could cause the backtracking, or even removal of RNAPII. This model would explain why we do not detect the association of downstream repair factors with DNA damage-stalled RNAPII.

We propose a model in which CSB is the first protein to be recruited to DNA damage-stalled RNAPII (Fig. 7c). This binding of CSB could bring about a conformational change, thereby exposing the newly identified CIM to facilitate efficient CSA recruitment through direct protein-protein contacts. Once bound, CSA targets UVSSA to DNA damage-stalled RNAPII, and this interaction is stabilized by CSB. UVSSA, in turn, mediates the recruitment of the TFIIH complex in a cooperative manner that is stabilized by both CSB and CSA. Although both CS proteins could interact with TFIIH, it is likely that only CSA contributes directly to this stabilization, while CSB contributes indirectly through ensuring the association of CSA itself and stabilizing the interaction between CSA and UVSSA. At the stage when TFIIH is bound, it seems likely that RNAPII and

CSB/CSA/UVSSA are displaced and that the TCR-specific pre-incision complex is assembled starting with XPA. It is interesting to note that the yeast orthologue of CSB, RAD26, is bound to the DNA upstream of RNAPII [54], while human TFIIH in the transcription pre-initiation complex (PIC) is bound downstream of RNAPII [55]. If TFIIH is recruited to the same side of RNAPII during TCR, it suggests that CSB/CSA/UVSSA extend from the upstream to the downstream DNA around RNAPII to position TFIIH. It will be very interesting to gain structural insights into these molecular events. In conclusion, our findings reveal the recruitment mechanism of the TFIIH complex to DNA damage-stalled RNAPII, which involves the sequential and cooperative assembly of the CSB, CSA and UVSSA proteins.

Methods

Cell lines. Cell lines (Listed in Supplementary Table 1) were cultured at 37°C in an atmosphere of 5% CO₂ in DMEM (Thermo Fisher Scientific) supplemented with penicillin/streptomycin (Sigma) and 10% Fetal bovine serum (FBS; Bodinco BV). Sf9 cells were cultured in ESF 921 insect cell culture medium (Fisher Scientific). U2OS 2-6-3 cells containing 200 copies of a LacO-containing cassette (~4 Mbp) were a gift from Susan Janicki [40]. UVSSA-deficient KPS3-hTERT cells and their UVSSA-rescued counterparts were a gift from Tomoo Ogi [17]. U2OS Flp-In/T-REx cells, which were generated using the Flp-In™/T-REX™ system (Thermo Fisher Scientific), were a gift from Daniel Durocher [51].

Generation of knockout cell lines. To generate stable knockouts, U2OS Flp-In/T-REx cells were co-transfected with pLV-U6g-PPB encoding a guide RNA from the LUMC/Sigma-Aldrich sgRNA library (see Supplementary Table 4 for plasmids, Supplementary Table 2 for sgRNA sequences) together with an expression vector encoding Cas9-2A-GFP (pX458; Addgene #48138) using lipofectamine 2000 (Invitrogen). Transfected cells were selected on puromycin (1 µg/mL) for 3 days, plated at low density after which individual clones were isolated. To generate double knockouts, single knockout clones were transfected with pLV-U6g-PPB encoding a sgRNA together with pX458 encoding Cas9, cells were FACS sorted on BFP/GFP, plated at low density after which individual clones were isolated. Isolated knockout clones were verified by Western blot analysis and/or sanger sequencing. The absence of Cas9 integration/stable expression was confirmed by western blot analysis.

PCR analysis of knockout clones. Genomic DNA was isolated by resuspending cell pellets in WCE buffer (50mM KCL, 10mM Tris pH 8.0, 25 mM MgCl₂ 0.1 mg/mL gelatin, 0.45% Tween-20, 0.45% NP-40) containing 0.1 mg/mL Proteinase K (EO0491; Thermo Fisher Scientific) and incubating for 1h at 56°C followed by a 10 min heat inactivation of Proteinase K by 96°C. Fragments of approximately 1kb, containing the sgRNA sequence, were amplified by PCR (sequencing primers are listed in Supplementary Table 3) followed by Sanger sequencing using either the forward or the reverse primer.

Generation of stable cell lines. Selected knockout clones of CSB, CSA, and UVSSA (see Supplementary Table 1) were subsequently used to stably express GFP-CSB^{WT}, GFP-CSB^{ΔCIM}, CSA^{WT}-GFP, GFP-UVSSA^{WT}, GFP-UVSSA^{ΔCIR}, and GFP-UVSSA^{ΔTIR} by co-transfection of pCDNA5/FRT/TO-Puro plasmid encoding these CSB, CSA, and UVSSA variants (2 µg), together with pOG44 plasmid encoding the Flp recombinase (0.5 µg). After selection on 1 µg/mL puromycin and 4 µg/mL blasticidin S, single clones were isolated and expanded. Clones were selected based on their near-endogenous expression level compared to parental U2OS Flp-In/T-REx cells. Expression of these GFP-tagged TCR proteins was induced by the addition of 2 µg/ml Doxycycline for 24 hrs.

Plasmid constructs. The Neomycin resistance gene in pcDNA5/FRT/TO-Neo (Addgene #41000) was replaced with a Puromycin resistance gene. Fragments spanning GFP-N1 (clontech) and GFP-C1 (clontech) including the multiple cloning site were inserted into pcDNA5/FRT/TO-puro. CSB^{WT}, CSA^{WT}, and UVSSA^{WT} were amplified by PCR (see Supplementary Table 5 for primers) and inserted into pcDNA5/FRT/TO-Puro-GFP-N1 or pcDNA5/FRT/TO-Puro-GFP-C1 and in mCherry-LacR-NLS-C1/C3. Deletion constructs of CSB and UVSSA were generated by site-directed mutagenesis PCR. All sequences were verified by sequencing.

Illudin S survival assay. Knockout and rescue cell lines were trypsinized, seeded at low density and

mock-treated or exposed to a dilution series of Illudin S (Santa cruz; sc-391575) for 72 h (30, 60, 100 pg/mL or 50, 100, and 200 pg/mL). On day 10, the cells were washed with 0.9% NaCl and stained with methylene blue. Colonies of more than 20 cells were scored.

Immunoprecipitation for Co-IP. Cells were UV irradiated (20 J/m²) or mock treated and harvested 1 h after UV. Chromatin-enriched fractions were prepared by incubating the cells for 20 min on ice in IP buffer (IP-130 for endogenous RNAPII IP and IP-150 for GFP-IP), followed by centrifugation, and removal of the supernatant. For endogenous RNA pol II IPs the chromatin-enriched cell pellets were lysed in IP-130 buffer (30 mM Tris pH 7.5, 130 mM NaCl, 2 mM MgCl₂, 0.5% Triton X-100, protease inhibitor cocktail (Roche), 250 U/mL Benzonase[®] Nuclease (Novagen), and 2 µg RNAPII-S2 (ab5095, Abcam) for 2-3 h at 4 °C. For GFP IPs the chromatin-enriched cell pellets were lysed in IP-150 buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5% NP-40, 2 mM MgCl₂, protease inhibitor cocktail (Roche), and 500 U/mL Benzonase[®] Nuclease (Novagen) for 1 h at 4 °C. Protein complexes were pulled down by 1.5 h incubation with Protein A agarose beads (Millipore) or GFP-Trap[®] A beads (Chromotek). For subsequent analysis by western blotting, the beads were washed 6 times with IP-130 buffer for endogenous RNAPII IP and EBC-2 buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5% NP-40, and protease inhibitor cocktail (Roche)) for GFP-IPs. The samples were prepared by boiling in Laemml-SDS sample buffer. Unless indicated otherwise, all IP experiments were performed on the chromatin fraction. At least two independent replicates of each IP experiment were performed.

Generation of mass spectrometry samples. For the generation of mass spectrometry samples, the beads were washed 4 times with EBC-2 buffer without NP-40 and 2 times with 50 mM ammonium bicarbonate followed by overnight digestion using 2.5 µg trypsin at 37 °C under constant shaking. The bead suspension was loaded onto a 0.45 µm filter column (Millipore) to elute the peptides. The peptides were passed through a C-18 stage tips for desalting. The stagetips were activated by washing with methanol followed by washing with buffer B (80% Acetonitrile and 0.1% formic acid) and 0.1% formic acid. Peptides were acidified with 2% Trifluoroacetic acid and loaded on the stagetips. The peptides were eluted twice with 25 µl 60% Acetonitrile/ 0.1% Formic acid and lyophilized. At least three biological repeats for each condition were performed.

Mass spectrometry. Mass spectrometry analysis was performed on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Germany) coupled to an EASY-nanoLC 1000 system (Proxeon, Odense, Denmark). The mass spectrometer was operated in positive-ion mode at 2.9 kV with the capillary heated to 250 °C in a Data-Dependent Acquisition (DDA) mode with a top 7 method, Precursor ions with a charge state of 1 and greater than 6 were excluded from triggering MS/MS events. For the UV-dependent UVSSA interactors (Fig. 5b), samples were analysed essentially as previously described [56]. Digested peptides were separated using a 15 cm fused silica capillary (ID: 75 µm, OD: 375 µm, Polymicro Technologies, California, US) in-house packed with 1.9 µm C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 95% acetonitrile with 0.1% formic acid at a flow rate of 200 nl/min for 65 mins. Full scan MS spectra were obtained with a resolution of 70,000, a target value of 3x10⁶ and a scan range from 400 to 2,000 m/z. Maximum Injection Time (IT) was set to 50 ms. Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded with a resolution of 35,000, a maximum IT of 120 ms, a target value of 1x10⁵ and a normalized collision energy of 25%. Minimum AGC target was set to 103 The precursor ion masses selected for MS/MS analysis were subsequently dynamically excluded from MS/MS analysis for 60 s. For interactors of the CSB and UVSSA mutant proteins (Supplementary Fig. 6a-c and Fig. 8a-e), mass spectrometry analysis was performed as previously described [56]. Peptides were separated in a pre-cut 25 cm silica emitter (FS360-75-15-N-5-C25, MS Wil B.V, The Netherlands), in-house packed with 1.9 µm C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 95% acetonitrile with 0.1% formic acid at a flow rate of 200 nl/min for 125 mins. Full scan MS spectra were obtained with a resolution of 70,000, a target value of 3x10⁶ and a scan range from 300 to 1,600 m/z. Maximum Injection Time (IT) was set to 250 ms. Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded with a resolution of 35,000, a maximum IT of 120 ms, a target value of 1x10⁵ and a normalized collision energy of 25%. Minimum AGC target was set to 104. Dynamic exclusion was set to 40 s.

Mass spectrometry data analysis. Raw mass spectrometry files were analysed with MaxQuant software (version 1.5.3.30) as described [57], with the following modifications from default settings: the maximum number of mis-cleavages by trypsin was set to 4, Label Free Quantification (LFQ)

was enabled thereby disabling the Fast LFQ feature. Match-between-runs feature was enabled with a match time window of 0.7 minutes and an alignment time window of 20 minutes. We performed the search against an *in silico* digested UniProt reference proteome for *Homo sapiens* (14th December 2017). Analysis output from MaxQuant was further processed in the Perseus (version 1.5.5.3) computational platform [58]. Proteins identified as common contaminants, only identified by site and reverse peptide were filtered out, and then all the LFQ intensities were log₂ transformed. Different biological repeats of each condition were grouped and only protein groups identified in all biological replicates in at least one condition were included for further analysis. Missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (log₂) and a randomized 0.3 width (log₂) considering total matrix values. Volcano plots were generated and Student's T-tests were performed to compare the different conditions. Spreadsheets from the statistical analysis output from Perseus were further processed in Microsoft Excel for comprehensive visualization and analysis of the data. A web app (VolcanoR) was made with R/Shiny for generating and sharing interactive volcano plots.

Western blot. 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 2 h at RT. The membrane was then probed with antibodies (listed in Supplementary Table 6) as indicated.

Slot blot for CPDs after IP. Cells were UV irradiated (20 J/m²) or mock treated and crosslinked with 0.5 mg/mL disuccinimidyl glutarate (DSG; Thermo Fisher) in PBS for 45 min at room temperature. Cells were washed with PBS and crosslinked with 1% PFA for 20 min at room temperature. Fixation was stopped by adding 1.25 M Glycine in PBS to a final concentration of 0.1 M for 3 minutes at room temperature. Cells were washed with cold PBS and lysed and collected in a buffer containing 0.25% Triton X-100, 10 mM EDTA (pH 8.0), 0.5 mM EGTA (pH 8.0) and 20 mM Hepes (pH 7.6). Chromatin was pelleted in 5 min at 400 g and incubated in a buffer containing 150 mM NaCl, 1 mM EDTA (pH 8.0), 0.5 mM EGTA (pH 8.0) and 50 mM Hepes (pH 7.6) for 10 minutes at 4 °C. Chromatin was again pelleted for 5 min at 400 g and resuspended in ChIP-buffer (0.15% SDS, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA (pH 8.0), 0.5 mM EGTA (pH 8.0) and 20 mM Hepes (pH 7.6)) to a final concentration of 20x10⁶ cells/mL. Chromatin was sonicated to approximately 1 nucleosome length using a Bioruptor waterbath sonicator (Diagenode). Chromatin of 20x10⁶ cells was incubated with 6 µg RNAPII-S2 (ab5095, Abcam) overnight at 4 °C, followed by a 1.5 hrs protein-chromatin pull-down with a 1:1 mix of protein A and protein G Dynabeads (Thermo Fisher; 10001D and 10003D). The beads were washed extensively, followed by decrosslinking for 4 hrs at 65 °C in the presence of proteinase K and RNase A. The DNA was purified and concentrated using MinElute columns (Qiagen) and the concentration was measured using a dsDNA HS Qubit assay (Thermo Fisher Scientific). 15 ng of DNA for each sample was denatured for 15 min at 98 °C and blotted on Hybond N+ (RPN203B, GE healthcare). The membrane was dried for at least 1 hr at 57 °C and blocked overnight in 5% milk in PBS +0.1% tween. CPDs were detected using anti-CPD mouse monoclonal antibody followed by incubation with an anti-mouse HRP conjugated antibody (Supplementary Table 6). CPD signals were detected using an ECL reagent (Sigma, GERPN2232).

Chromatin tethering. U2OS 2-6-3 cells containing 200 copies of a LacO-containing cassette were co-transfected with lipofectamine 2000 (Invitrogen) and plasmid DNA for 6 h at 37 °C in an atmosphere of 5% CO₂. 24 h after transfection the cells were fixed with 4% paraformaldehyde (Sigma) in PBS for 15 min. The cells were either permeabilized with 0.5% triton X-100 (Sigma) in PBS for 10 min and mounted in poly mount (Polysciences; 18606) or subjected to immunofluorescent labeling.

Immunofluorescent labeling. Cells were permeabilized with 0.5% triton X-100 (Sigma) in PBS for 10 min, followed by treatment with 100 mM glycine in PBS for 10 min to block unreacted aldehyde groups. Cells were rinsed with PBS and equilibrated in wash buffer (WB: PBS containing 0.5% BSA, and 0.05% Tween-20 (Sigma-Aldrich)) for 10 min. Antibody steps and washes were in WB. The primary antibody rabbit-p89 (1/100; Santa Cruz; SC-293; S19) was incubated for 2 h at RT. Detection was done using goat-rabbit Ig coupled to Alexa 488 (1:1000; Invitrogen). Cells were incubated with 0.1 µg/mL DAPI and mounted in Poly mount (Polysciences; 18606).

Microscopic analysis of fixed cells. Images of fixed samples were acquired on a Zeiss AxioImager M2 or D2 widefield fluorescence microscope equipped with a 63x PLAN APO (1.4 NA) oil-immersion

objectives (Zeiss) and an HXP 120 metal-halide lamp used for excitation. Fluorescent probes were detected using the following filters: DAPI (excitation filter: 350/50 nm, dichroic mirror: 400 nm, emission filter: 460/50 nm), GFP/Alexa 488 (excitation filter: 470/40 nm, dichroic mirror: 495 nm, emission filter: 525/50 nm), mCherry (excitation filter: 560/40 nm, dichroic mirror: 585 nm, emission filter: 630/75 nm). Images were recorded using ZEN 2012 software (blue edition, version 1.1.0.0).

Genome-wide XR-sequencing. XR-seq was performed as previously described [47, 52]. Briefly, cells were harvested 3h after treatment with 20 J/m² UVC (254 nm). Primary excision products were pulled down by TFIIH coimmunoprecipitation with anti-p62 and anti-p89 antibodies (Santa Cruz Biotechnology sc25329 and sc271500), and ligated to both 5' and 3' adaptors. Ligation products containing CPD were purified by immunoprecipitation with the anti-CPD antibody (Cosmo Bio NM-DND-001) and repaired in vitro by *Drosophila melanogaster* CPD photolyase. Repaired DNA was PCR-amplified with Index primers and purified by 10% native polyacrylamide gels. Libraries were pooled and sequenced in a single HiSeq 2500 lane producing at least 10 million single-end 50 nt reads per sample. Quality score for each nucleotide was analyzed using the Fastx-toolkit to ensure only high-quality reads are processed. Adapter sequence was trimmed from each read using Trimmomatic (version 0.36) [59]. Reads were aligned to the genome using Bowtie (version 1.2.2) [60]. Following alignment, reads that were mapped to chromosome Y or mitochondrial chromosome were filtered (U2OS cell line is derived from female bone tissue) and PCR duplicates were removed using PicardCommandLine MarkDuplicates (version 2.8.1). There were high levels of PCR duplicates due to low efficiency of excised oligo recovery, but these were sufficient for analysis of TCR. To plot average XR-seq signal along genes, the genes annotation file was downloaded from Ensembl, assembly GRCh38, release 96. Non-overlapping regions around the TSS were obtained using custom scripts and BEDTools (version 2.26.0) slop and merge commands [61]. All samples were converted to BED format using bedtools bamtobed command. Strand-specific profiles over the TSS were created using the R version 3.5.1) Bioconductor genomation package (version 1.14.0) [62]. For U2OS (FRT)^{WT}, UVSSA-KO, and GFP-UVSSA^{WT} the experiment was done in duplicate. For CSA-KO, GFP-UVSSA^{ΔCIR} and GFP-UVSSA^{ΔTIR} the experiment was done once. (Sequence depth is available at: <https://doi.org/10.1038/s41467-020-15903-8> (Supplementary Table 8))

Protein expression and purification. Coding sequences of *Xenopus laevis* CSB and CSA-DDB1-CUL4-RBX1 (CRL4^{CSA}), as well as human CSB were amplified from cDNA clones or ordered as codon-optimized gene blocks from Integrated DNA Technologies. All open reading frames were cloned into pAceBac1 (pAB1) or pIDC vectors containing the indicated affinity tags (Supplementary Table 4). For the generation of CRL4^{CSA}, CSA/DDB1 and CUL4A/RBX1 heterodimers were cloned into separate vectors, respectively. To obtain bacmids for insect cell expression, plasmids were transformed into chemically competent DH10Bac cells and purified using ZR BAC DNA miniprep kit (Zymo Research). Baculoviruses encoding CSB variants, CSA/DDB1, or CUL4A/RBX1 were amplified in three stages (P1, P2, and P3) in Sf9 cells (Expression Systems). Protein expression was performed for 72 h in 500 mL Sf9 cells per construct infected with 10 mL P2 or P3 baculovirus. Cells were cultured at 27 °C in ESF 921 insect cell culture medium (Fisher Scientific), pelleted at 1,000 xg for 15 min, frozen in liquid nitrogen, and stored at -80 °C. Protein purifications were performed at 4 °C. Cell pellets were resuspended in a final volume of 50 mL Wash Buffer (50 mM HEPES [pH 7.5], 300 mM NaCl, 10% glycerol) containing 0.1% NP-40 and one EDTA-free cOmplete protease inhibitor tablet (Roche). Cells were lysed by sonication and cleared by centrifugation for 1 h at 30,000xg. The clarified lysate was incubated with 0.3-0.6 mL pre-equilibrated anti-FLAG M2 Affinity Gel (Sigma) for 1 h at 4 °C on a rotating wheel. The resin was washed extensively with Wash Buffer, and proteins were eluted with Wash Buffer containing 0.2 mg/mL 3xFLAG peptide (Sigma). CSB proteins were further purified by gel filtration (Superdex 200 Increase) in Wash Buffer containing 2 mM DTT, and pooled peak fractions were concentrated with 5 mL 10 MWCO spin concentrators (Millipore), frozen in liquid nitrogen, and stored at -80 °C. Eluted CSA-StrepII/FLAG-DDB1 complex was applied to 0.3 mL pre-equilibrated Strep-Tactin XT Superflow high capacity resin in a disposable gravity-flow column and washed 5x with 0.6 mL Wash Buffer. FLAG peptide-eluted FLAG-CUL4A/RBX1 complex was incubated with the immobilized CSA-StrepII/FLAG-DDB1 complex for 1 h at 4 °C to assemble CRL4^{CSA}. The resin was washed 5x with 0.6 mL Wash Buffer to remove excess FLAG-CUL4A/RBX1, and CRL4^{CSA} was eluted with BXT Buffer (Iba-lifesciences), which contains 50 mM biotin. Pooled fractions were dialyzed O/N into 0.5x Wash Buffer containing 2 mM DTT, concentrated with 0.5 mL 3 MWCO spin concentrators (Millipore), frozen in liquid nitrogen, and stored at -80 °C.

Pull-down using immobilized CSB proteins. Purified FLAG-tagged CSB proteins were immobilized

on pre-equilibrated anti-FLAG M2 Magnetic Beads (Sigma) for 2 h at 4 °C. The beads were washed 3x with 0.3 mL Pull-down Buffer (20 mM HEPES [pH 7.5], 100 mM KCl, 5 mM MgCl₂, 0.5 mM EDTA, 0.25 mg/mL BSA, 0.03% Tween) and incubated with *Xenopus laevis* nucleoplasmic extract (NPE) for 1 h at 4 °C. The beads were washed 3x with 0.3 mL Pull-down Buffer and resuspended in Laemmli-SDS sample buffer. Samples were resolved by SDS-PAGE and analyzed by Western blot.

In vitro ubiquitylation assay. Purified xCRL4^{CSA} was neddylated in vitro using the NEDD8 Conjugation Initiation Kit (Boston Biochem) according to the manufacturer's protocols, except using 0.5x Uba3, 0.5x UbcH12, and 0.33x NEDD8 as compared to the recommended final concentrations. The reaction was incubated for 25 min at RT immediately prior to the in vitro ubiquitylation reaction, which contained the following final concentrations in Ubiquitylation Buffer (40 mM Tris pH 7.5, 10 mM MgCl₂, 0.6 mM DTT): 100 nM E1 (Enzo Life Sciences), 2.5 μM UBE2D2 (Boston Biochem), approximately 50 nM neddylated xCRL4^{CSA}, 50 μM ubiquitin, 10 mM ATP, and 200-250 nM CSB protein. Reaction were incubated for indicated times at RT and stopped in Laemmli-SDS sample buffer prior to SDS-PAGE and Western blot analysis.

Acknowledgments

The authors acknowledge Jean-Marc Egly and Rick Wood for their generous gift of TFIIH and XPA antibodies, respectively. Tom Misteli provided LacR-NLS plasmid, Tomoo Ogi provided UVSSA-deficient KPS3-hTERT cells, Susan Janicki provided U2OS 2-6-3 cells. This work was funded by an LUMC Research Fellowship and an NWO-VIDI grant (ALW.016.161.320) to MSL, an ERC starting grant (310913) to ACOV, a Young Investigator Grant from the Dutch Cancer Society (KWF-YIG: 11367) to RG-P, and Israel Science Foundation grants (1710/17 and 1762/17) administered by the Israeli Academy for Science and humanities and The Israel Cancer Association grant (20191630) to SA. SA is the recipient of the Jacob and Lena Joels memorial fund senior lectureship. JCW was supported by NIH grant HL098316 and is a Howard Hughes Medical Institute (HHMI) Investigator and an American Cancer Society Research Professor. TETM was supported by an EMBO Long-term fellowship (ALTF 1316-2016) and an HHMI fellowship of The Jane Coffin Childs Memorial Fund for Medical Research.

Supplementary tables

Supplementary Table 1: Cell lines

Cell line	Origin
KPS3-hTERT	[17]
KPS3-hTERT + UVSSA	[17]
Sf9	cat # 94-001S
U2OS (FRT)	[51]
U2OS (FRT) CSA-KO (2-4)	This study
U2OS (FRT) CSA-KO (2-4) + CSA-GFP-5	This study
U2OS (FRT) CSB-KO (1-12)	This study
U2OS (FRT) CSB-KO (1-12) + GFP-CSB Δ CIM-4	This study
U2OS (FRT) CSB-KO (1-12) + GFP-CSB-3	This study
U2OS (FRT) UVSSA-KO (1-8)	This study
U2OS (FRT) UVSSA-KO (1-8) / CSA (2-40) + GFP-UVSSA-3	This study
U2OS (FRT) UVSSA-KO (1-8) / CSB-KO (1-12) + GFP-UVSSA-3	This study
U2OS (FRT) UVSSA-KO (1-8) + GFP-UVSSA Δ CIR-1	This study
U2OS (FRT) UVSSA-KO (1-8) + GFP-UVSSA Δ TIR-6	This study
U2OS (FRT) UVSSA-KO (1-8) + GFP-UVSSA-3	This study
U2OS (FRT) XPA-KO (2-8)	This study
U2OS (FRT) XPC-KO (2-7)	This study
U2OS 2-6-3	[40]

Supplementary Table 2: Sequences of sgRNAs

Gene	Sequence	Identifier
CSB/ERCC6	5-AGACAGAATGATCCGATGAGGGG-3	sgML#003
CSA/ERCC8	5-CCAGACTTCAAGTCACAAAGTTG-3	sgML#018
UVSSA	5-AGAGAGCTGCTTTAGGCTGCTGG-3	sgML#019
XPA	5-CCTGTGTCAATTATCTTTGGGGC-3	sgML#002
XPC	5-TGGGGGTTTCTCATCTTCAAAGG-3	sgML#014

Supplementary Table 3: Sequencing primers to validate KO cell lines

Gene	Sequence	Identifier
CSB/ERCC6	5-GTAGGGGCCAGTTGTIAGAATGTAA-3	oML#078
	5-CTCACATTCTGAATGACTTGGCTA-3	oML#079
CSA/ERCC8	5-CAGTCTGTGTCCAGTTTCTGTG-3	oML#084
	5-CATATTTGTTATGTGTTTCTTTGAG-3	oML#085
	5-GTACATACATACATACACATTTACCAATAC-3	oML#100
	5-CTGAGAAAAAATGTACCTAAATATTAAG-3	oML#101
UVSSA	5-ACCCAGAGGTACACAGAGATTG-3	oML#090
	5-GCTCTTAGAAGTGTCCCTGTG-3	oML#091
	5-ATCAGGAGGCTGAGGCGGCTG-3	oML#076
	5-AGGAGCCTACCCGGGAGCCGGG-3	oML#077

Supplementary Table 4: Plasmids

Plasmid	Origin
pcDNA5/FRT/TO-Neo	Addgene #41000
pcDNA5/FRT/TO-Puro	This study
pcDNA5/FRT/TO-Puro-CSA ^{WT} -GFP	This study
pcDNA5/FRT/TO-Puro-GFP-C1	This study
pcDNA5/FRT/TO-Puro-GFP-CSB ^{ΔCIM}	This study
pcDNA5/FRT/TO-Puro-GFP-CSB ^{WT}	This study
pcDNA5/FRT/TO-Puro-GFP-N1	This study
pcDNA5/FRT/TO-Puro-GFP-UVSSA ^{ΔCIR}	This study
pcDNA5/FRT/TO-Puro-GFP-UVSSA ^{ΔTIR}	This study
pcDNA5/FRT/TO-Puro-GFP-UVSSA ^{WT}	This study
pEGFP-C1	Clontech
pEGFP-N1	Clontech
pLV-U6g-PPB	LUMC/Sigma-Aldrich sgRNA library
pmCherry-LacR-UVSSA ^{ΔTIR}	This study
pmCherry-LacR-C1	[42]
pmCherry-LacR-C3	This study
pmCherry-LacR-CSB ^N	This study
pmCherry-LacR-CSB ^M	This study
pmCherry-LacR-CSB ^C	This study
pmCherry-LacR-CSB ^{ΔN}	This study
pmCherry-LacR-CSB ^{ΔC}	This study
pmCherry-LacR-CSB ¹²²¹⁻¹³⁰⁵	This study
pmCherry-LacR-CSB ¹⁴⁰⁰⁻¹⁴⁹³	This study
pmCherry-LacR-CSB ¹⁴¹⁷⁻¹⁴⁹³	This study
pmCherry-LacR-CSB ¹³⁰⁶⁻¹³⁹⁹	This study
pmCherry-LacR-CSB ^{Δ1306-1300}	This study
pmCherry-LacR-CSB ^{Δ1306-1352}	This study
pmCherry-LacR-CSB ^{Δ1353-1399}	This study
pmCherry-LacR-CSB ^{Δ1400-1428}	This study
pmCherry-LacR-CSB ^{Δ1353-1368}	This study
pmCherry-LacR-CSB ^{Δ1369-1384}	This study
pmCherry-LacR-CSB ^{Δ1385-1399}	This study
pmCherry-LacR-CSB ^{WT}	This study
pmCherry-LacR-NLS	[63]
pmCherry-LacR-UVSSA ^{ΔCIR}	This study
pmCherry-LacR-UVSSA ^{WT}	This study
pOG44	Thermo Fisher
pX458	Addgene #48138
pTM58_pAB1_FLAG-xlDDB1_x_(pIDC_xlCSA-StrepII)x2	This study
pTM65_pAB1_FLAG-xlCSB ^{WT}	This study
pTM67_pAB1_FLAG-xlCUL4A_xlRBX1	This study
pTM141_pAB1_FLAG-xlCSB ^{ΔCIM}	This study
pTM142_pAB1_FLAG-hsCSB ^{WT}	This study
pTM143_pAB1_FLAG-hsCSB ^{ΔCIM}	This study

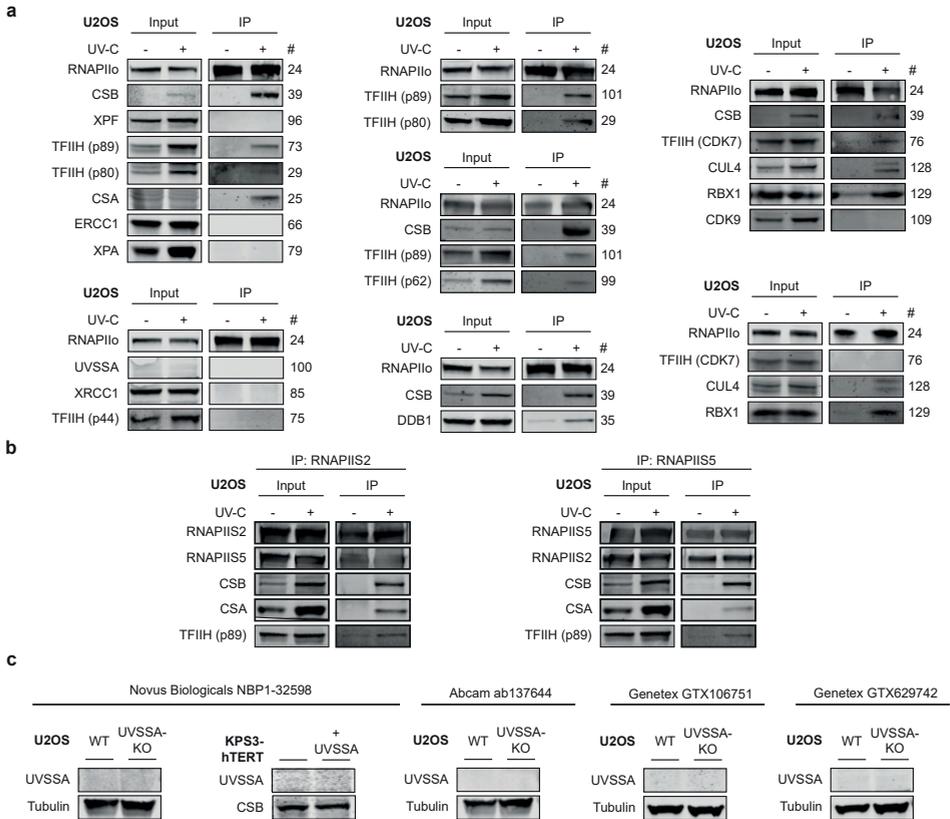
Supplementary Table 5: Primers for cloning

Gene	Sequence	Identifier
CSB	5-TTAAGTCGACCCAAATGAGGGAATCCCCAC-3	oML#375
WT	5-AATTGCGGCCGCTTAGCAGTATTCTGGCTTGAGTTTC-3	oML#376
CSA	5-CACAATGCTAGCGCCACCATGCTGGGGTTTTGTCCG-3	oML#041
WT	5-GCATGGTGAAC TACCGGTGCTCTCTTCATCACTGCTG-3	oML#042
UVSSA	5-ACAATTGAATTCGATGGATCAGAACTTTCGAAG-3	oML#035
WT	5-GTGTAAAGATCTCTAGTTCAGTGGCTAGTTAAAC-3	oML#036
CSB	5-TCCAGCCTCGAGGTCCAAATGAGGGAATCCCCCACTC-3	oML#173
ΔC	5-TCAGGTCGGATCCTTATCGAGTTCCTTCAAAC TTGGCGTCTC-3	oML#174
CSB	5-TCCAGCCTCGAGGTCCAAATGAGGGAATCCCCCACTC-3	oML#173
-N	5-GCATCAGGTCGGATCCTTAATCTCCATCATCTCGGTATCTTCCCAC-3	oML#178
CSB	5-TCCAGCCTCGAGGTGATGGAGATGAAGATTATTATAAGCAGCGG-3	oML#175
ΔN	5-GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCAAATTC-3	oML#176
CSB	5-TCCAGCCTCGAGGTGATGGAGATGAAGATTATTATAAGCAGCGG-3	oML#175
-M	5-TCAGGTCGGATCCTTATCGAGTTCCTTCAAAC TTGGCGTCTC-3	oML#174
CSB	5-TCCAGCCTCGAGGTCCAAATCCACACCTGGTGAAGAAAAG-3	oML#177
-C	5-GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCAAATTC-3	oML#176
CSB	5-TCCAGCCTCGAGGTCCAAATCCACACCTGGTGAAGAAAAG-3	oML#177
1221-1305	5-GATGGAGGATCCTTACAGACACCGCTGACGAGAGAG-3	oML#196
CSB	5-TACAGCCTCGAGGTGGAGCAGTGTCTGGTGTTC-3	oML#197
1306-1399	5-GGCGATGGAGGATCCTTACAGGTGGTTTCTAGCTCTCATTTAGC-3	oML#198
CSB	5-TCCAGCCTCGAGGTATTCTGCCAGAGCGTTTAGAAAAGTGAAG-3	oML#199
1400-1493	5-GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCAAATTC-3	oML#176
CSB	5-TACATCCTCGAGGTCCCTGCTGCCACACACAG-3	oML#200
1417-1493	5-GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCAAATTC-3	oML#176
CSB	5-GCTCTCTCGTCAGCGGTGTCTGTGCCAGGATGGCATCATGAA-3	oML#226
Δ1306-1352	5-CCTTTTTCATGATGCCATCCTGGCACAGACACCGCTGACGAG-3	oML#227
CSB	5-CCTTCATCAACATCTCCAACAGAGAAGATTCTGCCAGAGCGTTAG-3	oML#232
Δ1353-1399	5-CACTTCTAAACGCTCTGGCAGAATCTTCTGTTGGAGATGTTG-3	oML#233
CSB	5-GAGGCTCTCTCGTCAGCGGTGTCTGATTCTGCCAGAGCGTTTAGAAAAGT-3	oML#224
Δ1306-1399	5-CTTTCACCTTCTAAACGCTCTGGCAGAATCAGACACCGCTGACGAGAG-3	oML#225
CSB	5-GCTAAAATGAGAGCTAGAAAACCACTGGTGGAGATGAGAAAATTCATC-3	oML#234
Δ1400-1428	5-GAAAAGCGATGAAGTTTCTCATCTCCACCAGGTGGTTTCTAGCTCTC-3	oML#235
CSB	5-CCTTCATCAACATCTCCAACAGAGAAGCATTTTACTGGAAGAGCAGAAG-3	oML#262
Δ1353-1368	5-CTGCATCTTCTGCTCTTCCACTAAAATGCTTCTGTTGGAGATGTTGA-3	oML#263
CSB	5-GAAAAAGGAGGGAAAAGATAATGTCCCTGAGGCTTCTCCTCACTCTTG-3	oML#264
Δ1369-1384	5-CATTTTAGCCAAGAGTGAGGAGGAAGCCTCAGGGACATTATCTTTTCC-3	oML#265
CSB	5-AGACTCTTCATCCGGGCCCTCATTCTGCCAGAGCGTTTAGA-3	oML#266
Δ1385-1399	5-CTTTCACCTTCTAAACGCTCTGGCAGAATGAGGGGCCCGGATGA-3	oML#267
UVSSA	5-CACAGACCCCGCACAGCCTCTGAGGCTGCTGGTGCCTTTTG-3	oML#128
Δ100-200	5-CAAAGTCAAAGGCACCCAGCAGCCTCAGAGGCTGTGCGGGG-3	oML#129
UVSSA	5-GGACAGAAGCCCTGGGGGATGCGGTGGTGCCTACGGCGTG-3	oML#138
Δ400-500		
LacR	5-ATTAACGCGTCAGTGGGCTGATC-3	oML#377
-C3	5-TAATAATAGATCTGAAACCTTCTCTTCTTAG-3	oML#378

Supplementary Table 6: Antibodies

Antibody	Host	Company (reference)	Use	Identifier
Cas9	Mouse	Cell Signalling, #14697 (7A9-3A3)	WB: 1:5000	aML#031
CDK7	Mouse	kindly provided by J.M. Egly (2F8)	WB: 1:2000	aML#076
CDK9	Rabbit	Bethyl (A303-493A)	WB: 1:2000	aML#109
CPD	Mouse	Cosmo Bio, CAC-NM-DND-001 (TDM-2)	WB: 1:1000	aML#020
CSA/ERCC8	Mouse	Santa Cruz, sc-376981 (D2)	WB: 1:500	aML#025
	Rabbit	Abcam, 137033 (EPR9237)	WB: 1:750	aML#028
CSB/ERCC6	Goat	Santa Cruz, sc-10459 (E-18)	WB: 1:1000	aML#039
	Rabbit	Santa Cruz, sc-25370 (H-300)	WB: 1:300	aML#003
CUL4	Mouse	Santa Cruz, sc-377188 (H11)	WB: 1:100	aML#128
DDB1	Goat	Abcam, ab9194	WB: 1:1000	aML#035
ERCC1	Mouse	Santa Cruz, sc-17809 (D10)	WB: 1:300	aML#066
FLAG	Rabbit	New England Peptide; antigen: C(dPEG4)DYKDDDDK	WB: 1:5000	
GFP	Mouse	Roche, #11814460001 (7.1 and 13.1)	WB: 1:1000	aML#011
	Rabbit	Abcam, ab290	WB: 1:1000	aML#044
Goat IgG (H+L) CF680	Donkey	Thermo fisher Scientific, A21084	WB: 1:10000	aML#037
Mouse IgG (H+L) CF770	Goat	Biotium, VWR #20077	WB: 1:10000	aML#009
Mouse IgG (HRP)	Goat	Abcam, ab6789	WB: 1:10000	aML#132
p44/GTF2H2	Mouse	kindly provided by J.M. Egly (1H5)	WB: 1:2000	aML#075
p62/GTF2H1	Mouse	kindly provided by J.M. Egly (3C9)	WB: 1:2000	aML#074
	Mouse	Santa Cruz, sc-48431 (G10)	WB: 1:500	aML#099
	Mouse	Santa Cruz, sc-25329 (H10)		
p80/XPD/ERCC2	Mouse	Abcam, ab54676	WB: 1:500	aML#029
p89/XPB/ERCC3	Mouse	Millipore, MABE1123 (15TF2-1B3)	WB: 1:2000	aML#101
	Mouse	kindly provided by J.M. Egly (1B3)	WB: 1:1000	aML#073
	Rabbit	Santa Cruz, sc-293 (S-19)	WB: 1:1000	aML#040
	Mouse	Santa Cruz, sc-271500 (G10)		
rabbit IgG (H+L) CF680	Goat	Biotium, VWR #20067	WB: 1:10000	aML#010
Rbx1	Mouse	Santa Cruz, sc-393640 (E11)	WB: 1:100	aML#129
RNAPII-S2	Rabbit	Abcam, ab5095	WB: 1:1000	aML#024
RNAPII-S5	Mouse	Abcam, ab5408 (4H8)	WB: 1:1000	aML#125
Tubulin	Mouse	Sigma, T6199 (DM1A)	WB: 1:1000	aML#008
UVSSA	Mouse	Genetex, GTX629742 (GT816)	WB: 1:500	aML#100
	Rabbit	Novus Biologicals, NBP1-32598	WB: 1:1000	aML#030
	Rabbit	Abcam ab137644	WB: 1:1000	aML#034
	Rabbit	Genetex, GTX106751	WB: 1:1000	aML#087
xICSA	Rabbit	New England Peptide; antigen: CHRTHINPAFEDAWSSSEDES	WB: 1:5000	
XPA	Rabbit	kindly provided by Rick Wood (CJ1)	WB: 1:10000	aML#079
XPC	Rabbit	Novus Biologicals, NB100-58801	WB: 1:2000	aML#077
XPF/ERCC4	Mouse	Santa Cruz, sc-136153 (3F2/3)	WB: 1:200	aML#096

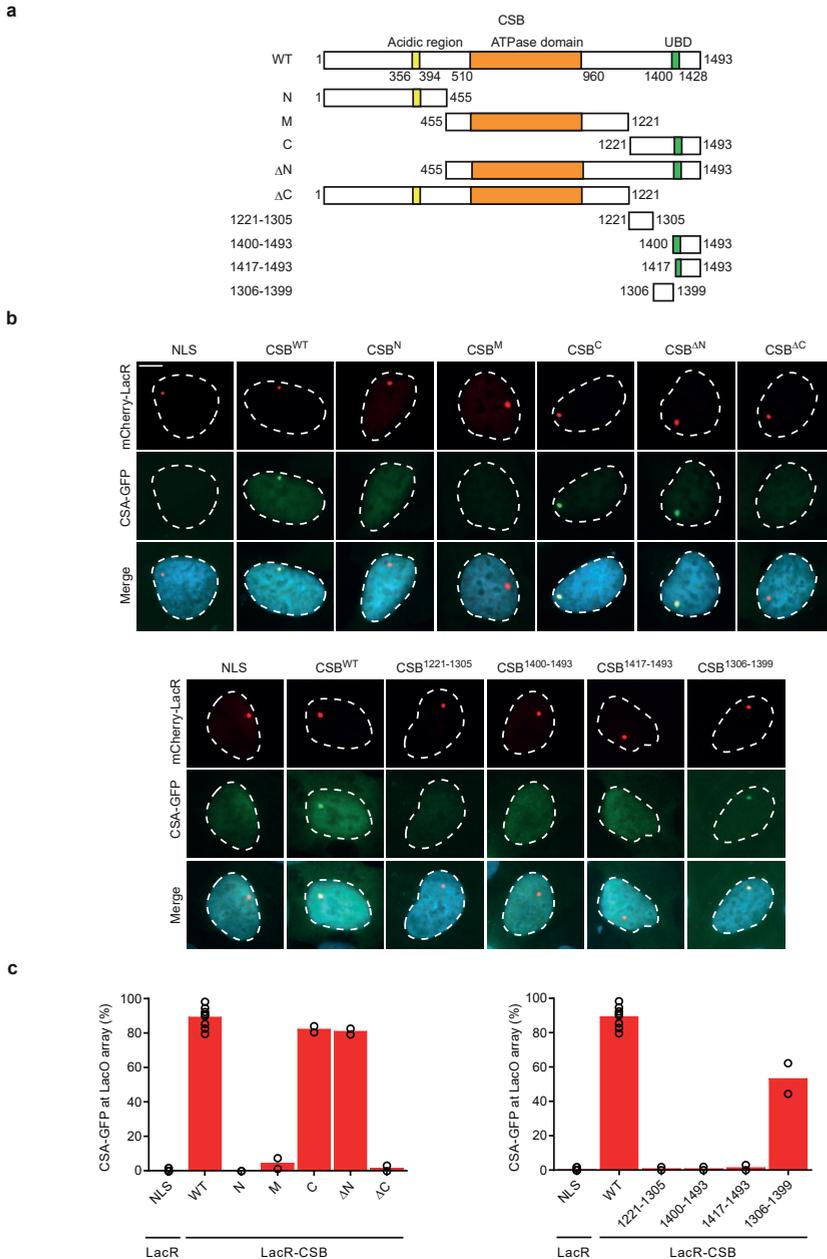
Supplementary figures



Supplementary Fig. 1: Testing of antibodies in IP and whole cell lysates. (a) Multiple endogenous RNAPII Co-IP experiments in WT cells for staining with various antibodies. Fig. 1b is composed of these Co-IPs. We were not always able to detect CDK7 probably due to weak interactions or inefficient antibody. (b) Endogenous RNAPII Co-IP using either a Ser2-phosphorylated RNAPII antibody or a Ser5-phosphorylated RNAPII antibody. (c) Testing of various UVSSA antibodies in KPS3-hTERT, KPS3-hTERT + UVSSA, U2OS (FRT) WT, and U2OS (FRT) UVSSA-KO cells. At least two independent replicates of each IP experiment were performed obtaining similar results.

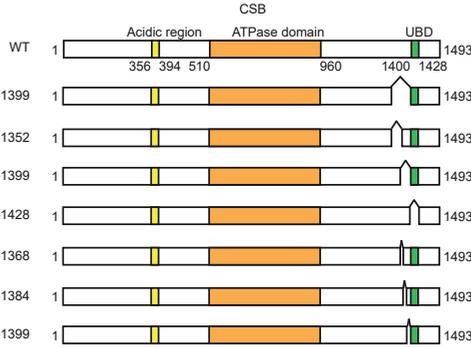


Supplementary Fig. 2: Sequence of CSB, CSA, and UVSSA-KO cells. (a) A schematic representation of CSB, CSA, and UVSSA including the location of the guide RNAs used for the generation of the CRISPR/Cas9-mediated KO. **(b)** Sequences of CSB, CSA, and UVSSA knockouts. **(c)** Western blot analysis of XPC and XPA knockouts ($n=2$).

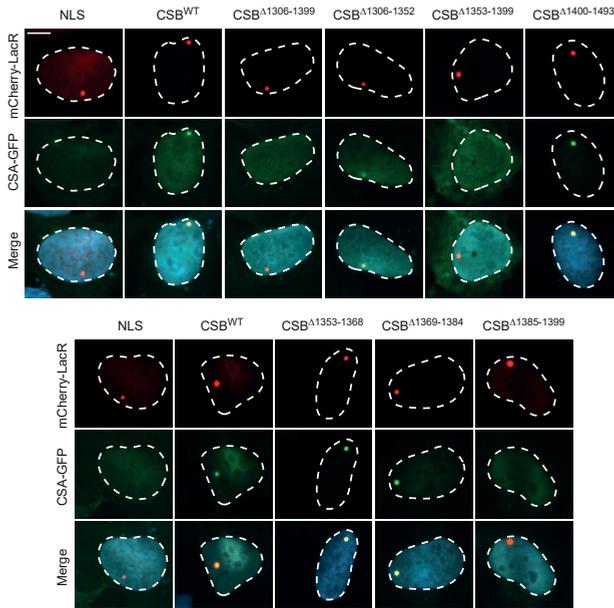


Supplementary Fig. 3: CSA interacts with the C-terminal region of CSB. (a) A schematic representation of CSB and its deletion mutants. (b) Recruitment of CSA-GFP to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins (scale bar = 5 μ m). (c) Quantification of CSA-GFP and mCherry-LacR-CSB co-localization at the LacO array. Each symbol represents the mean of an independent experiment ($n=2$ for all except LacR-NLS and LacR-CSB^{WT} which is $n=8$, >50 cells collected per experiment).

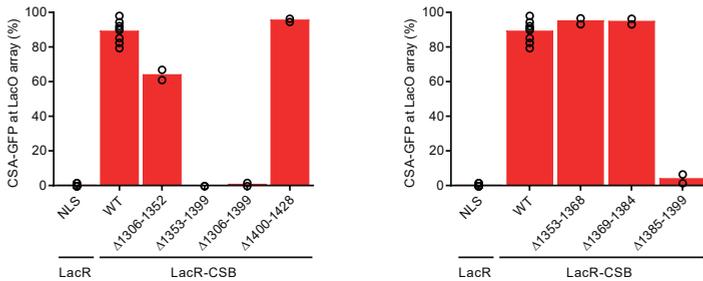
a



b



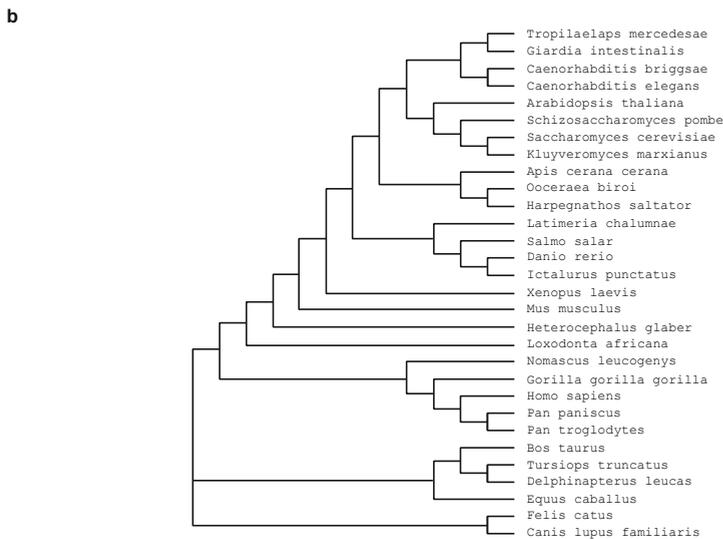
c



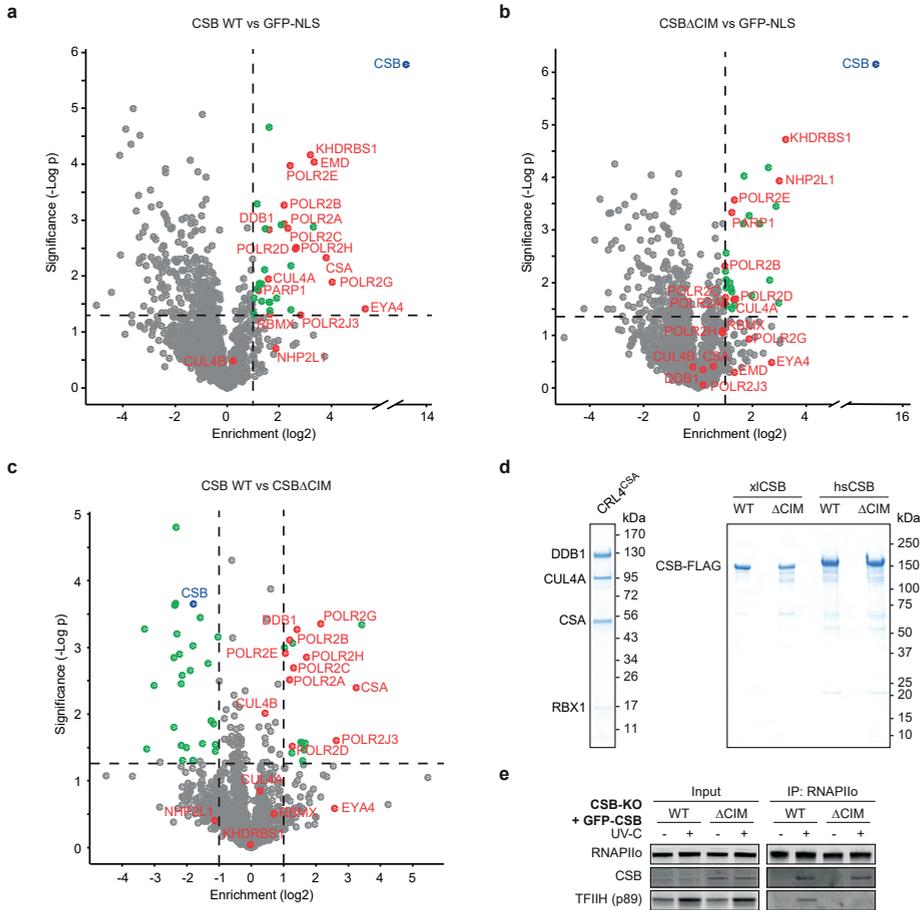
Supplementary Fig. 4: CSA interacts with amino acids 1385-1399 of CSB. (a) A schematic representation of CSB and its deletion mutants. (b) Recruitment of CSA-GFP to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins (scale bar = 5 μm). (c) Quantification of CSA-GFP and mCherry-LacR-CSB co-localization at the LacO array. Each symbol represents the mean of an independent experiment ($n=2$ for all except LacR-NLS and LacR-CSB^{WT} which is $n=8$, >50 cells collected per experiment).

a

		CIM	
1493	sp Q03468 ERCC6_HUMAN	ASSLLAKMRARNHL	Homo sapiens (Human)
1491	tr H2Q1W1 H2Q1W1_PANTR	ASSLLAKMRARNHL	Pan troglodytes (Chimpanzee)
1493	tr G3QVF5 G3QVF5_GORGO	ASSLLAKMRARNHL	Gorilla gorilla gorilla (Western lowland gorilla)
1492	tr G1S127 G1S127_NOMLE	ASSLLAKMRARNHL	Nomascus leucogenys (Northern white-cheeked gibbon)
1491	tr AOA2R8ZA95 AOA2R8ZA95_PANFA	ASSLLAKMRARNHL	Pan paniscus (Bonobo)
1482	tr E1BFL2 E1BFL2_BOVIN	SSLLAKMRARNHL	Bos taurus (Bovine)
1484	tr M3XEB4 M3XEB4_FELCA	PSSLLAKMRARNHL	Felis catus (Cat)
1486	tr E2QSK6 E2QSK6_CANLF	PSSLLAKMRARNHL	Canis lupus familiaris (Dog)
1460	tr F7D5S6 F7D5S6_HORSE	SSLLAKMRARNHL	Equus caballus (Horse)
1432	tr G3TCV9 G3TCV9_LOXAF	TSSLLAKMRARNHL	Loxodonta africana (African elephant)
1481	tr AOA0P6J577 AOA0P6J577_HETGA	TSSLLAKMRARNHL	Heterocephalus glaber (Naked mole rat)
1481	tr F8VPZ5 F8VPZ5_MOUSE	SSSLLAKMRARNHM	Mus musculus (Mouse)
1485	tr AOA2Y9MEF2 AOA2Y9MEF2_DELLE	SSSLLAKMRARNHL	Delphinapterus leucas (Beluga whale)
1462	tr AOA2U4AKE1 AOA2U4AKE1_TURTR	TSSSLLAKMRARNHL	Tursiops truncatus (Atlantic bottle-nosed dolphin)
1488	tr AOA1S3N477 AOA1S3N477_SALSA	SSSTLLAKMRARNYL	Salmo salar (Atlantic salmon)
1409	tr AOA2D0QRA6 AOA2D0QRA6_ICTFU	SSSLLAKMRARNHV	Ictalurus punctatus (Channel catfish)
1389	tr F1R294 F1R294_DANRE	SSSLLAKMRARNHL	Danio rerio (Zebrafish)
1370	tr AOA1L8FKT9 AOA1L8FKT9_XENLA	SSSLLAKMRARNHL	Xenopus laevis (African clawed frog)
1387	tr H3AWFO H3AWFO_LATCH	SSSLLAKMRARNHL	Latimeria chalumnae (West Indian ocean coelacanth)
1187	sp Q9ZV43 CHR8_ARATH	SAEFLNLRGSREQ	Arabidopsis thaliana (Mouse-ear cress)
1015	tr A8XNA8 A8XNA8_CAEBR	-----	Caenorhabditis briggsae (nematode)
957	tr Q93781 Q93781_CAELL	-----	Caenorhabditis elegans (nematode)
973	sp Q9UR24 RHP26_SCHPO	---TLALQKQR---	Schizosaccharomyces pombe (Fission yeast)
1085	sp P40352 RAD26_YEAST	NYDDGIT-FA--RSK	Saccharomyces cerevisiae (Baker's yeast)
1037	tr W0T437 W0T437_KLUMD	LKVKTIPSQE--KKK	Kluyveromyces marxianus (Yeast)
925	tr Q6WD94 Q6WD94_GIAIN	-----	Giardia intestinalis (intestinal parasite)
1222	tr AOA1V9XZ12 AOA1V9XZ12_9ACAR	SPSRPKGKR ^R SVAVL	Tropilaelaps mercedesae (bee mite)
1125	tr AOA2A3EJZ4 AOA2A3EJZ4_APICC	-----	Apis cerana cerana (Oriental honeybee)
1073	tr AOA026W8Z2 AOA026W8Z2_OOCBI	-----	Ooceraea biroi (Clonal raider ant)
1005	tr E2BDE2 E2BDE2_HARSA	-----	Harpegnathos saltator (Jerdon's jumping ant)

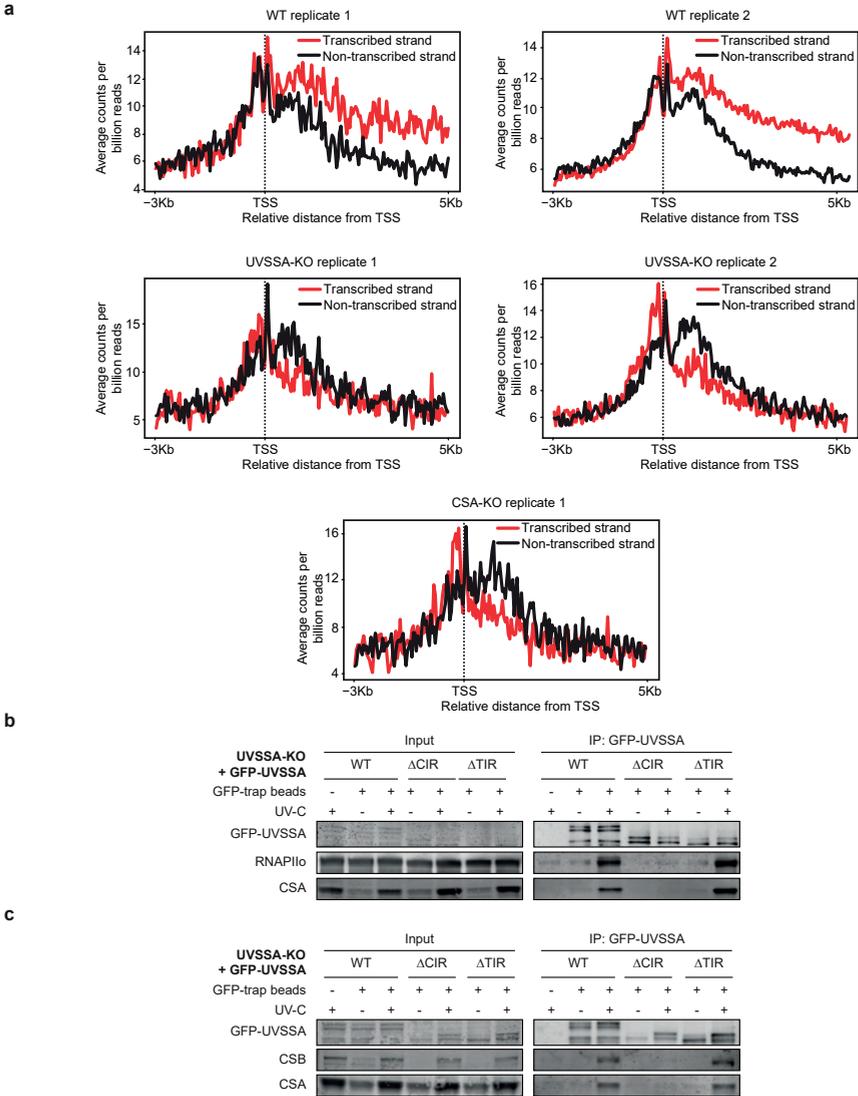


Supplementary Fig. 5: Alignment of CSB and CSB orthologues. (a) Alignment of the C-terminal CIM of CSB orthologues from a variety of different species. Sequences were aligned with ClustalW (b) A phylogenetic tree was constructed based on the alignment of CSB orthologues using ClustalW.

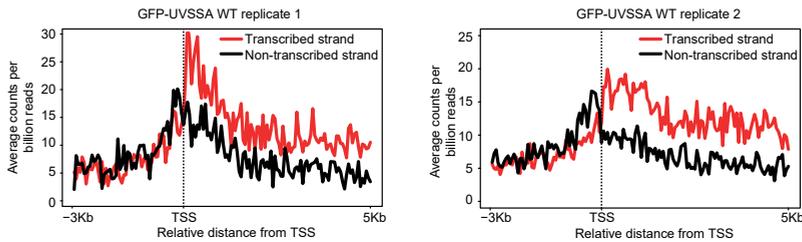


Supplementary Fig. 6: Mass-spectrometry after GFP-CSB pull-down, and purification of proteins.

(a-c) Volcano plots depicting mass spectrometry analysis after GFP immunoprecipitation on the soluble and chromatin fraction of UV-irradiated (20 J/m^2) cells comparing (a) GFP-CSB^{WT} with GFP-NLS (b) GFP-CSB Δ CIM with GFP-NLS, and (c) GFP-CSB^{WT} with GFP-CSB Δ CIM. The enrichment (\log^2) is plotted on the x-axis and the significance ($-\log p$ -value) is plotted on the y-axis. The $-\log p$ -value threshold was set to 1.3 ($p < 0.05$). The enrichment threshold was set to 1 and all significant hits are shown in green. Several selected hits are shown in red, and GFP-CSB is shown in blue. (d) Coomassie gels of recombinant xCRL4^{CSA} complex and xICSB or hsCSB variants. DDB1, CUL4A, and all CSB proteins contained an N-terminal FLAG-tag, and CSA contained a C-terminal Strep-tag II. (e) Endogenous RNAPII Co-IP in CSB-KO + GFP-CSB^{WT} and CSB-KO + GFP-CSB Δ CIM. At least two independent replicates of each IP experiment were performed obtaining similar results.



Supplementary Fig. 7: XR-seq in TCR-KO cells, and immunoprecipitation in UVSSA mutants. (a) Average CPD XR-seq repair signal 3 Kb upstream and 5 Kb downstream of the annotated TSS of 16,088 genes in two independent biological replicates of experiments in WT and UVSSA-KO cells, and in a single replicate of CSA-KO cells. Signal is plotted separately for the transcribed (red) and non-transcribed (black) strands. The bin size of 40 nt. **(b-c)** Co-IP of GFP-UVSSA^{WT}, GFP-UVSSA ^{Δ CIR}, and GFP-UVSSA ^{Δ TIR}. At least two independent replicates of each IP experiment were performed obtaining similar results.



Supplementary Fig. 9: XR-seq in GFP-UVSSA WT cells. Average CPD XR-seq repair signal 3 Kb upstream and 5 Kb downstream of the annotated TSS of 16,088 genes in two independent biological replicates of experiments in GFP-UVSSA WT cells. Signal is plotted separately for the transcribed (red) and non-transcribed (black) strands, with a bin size of 53 nt.

References

- [1] van der Weegen, Y. *et al.* The Cooperative action of CSB, CSA and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nat. Commun.* **11**, 2104 (2020).
- [2] 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**, 241-249 (1987).
- [3] Martejijn, 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**, 465-481 (2014).
- [4] Evans, E., Moggs, J. G., Hwang, J. R., Egly, J. M. & Wood, R. D. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.* **16**, 6559-6573 (1997).
- [5] Tapias, A. *et al.* Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J. Biol. Chem.* **279**, 19074-19083 (2004).
- [6] Huang, J. C., Svoboda, D. L., Reardon, J. T. & Sancar, A. Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc. Natl Acad. Sci. USA* **89**, 3664-3668 (1992).
- [7] Svoboda, D. L., Taylor, J. S., Hearst, J. E. & Sancar, A. DNA repair by eukaryotic nucleotide excision nuclease. Removal of thymine dimer and psoralen monoadduct by HeLa cell-free extract and of thymine dimer by *Xenopus laevis* oocytes. *J. Biol. Chem.* **268**, 1931-1936 (1993).
- [8] Hoffen, A. *et al.* Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells. *Nucleic Acids Res.* **21**, 5890-5895 (1993).
- [9] Spivak, G. *et al.* Ultraviolet-sensitive syndrome cells are defective in transcription-coupled repair of cyclobutane pyrimidine dimers. *DNA Repair* **1**, 629-643 (2002).
- [10] Nance, M. A. & Berry, S. A. Cockayne syndrome: review of 140 cases. *Am. J. Med. Genet.* **42**, 68-84 (1992).
- [11] Friedberg, E. C. Cockayne syndrome—a primary defect in DNA repair, transcription, both or neither? *Bioessays* **18**, 731-738 (1996).
- [12] Horibata, K. *et al.* Complete absence of Cockayne syndrome group B gene product gives rise to UV-sensitive syndrome but not Cockayne syndrome. *Proc. Natl Acad. Sci. USA* **101**, 15410-15415 (2004).
- [13] Nardo, T. *et al.* A UV-sensitive syndrome patient with a specific CSA mutation reveals separable roles for CSA in response to UV and oxidative DNA damage. *Proc. Natl Acad. Sci. USA* **106**, 6209-6214 (2009).
- [14] Itoh, T., Ono, T. & Yamaizumi, M. A new UV-sensitive syndrome not belonging to any complementation groups of xeroderma pigmentosum or Cockayne syndrome: siblings showing biochemical characteristics of Cockayne syndrome without typical clinical manifestations. *Mutat. Res.* **314**, 233-248 (1994).
- [15] Laugel, V. *et al.* Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. *Hum. Mutat.* **31**, 113-126 (2010).
- [16] Calmels, N. *et al.* Functional and clinical relevance of novel mutations in a large cohort of patients with Cockayne syndrome. *J. Med. Genet.* **55**, 329-343 (2018).
- [17] Nakazawa, Y. *et al.* Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase II processing in transcription-coupled nucleotide-excision repair. *Nat. Genet.* **44**, 586-592 (2012).

- 2
- [18] Zhang, X. *et al.* Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nat. Genet.* **44**, 593–597 (2012).
 - [19] 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**, 939–953 (1992).
 - [20] Gool, A. J. *et al.* The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex. *EMBO J.* **16**, 5955–5965 (1997).
 - [21] Tantin, D., Kansal, A. & Carey, M. Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Mol. Cell Biol.* **17**, 6803–6814 (1997).
 - [22] van den Boom, V. *et al.* DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J. Cell Biol.* **166**, 27–36 (2004).
 - [23] Svejstrup, J. Q. Rescue of arrested RNA polymerase II complexes. *J. Cell Sci.* **116**, 447–451 (2003).
 - [24] Tantin, D. RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62. *J. Biol. Chem.* **273**, 27794–27799 (1998).
 - [25] 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**, 357–367 (2003).
 - [26] Schwertman, P. *et al.* UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat. Genet.* **44**, 598–602 (2012).
 - [27] 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**, 555–564 (1995).
 - [28] Fischer, E. S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024–1039 (2011).
 - [29] 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**, 1429–1434 (2006).
 - [30] 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 (2017).
 - [31] 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).
 - [32] 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).
 - [33] 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**, 35118–35126 (2012).
 - [34] Compe, E. & Egly, J. M. Nucleotide excision repair and transcriptional regulation: TFIIH and beyond. *Annu. Rev. Biochem.* **85**, 265–290 (2016).
 - [35] Coin, F. *et al.* Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol. Cell* **31**, 9–20 (2008).
 - [36] Kocic, G. *et al.* Structural basis of TFIIH activation for nucleotide excision repair. *Nat. Commun.* **10**, 2885 (2019).
 - [37] Harlen, K. M. & Churchman, L. S. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **18**, 263–273 (2017).
 - [38] Schilbach, S. *et al.* Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature* **551**, 204–209 (2017).
 - [39] 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* **1**, 1027–1038 (2002).
 - [40] Janicki, S. M. *et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683–698 (2004).
 - [41] Luijsterburg, M. S. *et al.* PARP1 links CHD2-mediated chromatin expansion and H3.3 deposition to DNA repair by non-homologous end-joining. *Mol. Cell* **61**, 547–562 (2016).
 - [42] Luijsterburg, M. S. *et al.* A PALB2-interacting domain in RNF168 couples homologous recombination to DNA break-induced chromatin ubiquitylation. *eLife* **6** (2017).
 - [43] Anindya, R. *et al.* A ubiquitin-binding domain in Cockayne syndrome B required for transcription-coupled nucleotide excision repair. *Mol. Cell* **38**, 637–648 (2010).
 - [44] Batenburg, N. L. *et al.* ATM and CDK2 control chromatin remodeler CSB to inhibit RIF1 in DSB repair pathway choice. *Nat. Commun.* **8**, 1921 (2017).
 - [45] 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).
- [46] Volker, M. *et al.* Sequential assembly of the nucleotide excision repair factors in vivo. *Mol. Cell* **8**, 213–224 (2001).
- [47] 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**, 948–960 (2015).
- [48] Hu, J. *et al.* Genome-wide mapping of nucleotide excision repair with XR-seq. *Nat. Protoc.* **14**, 248–282 (2019).
- [49] 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. USA* **99**, 201–206 (2002).
- [50] 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**, 235–246 (2010).
- [51] Panier, S. *et al.* Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. *Mol. Cell* **47**, 383–395 (2012).
- [52] Adar, S., Hu, J., Lieb, J. D. & Sancar, A. Genome-wide kinetics of DNA excision repair in relation to chromatin state and mutagenesis. *Proc. Natl Acad. Sci. USA* **113**, E2124–E2133 (2016).
- [53] Nakazawa, Y. *et al.* Ubiquitination of DNA damage-stalled RNAPII promotes transcription-coupled repair. *Cell* **180**, 1228–1244 (2020).
- [54] Xu, J. *et al.* Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653–657 (2017).
- [55] He, Y. *et al.* Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016).
- [56] Salas-Lloret, D., Agabiti, G. & Gonzalez-Prieto, R. TULIP2: an improved method for the identification of ubiquitin E3-specific targets. *Front. Chem.* **7**, 802 (2019).
- [57] Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
- [58] Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. methods* **13**, 731–740 (2016).
- [59] Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- [60] Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- [61] Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- [62] Akalin, A., Franke, V., Vlahovicek, K., Mason, C. E. & Schubeler, D. Genomation: a toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics* **31**, 1127–1129 (2015).
- [63] Soutoglou, E. & Misteli, T. Activation of the cellular DNA damage response in the absence of DNA lesions. *Science* **320**, 1507–1510 (2008).