The sequential and cooperative action of CSB, CSA and UVSSA targets the TFIIH complex to DNA damage-stalled RNA polymerase II

Yana van der Weegen[1], Hadar Golan Berman[3], Tycho E.T. Mevissen[5], Katja Apelt[1], Román González-Prieto[2], Elisheva Heilbrun[3,4], Alfred C.O. Vertegaal[2], Diana van den Heuvel[1], Johannes C. Walter[5], Sheera Adar[3], and Martijn S. Luijsterburg[1,*]

[1] Department of Human Genetics, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands
[2] Department of Cell and Chemical Biology, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands
[3] Department of Microbiology and Molecular Genetics, The institute for medical research Israel-Canada, The Faculty of Medicine, The Hebrew University of Jerusalem, Ein Kerem, Jerusalem, 91120, Israel
[4] Department of Bioinformatics, School of life and health science, Jerusalem College of Technology, Jerusalem, 9372115, Israel.
[5] Howard Hughes Medical Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

Running title: Assembly mechanism of the human TCR complex

* Corresponding author: Martijn Luijsterburg (m.luijsterburg@lumc.nl)

Summary (149 words)

The response to DNA damage-stalled RNA polymerase II (RNAPIIo) 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 IIH (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 RNAPIIo-bound CSB, which recruits CSA through a newly identified CSA-interaction motif (CIM). Once recruited, CSA facilitates the association of UVSSA with stalled RNAPIIo. 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 reveal a sequential and highly cooperative assembly mechanism of TCR proteins and reveal the mechanism for TFIIH recruitment to DNA damage-stalled RNAPIIo to initiate repair.

Nucleotide excision repair (NER) is a versatile DNA repair pathway that removes a wide range of helix-distorting DNA lesions from our genome, including ultra-violet (UV) light–induced photolesions. Transcription-coupled repair (TCR) is a specialized NER sub-pathway that specifically removes DNA lesions from actively transcribed DNA strands[1]. It is believed that the TCR pathway is initiated by the stalling of elongating RNA polymerase II (RNAPIIo) at DNA lesions, which triggers the recruitment of the core NER machinery to repair these lesions[2]. After lesion recognition, the transcription factor IIH (TFIIH) complex is recruited to unwind the DNA[3, 4], followed by dual incision and the release of a 22-30 nucleotide-long DNA strand containing the lesion[5, 6]. The generated single-stranded DNA gap is filled by repair synthesis and the nick is sealed[2]. However, the mechanism through which TCR recognizes transcription-blocking lesions and recruits the repair machinery remains elusive.

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 the repair of transcription-blocking lesions through TCR[7, 8], the phenotypes are very different. CS is characterized by severe and progressive neurodegeneration[9, 10], while UV^{S}S shows a mild UV-sensitive phenotype[11-13]. The majority of CS patients carry mutations in the genes encoding either the CSB or CSA proteins[14, 15]. Patients with UV^{S}S carry mutations in the gene encoding the UVSSA protein[16, 17].

The 168 kDa CSB protein contains a central SWI2/SNF2-like DNA-dependent ATPase domain[18]. Biochemical experiments revealed that CSB resides in a complex with RNAPIIo[19, 20]. Indeed, live-cell imaging suggests that CSB monitors the progression of transcription elongation by continuously probing RNAPIIo complexes[21]. It has been suggested that CSB is involved in the removal or backtracking of RNAPII to make the DNA lesion accessible for repair proteins[22]. Although the association of CSB with RNAPII is sufficient to recruit TFIIH *in vitro*[23], whether additional factors are required to trigger the recruitment of the repair machinery *in vivo* remains unanswered.

In addition to CSB, the CSA and UVSSA proteins also associate with DNA damage-stalled RNAPIIo[16, 17, 24, 25]. The 44 kDa CSA protein contains seven WD40 repeats that form a seven bladed β-propeller[26]. Earlier work has shown that CSA is incorporated into a DDB1-CUL4-based E3 ubiquitin ligase complex[24, 27] that becomes transiently activated in response to UV irradiation and targets CSB for proteasomal degradation[28]. Current models suggest that CSA is dispensable for the recruitment of the excision repair machinery to stalled RNAPII[29], and that CSA is unlikely to recruit UVSSA to sites of UV-induced DNA damage[30]. Thus, the precise recruitment mechanism and the role of CSA in TCR is currently not clear.

The 81 kDa 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[16, 17, 25, 31]. However, ectopic expression of CSB in UVSSA-deficient cells did not rescue TCR, suggesting that UVSSA has additional functions in this repair mechanism[16]. Moreover, UVSSA was found to associate with RNAPII[17, 25], 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.

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 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 specifically 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[32]. Immunoprecipitation of RNAPIIo revealed a UV-specific association with the Cockayne syndrome proteins CSB and CSA, as well as with several subunits of the TFIIH complex (XPD/p80, XPB/p89, GTFH1/p62; **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[33].

Although the CS proteins and TFIIH readily assembled with RNAPIIo after UV irradiation, downstream repair proteins such as XPA, XPG, 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 RNAPIIo or in whole cell lysates due to a lack of specific antibodies (**Supplementary Fig 1b**). 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[22].

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 1c; 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 1d**)[34]. 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 1c, d**). We next applied our immunoprecipitation-based method in the different TCR-KO cells to establish how CSB and CSA recruitment to DNA damage-stalled RNAPIIo is regulated. CSB associated with RNAPIIo 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 RNAPIIo **(Fig 1e)**. The association of CSA with

stalled RNAPIIo was abolished in CSB-KO cells, but was not affected in cells lacking UVSSA (**Fig 1e**). Importantly, re-expressing GFP-tagged CSB in the CSB-KO cells restored the association between RNAPIIo and CSA (**Fig 1e, f**), confirming that CSB is required for the recruitment of CSA to damage-stalled RNAPIIo. The CSA protein is part of a DDB1-CUL4 E3 ubiquitin ligase complex[24, 27], and we therefore asked whether CSA associates with DNA damage-stalled RNAPIIo together with its E3 ubiquitin ligase partner DDB1. As an additional control we also included XPC-KO cells, which are deficient in global genome repair (GGR; **Supplementary Fig 2c**). Immunoprecipitation of RNAPIIo revealed a UV-specific interaction with DDB1 in WT, XPC-KO, and UVSSA-KO cells (**Fig 1g**). However, this interaction was completely abolished in CSA-KO and CSB-KO cells, showing that CSA indeed mediates the recruitment of the DDB1-CUL4 complex to lesion-stalled RNAPIIo (**Fig 1g**).

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[35]. 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[36, 37](**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 Figs 3, 4**). Fragments of CSB spanning the N-terminus or the central region containing the conserved ATPase/helicase domain (N, M, and ΔC) were unable to recruit CSA-GFP. Conversely, tethering of a LacR-tagged CSB region spanning the C-terminus (C and ΔN) triggered robust recruitment of CSA-GFP (**Fig 2a-d, Supplementary Fig 3a-c**). 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[38]) and a recently identified winged-helix domain (WHD; 1417-1493) that interacts with RIF1[39]. 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 3a-c**). Importantly, we found that tethering full-length CSB lacking this minimal interaction region (Δ1306-1399) 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 4a-c**). Moreover, while $CSB^{\Delta 1353\text{-}1368}$ and $CSB^{\Delta 1369\text{-}1384}$ 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 mediates the recruitment of 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 $CSB^{\Delta CIM}$ 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 $CSB^{\Delta CIM}$ even though equal amounts of CSB were immunoprecipitated (**Fig 3c**). Immunoprecipitation of endogenous RNAPIIo in these cell lines showed that both CSB^{WT} and $CSB^{\Delta CIM}$ associated equally with RNAPIIo after UV irradiation. However, $CSB^{\Delta CIM}$ 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-$CSB^{\Delta CIM}$ in the CSB-KO cells failed to restore the sensitivity to Illudin S, while expression of GFP-$CSB^{\Delta WT}$ 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 $CSB^{\Delta CIM}$ with ubiquitin, E1, E2, and the E3 ubiquitin ligase $CRL4^{CSA}$ consisting of *Xenopus laevis* CSA, DDB1, CUL4A, and RBX1 (**Supplementary Fig 6**). While $xlCRL4^{CSA}$ promoted the efficient ubiquitylation of $xlCSB^{WT}$, it did not ubiquitylate $xlCSB^{\Delta CIM}$ (**Fig 3f**). These data suggest that xlCSB uses its CIM to interact directly with xlCSA. Consistent with this interpretation, immobilized $xlCSB^{WT}$ but not $xlCSB^{\Delta CIM}$ interacted with endogenous xlCSA from *Xenopus* egg extracts (**Fig 3g**). Similar results were observed when xlCSB was substituted with hsCSB (**Fig 3f, g**). Collectively, these data demonstrate that CSA is recruited to DNA damage-stalled RNAPIIo by CSB through direct interactions with the newly identified C-terminal CIM in CSB.

UVSSA is recruited to DNA damage-stalled RNAPIIo by CSA

Previous studies have demonstrated that UVSSA associates with RNAPIIo, but due to conflicting results, it remains unclear if UVSSA recruitment to RNAPIIo is enhanced by UV irradiation and dependent on the CS proteins[17, 25,30]. Therefore, we monitored GFP-UVSSA recruitment to RNAPIIo 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 RNAPIIo 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 and CSB-KO cells (**Fig 4c**). Thus, GFP-UVSSA is UV-specifically targeted to DNA damage-stalled RNAPIIo in a manner that is dependent on the CS proteins[17]. Moreover, pull-down of GFP-UVSSA confirmed a robust UV-induced association with RNAPIIo, 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 RNAPIIo.

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

It has been shown that CSB, CSA, and UVSSA can each associate with TFIIH[23, 26, 40], but which of these proteins is responsible for the recruitment of TFIIH to DNA damage-stalled RNAPIIo to initiate repair is currently unknown. To directly asses if CSB and CSA are required for the recruitment of TFIIH, we monitored TFIIH (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 RNAPIIo revealed a UV-specific interaction with TFIIH in WT cells, while these interactions were severely reduced in the CSB-KO and CSA-KO cells (**Fig 5a**). Interestingly, TFIIH also failed to associate with RNAPIIo in CSB-KO cells complemented with GFP-CSB$^{\Delta CIM}$ (**Supplementary Fig 7a**), consistent with our findings that this mutant is not capable of recruiting CSA (**Fig 3c, g**). These initial results suggest that the TFIIH complex is recruited in a manner that requires both CS proteins.

UVSSA targets the TFIIH complex to stalled RNAPIIo in a CS protein-dependent manner

It has been reported that UVSSA can interact with TFIIH[16, 31, 40], 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-irradiated cells followed by mass spectrometry (MS). In the absence of UV-induced DNA damage, we identified 35 specific UVSSA interactors, including the known interactor USP7. However, we did not detect any significant interactions with RNAPII subunits or CSB in the chromatin fraction of unirradiated cells (**Supplementary Fig 7b; Supplementary Tables 1, 3**). 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; Supplementary Fig 7c; Supplementary Tables 2, 3**). 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[41].

Immunoprecipitation of endogenous RNAPIIo 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-KO cells, we found that the UV-induced interaction between RNAPIIo 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 RNAPIIo (**Fig 5e**). These findings demonstrate that CSB, CSA, and UVSSA are equally important for the recruitment of the TFIIH complex to DNA damage-stalled RNAPIIo.

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[42]. We generated nucleotide-resolution maps of UV-induced cyclobutane pyrimidine dimer (CPDs) repair in U2OS wild-type cells (**Fig 5f; Supplementary Fig 8a**), which revealed that CPD repair under these conditions is enriched on the transcribed strands within gene bodies consistent with TCR-mediated repair[42]. Importantly, the CPD repair bias in transcribed strands was completely lost in both CSA-KO (**Supplementary Fig 8a**) 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 the TFIIH complex to DNA damage-stalled RNAPIIo

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$^{\Delta 100\text{-}200}$) or the TFIIH complex (UVSSA$^{\Delta 400\text{-}500}$)[31] (**Fig 6a**). These separation-of-function mutants were characterized by our previously described chromatin-tethering approach. mCherry-LacR-UVSSA^{WT} clearly localized to the LacO array and triggered the robust recruitment of CSA-GFP and endogenous TFIIH. As expected, mCherry-LacR-UVSSA$^{\Delta 100\text{-}200}$ was unable to recruit CSA-GFP to the LacO array, but triggered robust TFIIH recruitment (**Fig 6b,c)**. In contrast, mCherry-LacR-UVSSA$^{\Delta 400\text{-}500}$ 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 RNAPIIo, CSB, CSA, and TFIIH. These interactors were virtually absent after pull-down of GFP-UVSSA$^{\Delta CIR}$ (**Fig 6e; Supplementary Fig 8b, 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 RNAPIIo, CSB, and CSA, while its interaction with TFIIH was completely abolished (**Fig 6e; Supplementary Fig 8b, c**).

We next set out to directly asses the ability of these UVSSA mutants to participate in TCR complex assembly. Immunoprecipitation of endogenous RNAPIIo showed a UV-specific association of RNAPIIo with CSB and CSA in both UVSSA^{WT} and mutant cell lines (**Fig 6f**). This is in line with our other data since CSB, CSA, and UVSSA associate sequentially with RNAPIIo, and UVSSA is the last TCR protein to be recruited. Moreover, endogenous RNAPIIo immunoprecipitation resulted in a UV-specific interaction with GFP-UVSSA^{WT} and GFP-UVSSA$^{\Delta TIR}$, whereas GFP-UVSSA$^{\Delta CIR}$ failed to associate with RNAPIIo. 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 RNAPIIo (**Fig 4c**). In addition, in both mutant cell lines the recruitment of TFIIH (p89) to DNA damage-stalled RNAPIIo was completely abolished (**Fig 6f**). These experiments strongly suggest that TFIIH is recruited to DNA damage-stalled RNAPIIo via direct protein-protein contacts with UVSSA. Importantly, the stable expression of GFP-UVSSA$^{\Delta CIR}$ and GFP-UVSSA$^{\Delta TIR}$ in UVSSA-KO cells failed to restore their sensitivity to Illudin S, which was almost fully restored by GFP-UVSSA^{WT} (**Fig 6g**).

Altogether, our data reveal a sequential and cooperative assembly mechanism of the human TCR complex, which involves the stepwise assembly of CSB, CSA, and UVSSA to target the TFIIH complex to DNA damage-stalled RNAPIIo to initiate DNA repair (**Fig 6h**).

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 eukaryotic TCR. Our findings suggest a highly cooperative recruitment mechanism that involves the sequential association of CSB, CSA and UVSSA to target the TFIIH complex to DNA damage-stalled RNAPIIo to initiate repair.

CSA recruitment by CSB is crucial for TCR

We show that both CSB and CSA associate with RNAPIIo in a manner that is strongly induced by UV irradiation. Importantly, we find that CSA recruitment is completely dependent on CSB. These findings are in line with earlier work showing that CSB facilitates the translocation of CSA to the nuclear matrix after UV irradiation[43]. Moreover, we demonstrate that CSA is required for the association of DDB1 with RNAPIIo, suggesting that CSA is recruited to DNA damage-stalled RNAPIIo as part of a $CRL4^{CSA}$ complex[24, 27]. Previous findings suggested that CSB dynamically associates with RNAPIIo under undamaged conditions and that this interaction is stabilized upon UV irradiation[21, 44]. While our method may not be sensitive enough to capture these transient interactions, our findings do support that the CSB-RNAPIIo interaction is stabilized after UV irradiation.

Earlier observations suggested that CSB physically interacts with CSA[26, 28], while other studies failed to detect this association[19, 20]. 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 RNAPIIo. 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 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. In line with this, previous studies revealed that the association of CSB with stalled RNAPIIo triggers a conformational change that repositions the N-terminus, thereby exposing residues in the C-terminus of CSB[44]. 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[38], we find that $CSB^{\Delta UBD}$ is fully functional in interacting with CSA. However, it is possible that the CIM and the UBD collaborate, as a tandem protein-interaction module[45], 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[27].

CSA recruits UVSSA to RNAPIIo in a UV-dependent manner

The recently identified UVSSA protein can be isolated as part of a chromatin-bound stalled RNAPIIo complex. Our current findings shed light on its recruitment mechanism by demonstrating that the association of UVSSA with RNAPIIo 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 RNAPIIo, suggesting that UVSSA is the last of these proteins to be recruited. Consistent with a reported association between CSA and UVSSA[31], we find that CSA targets UVSSA to DNA damage-stalled RNAPIIo 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 RNAPIIo, live-cell imaging experiments showed that UVSSA is recruited to sites of UV-C-induced laser damage independently of the CS proteins[25, 30]. There could be several reasons for these seemingly conflicting results. Firstly, the methodology is very different. We isolate RNAPIIo-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 stalled RNAPIIo. In line with this hypothesis, using similar conditions, GFP-CSA could not be detected at sites of local UV-C laser damage[30], even though CSA is essential for TCR and showed a robust association with stalled RNAPIIo under our conditions. Secondly, the time-frame during which UVSSA association is measured is different. While we isolate RNAPIIo-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 RNAPIIo 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 RNAPIIo.

TFIIH recruitment to DNA damage-stalled RNAPIIo is dependent on UVSSA

A major unresolved question is how the core NER machinery, likely starting with the TFIIH complex, is recruited to DNA damage-stalled RNAPIIo to initiate repair. Biochemical *in vitro* experiments have shown that the association of CSB with RNAPII is sufficient to recruit TFIIH[23]. In addition, CSA was shown to associate with the p44 subunit of TFIIH[26], while UVSSA can interact with the p62 subunit of TFIIH[40]. 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 RNAPIIo *in vivo*. Indeed, similar to previous results with CSB-deficient cells[42, 46], 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 stalled RNAPIIo. 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*[40]. Moreover, we found that the UVSSA$^{\Delta CIR}$ 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 RNAPIIo but also stabilizes the direct interaction between UVSSA and TFIIH, resulting in the recruitment of TFIIH to stalled RNAPIIo. 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 RNAPIIo *in vivo*, or whether other subunits and regions also contribute.

UVSSA: a NER-specific coupling factor?

Here we show that UVSSA is essential to bridge the TFIIH complex to CSB/CSA-bound RNAPIIo 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^{S}S). In line with this, CS fibroblasts are sensitive to oxidative damage, while UV^{S}S fibroblasts are not[12, 47]. Moreover, it was recently shown that CSB recruits the DNA repair protein XRCC1, which is involved in base excision repair (BER), to oxidative lesions in a transcription-dependent manner[48]. These findings suggest that the CS proteins are involved in transcription-dependent transactions in multiple DNA repair pathways through specific coupling factors. Here, we show that UVSSA is a NER-specific coupling factor. It would be interesting to explore if additional coupling factors exist that link the CS proteins to other DNA repair systems.

A model for TCR complex assembly

We propose a model in which CSB is the first protein to be recruited to DNA damage-stalled RNAPIIo (**Fig 6h**). 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 RNAPIIo, 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

RNAPIIo and CSB/CSA/UVSSA are displaced and that the TCR-specific pre-incision complex is assembled starting with XPA. In this regard it is interesting to note the yeast orthologue of CSB, RAD26, is bound to the DNA upstream of RNAPII[49], while human TFIIH in the transcription pre-initiation complex (PIC) is bound downstream of RNAPII[50]. 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.

Experimental Procedures

Cell lines. Cell lines (listed in **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). U2OS 2–6-3 cells containing 200 copies of a LacO-containing cassette (~4 Mbp) were a gift from Susan Janicki[35]. UVSSA-deficient KPS3-hTERT cells and their UVSSA-rescued counterparts were a gift from Tomoo Ogi[16]. 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[45].

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 **table 2** for plasmids, **table 3** 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 $MgCl_2$ 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 PCR amplified (sequencing primers are listed in **table 4**) followed by sanger sequencing using either the forward or the reversed primer.

Generation of stable cell lines. Selected knockout clones of CSB, CSA, and UVSSA (see **table 1**) were subsequently used to stably express GFP-CSB^{WT}, GFP-$CSB^{\Delta CIM}$, CSA^{WT}-GFP, GFP-$UVSSA^{WT}$, GFP-$UVSSA^{\Delta CIR}$, and GFP-$UVSSA^{\Delta TIR}$ 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 **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/m^2) 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 $MgCl_2$, 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 $MgCl_2$, 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, samples were prepared by boiling in Laemmli-SDS sample buffer. Unless indicated otherwise, all IP experiments were performed on the chromatin fraction.

Generation of mass spectrometry samples. For the generation of mass spectrometry samples the beads were washed 4 times with EBC-2 buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and protease inhibitor cocktail (Roche)) 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. Four biological repeats for each condition were performed.

Mass spectrometry. Mass spectrometry was performed essentially as previously described[51]. Samples were analyzed on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Germany) coupled to an EASY-nanoLC 1000 system (Proxeon, Odense, Denmark). 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. The mass spectrometer was operated in positive-ion mode at 2.9 kV with the capillary heated to 250°C. The mass spectrometer was operated in a Data-Dependent Acquisition (DDA) mode with a top 7 method. Full scan MS spectra were obtained with a resolution of 70,000, a target value of $3x10^6$ 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 $1x10^5$ and a normalized collision energy of 25%. The precursor ion masses selected for MS/MS analysis were subsequently dynamically excluded from MS/MS analysis for 60 sec. Precursor ions with a charge state of 1 and greater than 6 were excluded from triggering MS/MS events.

Mass spectrometry data analysis. Raw mass spectrometry files were analysed with MaxQuant software (v1.5.3.30 According to[52], with the following modifications from default settings: the maximum number of mis-cleveages by trypsin/p was set to 4, Label Free Quantification (LFQ) was enabled disabling the Fats 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 (v 1.5.5.3) computational platform[53]. 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 four 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.

Mass spectrometry data availability. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[54] partner repository with the dataset identifier PXD013572. For reviewing purposes, data can be downloaded using the following credentials: **Username:** reviewer15750@ebi.ac.uk, **Password:** J9ITSoH3

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 **table 6**) as indicated.

Chromatin tethering. U2OS 2–6-3 cells containing 200 copies of a LacO-containing cassette (Janicki et al., 2004) were co-transfected with lipofectamine 2000 (Invitrogen) and plasmid DNA for 6 h at 37 °C in an atmosphere of 5% CO_2. 24 h after transfection the cells were fixed with 4% paraformaldehyde (Sigma; 252549) 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 labeleing.

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.

Genome-wide XR-sequencing. XR-seq was performed as previously described[42, 46]. Briefly, cells were harvested 3h after treatment with 20J/m^2 UVC (254nm). 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 were 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 50nt 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[55] version 0.36. Reads were aligned to the genome using Bowtie[56]. 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 (http://broadinstitute.github.io/picard/). 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 slop and merge

commands[57]. All samples were converted to BED format using bedtools bamtobed command. Strand-specific profiles over the TSS were created using the R Bioconductor genomation package[58].

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 (Table 2). 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 27ºC in ESF 921 insect cell culture medium (Fisher Scientific), pelleted at 1,000xg for 15 min, frozen in liquid nitrogen, and stored at -80ºC. Protein purifications were performed at 4ºC. 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 4ºC 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) 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 -80ºC. 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 4ºC 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 -80ºC.

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 4ºC. The beads were washed 3x with 0.3 ml Pull-down Buffer (20 mM HEPES [pH 7.5], 100 mM KCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 0.25 mg/ml BSA, 0.03% Tween) and incubated with *Xenopus laevis* egg extract (HSS; high-speed supernatant) for 1 h at 4ºC. 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 xlCRL4^{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 $MgCl_2$, 0.6 mM DTT): 100 nM E1 (Enzo Life Sciences), 2.5 µM UBE2D2 (Boston Biochem), approximately 50 nM neddylated xlCRL4^{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.

Table 1: Cell lines

Cell lines	Origin
KPS3-hTERT	[16]
KPS3-hTERT + UVSSA	[16]
U2OS (FRT)	This study
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-CSBΔCIM-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) + GFP-UVSSAΔCIR-1	This study
U2OS (FRT) UVSSA-KO (1-8) + GFP-UVSSAΔTIR-6	This study
U2OS (FRT) UVSSA-KO (1-8) + GFP-UVSSA-3	This study
U2OS (FRT) UVSSA-KO (1-8) / CSA (2-4) + GFP-UVSSA-3	This study
U2OS (FRT) UVSSA-KO (1-8) / CSB-KO (1-12) + 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	[35]

Table 2: Plasmids

Plasmids	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$^{\Delta CIM}$	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$^{\Delta CIR}$	This study
pcDNA5/FRT/TO-Puro-GFP-UVSSA$^{\Delta TIR}$	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$^{\Delta TIR}$	This study
pmCherry-LacR-C1	[37]
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$^{\Delta N}$	This study
pmCherry-LacR-CSB$^{\Delta C}$	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$^{\Delta 1306-1300}$	This study
pmCherry-LacR-CSB$^{\Delta 1306-1352}$	This study
pmCherry-LacR-CSB$^{\Delta 1353-1399}$	This study
pmCherry-LacR-CSB$^{\Delta 1400-1428}$	This study
pmCherry-LacR-CSB$^{\Delta 1353-1368}$	This study
pmCherry-LacR-CSB$^{\Delta 1369-1384}$	This study
pmCherry-LacR-CSB$^{\Delta 1385-1399}$	This study
pmCherry-LacR-CSB^{WT}	This study
pmCherry-LacR-NLS	[59]
pmCherry-LacR-UVSSA$^{\Delta CIR}$	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$^{\Delta CIM}$	This study
pTM142_pAB1_FLAG-hsCSB^{WT}	This study
pTM143_pAB1_FLAG-hsCSB$^{\Delta CIM}$	This study

Table 3: Sequences of sgRNAs

sgRNAs		
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

Table 4: Sequencing primers to validate KO cell lines

Sequencing primers for knockouts		
CSB/ERCC6	5-GTAGGGGCCAGTTGTTAGAATGTAA-3	oML#078_sgML#003_CSB1_fw
	5-CTCACATTCTGAATGACTTGGCTA-3	oML#079_sgML#003_CSB1_rev
CSA/ERCC8	5-CAGTCTGTGTCCAGTTTCTGTG-3	oML#084_sgML#018_CSA_2FW
	5-CATATTTGTTATGTGTTTCTTTGAG-3	oML#085_sgML#018_CSA_2RV
	5-GTACATACATACATACACATTTACCAATAC-3	oML#100_sgML#018_CSA_2_Fw_Seq
	5-CTGAGAAAAAATGTACCTAAATATTAAG-3	oML#101_sgML#018_CSA_2_Rv_Seq
UVSSA	5-ACCCAGAGGTACACAGAGATTG-3	oML#090_sgML#019_UVSSA1_Fw
	5-GCTCTTAGAAGTGTCCCTGTG-3	oML#091_sgML#019_UVSSA1_Rv
	5-ATCAGGAGGCTGAGGCGGCTG-3	oML#076_sgML#020_UVSSA2_fw
	5-AGGAGCCTACCCGGGAGCCGGG-3	oML#077_sgML#020_UVSSA2_rev

Table 5: Primers

Primers		
CSB WT	TTAAGTCGACCCAAATGAGGGAATCCCCCAC	oML#375
	AATTGCGGCCGCTTAGCAGTATTCTGGCTTGAGTTTC	oML#376
CSA WT	CACAATGCTAGCGCCACCATGCTGGGGTTTTTGTCCG	oML#041
	GCATGGTGAACTACCGGTGCTCCTTCTTCATCACTGCTG	oML#042
UVSSA WT	ACAATTGAATTCGATGGATCAGAAACTTTCGAAG	oML#035
	GTGTAAAGATCTCTAGTTCAGTGCGTAGTTAAAC	oML#036
CSBΔC	TCCAGCCTCGAGGTCCAAATGAGGGAATCCCCCACTC	oML#173
	TCAGGTCGGATCCTTATCGAGTTCCTTCAAACTTGGCGTCTC	oML#174
CSB-N	TCCAGCCTCGAGGTCCAAATGAGGGAATCCCCCACTC	oML#173
	GCATCAGGTCGGATCCTTAATCTCCATCATCTCGGTATCTTCCCAC	oML#178
CSBΔN	TCCAGCCTCGAGGTGATGGAGATGAAGATTATTATAAGCAGCGG	oML#175
	GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCCAAATTC	oML#176
CSB-M	TCCAGCCTCGAGGTGATGGAGATGAAGATTATTATAAGCAGCGG	oML#175
	TCAGGTCGGATCCTTATCGAGTTCCTTCAAACTTGGCGTCTC	oML#174
CSB-C	TCCAGCCTCGAGGTCGAATTCCACACCTGGTGAAGAAAAG	oML#177
	GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCCAAATTC	oML#176
CSB 1221-1305	TCCAGCCTCGAGGTCGAATTCCACACCTGGTGAAGAAAAG	oML#177
	GATGGAGGATCCTTACAGACACCGCTGACGAGAGAG	oML#196
CSB 1306-1399	TACAGCCTCGAGGTGGAGCAGTGTCTGGTGTTCCC	oML#197
	GGCGATGGAGGATCCTTACAGGTGGTTTCTAGCTCTCATTTTAGC	oML#198
CSB 1400-1493	TCCAGCCTCGAGGTATTCTGCCAGAGCGTTTAGAAAGTGAAAG	oML#199
	GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCCAAATTC	oML#176
CSB 1417-1493	TACATCCTCGAGGTGCCCTGCTGCCCACCACAG	oML#200
	GCATCAGGTCGGATCCTTAGCAGTATTCTGGCTTGAGTTTCCAAATTC	oML#176
CSBΔ1306-1352	GCTCTCTCGTCAGCGGTGTCTGTGCCAGGATGGCATCATGAA	oML#226
	CCTTTTTCATGATGCCATCCTGGCACAGACACCGCTGACGAG	oML#227
CSBΔ1353-1399	CCTTCATCAACATCTCCAACAGAGAAGATTCTGCCAGAGCGTTTAG	oML#232
	CACTTTCTAAACGCTCTGGCAGAATCTTCTCTGTTGGAGATGTTG	oML#233
CSBΔ1306-1399	GAGGCTCTCTCGTCAGCGGTGTCTGATTCTGCCAGAGCGTTTAGAAAGTG	oML#224
	CTTTCACTTTCTAAACGCTCTGGCAGAATCAGACACCGCTGACGAGAG	oML#225
CSBΔ1400-1428	GCTAAAATGAGAGCTAGAAACCACCTGGTGGAGATGAGAAACTTCATC	oML#234
	GAAAGCGATGAAGTTTCTCATCTCCACCAGGTGGTTTCTAGCTCTC	oML#235
CSBΔ1353-1368	CCTTCATCAACATCTCCAACAGAGAAGCATTTTAGTGGAAGAGCAGAAG	oML#262
	CTGCATCTTCTGCTCTTCCACTAAAATGCTTCTCTGTTGGAGATGTTGA	oML#263
CSBΔ1369-1384	GAAAAAGGAGGGAAAAGATAATGTCCCTGAGGCTTCCTCCTCACTCTTG	oML#264
	CATTTTAGCCAAGAGTGAGGAGGAAGCCTCAGGGACATTATCTTTTCC	oML#265
CSBΔ1385-1399	AGACTCTTCATCCGGGCCCCTCATTCTGCCAGAGCGTTTAGA	oML#266
	CTTTCACTTTCTAAACGCTCTGGCAGAATGAGGGGCCCGGATGA	oML#267
UVSSA Δ100-200	CACAGACCCCGCACAGCCTCTGAGGCTGCTGGTGCCTTTTG	oML#128
	CAAAGTCAAAAGGCACCAGCAGCCTCAGAGGCTGTGCGGGG	oML#129
UVSSA Δ400-500	GGACAGAAGCCCTGGGGGATGCGGTGGTGCCCTACGGCGTG	oML#138
LacR-C3	ATTAAAACGCGTCAGTGGGCTGATC	oML#377
	TAATAATAGATCTGAAACCTTCCTCTTCTTCTTAG	oML#378

Table 6: Antibodies

Antibodies	Host		Clone	WB	
Cas9	Mouse	Cell Signaling technology, #14697	7A9 and 3A3	1/5000	aML#031
CSA/ERCC8	Mouse	Santa Cruz, sc-376981	D2	1/500	aML#025
CSA/ERCC8	Rabbit	Abcam, 137033	EPR9237	1/750	aML#028
CSB/ERCC6	Goat	Santa Cruz, SC-10459	E-18	1/1000	aML#039
DDB1	Goat	Abcam, ab9194		1/1000	aML#035
ERCC1	Mouse	Santa Cruz, sc-17809	D10	1/300	aML#066
FLAG	Rabbit	New England Peptide; antigen: C(dPEG4)DYKDDDDK		1/5000	
GFP	Mouse	Roche, #11814460001	7.1 and 13.1	1/1000	aML#011
GFP	Rabbit	Abcam, ab290		1/1000	aML#044
Goat IgG (H+L) CF680	Donkey	Thermo fisher Scientific, A21084		1/10000	aML#037
Mouse IgG (H+L) CF770	Goat	Biotium, VWR #20077		1/10000	aML#009
p44/ GTF2H2	Mouse	kindly provided by J.M. Egly	1H5	1/2000	aML#075
p62/GTF2H1	Mouse	kindly provided by J.M. Egly	3C9	1/2000	aML#074
p62/GTF2H1	Mouse	Santa Cruz, sc-48431	G10	1/500	aML#099
p80/XPD/ ERCC2	Mouse	Abcam, ab54676		1/500	aML#029
p89	Mouse	Millipore, MABE1123	15TF2-1B3	1/2000	aML#101
p89/XPB/ERCC3	Mouse	kindly provided by J.M. Egly	1B3	1/1000	aML#073
p89/XPB/ERCC3	Rabbit	Santa Cruz, SC-293	S-19	1/1000	aML#040
rabbit IgG (H+L) CF680	Goat	Biotium, VWR #20067		1/10000	aML#010
RNAPII-S2	Rabbit	Abcam, ab5095		1/1000	aML#024
Tubulin	Mouse	Sigma, T6199	DM1A	1/1000	aML#008
UVSSA	Mouse	Genetex, GTX629742	GT816	1/500	aML#100
UVSSA	Rabbit	Novus Biologicals, NBP1-32598		1/1000	aML#030
UVSSA	Rabbit	Abcam ab137644		1/1000	aML#034
UVSSA	Rabbit	Genetex, GTX106751		1/1000	aML#087
xlCSA	Rabbit	New England Peptide; antigen: CHRTHINPAFEDAWSSSEDES		1/5000	
XPA	Rabbit	kindly provided by Rick Wood	CJ1	1/10000	aML#079
XPC	Rabbit	Novus Biologicals, NB100-58801		1/2000	aML#077
XPF/ ERCC4	Mouse	Santa Cruz, sc-136153	3F2/3	1/200	aML#096
XPG/ ERCC5	Rabbit	Novus Biologicals, NB100-74611		1/1000	aML#046

Author Contributions

YvdW generated knockout cells, constructs and stable cell-lines, performed LacR-based tethering assays, clonogenic survivals, PCR and Western blot analysis to validate knockouts, Co-IP experiments for Western blot analysis, Co-IP experiments for mass spectrometry, and wrote the paper. KA generated stable cell-lines, Western blot analysis to validate knockouts, and Co-IP experiments. RG-P and ACOV analyzed the mass spectrometry samples. HG performed XR-seq. HG, EH and SA analyzed the XR-seq samples. TETM generated recombinant CSB proteins and xlCRL4^{CSA}, and performed pull-down and *in vitro* ubiquitylation assays. DvdH generated knockout cells, constructs, and performed Western blot analysis to validate knockouts, and Co-IP experiments. JCW supervised TETM. MSL supervised the project and wrote the paper.

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 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.

Supplemental Tables

Supplementary Table 1. List of UVSSA-interacting proteins identified by mass spectrometry. UVSSA-KO cells complemented with GFP-UVSSA^{WT} were subjected to immunoprecipitation using GFP Trap beads or block beads (BB) in triplicate. Following trypsin digestion and desalting, eluted peptides were analyzed on a Q-Exactive Orbitrap mass spectrometer. Raw MS files were analyzed with the MaxQuant software suite. The difference, significance, and number of unique peptides are indicated. Hits with a $\log^2$ difference above 1 are considered significantly enriched.

Supplementary Table 2. List of UV-induced UVSSA-interacting proteins identified by mass spectrometry. UVSSA-KO cells complemented with GFP-UVSSA^{WT} were mock-treated or UV irradiated (20 J/m^2) and subjected to immunoprecipitation using GFP Trap beads in triplicate. Following trypsin digestion and desalting, eluted peptides were analyzed on a Q-Exactive Orbitrap mass spectrometer. Raw MS files were analyzed with the MaxQuant software suite. The difference, significance, and number of unique peptides are indicated. Hits with a $\log^2$ difference above 1 are considered significantly enriched.

Supplementary Table 3. List of UV-induced UVSSA-interacting proteins identified by mass spectrometry. Raw data file of the mass spectrometry samples shown in Supplementary Table 1 and 2.

References

1. 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).
2. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* **15**, 465-481 (2014).
3. 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).
4. Tapias, A. *et al.* Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J Biol Chem* **279**, 19074-19083 (2004).
5. 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 U S A* **89**, 3664-3668 (1992).
6. 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).
7. van 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).
8. Spivak, G. *et al.* Ultraviolet-sensitive syndrome cells are defective in transcription-coupled repair of cyclobutane pyrimidine dimers. *DNA Repair (Amst)* **1**, 629-643 (2002).
9. Nance, M.A. & Berry, S.A. Cockayne syndrome: review of 140 cases. *American journal of medical genetics* **42**, 68-84 (1992).
10. Friedberg, E.C. Cockayne syndrome--a primary defect in DNA repair, transcription, both or neither? *Bioessays* **18**, 731-738 (1996).
11. 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 U S A* **101**, 15410-15415 (2004).
12. 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 U S A* **106**, 6209-6214 (2009).
13. 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).
14. Laugel, V. *et al.* Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. *Human mutation* **31**, 113-126 (2010).
15. Calmels, N. *et al.* Functional and clinical relevance of novel mutations in a large cohort of patients with Cockayne syndrome. *Journal of medical genetics* **55**, 329-343 (2018).
16. Nakazawa, Y. *et al.* Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair. *Nat Genet* **44**, 586-592 (2012).
17. 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).
18. 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).
19. van 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).
20. 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).
21. van den Boom, V. *et al.* DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J Cell Biol* **166**, 27-36 (2004).
22. Svejstrup, J.Q. Rescue of arrested RNA polymerase II complexes. *J Cell Sci* **116**, 447-451 (2003).
23. 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).

24. 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).
25. Schwertman, P. *et al.* UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat Genet* **44**, 598-602 (2012).
26. 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).
27. Fischer, E.S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024-1039 (2011).
28. 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 & development* **20**, 1429-1434 (2006).
29. 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).
30. Wienholz, F. *et al.* FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res* (2019).
31. 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).
32. 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).
33. Schilbach, S. *et al.* Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature* **551**, 204-209 (2017).
34. Jaspers, N.G. *et al.* Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. *DNA Repair (Amst)* **1**, 1027-1038 (2002).
35. Janicki, S.M. *et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683-698 (2004).
36. 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).
37. Luijsterburg, M.S. *et al.* A PALB2-interacting domain in RNF168 couples homologous recombination to DNA break-induced chromatin ubiquitylation. *eLife* **6** (2017).
38. 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).
39. 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).
40. 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).
41. Volker, M. *et al.* Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* **8**, 213-224 (2001).
42. 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 & development* **29**, 948-960 (2015).
43. Kamiuchi, S. *et al.* Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proc Natl Acad Sci U S A* **99**, 201-206 (2002).
44. 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).
45. Panier, S. *et al.* Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. *Mol Cell* **47**, 383-395 (2012).
46. 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 U S A* **113**, E2124-2133 (2016).
47. D'Errico, M. *et al.* The role of CSA in the response to oxidative DNA damage in human cells. *Oncogene* **26**, 4336-4343 (2007).
48. Menoni, H. *et al.* The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage. *Nucleic Acids Res* **46**, 7747-7756 (2018).

49. Xu, J. *et al.* Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653-657 (2017).
50. He, Y. *et al.* Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359-365 (2016).
51. Kumar, R., Gonzalez-Prieto, R., Xiao, Z., Verlaan-de Vries, M. & Vertegaal, A.C.O. The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. *Nat Commun* **8**, 1809 (2017).
52. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature protocols* **11**, 2301-2319 (2016).
53. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature methods* **13**, 731-740 (2016).
54. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450 (2019).
55. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
56. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25 (2009).
57. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
58. 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).
59. Soutoglou, E. & Misteli, T. Activation of the cellular DNA damage response in the absence of DNA lesions. *Science* **320**, 1507-1510 (2008).

Figure 1

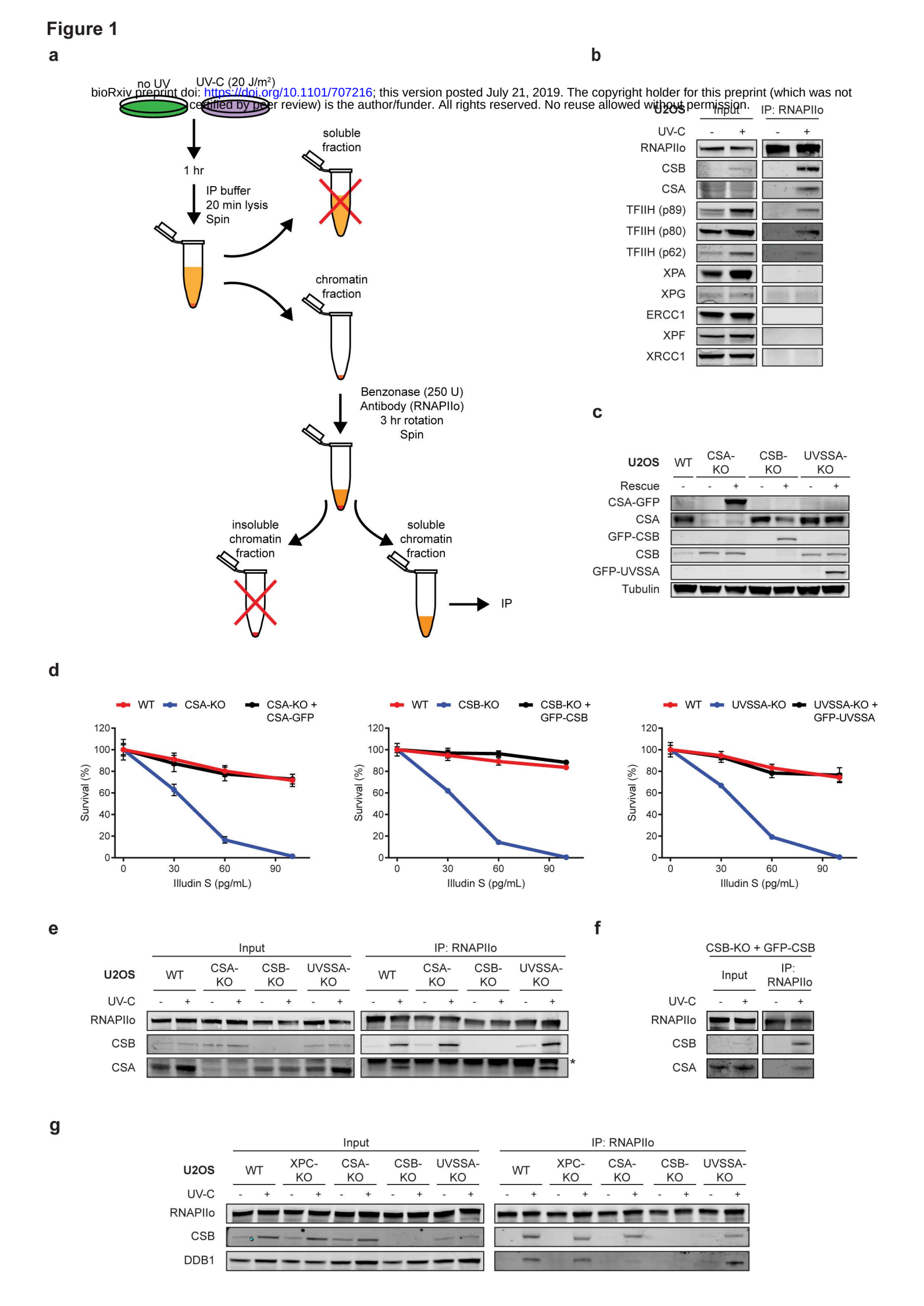


Figure 1. CSA is recruited to DNA damage-stalled RNAPIIo by CSB. (a) Outline of a new IP method to isolate RNAPIIo and associated proteins on mock-treated or UV-irradiated (20J/m^2) U2OS (FRT) cells. (b) Endogenous RNAPII Co-IP on WT cells (see also Supplementary Fig 1a). (b) Western blot analysis of CSA, CSB, and UVSSA knockout cells complemented with inducible GFP-tagged versions of these proteins. See Supplementary Figure 2 for validation of knockouts by sequencing. (d) Clonogenic Illudin S survival of WT, CSA, CSB, and UVSSA knockout and rescue cell lines. Data represent mean ± SEM of two independent experiments. Endogenous RNAPII Co-IP on (e) WT, CSA, CSB, and UVSSA knockout cells, (f) CSB-KO stably expressing GFP-CSB, and (g) WT, XPC, CSA, CSB, and UVSSA knockout cells. The asterisk in panel e indicates the heavy chain of the RNAPII antibody.

Figure 2

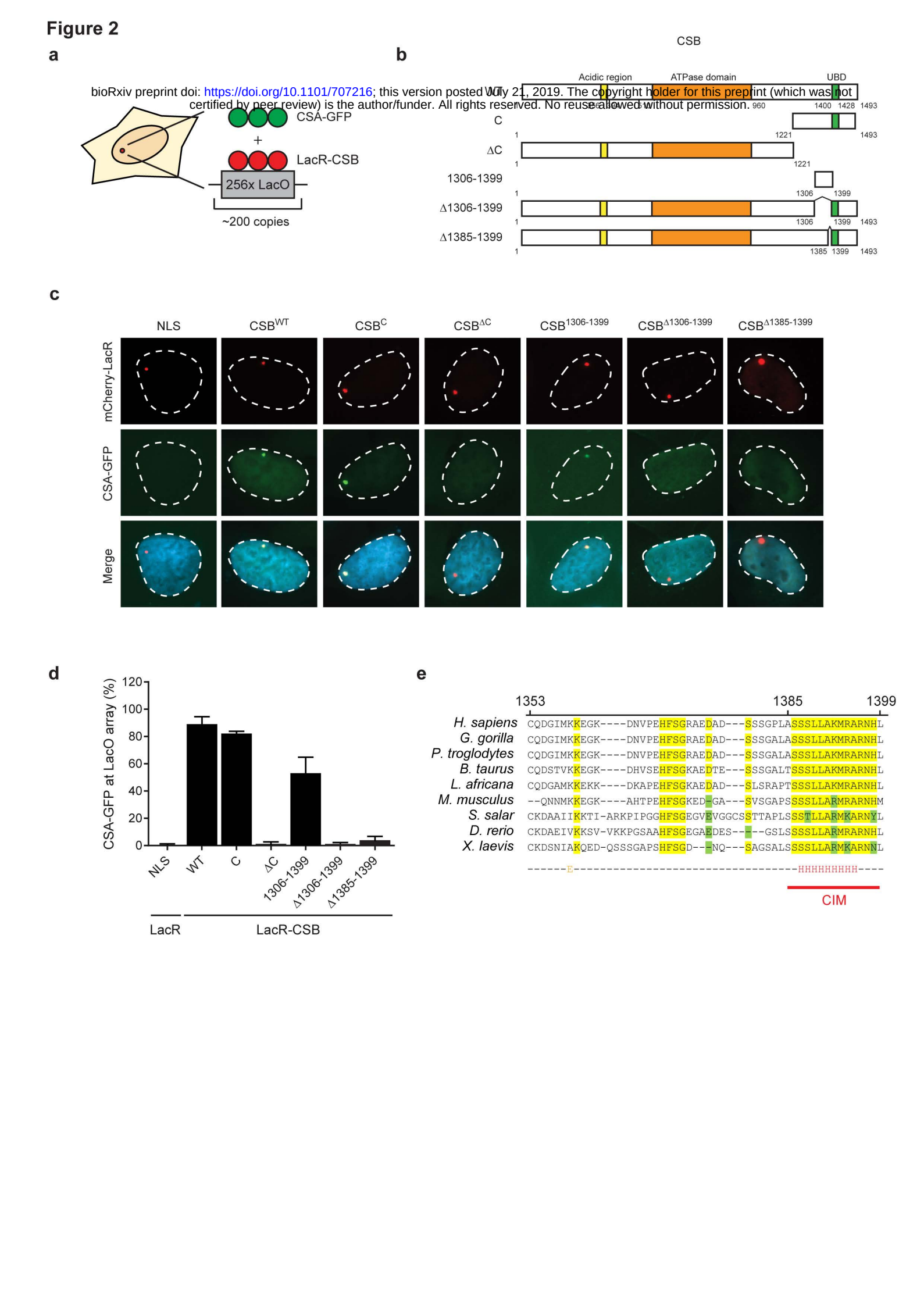


Figure 2. CSA interacts with the newly identified C-terminal CIM of CSB. (a) Outline of the chromatin-tethering approach in U2OS 2-6-3 cells. (b) A schematic representation of CSB and its deletion mutants. (c) Recruitment of CSA-GFP to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins. (d) Quantification of CSA-GFP and mCherry-LacR-CSB co-localization at the LacO array. Values represent the mean ± SD of >50 cells collected in two independent experiments. (e) Sequence alignment of CSB orthologues. See Supplementary Figures 3 and 4 for additional mutants, and Supplementary Figure 5 for additional alignments.

Figure 3

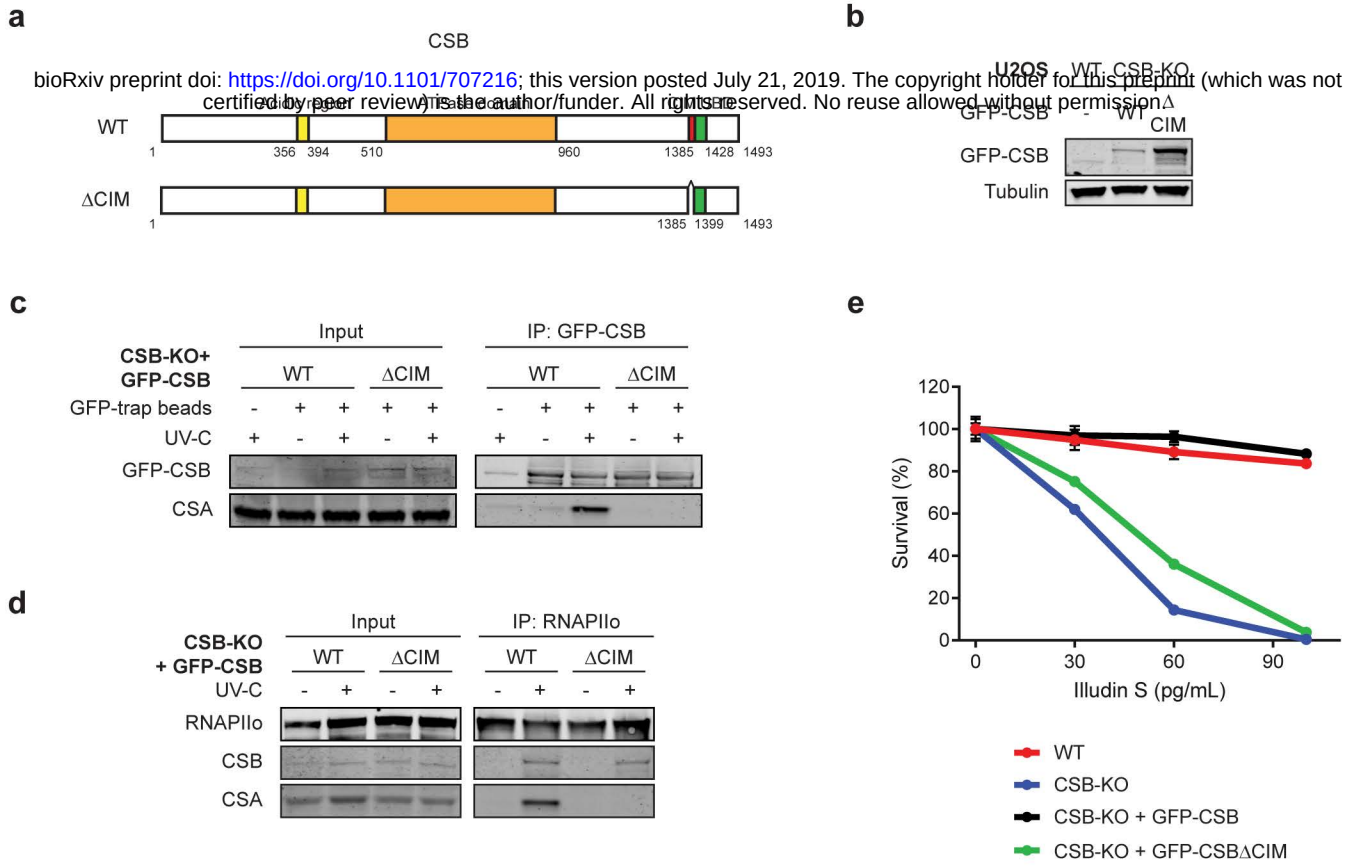


Figure 3. The CIM of CSB mediates the recruitment of CSA to damage-stalled RNAPIIo. (a) A schematic representation of CSB and the CSB$^{\Delta CIM}$ mutant. (b) Western blot analysis of U2OS (FRT) and CSB-KO complemented with either GFP-CSB^{WT} or GFP-CSB$^{\Delta CIM}$. (c) Co-IP of GFP-CSB^{WT} and GFP-CSB$^{\Delta CIM}$ on the combined soluble and chromatin fraction. (d) Endogenous RNAPII Co-IP in GFP-CSB^{WT} and GFP-CSB$^{\Delta CIM}$ cell lines. See also Supplementary Figure 7a for additional Co-IP data. (e) Clonogenic Illudin S survival of WT and CSB-KO cell lines and the GFP-tagged CSB rescue cell lines. Data represent mean ± SEM of two independent experiments. Note that the same survival data for WT, CSB-KO and CSB-KO + GFP-CSB is also shown in Fig. 1d. (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* ubiquitination reactions were stopped and blotted with anti-FLAG (top three panels) or anti-xlCSA (bottom panel) antibodies See also Supplementary Fig 6. (g) Immobilized recombinant CSB variants were incubated with *Xenopus laevis* egg extract (HSS), recovered, and blotted with anti-FLAG (top panel) or anti-xlCSA (bottom panel) antibody.

Figure 4

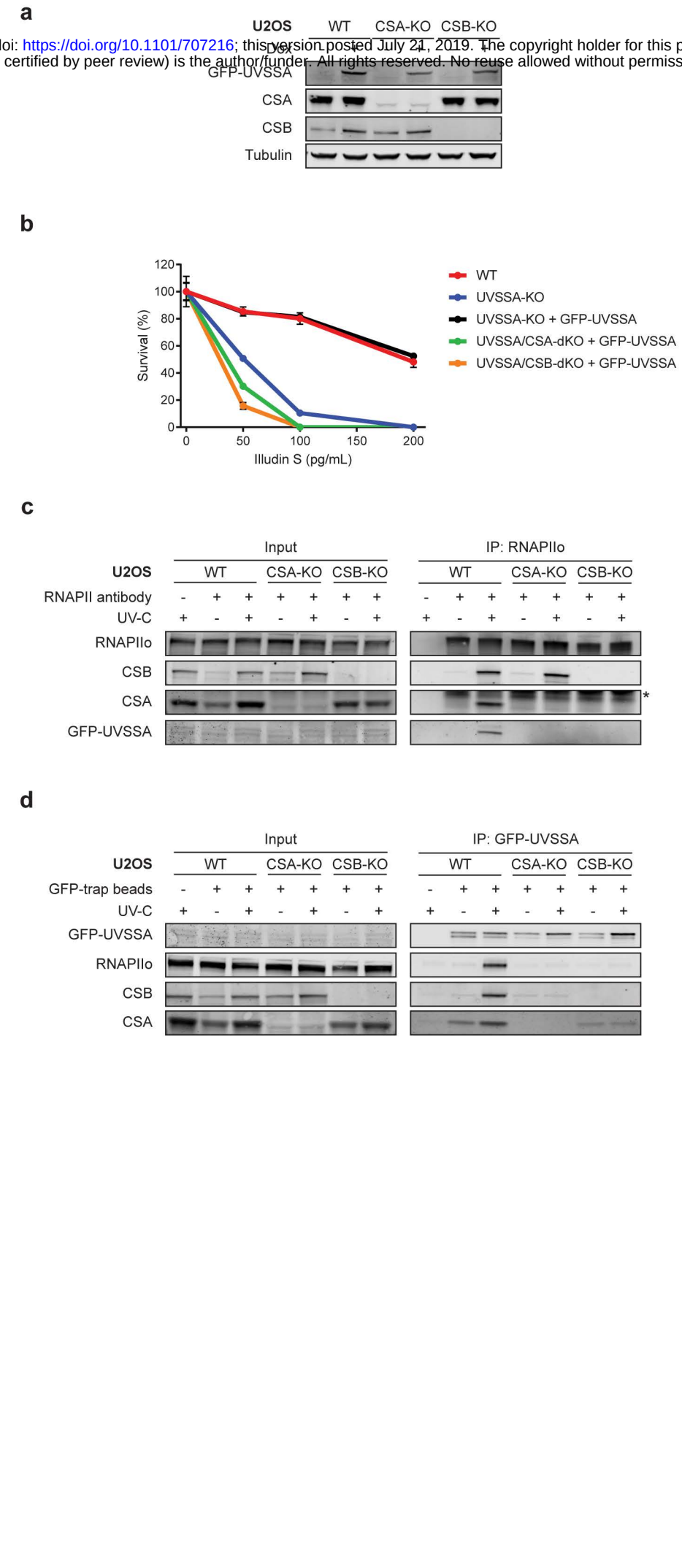


Figure 4. UVSSA is recruited to DNA damage-stalled RNAPIIo by CSA. (a) Western blot analysis of UVSSA-KO, UVSSA/CSA-dKO, and UVSSA/CSB-dKO complemented with GFP-UVSSA. (b) Clonogenic Illudin S survival of WT, UVSSA-KO, UVSSA/CSA-dKO, and UVSSA/CSB-dKO cell lines complemented with GFP-UVSSA. Data represent mean ± SEM of two independent experiments. (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 panel c indicates the heavy chain of the RNAPII antibody.

Figure 5

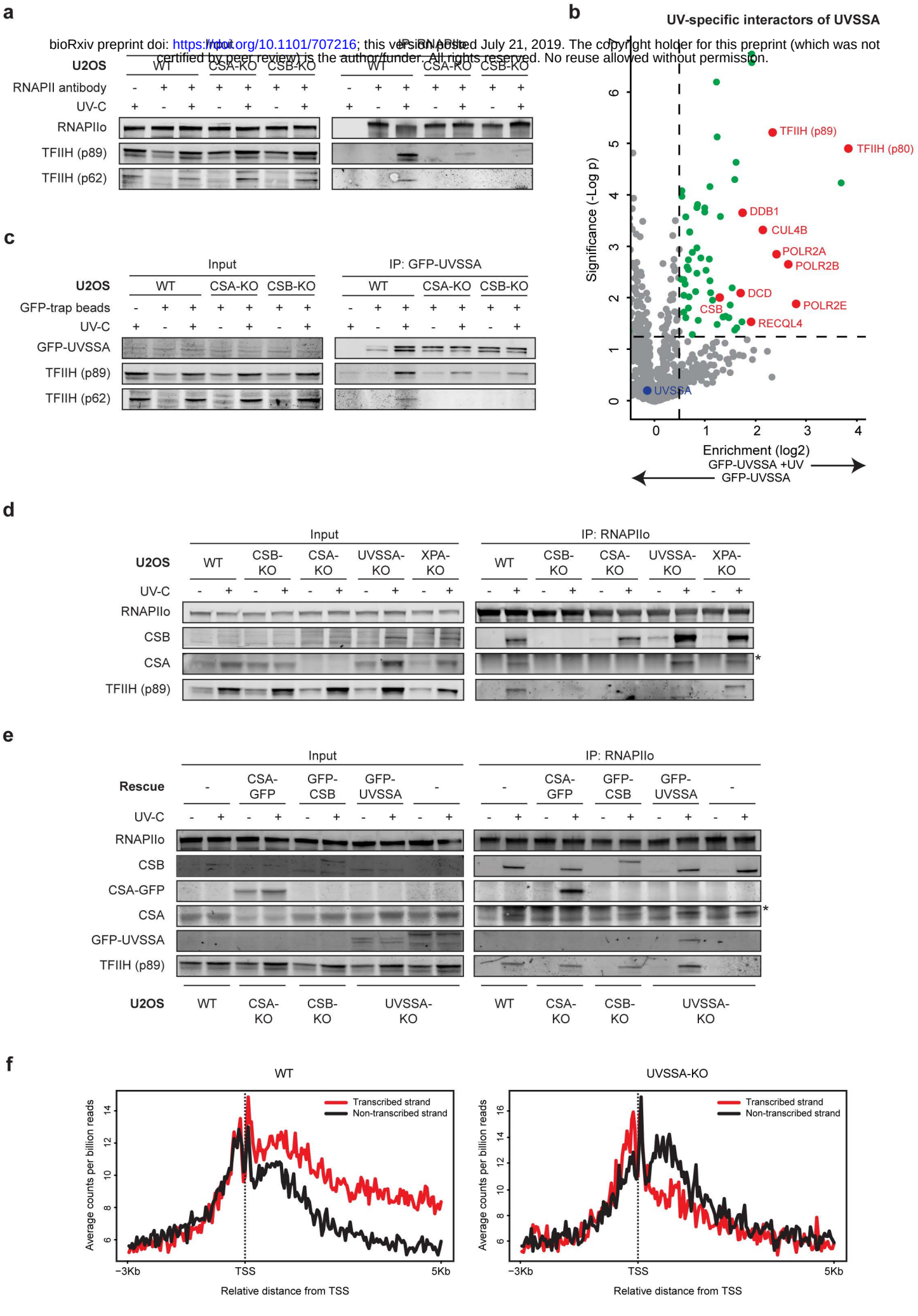


Figure 5. CSA, CSB, and UVSSA are equally important for TFIIH recruitment. (a) Endogenous RNAPII Co-IP in UVSSA-KO, UVSSA/CSA-dKO, and UVSSA/CSB-dKO complemented with GFP-UVSSA. (b) Volcano plot depicting the statistical differences of the MS analysis on GFP-UVSSA pull-down in mock-treated and UV-irradiated samples. The enrichment (log^2) is plotted on the x-axis and the significance (t-test $-log^2$ p-value) is plotted on the y-axis. All significantly UV-induced hits are indicated in green. Several selected hits are shown in red (See also Supplementary Fig 7b, c and Supplementary table 1-3 for additional MS data analysis). (c) Co-IP of GFP-UVSSA in UVSSA-KO and UVSSA-dKO cells complemented with GFP-UVSSA. (d) Endogenous RNAPII Co-IP in WT, CSB-KO, CSA-KO, UVSSA-KO and XPA-KO cells. (e) Endogenous RNAPII Co-IP in WT and UVSSA-KO cells and CSA-KO, CSB-KO, and UVSSA-KO cells complemented with GFP-tagged versions of these proteins. The asterisk in panels d and e indicates the heavy chain of the RNAPII antibody. (f) CPD XR-seq repair signal 3 Kb upstream and 5 Kb downstream of the annotated TSS of 16.088 genes in WT and UVSSA-KO cells. Signal is plotted separately for the transcribed (red) and non-transcribed (black) strands. The data represent the average of two independent experiments with a bin size of 40 nt. See also Supplementary Figure 8a for additional XR-seq data.

Figure 6

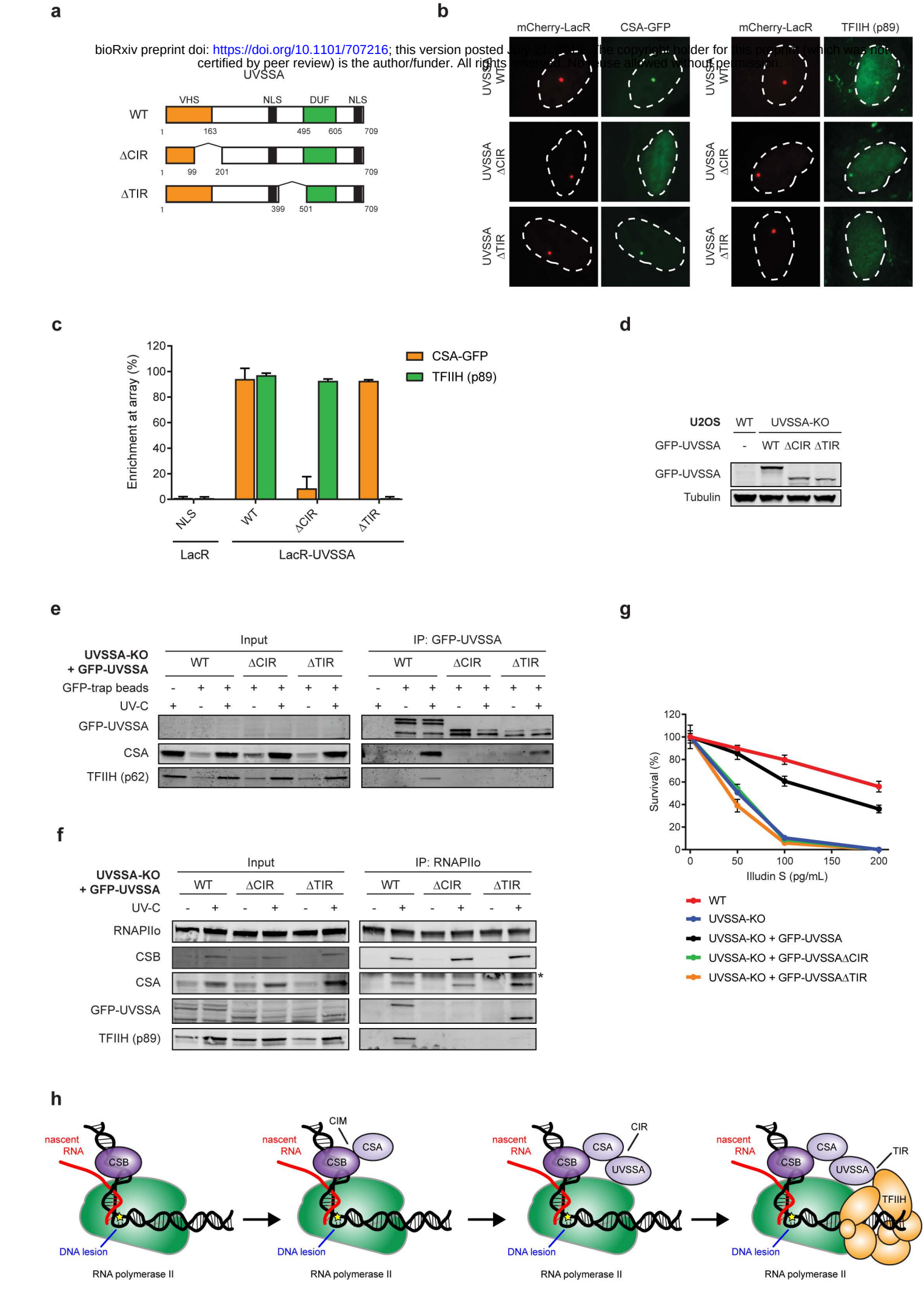
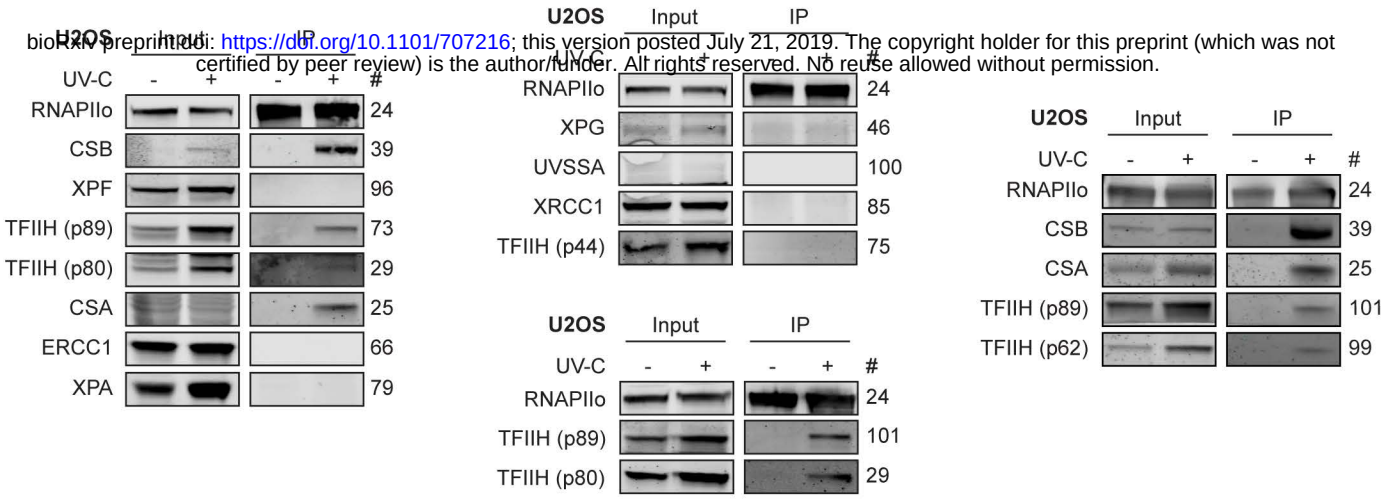


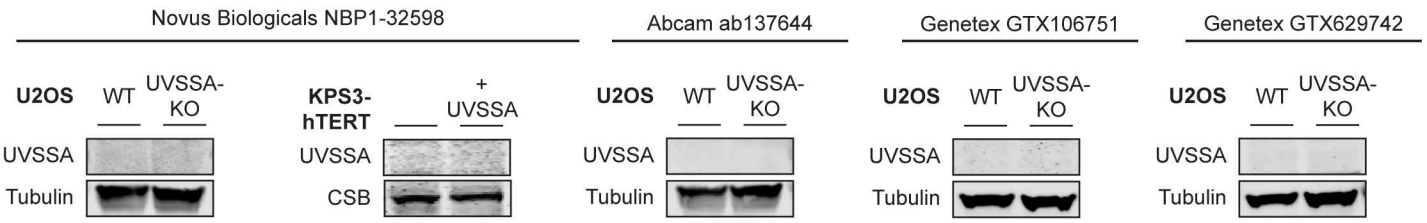
Figure 6. UVSSA is the key protein that recruits TFIIH. (a) A schematic representation of UVSSA WT and deletion mutants. The CSA-interacting region (CIR) and TFIIH-interacting region (TIR) are indicated. (b) Recruitment of CSA-GFP and TFIIH (p89) to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins. (c) Quantification of CSA-GFP and endogenous TFIIH (p89) co-localization at the LacO array. Values represent the mean ± SD of >50 cells collected in two independent experiments. (d) Western blot analysis of U2OS (FRT) and UVSSA-KO cells complemented with GFP-UVSSA^{WT}, GFP-UVSSA$^{\Delta CIR}$, and GFP-UVSSA$^{\Delta TIR}$. (e) Co-IP of GFP-UVSSA^{WT}, GFP-UVSSA$^{\Delta CIR}$, and GFP-UVSSA$^{\Delta TIR}$. (f) Endogenous RNAPII Co-IP in GFP-UVSSA^{WT}, GFP-UVSSA$^{\Delta CIR}$, and GFP-UVSSA$^{\Delta TIR}$ cell lines. See also Supplementary Figure 8b, c for additional Co-IP data. (g) Clonogenic Illudin S survival of WT and UVSSA-KO cell lines and the GFP-tagged UVSSA rescue cell lines. Data represent mean ± SEM of two independent experiments. (h) Model of how the assembly of CSB, CSA, and UVSSA targets the TFIIH complex to DNA damage-stalled RNAPIIo.

Figure S1

a



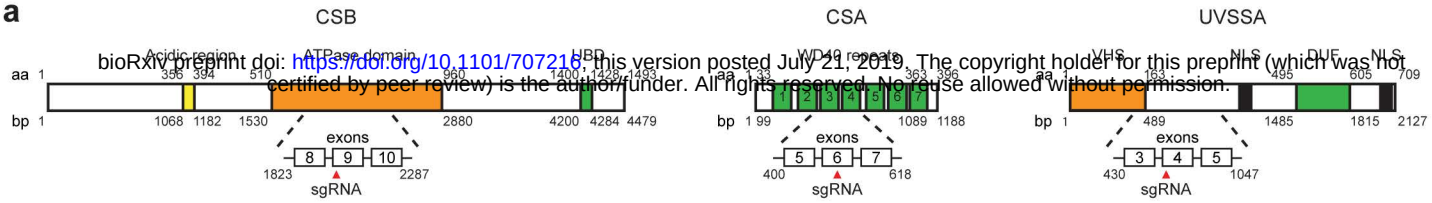
b



Supplementary Figure 1. Testing of antibodies in IP and whole cell lysates. (a) Various examples of endogenous RNAPII Co-IP experiments in WT cells. (b) Testing of various UVSSA antibodies in KPS3-hTERT, KPS3-hTERT + UVSSA, U2OS (FRT) WT, and U2OS (FRT) UVSSA-KO cells.

Figure S2

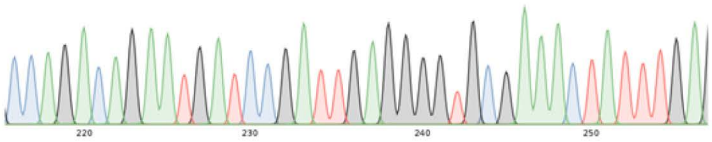
a



b

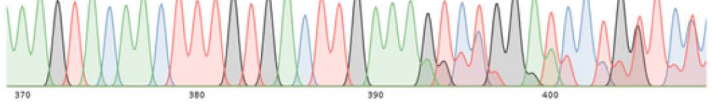
U2OS (FRT) CSB-KO Homozygote: 1 nucleotide insertion

```
1 ACTTTCCCGGGAAGATGAAGTCAAAGAGCGACCACAGCTCTCGGAGGTTATTTTGCATCGGTGAGCCAGACAGAATGATCCGA-TGAGGGGTGCGAAACT  99
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
1 ACTTTCCCGGGAAGATGAAGTCAAAGAGCGACCACAGCTCTCGGAGGTTATTTTGCATCGGTGAGCCAGACAGAATGATCCGATTGAGGGGTGCGAAACT 100
```



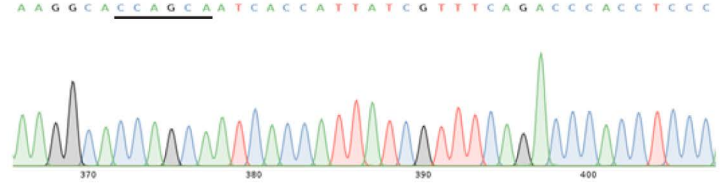
U2OS (FRT) CSA-KO Heterozygote: 1 nucleotide insertion and 2 nucleotide insertion

```
1 AGTTGGTACTAGAGGACCCAAAGTACAACTTTGTGACTTG--AAGTCTGGATCCTGTTCTCACATTCTACAGGGTATTTTTATTTTATTTCAAACGGCAA  98
1 AGTTGGTACTAGAGGACCCAAAGTACAACTTTGTGACTTG-AAAGTCTGGATCCTGTTCTCACATTCTACAGGGTATTTTTATTTTATTTCAAACGGCAA  99
1 AGTTGGTACTAGAGGACCCAAAGTACAACTTTGTGACTTGAAAAGTCTGGATCCTGTTCTCACATTCTACAGGGTATTTTTATTTTATTTCAAACGGCAA 100
```



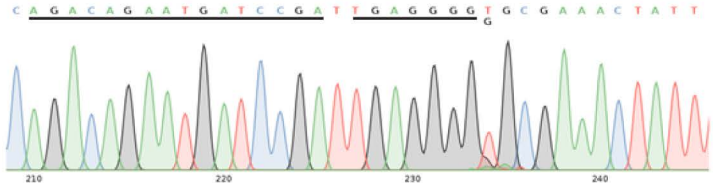
U2OS (FRT) UVSSA-KO Homozygote: 225 nucleotide insertion (in frame with multiple stop codons)

```
  1 GATTCCGTCTCCGGGTTCGGGTCAAAGTCAAAAGGCACCAGCA----------------------------------------------------------  43
    |||||||||||||||||||||||||||||||||||||||||||
  1 GATTCCGTCTCCGGGTTCGGGTCAAAGTCAAAAGGCACCAGCAATCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGG 100

 43 ----------------------------------------------------------------------------------------------------  43

101 GTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACAAAAGGCCGGCGGCCACGAAAAA 200

 43 ----------------------------------------------------------GCCTAAAGCAGCTCTCTACCTCCGTCAAGCAG  75
                                                                ||||||||||||||||||||||||||||||||
201 GGCCGGCCAGGCAAAAAAGAAAAAGGAAACGCGAATTTTAACAAAATATTAACGTTTACAATTTTATGGCCTAAAGCAGCTCTCTACCTCCGTCAAGCAG 300
```



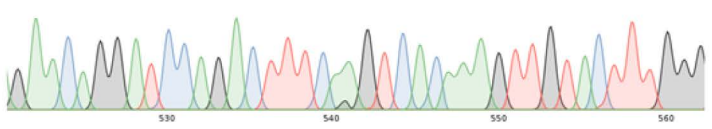
U2OS (FRT) UVSSA-KO/CSB-KO +GFP-UVSSA Homozygote: 1 nucleotide insertion

```
1 GCGACCACAGCTCTCGGAGGTTATTTTGCATCGGTGAGCCAGACAGAATGATCCGA-TGAGGGGTGCGAAACTATTTGAGGAAAGGAAGCACCTTTTTAT  99
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 GCGACCACAGCTCTCGGAGGTTATTTTGCATCGGTGAGCCAGACAGAATGATCCGATTGAGGGGTGCGAAACTATTTGAGGAAAGGAAGCACCTTTTTAT 100
```

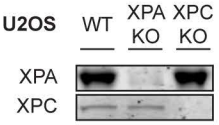


U2OS (FRT) UVSSA-KO/CSA-KO +GFP-UVSSA Homozygote: 1 nucleotide insertion

```
1 GTTGCCGTTTGAAATAAAATAAAAATACCCTGTAGAATGTGAGAACAGGATCCAGAC-TTCAAGTCACAAAGTTGTACTTTGGGTCCTCTAGTACCAACT  99
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 GTTGCCGTTTGAAATAAAATAAAAATACCCTGTAGAATGTGAGAACAGGATCCAGACTTTCAAGTCACAAAGTTGTACTTTGGGTCCTCTAGTACCAACT 100
```



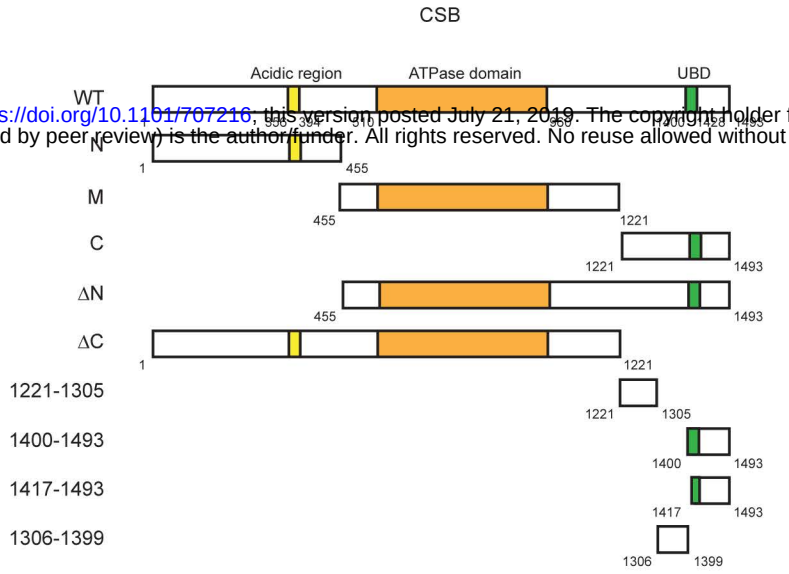
c



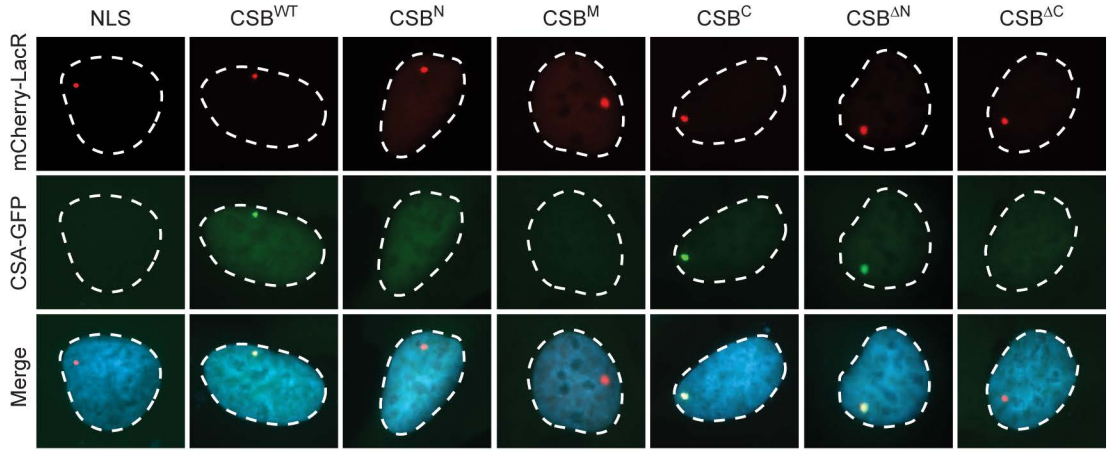
Supplementary Figure 2. **Sequence of CSB, CSA, and UVSSA-KO cells**. (a) A schematic representation of CSB, CSA, and UVSSA including the location of the guide RNAs used for the generation of the CRISPR/Cas9-mediated KO. (b) Sequences of CSB, CSA, and UVSSA knockouts. (c) Western blot analysis of XPC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

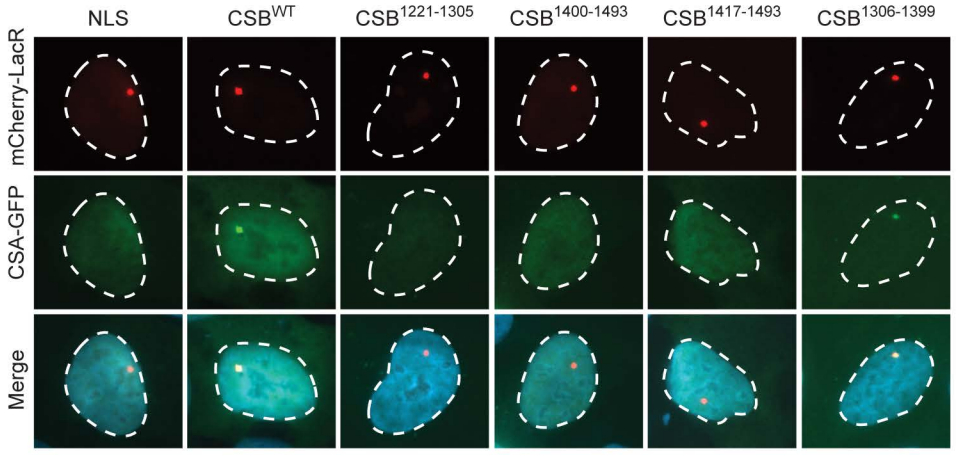
Figure S3

a

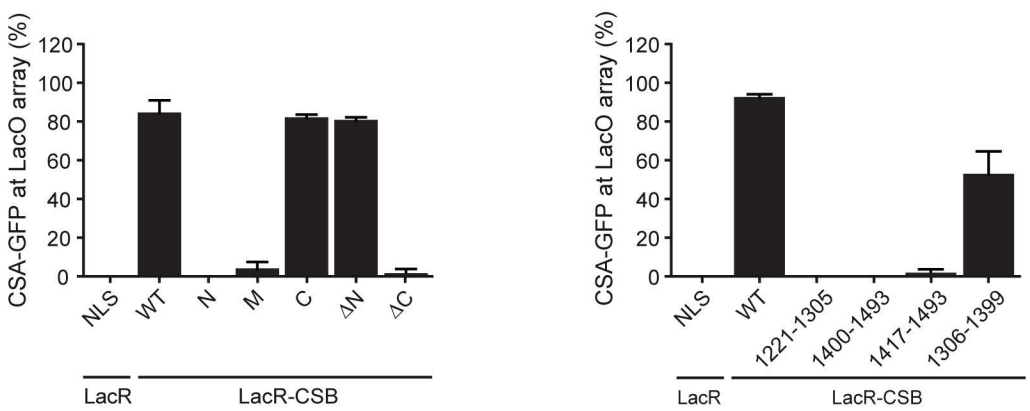


b





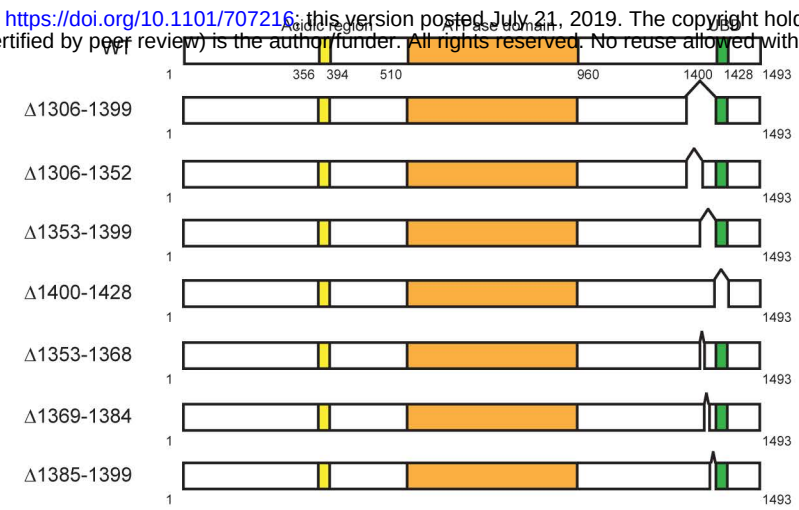
c



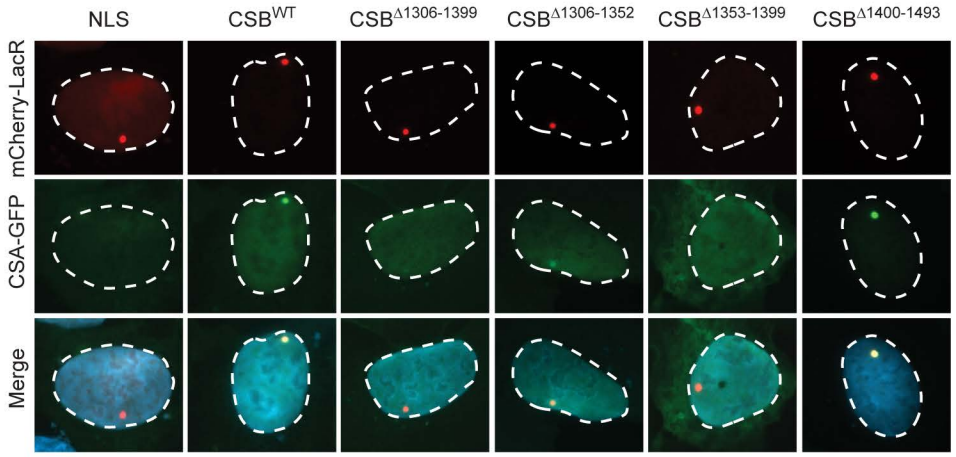
Supplementary Figure 3. CSA interacts with the C-terminal region of CSB. (a) A schematic representation of CSB and its deletion mutants. (b) Recruitment of CSA-GFP to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins. (c) Quantification of CSA-GFP and mCherry-LacR-CSB co-localization at the LacO array. Values represent the mean ± SD of >50 cells collected in two independent experiments.

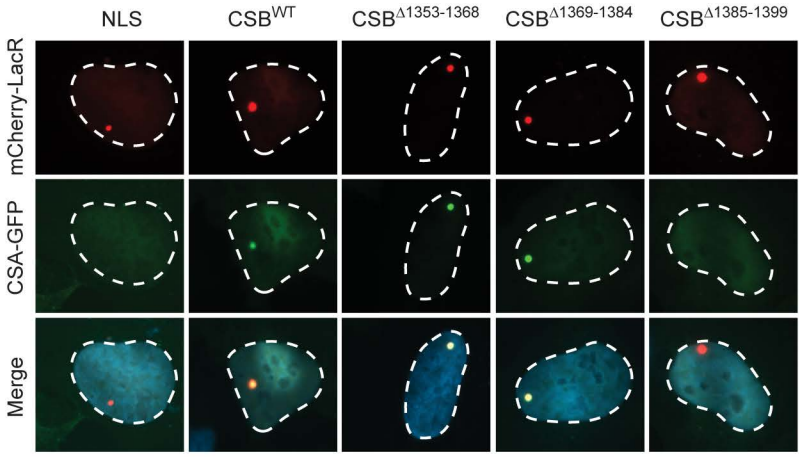
Figure S4

a

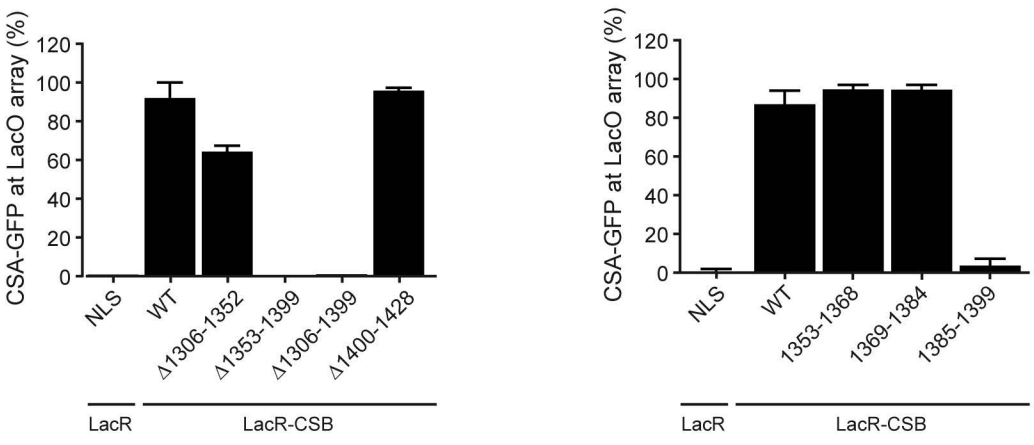


b





c



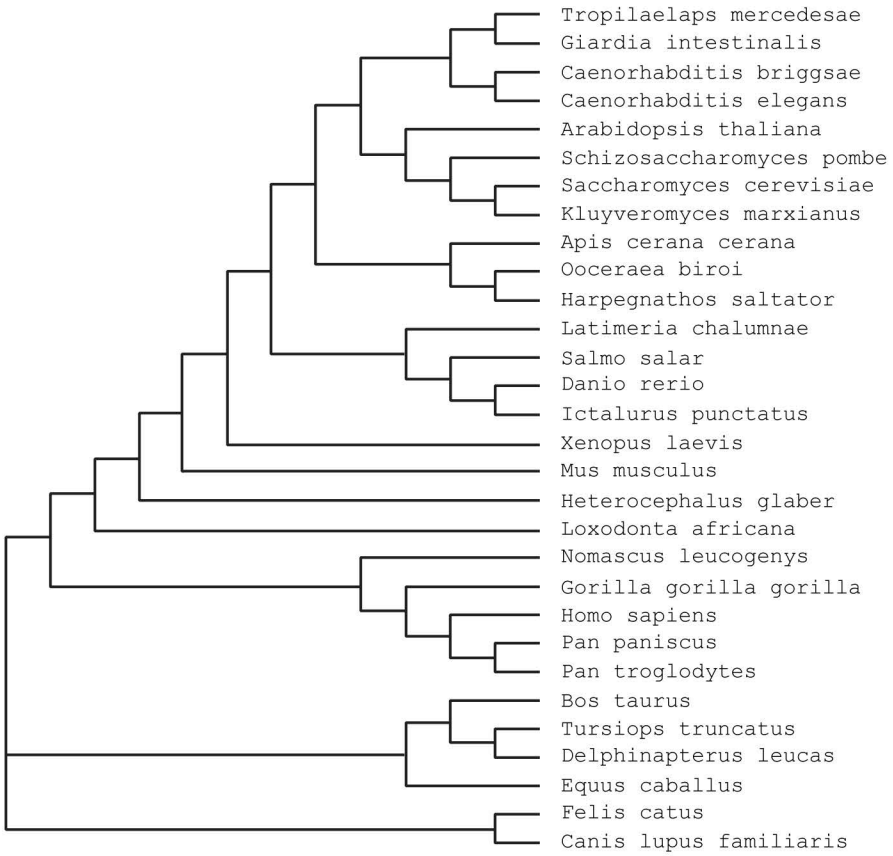
Supplementary Figure 4. CSA interacts with amino acids 1385-1399 of CSB. (a) A schematic representation of CSB and its deletion mutants. (b) Recruitment of CSA-GFP to the LacO array upon tethering of the indicated mCherry-LacR fusion proteins. (c) Quantification of CSA-GFP and mCherry-LacR-CSB co-localization at the LacO array. Values represent the mean ± SD of >50 cells collected in two independent experiments.

Figure S5

a

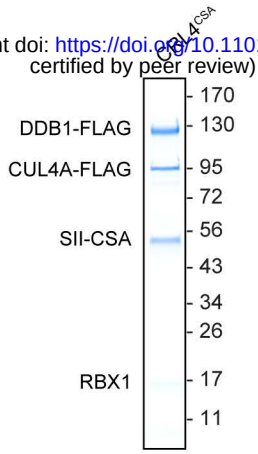
1493	sp|Q03468|ERCC6_HUMAN	ASSSLLAKMRARNHL	Homo sapiens (Human)
1491	tr|H2Q1W1|H2Q1W1_PANTR	ASSSLLAKMRARNHL	Pan troglodytes (Chimpanzee)
1493	tr|G3QVF5|G3QVF5_GORGO	ASSSLLAKMRARNHL	Gorilla gorilla gorilla (Western lowland gorilla)
1492	tr|G1S127|G1S127_NOMLE	ASSSLLAKMRARNHL	Nomascus leucogenys (Northern white-cheeked gibbon)
1491	tr|A0A2R8ZA95|A0A2R8ZA95_PANPA	ASSSLLAKMRARNHL	Pan paniscus Bonobo)
1482	tr|E1BFL2|E1BFL2_BOVIN	TSSSLLAKMRARNHL	Bos taurus (Bovine)
1484	tr|M3XEB4|M3XEB4_FELCA	PSSSLLAKMRARNHL	Felis catus (Cat)
1486	tr|E2QSK6|E2QSK6_CANLF	PSSSLLAKMRARNHL	Canis lupus familiaris (Dog)
1460	tr|F7D5S6|F7D5S6_HORSE	TSSSLLAKMRARNHL	Equus caballus (Horse)
1432	tr|G3TCV9|G3TCV9_LOXAF	TSSSLLAKMRARNHL	Loxodonta africana (African elephant)
1481	tr|A0A0P6J577|A0A0P6J577_HETGA	TSSSLLARMRARNHL	Heterocephalus glaber (Naked mole rat)
1481	tr|F8VPZ5|F8VPZ5_MOUSE	SSSSLLARMRARNHM	Mus musculus (Mouse)
1485	tr|A0A2Y9MEF2|A0A2Y9MEF2_DELLE	TSSSLLAKMRARNHL	Delphinapterus leucas (Beluga whale)
1462	tr|A0A2U4AKE1|A0A2U4AKE1_TURTR	TSSSLLAKMRARNHL	Tursiops truncatus (Atlantic bottle-nosed dolphin)
1488	tr|A0A1S3N477|A0A1S3N477_SALSA	SSSTLLARMKARNYL	Salmo salar (Atlantic salmon)
1409	tr|A0A2D0QRA6|A0A2D0QRA6_ICTPU	SSSSLLARMRARNHV	Ictalurus punctatus (Channel catfish)
1389	tr|F1R294|F1R294_DANRE	SSSSLLARMRARNHL	Danio rerio (Zebrafish)
1370	tr|A0A1L8FKT9|A0A1L8FKT9_XENLA	SSSSLLARMKARNNL	Xenopus laevis (African clawed frog)
1387	tr|H3AWF0|H3AWF0_LATCH	SSSSLLANMRARNHL	Latimeria chalumnae (West Indian ocean coelacanth)
1187	sp|Q9ZV43|CHR8_ARATH	SSAELLNRIRGSREQ	Arabidopsis thaliana (Mouse-ear cress)
1015	tr|A8XNA8|A8XNA8_CAEBR	---------------	Caenorhabditis briggsae (nematode)
957	tr|Q93781|Q93781_CAEEL	---------------	Caenorhabditis elegans (nematode)
973	sp|Q9UR24|RHP26_SCHPO	---TLLARLKQRR--	Schizosaccharomyces pombe (Fission yeast)
1085	sp|P40352|RAD26_YEAST	NYDDGIT-FA--RSK	Saccharomyces cerevisiae (Baker's yeast)
1037	tr|W0T437|W0T437_KLUMD	LKVKTLPSQE--KKK	Kluyveromyces marxianus (Yeast)
925	tr|Q6WD94|Q6WD94_GIAIN	---------------	Giardia intestinalis (intestinal parasite)
1222	tr|A0A1V9XZ12|A0A1V9XZ12_9ACAR	SPSRPKGKRRSVAVL	Tropilaelaps mercedesae (bee mite)
1125	tr|A0A2A3EJZ4|A0A2A3EJZ4_APICC	---------------	Apis cerana cerana (Oriental honeybee)
1073	tr|A0A026W8Z2|A0A026W8Z2_OOCBI	---------------	Ooceraea biroi (Clonal raider ant)
1005	tr|E2BDE2|E2BDE2_HARSA	---------------	Harpegnathos saltator (Jerdon's jumping ant)

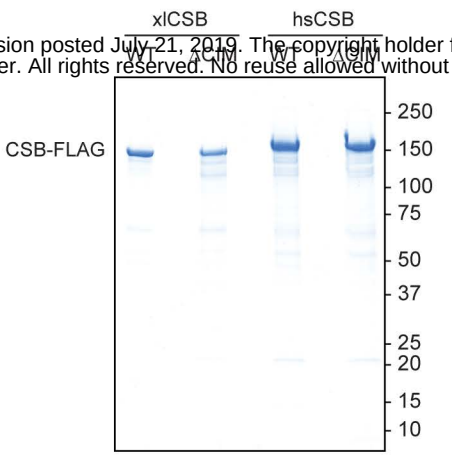
b



Supplementary Figure 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.

Figure S6

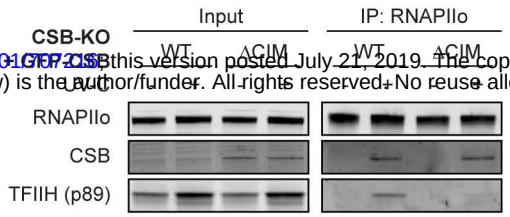




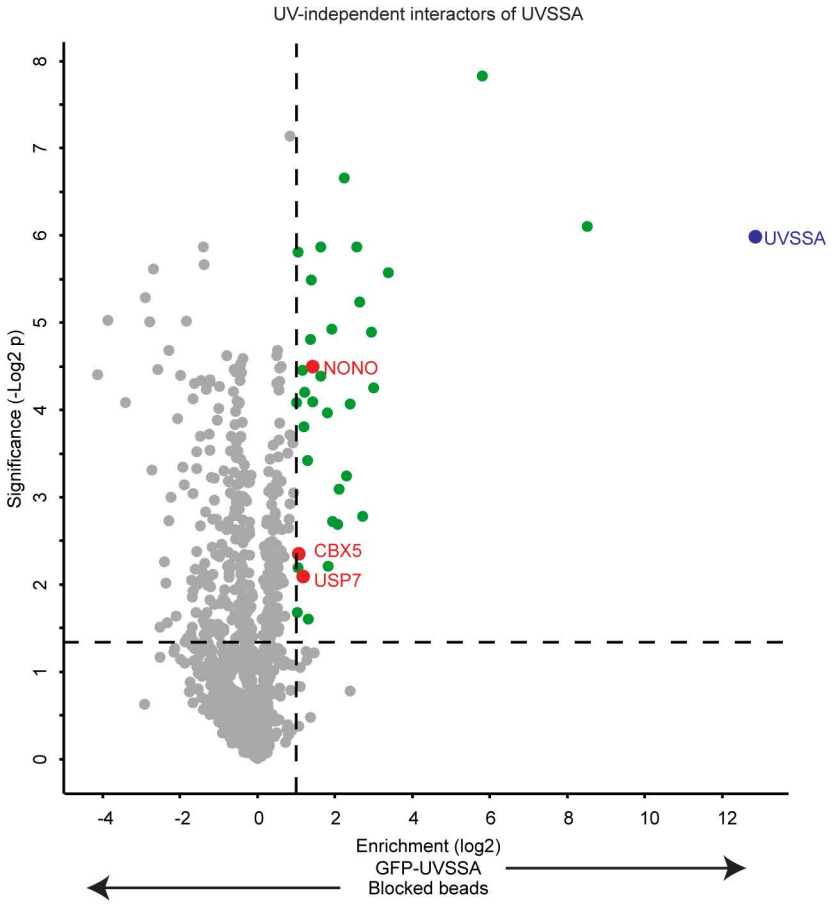
Supplementary Figure 6. Purified CRL4^{CSA} and CSB protein. Coomassie gels of recombinant xlCRL4^{CSA} complex and xlCSB or hsCSB variants. DDB1, CUL4A, and all CSB proteins contained an N-terminal FLAG-tag, and CSA contained a C-terminal Strep-tag II.

Figure S7

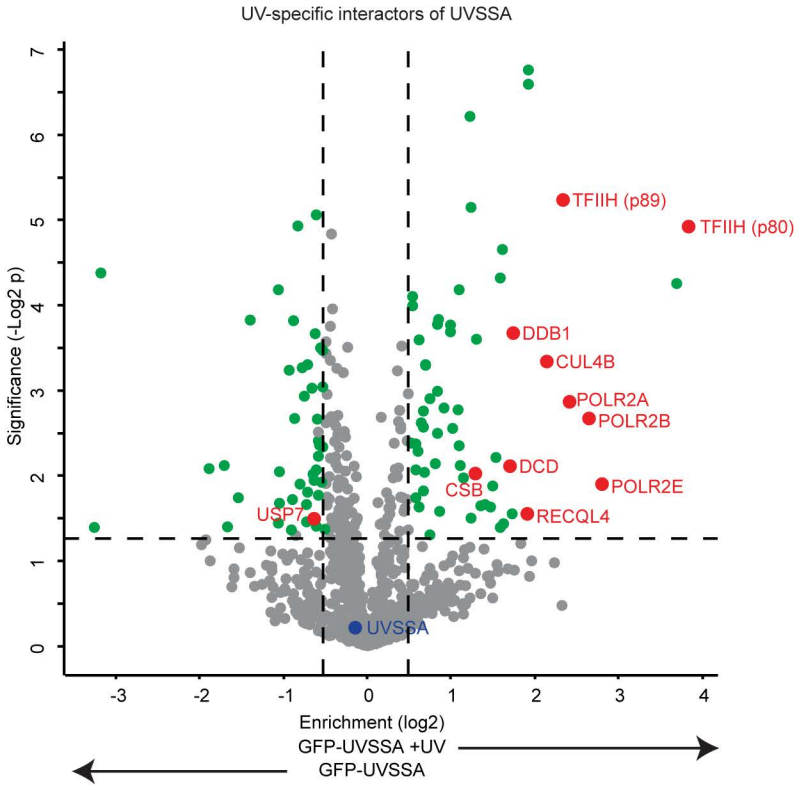
a



b



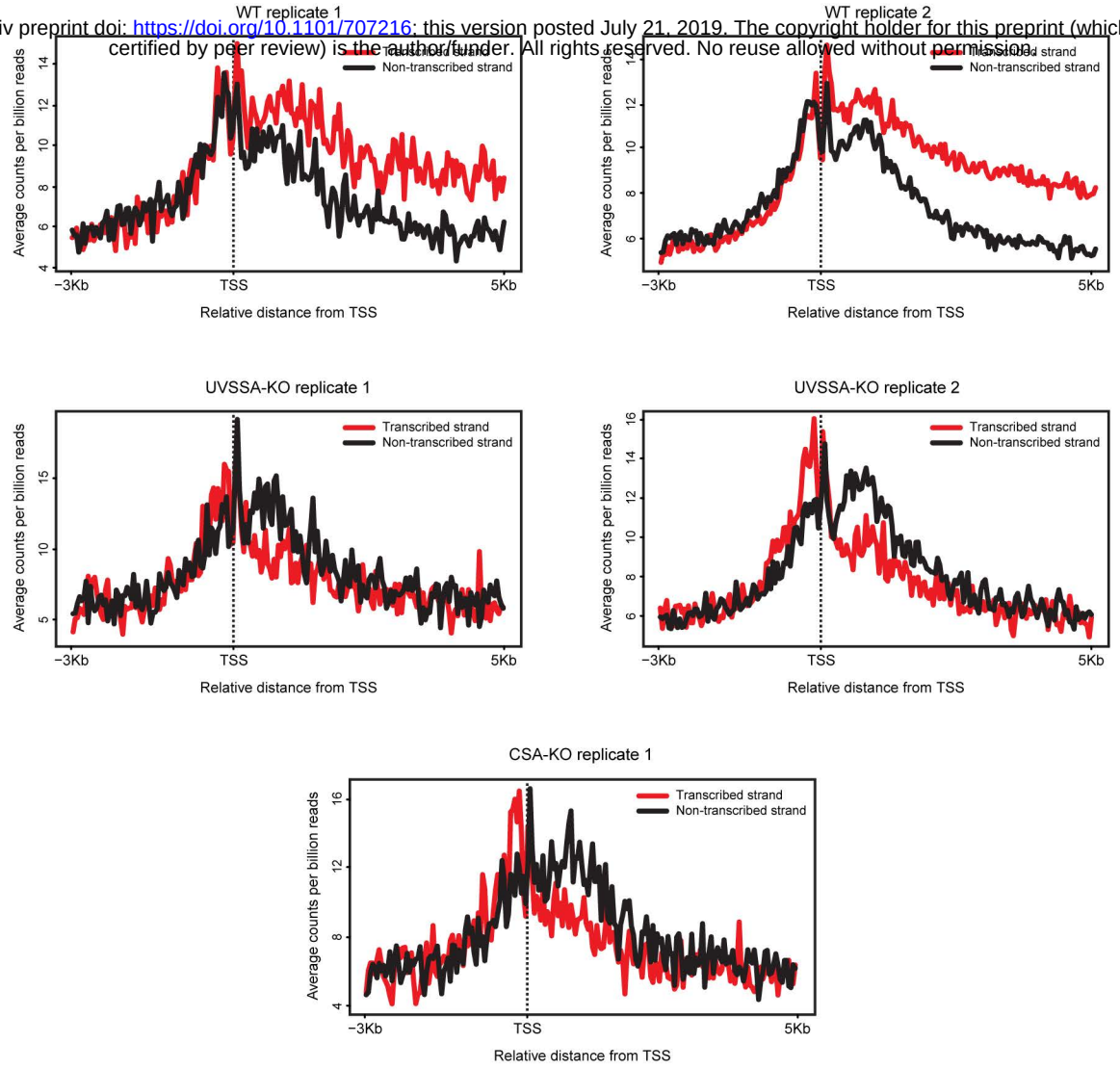
c



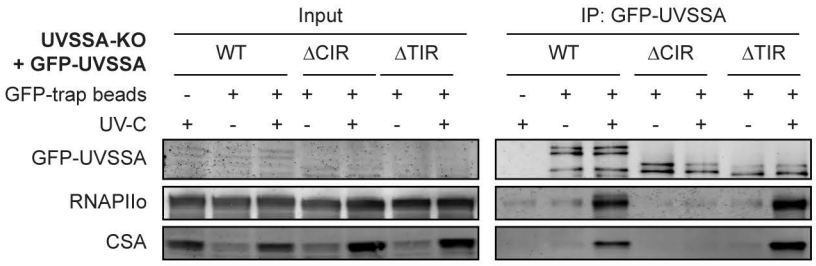
Supplementary Figure 7. Mass spectrometry after GFP-UVSSA pull-down. (a) Endogenous RNAPII Co-IP in CSB-KO + GFP-CSB^{WT} and CSB-KO + GFP-CSB$^{\Delta CIM}$ (b-c) Volcano plots depicting mass spectrometry analysis comparing (b) GFP-UVSSA pull-down versus block-beads control in mock-treated cells samples, and (c) GFP-UVSSA pull-down in mock-treated versus UV-irradiated (20 J/m^2) cells samples. The enrichment (log^2) is plotted on the x-axis and the significance ($-log^2$ p-value) is plotted on the y-axis. The $-log^2$ p-value threshold was set to 1.3 ($p<0.05$). The enrichment threshold was set to 1 in GFP-UVSSA versus blocked beads and 0.5 in UV-treated GFP-UVSSA vs GFP-UVSSA. All significantly significant hits are shown in green. Several selected hits are shown in red.

Figure S8

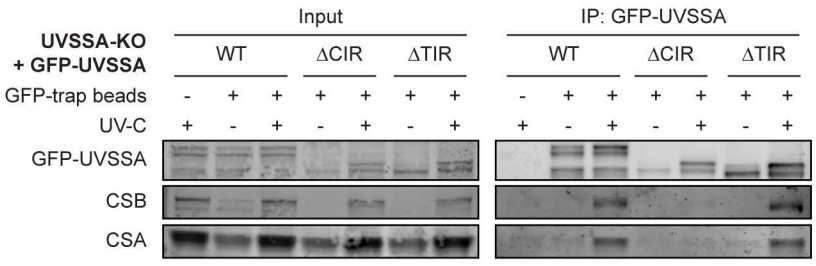
a



b



c



Supplementary Figure 8. 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$^{\Delta CIR}$, and GFP-UVSSA$^{\Delta TIR}$.