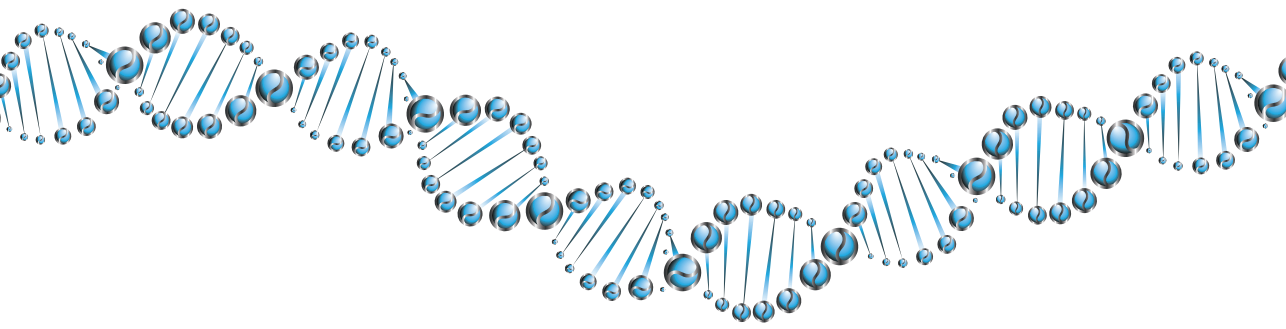
Unravelling molecular mechanisms in transcription-coupled nucleotide excision repair

Weegen, Y. van der

4

The identification of candidate TCR proteins and potential CRL4^{CSA} substrates using a proximity -dependent labeling approach

Yana van der Weegen, Noud H.M. Klaassen, Román González-Prieto, Alfred C.O. Vertegaal, and Martijn S. Luijsterburg



Abstract

Transcription-blocking DNA lesions are efficiently repaired by transcription-coupled nucleotide excision repair (TCR). Ubiquitylation of TCR proteins by the DDB1-CUL4A-RBX1-CSA ($CRL4^{CSA}$) ubiquitin ligase complex plays a key role in regulating TCR. However, the full repertoire of relevant substrates of the $CRL4^{CSA}$ complex remain the be identified. In this study, we captured DNA damage-induced CSA-proximal proteins using a proximity-dependent biotin identification (BioID) method. These proteins included factors that are known to interact through direct protein-protein contacts with CSA, such as the NER proteins UVSSA and TFIIH. In addition, we identified CDYL and HTATSF1 as new putative CSA-associated proteins, but whether they are involved in TCR and/or substrates of the $CRL4^{CSA}$ complex remains to be elucidated.

Introduction

Transcription-coupled nucleotide excision repair (TCR) removes transcription-blocking DNA lesions from the genome [1]. TCR is initiated when elongating RNA polymerase II (RNAPIIo) is unable to progress past bulky helix-distorting lesions, resulting in the stalling of RNAPIIo and the subsequent recruitment of the TCR proteins, starting with Cockayne syndrome protein B (CSB) [2, 3]. CSB, in turn, mediates the recruitment of CSA, via a CSA-interaction motif (CIM) located in the C-terminus of CSB, after which CSA recruits the UV-stimulated scaffold protein A (UVSSA) to DNA damage-stalled RNAPIIo. The sequential and cooperative assembly of the TCR complex leads to the association of the Transcription factor IIH (TFIIH) complex to initiate repair [2].

The CSA protein contains a seven-bladed WD40 propeller and functions as the substrate-recognition subunit of a DDB1-CUL4A-RBX1 ubiquitin ligase complex ($CRL4^{CSA}$) by directly interacting with DDB1 through a helix-loop-helix motif (aa 1-29; Fig. 1a, b) [4, 5]. The COP9 signalosome (CSN) regulates the ubiquitin ligase activity of CRL complexes by removal of the inhibitory ubiquitin-like NEDD8 modifier from cullins. Recruitment of the $CRL4^{CSA}$ complex to lesion-stalled RNAPIIo causes the CSN complex to dissociate from $CRL4^{CSA}$, rendering the $CRL4^{CSA}$ complex active [5, 6]. The DNA damage-induced activation of $CRL4^{CSA}$ results in CSA auto-ubiquitylation, as well as ubiquitylation of CSB, RNAPIIo, and UVSSA [5–9]. It is therefore not surprising that mutations in CSA that disrupt its binding to DDB1 causes defects in TCR [10, 11]. Interestingly, the CIM in CSB is located next to a ubiquitin-binding domain (UBD) [2], it is therefore likely that the UBD binds to auto-ubiquitylated CSA thereby stabilizing the interaction between CSB and CSA. Recent studies demonstrated that ubiquitylation of RNAPII and UVSSA are particularly important for the assembly of the TCR complex. Ubiquitylation of RNAPII (RPB1-K1268) stimulates the association of ubiquitylated UVSSA (UVSSA-K414) with RNAPII and, in turn, UVSSA-K414 ubiquitylation is required for the transfer of TFIIH onto lesion-stalled RNAPIIo [7]. Importantly, we have recently identified ELOF1 as a novel TCR factor that interacts with $CRL4^{CSA}$ through direct protein-protein contacts and facilitates $CRL4^{CSA}$-depedent RNAPII ubiquitylation [12]. Moreover, the deubiquitylating enzyme (DUB) ubiquitin-specific peptidase 7 (USP7) interacts with UVSSA, thereby protecting CSB, CSA, RNAPII, and UVSSA from proteasomal degradation [9, 12, 13].

Although these recent findings identify a central role for ubiquitylation during TCR complex assembly, the full repertoire of relevant substrates of the $CRL4^{CSA}$ complex remain the be identified. In this study, we established a proximity-dependent biotin identification (BioID) method to identify novel CSA-interacting proteins, including transient interactors such as potential substrates of the $CRL4^{CSA}$ complex.

Results

Establishing a proximity-dependent biotin identification (BioID) method

The $CRL4^{CSA}$ ubiquitin ligase complex targets proteins for ubiquitylation in response to DNA damage. However, not many $CRL4^{CSA}$ substrates have been identified since identification of substrates of an E3 ubiquitin ligase complex can be challenging due to the transient nature of the interaction between the substrate and the ubiquitin ligase complex. Regular immunoprecipitation approaches rely on relatively stable protein-protein interactions; therefore, substrates of ubiquitin ligase complexes may often go undetected using these methods. To circumvent this problem, we set out to establish a proximity-dependent biotin identification (BioID) method utilizing *CSA* knockout (KO) cells stably expressing inducible versions of CSA fused to a promiscuous mutant of the *E.coli* BirA biotin ligase. Proteins that are in close proximity (~10 nm) will be conjugated with biotin allowing their purification by streptavidin immunoprecipitation and identification by mass spectrometry (MS) (Fig. 1a). Importantly, biotinylation by BirA does not require a direct or stable interaction between the BirA fusion protein and the biotinylated protein, making it particularly suitable for the detection of weak or transient protein-protein interactions that cannot be identified using traditional co-immunoprecipitation approaches. However, since CSA-BirA proximal proteins are not necessarily CSA interactors or substrates of the $CRL4^{CSA}$ complex, it is essential that proteins identified by BioID are validated to determine the nature of their association.

In order for the identification of potential substrates of the $CRL4^{CSA}$ complex via BioID to be successful, the BirA biotin ligase must face the substrate-binding surface of CSA. Therefore, we generated fusion proteins in which BirA was inserted in three positions that contain loops that protrude from the highly folded CSA WD40 propeller structure (235, 290, and 355) based on the available cryo-EM structure of the CSA-DDB1 complex (PDB: 4A11) [5]. In addition, we fused BirA to the unstructured C-terminus of CSA (Fig. 1b, c). To attempt to maintain CSA protein folding and functionality, we flanked the *BirA* gene by a linker sequence encoding flexible amino acids Glycine and Serine. We stably expressed the CSA-BirA fusion proteins in U2OS *CSA*-KO cells equipped with the Flp-In/T-Rex system to allow doxycycline-inducible expression of the CSA-BirA fusion proteins, which was confirmed by western blot analysis (Fig. 2a). As expected, clonogenic survival assays revealed that the *CSA*-KO cells were highly sensitive to Illudin S (Fig. 2b), which is a compound that generates transcription-blocking DNA lesions that are exclusively repaired by TCR [14]. Importantly, while the CSA-BirA internal fusion proteins (235, 290, and 355) were unable to restore the illudin S sensitive phenotype of *CSA*-KO cells, the CSA-BirA-C1 fusion protein fully restored Illudin S resistance (Fig. 2b, c). These results demonstrate that the internal CSA-BirA

fusion proteins are not functional, but that the CSA-BirA-C1 fusion protein is fully functional and is able to complement the genetic deletion of endogenous CSA.

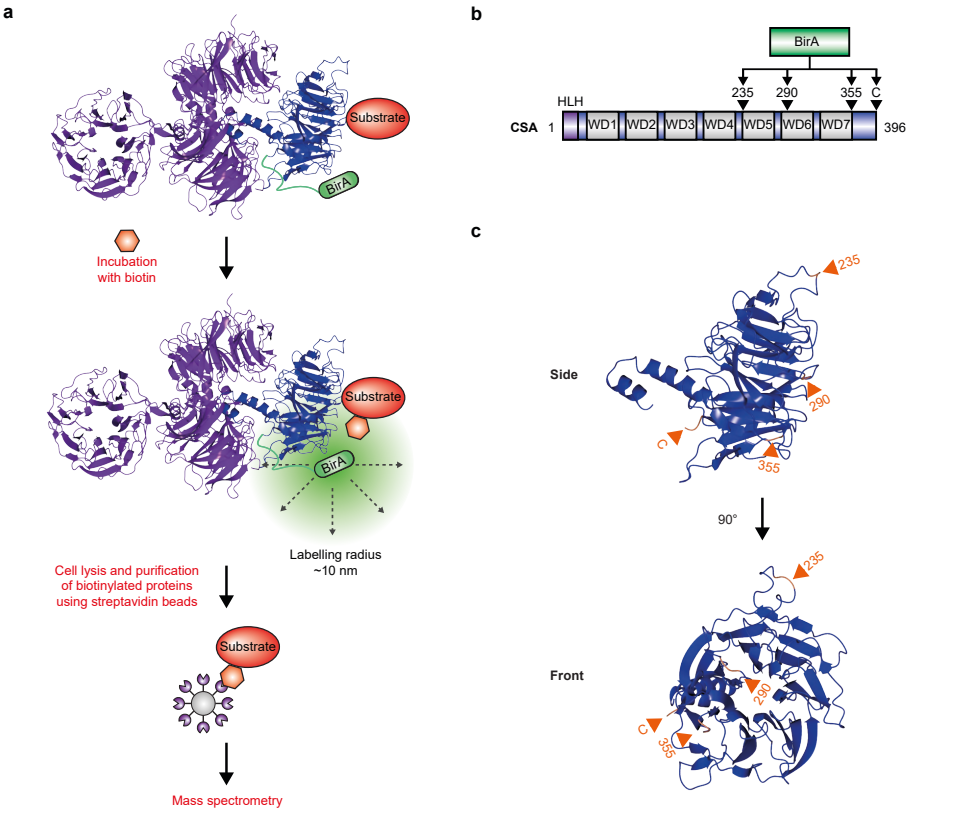


Fig. 1: Establishing a proximity-dependent Biotin identification (BioID) method. (**a**) Schematic representation of the BioID method. (**b**) Domain architecture of CSA, showing the positions in the *CSA* gene where BirA was inserted (235, 290, 355, C). (**c**) Side-view and front-view of the CSA structure. BirA was fused to the C-terminus as well as inserted in three positions that contain loops that protrude from CSA (235, 290, and 355). The positions where the *BirA* gene was inserted are indicated in orange. The available cryo-EM structure (PDB: 4A11) [5] of DDB1 (purple) bound to CSA (blue) was used in this figure.

To further confirm that fusion of BirA to the C-terminus of CSA does not interfere with the function of CSA, we assessed its ability to interact with elongating (RNAPIIo) in response to Illudin S-induced DNA damage. To this end, we immunoprecipitated endogenous RNAPIIo in the absence and presence of Illudin S, by employing a procedure we recently established to isolate intact TCR complexes [2]. In order to determine the optimal Illudin S treatment conditions, we tested a variety of Illudin S concentrations (10, 20 and 30 ng/mL) and incubation times (3, 6, and 17 hrs). RNAPIIo interacted with CSB and CSA-BirA-C1 in a DNA damage-dependent manner (Fig. 2d). Importantly, the DNA damage-induced association of CSB and CSA-BirA-C1 to RNAPIIo was similar at all time points and Illudin S concentrations analyzed except 10 ng/mL Illudin S treatment for 17 h, which showed weaker interactions compared to the other

conditions (Fig. 2d). It is important to note that labelling with Biotin requires long incubation times (>12 h) [15] and since we wanted to simultaneously treat cells with Biotin and Illudin S we decided to continue with 20 ng/mL Illudin S for 17 h.

Identification of CRL4^{CSA} substrates by BioID

The ubiquitin ligase function of the CRL4^{CSA} complex is activated upon CSB-dependent recruitment to DNA damage-stalled RNAPIIo [2, 5]. To identify CSA-interacting proteins that are dependent on its recruitment to lesion-stalled RNAPIIo, we knocked out *CSB*, using CRISPR-Cas9-mediated genome editing, in *CSA*-KO cells expressing CSA-BirA-C1. Knockout of *CSB* was verified by western blot analysis (Fig. 2e) and clonogenic survival assays in the presence of Illudin S (Fig. 2f).

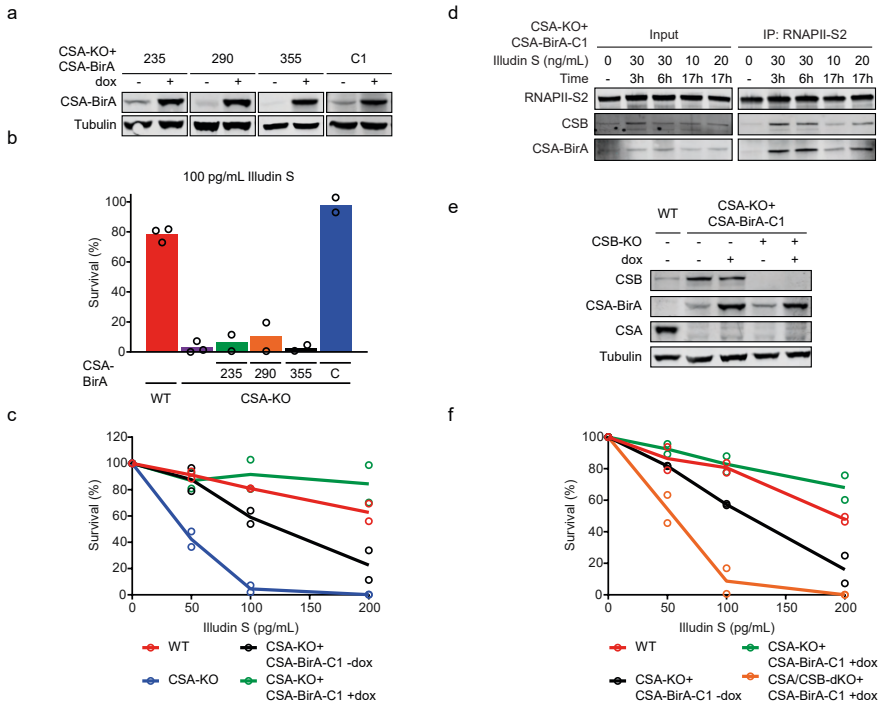


Fig. 2: The CSA-BirA-C1 fusion protein is fully functional. (**a**) Western blot analysis of U2OS (FRT) *CSA*-KO cells complemented with doxycycline (dox)-inducible CSA-BirA fusion proteins. (**b**) Clonogenic Illudin S survival of U2OS (FRT) WT, *CSA*-KO, and BirA-tagged CSA rescue cell lines. Each symbol represents the mean of an independent experiment (n=3 for U2OS (FRT) WT and *CSA*-KO, n=2 for the CSA rescue cell lines), each containing 2 technical replicates. (**c**) Clonogenic Illudin S survival of U2OS (FRT) WT, *CSA*-KO, and *CSA*-KO complemented with the C-terminal CSA-BirA fusion protein, in the presence and absence of dox. Each symbol represents the mean of an independent experiment (n=2), each containing 2 technical replicates. (**d**) Endogenous RNAPIIo Co-Immunoprecipitation (IP) on U2OS (FRT) *CSA*-KO cells complemented with CSA-BirA-C1 after mock treatment or Illudin S treatment (n=1). (**e**) Western blot analysis of U2OS (FRT) WT, *CSA*-KO and *CSA*/*CSB*-dKO complemented with CSA-BirA-C1. (**f**) Clonogenic Illudin S survival of U2OS (FRT) WT, *CSA*-KO and *CSA*/*CSB*-dKO complemented with CSA-BirA-C1. Each symbol represents the mean of an independent experiment (n=2), each containing 2 technical replicates.

Having generated *CSA*-KO cells and *CSA*/*CSB*-dKO cells expressing a functional CSA-BirA-C1 fusion, we subsequently validated if our BioID method works. To this end we incubated cells with biotin (50 μM) in the presence or absence of Illudin S (20 ng/mL) for 17 h, followed by purification of biotinylated proteins by streptavidin immunoprecipitation and identification by western blot analysis. While WT cells showed biotinylation of CSA-BirA, CSB, and the TFIIH complex (p89 and p62) in response to Illudin S treatment, biotinylation of these proteins was completely abolished in cells lacking *CSB* (Fig. 3a). This was expected since CSA interacts with the TFIIH complex [16] and recruitment of the TFIIH complex to lesion-stalled RNAPIIo is dependent on both CSB and CSA [2]. These initial results demonstrate that under our conditions we can capture CSA-interacting proteins that are dependent on the recruitment of $CRL4^{CSA}$ complex to lesion-stalled RNAPIIo.

Next, we purified biotinylated proteins by streptavidin immunoprecipitation followed by unbiased MS. The MS analysis of our BioID experiment in *CSA*-KO cells expressing CSA-BirA-C1 revealed that CSA-BirA biotinylates many proteins in response to Illudin S treatment, including UVSSA and the p44 subunit of the TFIIH complex (Fig. 3b). This is in line with earlier work showing that CSA interacts through direct protein-protein contacts with both UVSSA [2] and the p44 subunit of TFIIH [16]. We previously showed that the TCR proteins are recruited in a highly cooperative manner in which CSB stabilizes the interaction between CSA and UVSSA, while CSA is required to stabilize the interaction between UVSSA and TFIIH [2]. Indeed, knockout of *CSB* resulted in reduced biotinylation of UVSSA and complete loss of TFIIH (p44) biotinylation (Fig. 3c, d).

In addition, we observed increased biotinylation of the COP9 signalosome (CSN) complex upon Illudin S treatment in cells lacking *CSB* compared to WT cells (Fig. 3d). Recruitment of the $CRL4^{CSA}$ complex to lesion-stalled RNAPIIo causes the CSN complex to dissociate from $CRL4^{CSA}$ [5, 6]. In the absence of CSB, the $CRL4^{CSA}$ complex is not recruited to lesion stalled RNAPIIo, possibly explaining the increased biotinylation seen in *CSB*-KO cells. Moreover, the transcription elongation factor HIV Tat-specific factor 1 (HTATSF1) is biotinylated upon Illudin S treatment in a CSB-dependent manner (Fig. 3b, d). This is consistent with earlier work showing that HTATSF1 interacts with CSB [17], therefore it is likely that HTATSF1 is brought in close proximity to CSA-BirA by interacting with CSB.

Interestingly, chromodomain Y-like protein (CDYL) was one of the most efficiently biotinylated proteins upon Illudin S treatment in WT cells but not in cells lacking *CSB* (Fig. 3b-d). Streptavidin immunoprecipitation of biotinylated proteins followed by western blot analysis indeed confirmed DNA damage-specific biotinylation of CDYL, as well as TFIIH subunits, in a manner that was fully dependent on CSB (Fig. 3e). These findings demonstrate that CDYL is brought into close proximity of CSA upon recruitment of the $CRL4^{CSA}$ complex to lesion-stalled RNAPIIo, indicating that CDYL might be involved in the repair of transcription-blocking DNA lesions. Nevertheless, it is important to emphasize that with the BioID method you do not only capture proteins that directly interact with the target, in this case CSA, but also proteins that are in proximity of the target. Therefore, CDYL may directly or indirectly interact with CSA, and potentially be a substrate of the $CRL4^{CSA}$ complex.

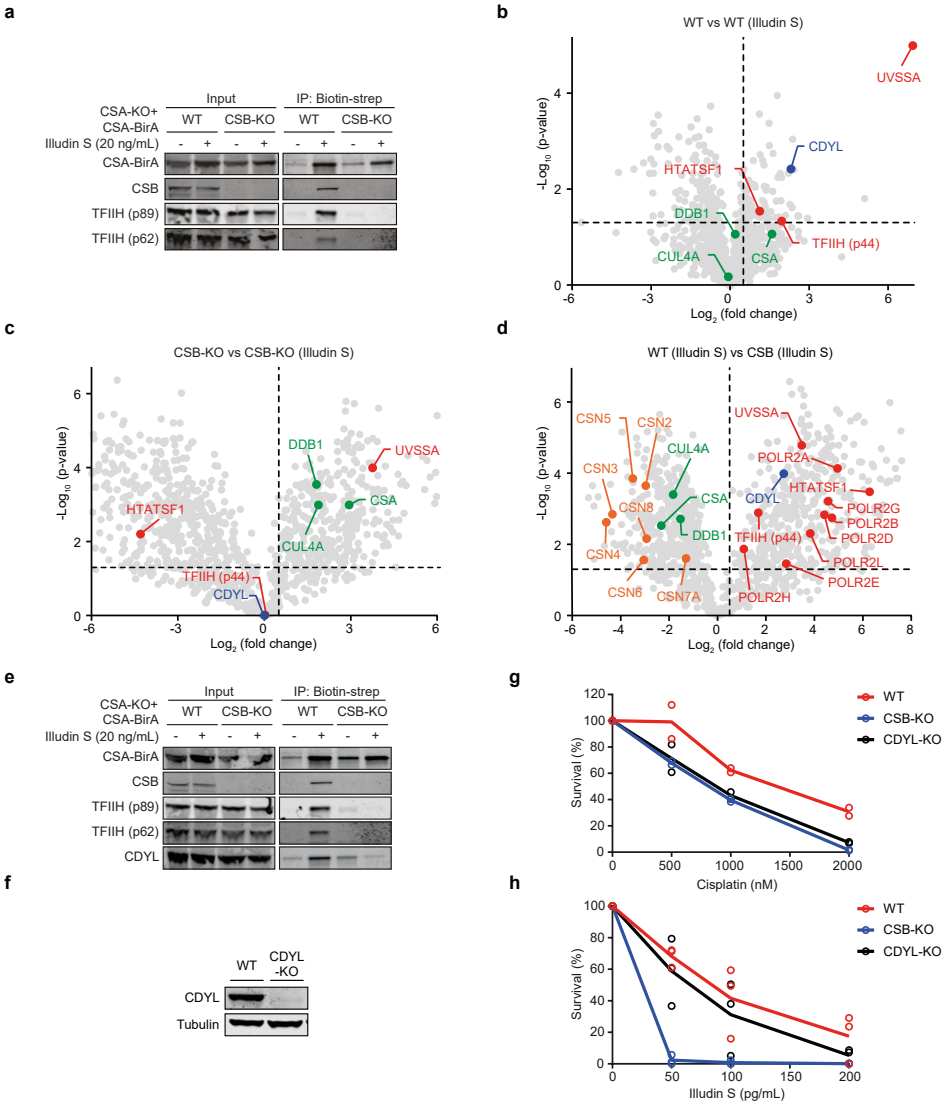


Fig. 3: With our BioID method we can identify CSA-proximal proteins. (**a**) Biotin-streptavidin immunoprecipitation on U2OS (FRT) *CSA*-KO and *CSA*/*CSB*-dKO complemented with CSA-BirA-C1 after mock treatment or Illudin S treatment (n=2). (**b-d**) Volcano plot depicting statistical difference between 3 replicates after biotin-streptavidin immunoprecipitation comparing (**b**) WT mock with illudin S (**c**) *CSB*-KO mock with illudin S, (**d**) WT illudin S with *CSB*-KO illudin S. The fold change ($\log_2$) is plotted on the x-axis and the significance (t-test -$\log_{10}$(P value)) is plotted on the y-axis. $CRL4^{CSA}$ proteins are indicated in green, proteins known to associate with the TCR complex are indicated in red, subunits of the COP9 signalosome are indicated in orange, and interesting hits are indicated in blue. (**e**) Biotin-streptavidin immunoprecipitation on U2OS (FRT) *CSA*-KO and *CSA*/*CSB*-dKO complemented with CSA-BirA-C1 after mock treatment or Illudin S treatment (n=2). (**f**) Western blot analysis of U2OS (FRT) WT and *CDYL*-KO cells. (**g-h**) Clonogenic survival of U2OS (FRT) WT, *CSB*-KO and *CDYL*-KO after treatment with (**g**) cisplatin (**h**) illudin S. Each symbol represents the mean of an independent experiment (n=2 for cisplatin, n=3 for illudin S), each containing 2 technical replicates.

Validation of CDYL as a TCR protein

To further investigate if CDYL is involved in TCR, we generated *CDYL*-KO cells using CRISPR/Cas9, which were validated by western blot analysis (Fig. 3f). Previous studies reported that depletion of CSB [18] as well as CDYL [19] sensitizes cells to cisplatin. To this end, we performed clonogenic survival assays in the presence of increasing concentrations of cisplatin. Our results confirm that both *CSB*-KO cells and *CDYL*-KO cells are sensitive to cisplatin compared to WT cells (Fig. 3g). The repair of cisplatin-induced DNA lesions requires genes from a variety of DNA repair pathways, including NER and homologous recombination (HR) [20]. To more specifically determine if CDYL is required for TCR we exposed cells to Illudin S, which generates transcription-blocking DNA lesions that are repair by TCR [14]. Clonogenic survival assays revealed that *CDYL*-KO cells are not nearly as sensitive to Illudin S as *CSB*-KO cells (Fig. 3h). This is consistent with recently published genome-wide CRISPR screens, which showed that loss of *CDYL* did not cause any sensitivity to UV irradiation [21] or Illudin S treatment [12]. Importantly, loss of *CSB*, *CSA*, and *UVSSA*, but not *CDYL*, conferred resistance to trabectedin [21]. A hallmark of TCR-deficient cells is resistance to trabectedin since TCR genes are required for the cytotoxic effects of trabectedin [22]. These combined findings indicate that CDYL is not essential for TCR, however, at this time the data on the involvement of CDYL in TCR is fairly limited and therefore additional experiments will be required to determine if CDYL plays a role in TCR.

Discussion

The cooperative and sequential assembly of the TCR proteins, CSB, CSA, and UVSSA, target the TFIIH complex to lesion-stalled RNAPIIo to initiate repair [2]. However, it has become increasingly clear that in addition to direct protein-protein contacts, the DNA damage-induced ubiquitylation of TCR factors plays a key role in the regulation of TCR. The $CRL4^{CSA}$ ubiquitin ligase complex, which is activated upon CSA auto-ubiquitylation, has been implicated in the ubiquitylation of multiple TCR proteins, including RNAPII and CSB [2, 5–7, 17]. Although recent findings emphasize the vital role of ubiquitylation by the $CRL4^{CSA}$ complex during TCR, it is unlikely that all relevant substrates of the $CRL4^{CSA}$ complex are identified since identification of such substrates can be challenging due to the transient nature of the ligase-substrate interaction. Here, we demonstrate that CSA-proximal proteins can be captured with BioID and therefore the BioID method can be used to identify novel TCR proteins as well as potential substrates of the $CRL4^{CSA}$ complex.

We provide several lines of evidence showing that our BioID method enables us to map specific interactions with the $CRL4^{CSA}$ complex (Fig. 3b-d). Firstly, the DNA damage-induced biotinylation of the CSA-interacting proteins UVSSA and TFIIH was dependent on CSB. UVSSA is known to be ubiquitylated in response to UV irradiation [8, 9], but whether the $CRL4^{CSA}$ ubiquitin ligase complex is responsible for UVSSA ubiquitylation remains to be elucidated. UVSSA contains a CSA-interaction region (CIR) and a TFIIH-interaction region (TIR). Interestingly, deletion of the CIR from UVSSA does not only prevent UVSSA from interacting with CSA, it also prevents UVSSA from interacting with the TFIIH complex [2]. Recent studies have shown that UVSSA ubiquitylation is required for recruitment

of the TFIIH complex to lesion-stalled RNAPIIo [7]. We speculate that the inability of UVSSA ΔCIR to interact with TFIIH is caused by defective UVSSA ubiquitylation by the $CRL4^{CSA}$ complex. In this regard, it would be interesting to examine if mutations that disrupt the interaction between CSA and UVSSA result in impaired UVSSA ubiquitylation. Secondly, we demonstrated increased biotinylation of the CSN complex in *CSB*-deficient cells compared to WT cells, which is consistent with earlier findings showing that recruitment of the $CRL4^{CSA}$ complex to lesion-stalled RNAPIIo causes the CSN complex to dissociate from $CRL4^{CSA}$ [5, 6]. Finally, the CSB-interacting protein, HTATSF1, is biotinylated upon Illudin S treatment in a CSB-dependent manner, thereby demonstrating that we can capture CSA-proximal proteins that do not interact through direct protein-protein contacts with CSA. HTATSF1 is a general transcription factor that associates with RNAPII and stimulates transcription elongation [23]. In addition, genome-wide siRNA experiments showed that knockdown of HTATSF1 results in reduced transcription in response to UV irradiation [24], suggesting that HTATSF1 might also be involved in transcriptional recovery following genotoxic stress. It is noteworthy, that HTATSF1 is a general transcription factor and might therefore have a general role in the restart of transcription and not specifically in transcription-restart following repair of transcription-blocking DNA lesions. Thus, future studies should examine the ability of HTATSF1-depleted cells to restart transcription in the absence of DNA damage, for example by measuring transcription restart following treatment with the reversible transcription inhibitor 5,6-dichlorobenzimidazole1-β-D-ribofuranoside (DRB). If HTATSF1-depleted cells are able to restart transcription following DRB treatment, it would indicate that HTATSF1 is specifically involved in the regulation of transcription following DNA damage. Moreover, recently published genome-wide CRISPR screens showed that loss of *HTATSF1* only caused mild sensitivity to UV irradiation [21] or Illudin S treatment [12], which suggests that HTATSF1 might not be directly involved in the removal of DNA lesions but rather in regulating transcription restart once repair is finished. This assumption could be addressed by using the strand-specific ChIP-seq (TCR-seq) method to quantify the genome-wide TCR kinetics of HTATSF1-depleted cells. It should be noted that the genome-wide CRISPR screen showed reduced viability of *HTATSF1*-depleted cells in the absence of genotoxic stress [12], indicating that *HTATSF1* is an essential gene. In conclusion, data on the involvement of HTATSF1 in TCR is fairly limited and all originates from genome-wide screens, making future research imperative to determine if, in addition to its role in transcription elongation, HTATSF1 is involved in the transcription-related DNA damage response.

CDYL was one of the most efficiently biotinylated proteins upon Illudin S treatment in WT cells but not in cells lacking *CSB* (Fig. 3b-d). CDYL is a transcriptional co-repressor that contains a N-terminal chromodomain and a C-terminal acetyl-CoA-binding domain (Fig. 4a) [25]. Earlier work has shown that CDYL promotes transcription repression at sites of DNA double-strand breaks (DSBs) [19], a role for CDYL in TCR has not yet been described. We demonstrate that *CDYL*-KO cells are only mildly sensitive to transcription-blocking DNA damage induced by Illudin S compared to TCR-deficient *CSB*-KO cells (Fig. 3h). Moreover, given that CDYL is a transcriptional co-repressor, it is unlikely that CDYL is required for transcription restart following DNA repair. However, in

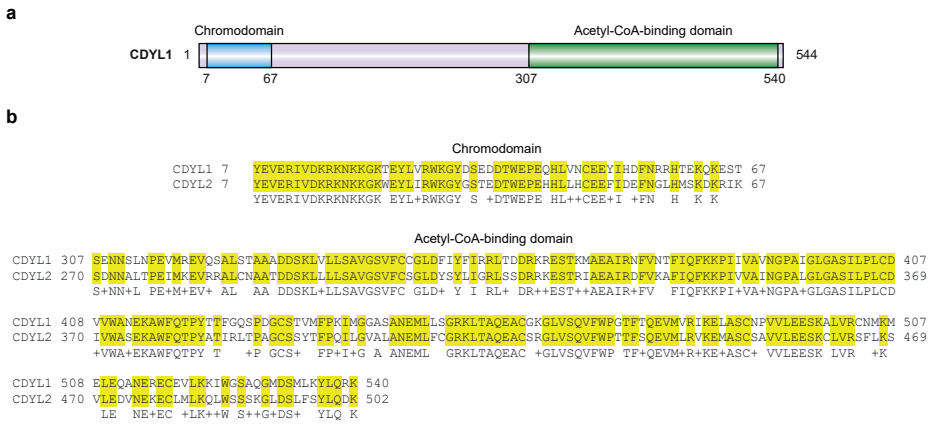


Fig. 4: Alignment of CDYL1 and CDYL2 paralogues. (**a**) A schematic representation of the domain architecture of CDYL1. The chromodomain is depicted in blue and the Acetyl-CoA-binding domain in green. (**b**) Alignment of the chromodomain and the Acetyl-CoA-binding domain of human CDYL1 and CDYL2. Identical residues are indicated in yellow.

the evolutionary timeline, genomic duplication events resulted in the formation of CDYL1 and CDYL2 paralogues [26]. Proteins encoded by the *CDYL1* and *CDYL2* genes share a 53.56 % sequence identity, that increases to ~70% at the chromodomain and the acetyl-CoA-binding domain (Fig. 4b). It is interesting to note that we did not find CDYL2 in our BioID. Studies have shown that many duplicate gene pairs are functionally redundant, meaning that the protein encoded by one gene can function in place of the other [27–29]. Therefore, it is possible that *CDYL1*-KO cells are not sensitive to Illudin S because *CDYL2* compensates for the loss of *CDYL1*. In this context, it would be interesting to explore if knockout of *CDYL1* and *CDYL2* together causes an additive sensitivity to Illudin S compared to either single knockout. In addition, it will be important to study whether combined loss of *CDYL1*/*CDYL2* affects the restart of transcription following genotoxic stress.

The chromodomain of CDYL1 and CDYL2 preferentially interacts with tri-methylated histone H3K9 (H3K9me$_3$) [30, 31]. *CSB*-deficient cells show reduced H3K9me$_3$ methylation, due to decreased expression of the methyltransferases Suppressor Of Variegation 3-9 Homolog 1 (SUV39H1) and SET Domain Bifurcated Histone Lysine Methyltransferase 1 (SETDB1) [32]. Moreover, the local deposition of H3K9me$_3$ is required for DSB repair and maintaining genomic stability [33]. In line with this, both CSB and CDYL1 have been shown to promote DSB repair via HR [19, 34]. Further studies should investigate the link between CDYL1 and CSB in the repair of DSBs, as well as transcription-blocking DNA lesions, by assessing if *CDYL1*/*CDYL2*/*CSB* triple knockout cells show increased sensitivity to DNA-damaging agents, such as cisplatin, illudin S, and UV.

Chaperone chromatin assembly factor 1 (CAF-1) associates with UV-damaged chromatin where it promotes incorporation of newly synthesized histone H3.1 after lesions have been repaired by NER [35, 36]. Interestingly, during DNA replication CDYL1 interacts with CAF-1 and recruits the histone-modifying

enzymes G9a, SETDB1, and EZH2 to restore H3K9me$_{2/3}$ and H3K27me$_{2/3}$ marks on newly deposited histones [37]. The methyltransferase G9a does not only methylate histone H3 but also methylates non-histone proteins, including CDYL1 and CSB [31]. Intriguingly, methylation of CDYL1 abolishes the association of CDYL1 with H3K9me$_3$, suggesting that G9a methylates CDYL1 to regulate its binding to chromatin [31]. In addition, it has been shown that, in the absence of DNA damage, CSB interacts with G9a and stimulates the methyltransferase activity of G9a [38]. It is tempting to hypothesize that in context of TCR, CSB is not only involved in the early repair steps but might also work together with CDYL1 and CAF-1 to preserve original histone marks after repair is finished. Therefore, it would be interesting to examine if CAF-1 recruitment to UV-damaged chromatin, and the subsequent incorporation of newly synthesized histone H3.1 is dependent on CDYL1/2, and if depletion of CDYL1/2 affect histone modifications upon UV irradiation.

In conclusion, our BioID approach yielded known TCR proteins and potential new TCR candidates, amongst which CDYL and HTATSF1, that await further validation.

Methods

Cell lines. All cell lines (listed in Supplementary Table 1) were cultured at 37°C in an atmosphere of 5% CO_2 in DMEM (Thermo Fisher Scientific) supplemented with penicillin/streptomycin (Sigma) and 10% fetal bovine serum (FBS; Bodinco BV or Thermo Fischer Scientific (Gibco)). U2OS Flp-In/T-REx cells (from here on out called U2OS (FRT)), which were generated using the Flp-InTM/T-RExTM system (Thermo Fisher Scientific), were a gift from Daniel Durocher [39].

Generation of knock-out cells. U2OS (FRT) cells were co-transfected with Cas9-2A-GFP (pX458; Addgene #48138) together with pLV-U6g-PPB encoding a guide RNA from the LUMC/Sigma-Aldrich sgRNA library using lipofectamine 2000 (Invitrogen) (sgRNAs are listed in Supplementary Table 2 and plasmids in Supplementary Table 3). Transfected cells were selected on puromycin (1 μg/mL) for 3 days and seeded 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

PCR analysis of knockout clones. Genomic DNA was isolated by resuspending cell pellets in WCE buffer (50mM KCl, 10mM Tris pH 8.0, 25mM MgCl2 0.1 mg/mL gelatin, 0.45% Tween-20, 0.45% NP-40) containing 0.1 mg/mL Proteinase K (EO0491; Thermo Fisher Scientific) and incubating for 1 h at 56°C followed by a 10 min heat inactivation of Proteinase K by 96°C. Fragments of ~1 kb, containing the sgRNA sequence, were amplified by PCR (Sequencing primers are listed in as Supplementary Table 4), followed by Sanger sequencing using either the forward or the reverse primer.

Plasmids. CSA-GFP in pcDNA5/FRT/TO-puro-CSA-GFP [2] was replaced with CSA$^{1\text{-}235}$, CSA$^{1\text{-}290}$, CSA$^{1\text{-}355}$, or CSA^{WT} without GFP (Primers for cloning are listed in Supplementary Table 5). BirA spanning a flexible linker (GSGGSGGGSGG) was inserted into the CSA plasmids to generate pcDNA5/FRT/TO-puro-CSA$^{1\text{-}235}$-BirA, CSA$^{1\text{-}290}$-BirA, CSA$^{1\text{-}355}$-BirA, and CSA^{WT}-BirA, Fragments spanning CSA$^{235\text{-}396}$, CSA$^{290\text{-}396}$, and CSA$^{355\text{-}396}$ were PCR amplified and inserted into the CSA-BirA fusion plasmids to generate internal in-frame CSA fusion proteins with BirA located on position 235, 290, 355. All sequences were verified by Sanger sequencing.

Generation of stable cell lines. U2OS (FRT) *CSA*-KO clone 2-16 was selected and subsequently used to stably express inducible BirA-tagged proteins by co-transfecting pCDNA5/FRT/TO-Puro plasmid encoding BirA-tagged fusion proteins (5 μg), together with pOG44 plasmid encoding the Flp recombinase (0.5 μg) using lipofectamine 2000 (Invitrogen) (Plasmids are listed in Supplementary Table 3). Cells expressing inducible BirA-tagged proteins were selected by incubation with 1 μg/mL

puromycin and 4 μg/mL blasticidin S. Expression of these BirA-tagged proteins was induced by the addition of 2 μg/mL doxycycline for 24 h.

Western blotting. 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). Membranes were blocked with blocking buffer (Rockland, MB-070-003) for 1 h at RT. Membranes were then probed with antibodies as indicated (Antibodies are listed in Supplementary Table 6).

Clonogenic survival assays. Cells were seeded at low density and mock-treated or exposed to a dilution series of Illudin S (Santa Cruz; sc-391575) or Cisplatin (Thermo Fisher Scientific; 22-515-0) for 72 h. 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 mock treated or Illudin S treated (10, 20 and 30 ng/mL) for different periods of time (3, 6, 17 h). Chromatin-enriched fractions were prepared by incubating the cells for 20 min on ice in IP-130 buffer (30 mM Tris pH 7.5, 130 mM NaCl, 2 mM $MgCl_2$, 0.5% Triton X-100, and protease inhibitor cocktail (Roche)), followed by centrifugation, and removal of the supernatant. The chromatin-enriched cell pellets were lysed in IP-130 buffer containing 250 U/mL Benzonase® Nuclease (Novagen) and 2 μg RNAPII-S2 (ab5095, Abcam) for 2-3 h at 4°C. Protein complexes were pulled down by incubation with Protein A agarose beads (Millipore) for 1.5 h at 4°C, after which the beads were washed 6 times with IP-130 buffer. The samples were prepared by boiling in Laemmli-SDS sample buffer. Bound proteins were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotted with the indicated antibodies.

BioID. Cells were treated with 50 μM Biotin (Sigma #B4501) in the absence or presence of 20 ng/mL Illudin S (Santa Cruz; sc-391575) for 17 h. Cells were lysed in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1% deoxycholate, 2 mM $MgCl_2$, 0.% NP-40, protease inhibitor cocktail (Roche), and 500 U/mL Benzonase® Nuclease (Novagen)) for 1 h at 4°C. Cells were centrifuged and biotinylated proteins were pulled down by incubating the supernatant with streptavidin-sepharose beads (Millipore) overnight at 4 °C. After pulldown, the beads were washed 6 times with wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1% deoxycholate, 1 mM EDTA, 0.5% NP-40, protease inhibitor cocktail (Roche)). Samples for analysis by western blotting were prepared by boiling the beads in Laemmli-SDS sample buffer saturated with 0.2 mg/mL Biotin. Biotinylated proteins were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotted with the indicated antibodies.

Mass spectrometry. For the generation of mass spectrometry samples, the beads were washed 3 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 desalted using C-18 stage tips by washing with 0.1% formic acid. Finally, peptides were eluted with 0.1% formic acid/60% acetonitrile and lyophilized as described [40].

Mass spectrometry was performed essentially as previously described [41]. 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 125 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 files were analysed with MaxQuant software (v1.5.3.30) as previously described [42], 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 Fast LFQ feature. Match-between-runs feature was enabled with a match time window of 0.7 minutes and an alignment time window of 20 minutes. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens (14th December 2017). Analysis output from MaxQuant was further processed in the Perseus (v 1.5.5.3) computational platform [43]. Proteins identified as common contaminants, only identified by site and reverse peptide were filtered out, and then all the LFQ intensities were $\log_2$ transformed. Different biological repeats of each condition were grouped and only protein groups identified in all three 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 ($\log_2$) and a randomized 0.3 width ($\log_2$) 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.

Supplementary tables

Supplementary Table 1: Cell lines

Cell line	Origin
U2OS (FRT)	[39]
U2OS (FRT) CSB-KO (1-12)	[2]
U2OS (FRT) CSA-KO (2-16)	This study
U2OS (FRT) CSA-KO (2-16) + CSA-BirA-235 (13)	This study
U2OS (FRT) CSA-KO (2-16) + CSA-BirA-290 (11)	This study
U2OS (FRT) CSA-KO (2-16) + CSA-BirA-355 (2)	This study
U2OS (FRT) CSA-KO (2-16) + CSA-BirA-C1 (17)	This study
U2OS (FRT) CSA-KO (2-16)/CSB-KO (1-28) + CSA-BirA-C1 (17)	This study
U2OS (FRT) CDYL-KO (2-17)	This study

Supplementary Table 2: Sequences of sgRNAs

Gene	Sequence	Gene_ID	Exon_ID	Exon
CSB_1	CTCATCGGATCATTCTGTCT	ENSG00000225830	ENSE00002514316	10
CSA_2	CAACTTTGTGACTTGAAGTC	ENSG00000049167	ENSE00003589574	6
CDYL1_2	AACATACAGACATCTGTTAC	ENSG00000153046	ENSE00003631402	3

Supplementary Table 3: Plasmids

Plasmid	Origin
pcDNA5/FRT/TO-Puro-CSA-GFP	[2]
pcDNA5/FRT/TO-puro-CSA-BirA-235	This study
pcDNA5/FRT/TO-puro-CSA-BirA-290	This study
pcDNA5/FRT/TO-puro-CSA-BirA-355	This study
pcDNA5/FRT/TO-puro-CSA-BirA-C1	This study
pUC57-BirA	This study (Genscript)
pLV-U6g-PPB	LUMC/Sigma-Aldrich sgRNA library
pX458	Addgene #48138
pOG44	Thermo Fisher

Supplementary Table 4: Primers for sequencing

Gene	Sequence	Identifier
CDYL	5-GTCTGAAGTGGTGGCTGCC-3	oML#304
	5-GTGCCCGCCAGATGAAGAG-3	oML#305
	5-TCGACTAGATGGCCTTGACTC-3	Oml#326
	5-CAGAGGCACAGAGGCCTG-3	oML#327
CSA	5-CAGTCTGTGTCCAGTTTCTGTG-3	oML#084
	5-CATATTTGTTATGTGTTTCTTTGAG-3	oML#085
CSB	5-GTAGGGGCCAGTTGTTAGAATGTAA-3	oML#078
	5-CTCACATTCTGAATGACTTGGCTA-3	oML#079

Supplementary Table 5: Primers for cloning

Gene	Sequence	Identifier
CSA 175-235	5-TGCGACGGATCCTGTTCTCACATTCTACAGGGTC-3	oML#054
	5-ATATGAGTTAACAATTGTTCCGGATGACTTTTTCCCATTATGTTGATC-3	oML#055
CSA 175-290	5-TGCGACGGATCCTGTTCTCACATTCTACAGGGTC-3	oML#054
	5-ATATGAGTTAACAATTGTTCCGGAATTACAAACTTTTCCATAGTTCACAAG-3	oML#060
CSA 175-355	5-TGCGACGGATCCTGTTCTCACATTCTACAGGGTC-3	oML#054
	5-ATATGAGTTAACAATTGTTCCGGATCTGCTACCACTATAAAGTTCC-3	oML#062
CSA 175-396	5-TGCGACGGATCCTGTTCTCACATTCTACAGGGTC-3	oML#054
	5-CAAGATGTTAACAATTGTTCCGGATCCTTCTTCATCACTGCTG-3	oML#064
CSA 235-396	5-ACTGCTGCTGCTGGATCCGGACAAGCTGTTGAATCAGCAAAC-3	oML#058
	5-CAAGATGTTAACTTATGCGTAATCCGGTACATCGTAAGGGTATC CTTCTTCATCACTGCTG-3	oML#059
CSA 290-396	5-GGCTGTCGCGGCAGTGTCCGGAAACAGTAAAAAAGGATTGAAATTCAC-3	oML#061
	5-CAAGATGTTAACTTATGCGTAATCCGGTACATCGTAAGGGTATC CTTCTTCATCACTGCTG-3	oML#059
CSA 355-396	5-TGTCGCGGCAGTGTCCGGAGACTGCAACATTCTGGCTTG-3	oML#063
	5-CAAGATGTTAACTTATGCGTAATCCGGTACATCGTAAGGGTATC CTTCTTCATCACTGCTG-3	oML#059

Supplementary Table 6: Antibodies

Antibody	Host	Company (reference)	Use	Identifier
CSA/ERCC8	Rabbit	Abcam, 137033 (EPR9237)	WB: 1:500	aML#028
CSB/ERCC6	Goat	Santa Cruz, sc-10459 (E-18)	WB: 1:1000	aML#039
RNAPII-S2	Rabbit	Abcam, ab5095	WB: 1:1000	aML#024
Tubulin	Mouse	Sigma, T6199 (DM1A)	WB: 1:1000	aML#008
p89/XPB/ERCC3	Mouse	kindly provided by J.M. Egly (1B3)	WB: 1:1000	aML#073
p62/GTF2H1	Mouse	Kindly provided by J.M. Egly (3C9)	WB: 1:2000	aML#074
CDYL	Rabbit	Sigma, HPA035578	WB: 1:750	aML#095

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] van der Weegen, Y. *et al.* The Cooperative action of CSB, CSA and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nat. Commun.* **11**, 2104 (2020).

[3] 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).

[4] 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).

[5] Fischer, E. S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024–1039 (2011).

[6] Groisman, R. *et al.* CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes Dev.* **20**, 1429–1434 (2006).

[7] Nakazawa, Y. *et al.* Ubiquitination of DNA damage-stalled RNAPII promotes transcription-coupled repair. *Cell* **180**, 1228–1244 (2020).

[8] Higa, M., Tanaka, K. & Saijo, M. Inhibition of UVSSA ubiquitination suppresses transcription-coupled nucleotide excision repair deficiency caused by dissociation from USP7. *FEBS J* **285**, 965–976 (2018).

[9] Schwertman, P. *et al.* UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat. Genet.* **44**, 598–602 (2012).

[10] Jin, J., Arias, E. E., Chen, J., Harper, J. W. & Walter, J. C. A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol. Cell* **23**, 709–21 (2006).

[11] Cao, H., Williams, C., Carter, M. & Hegele, R. A. CKN1 (MIM 216400): mutations in Cockayne syndrome type A and a new common polymorphism. *J. Hum. Genet.* **49**, 61–63 (2004).

[12] van der Weegen, Y. *et al.* ELOF1 is a transcription-coupled DNA repair factor that directs RNA polymerase II ubiquitylation. *Nat. Cell Biol.* **23**, 595–607 (2021).

[13] 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).

[14] Jaspers, N. G. *et al.* Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. *DNA Repair* **1**, 1027–1038 (2002).

[15] Roux, K. J., Kim, D. I., Burke, B. & May, D. G. BioID: A Screen for Protein-Protein Interactions. *Curr. Protoc. Protein Sci.* **91**, 19.23.1–19.23.15 (2018).

[16] 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).

[17] Liebelt, F. *et al.* Transcription-coupled nucleotide excision repair is coordinated by ubiquitin and SUMO in response to ultraviolet irradiation. *Nucleic Acids Res.* **48**, 231–248 (2020).

[18] Enoiu, M., Jiricny, J. & Scharer, O. D. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res.* **40**, 8953–8964 (2012).

[19] Abu-Zhayia, E. R., Awwad, S. W., Ben-Oz, B. M., Khoury-Haddad, H. & Ayoub, N. CDYL1 fosters double-strand break-induced transcription silencing and promotes homology-directed repair. *J. Mol. Cell Biol.* **10**, 341–357 (2018).

[20] Dronkert, M. L. & Kanaar, R. Repair of DNA interstrand cross-links. *Mutat. Res.* **486**, 217–247 (2001).

[21] Olivieri, M. *et al.* A Genetic Map of the Response to DNA Damage in Human Cells. *Cell* **182**, 481–496 (2020).

[22] Takebayashi, Y. *et al.* Antiproliferative activity of ecteinascidin 743 is dependent upon transcription-coupled nucleotide-excision repair. *Nat. Med.* **7**, 961–966 (2001).

[23] Lu, X. *et al.* Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. *Nucleic Acids Res.* **44**, 6853–6867 (2016).

[24] Boeing, S. *et al.* Multiomic Analysis of the UV-Induced DNA Damage Response. *Cell Rep.* **15**, 1597–1610 (2016).

[25] Caron, C. *et al.* Cdyl: a new transcriptional co-repressor. *EMBO Rep.* **4**, 877–882 (2003).

[26] Dorus, S., Gilbert, S. L., Forster, M. L., Barndt, R. J. & Lahn, B. T. The CDY-related gene family: coordinated evolution in copy number, expression profile and protein sequence. *Hum. Mol. Genet.* **12**, 1643–1650 (2003).

[27] Nieto Moreno, N. *et al.* GSK-3 is an RNA polymerase II phospho-CTD kinase. *Nucleic Acids Res.*

48, 6068–6080 (2020).
[28] Qian, W., Liao, B. Y., Chang, A. Y. & Zhang, J. Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet.* **26**, 425–430 (2010).
[29] Ng, J. M. *et al.* A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev.* **17**, 1630–1645 (2003).
[30] Dong, C. *et al.* Structural Basis for the Binding Selectivity of Human CDY Chromodomains. *Cell Chem. Biol.* **27**, 827–838 e7 (2020).
[31] Rathert, P. *et al.* Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* **4**, 344–346 (2008).
[32] Lee, J. H. *et al.* Cockayne syndrome group B deficiency reduces H3K9me3 chromatin remodeler SETDB1 and exacerbates cellular aging. *Nucleic Acids Res.* **47**, 8548–8562 (2019).
[33] Ayrapetov, M. K., Gursoy-Yuzugullu, O., Xu, C., Xu, Y. & Price, B. D. DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin. *Proc. Natl. Acad. Sci. USA* **111**, 9169–9174 (2014).
[34] Batenburg, N. L., Thompson, E. L., Hendrickson, E. A. & Zhu, X. D. Cockayne syndrome group B protein regulates DNA double-strand break repair and checkpoint activation. *EMBO J.* **34**, 1399–1416 (2015).
[35] Polo, S. E., Roche, D. & Almouzni, G. New histone incorporation marks sites of UV repair in human cells. *Cell* **127**, 481–493 (2006).
[36] Green, C. M. & Almouzni, G. Local action of the chromatin assembly factor CAF-1 at sites of nucleotide excision repair in vivo. *EMBO J.* **22**, 5163–5174 (2003).
[37] Liu, Y. *et al.* Chromodomain protein CDYL is required for transmission/restoration of repressive histone marks. *J. Mol. Cell Biol.* **9**, 178–194 (2017).
[38] Yuan, X., Feng, W., Imhof, A., Grummt, I. & Zhou, Y. Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a. *Mol. Cell* **27**, 585–595 (2007).
[39] Panier, S. *et al.* Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. *Mol. Cell* **47**, 383–395 (2012).
[40] Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–906 (2007).
[41] Kumar, R., González-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).
[42] Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
[43] Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. methods* **13**, 731–740 (2016).