Supplementary Information

CSB Interacts with SNM1A and Promotes DNA Interstrand Crosslink Processing

Teruaki Iyama, Sook Y. Lee, Brian R. Berquist, Opher Gileadi, Vilhelm A. Bohr, Michael M. Seidman, Peter J. McHugh, David M. Wilson III

Table S1. Synthetic Oligonucleotides. Oligonucleotide name and nucleotide sequence are designated. P = phosphate; B = biotin.

Oligonucleotide name	**Nucleotide sequence**
linker Fw1	5′-CCGGAGGTGGAGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCGC-3′
linker Rv1	5′-GATCGCGATCCGCCACCGCCAGAGCCACCTCCGCCTGAACCGCCTCCACCT-3′
linker Fw2	5′-CCGGAGGTGGAGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCG-3′
linker Rv2	5′-GATCCGATCCGCCACCGCCAGAGCCACCTCCGCCTGAACCGCCTCCACCT-3′
CSB XhoI-Fw	5′-AGTCCTCGAGGAATGCCAAATGAGGGAATC-3′
CSB BamHI-Rv	5′-AAGCAAGGATCCTTAGCAGTATTCTGGCTTGAG-3′
SNM1A XhoI-Fw	5′-ATCTCGAGACCATGTTAGAAGACATTTCCG-3′
SNM1A BamHI-Rv	5′-GGAGGATCCATCAATATCCAGCTTC-3′
5'-60mer	5'-P-TTATAAATAAAATATTAAAACCAATTTGATCATCTATTATATATTTATTAATTTATTTTT-3′
5′-20mer	5′-P-ATAATTTGATCATCTATTAT-3′
5′-Biotin-21mer	5′-B-TATAATAGATGATCAAATTAT-P-3′
5'-Biotin-61mer	5'-B-TAAAAATAAATTAATAAATATATAATAGATGATCAAATTGGTTTTAATATTTTATTTATAA-P-3′
Fluorescent-20mer	5'-P-A[FluoresceinT]AATTTGA[BHQ1-T]CATCTATTAT-3'
5′-Biotin-20mer	5′-B-ATAATAGATGATCAAATTAT-3′
dtCSB BamHI-Fw	5′-ATCGGATCCACCATGGCCTACCCATACGATGTTCCAGATTACGCTAGCCCAAATGAGGGA-3′
dtCSB XhoI-Rv	5′-CTGCTCGAGTTAGTGATGGTGATGGTGGTGACGACCTTCGATGCAGTATTCTGGCTTGAG-3′

Table S2. Summary of the yeast two-hybrid hits using CSB as bait. The 59 independent clones first isolated are listed, including gene name (if known) and NCBI accession number. The qualitative results of the β-galactosidase (LacZ) assay and growth on minimal medium lacking histidine and adenine (SD-ade, stringent selection) are summarized (see Figure S1). Based on these data, an overall rating of interaction strength is noted. N/A = not applicable.

Clone No.	Homo Sapiens, Gene Name	Accession No. (NCBI)	LacZ Positive	Growth on SD-ade	Interaction Strength
1	NODAL modulator 1 (NOMO1), mRNA	NM_014287.3	++	-	intermediate
2	Chromosome 8 genomic contig, GRCh37.p13 Primary Assembly	NT_167187.1	+	-	weak
3	Latent transforming growth factor beta binding protein 1 (LTBP1), transcript variant 2, mRNA	NM_000627.3	+	-	weak
4	Ankyrin repeat domain 7 (ANKRD7), mRNA	NM_019644.3	-	-	weak
5	DnaJ (Hsp40) homolog, subfamily C, member 13 (DNAJC13), mRNA	NM_015268.3	-	-	weak
6	Chromosome 19 genomic contig, GRCh37.p13 Primary Assembly	NT_011295.11	++	+	strong
7	HECT domain containing E3 ubiquitin protein ligase 4 (HECTD4), mRNA	NM_001109662.3	+	-	weak
8	Unknown sequence	N/A	++	-	intermediate
9	SHQ1, H/ACA ribonucleoprotein assembly factor (SHQ1), mRNA	NM_018130.2	-	-	weak
10	PDZ and LIM domain 1 (PDLIM1), mRNA	NM_020992.3	+	-	weak
11	Chromosome 20 genomic scaffold, alternate assembly CHM1_1.1	NW_004929418.1	+	-	weak
12	Solute carrier family 27 (fatty acid transporter), member 2 (SLC27A2), transcript variant 1, mRNA	NM_003645.3	+	-	weak
13	Zinc finger, MYND-type containing 12 (ZMYND12), transcript variant 1, mRNA	NM_032257.4	-	-	weak
14	Tryptophan rich basic protein (WRB), transcript variant 1, mRNA	NM_004627.4	+	-	weak
15	Ornithine decarboxylase 1 (ODC1), mRNA	NM_002539.1	+	(+)	weak
16	HECT domain containing E3 ubiquitin protein ligase 4 (HECTD4), mRNA	NM_001109662.3	+	-	weak

17	Serum/glucocorticoid regulated kinase 1 (SGK1), transcript variant 1, mRNA	NM_005627.3	+	-	weak
18	Hydroxysteroid (17-beta) dehydrogenase 11 (HSD17B11), mRNA	NM_016245.3	+	-	weak
19	Ribosomal protein L26-like 1 (RPL26L1), mRNA	NM_016093.2	-	-	weak
20	Unknown sequence	N/A	++	-	intermediate
21	DEAD (Asp-Glu-Ala-Asp) box helicase 5 (DDX5), mRNA	NM_004396.3	+	-	weak
22	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 55kDa (TAF7), mRNA	NM_005642.2	+	-	weak
23	Chromosome 15 genomic contig, GRCh37.p13 Primary Assembly	NT_010274.17	+	-	weak
24	FK506 binding protein 5 (FKBP5), transcript variant 1, mRNA	NM_004117.3	+	-	weak
25	Zinc finger, AN1-type domain 5 (ZFAND5), transcript variant a, mRNA	NM_001102420.2	+	-	weak
26	3-hydroxyisobutyryl-CoA hydrolase (HIBCH), transcript variant 1, mRNA	NM_014362.3	++	-	intermediate
27	Zinc finger protein 236 (ZNF236), mRNA	NM_007345.3	+	-	weak
28	Deoxynucleotidyltransferase, terminal, interacting protein 2 (DNTTIP2), mRNA	XR_246259.1	+++	++	strong
29	Glutathione S-transferase pi 1 (GSTP1), mRNA	NM_000852.3	+	-	weak
30	Signal recognition particle 9kDa (SRP9), transcript variant 1, mRNA	NM_001130440.1	++	-	intermediate
31	Mitochondrial ribosomal protein L51 (MRPL51), mRNA	NM_016497.3	+	-	weak
32	Family with sequence similarity 49, member B (FAM49B), transcript variant 1, mRNA	NM_001256763.1	+	-	weak
33	Chromosome 20 genomic contig, GRCh37.p13 Primary Assembly	NT_011362.10	++	-	intermediate
34	Chromosome 15 genomic contig, GRCh37.p13 Primary Assembly	NT_010274.17	-	-	weak
35	HIV-1 Tat specific factor 1 (HTATSF1), transcript variant 1, mRNA	NM_001163280.1	++	-	intermediate
36	Inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2), mRNA	NM_002223.2	-	(+)	weak
37	Kruppel-like factor 3 (basic) (KLF3), mRNA	NM_016531.5	+	-	weak

38	Chromosome 9 genomic contig, GRCh37.p13 Primary Assembly	NT_008413.18	+	-	weak
39	Host cell factor C2 (HCFC2), mRNA	NM_013320.2	+	-	weak
40	Membrane-spanning 4-domains, subfamily A, member 6E (MS4A6E), mRNA	NM_152851.2	+	-	weak
41	Chromosome 19 genomic contig, GRCh37.p13 Primary Assembly	NT_011295.11	++	-	intermediate
42	Fas cell surface death receptor (FAS), transcript variant 1, mRNA	NM_000043.4	++	-	intermediate
43	Mitochondrial ribosomal protein L3 (MRPL3), mRNA	NM_007208.3	++	-	intermediate
44	DNA cross-link repair 1A (DCLRE1A), transcript variant 1, mRNA	NM_001271816.1	+++	+++	strong
45	Chromosome 15 genomic contig, GRCh37.p13 Primary Assembly	NT_010194.17	++	-	intermediate
46	NODAL modulator 1 (NOMO1), mRNA	NM_014287.3	+	-	weak
47	Vesicle-associated membrane protein 8 (VAMP8), mRNA	NM_003761.4	++	-	intermediate
48	Chromosome 9 genomic contig, GRCh37.p13 Primary Assembly	NT_008413.18	-	-	weak
49	Saccharopine dehydrogenase (putative) (SCCPDH), mRNA	NM_016002.2	+	-	weak
50	SEC23 interacting protein (SEC23IP), transcript variant 1, mRNA	NM_007190.3	+	-	weak
51	Chromosome 7 genomic contig, GRCh37.p13 Primary Assembly	NT_007933.15	+	-	weak
52	CD93 molecule (CD93), mRNA	NM_012072.3	-	-	weak
53	Chromosome 3 genomic contig, GRCh37.p13 Primary Assembly	NT_005612.16	-	-	weak
54	Zinc finger protein 227 (ZNF227), mRNA	NM_182490.1	+	-	weak
55	Chromosome 3 genomic contig, GRCh37.p13 Primary Assembly	NT_005612.16	-	-	weak
56	Leo1, Paf1/RNA polymerase II complex component, homolog (S. cerevisiae) (LEO1), mRNA	NM_138792.2	+++	++	strong
57	Chromosome 4 genomic contig, GRCh37.p13 Primary Assembly	NT_016354.19	-	-	weak
58	Family with sequence similarity 76, member A (FAM76A), transcript variant 1, mRNA	NM_001143912.1	-	-	weak
59	Charged multivesicular body protein 5 (CHMP5), transcript variant 1, mRNA	NM_016410.5	+++	+++	strong

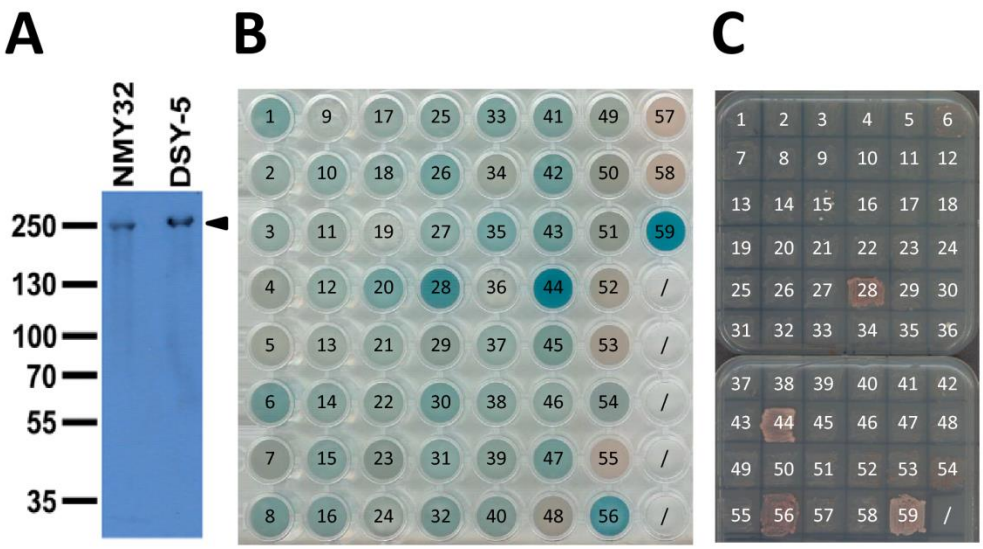

Figure S1. Summary of the yeast two hybrid screen. The bait construct was made by subcloning the *CSB* cDNA into the vector pLexA-DIR. Following transformation of the bait construct using standard procedures, correct expression of the bait fusion protein was verified by western blotting of cell extracts (from the two yeast strains indicated) with a mouse monoclonal antibody directed against the LexA DNA binding domain (panel A). The absence of self-activation was verified by co-transformation of the bait plasmid with a control prey and demonstration of the lack of histidine/adenine prototrophy and β-galactosidase activity. For the yeast two-hybrid screen, the bait was co-transformed together with a normalized human universal cDNA library into the strain NMY32 (MATa his3Δ200 trp1-901 leu2-3,112 (lexAop)8-ADE2 LYS2::(lexAop)4-HIS3 URA3::(lexAop)8-lacZ GAL4). A total of 2.5 x 10^6 transformants were screened, yielding 59 transformants that grew on minimal medium lacking the amino acids tryptophan, leucine and histidine (selective medium). Library plasmids were isolated from positive clones, and the identity of positive interactors was determined by sequencing (summarized in Table S2). Positive transformants were then tested for β-galactosidase activity using a PXG β-galactosidase assay. Forty-six of the 59 initial positives showed β-galactosidase activity and were considered real positives (panel B). As a means of evaluating the strength of the interactions, growth was determined in the absence of both histidine and adenine following plasmid co-transformation; seven of the total interactors were shown to yield a positive result under these more stringent conditions (panel C). See Table S2 for complete summary.

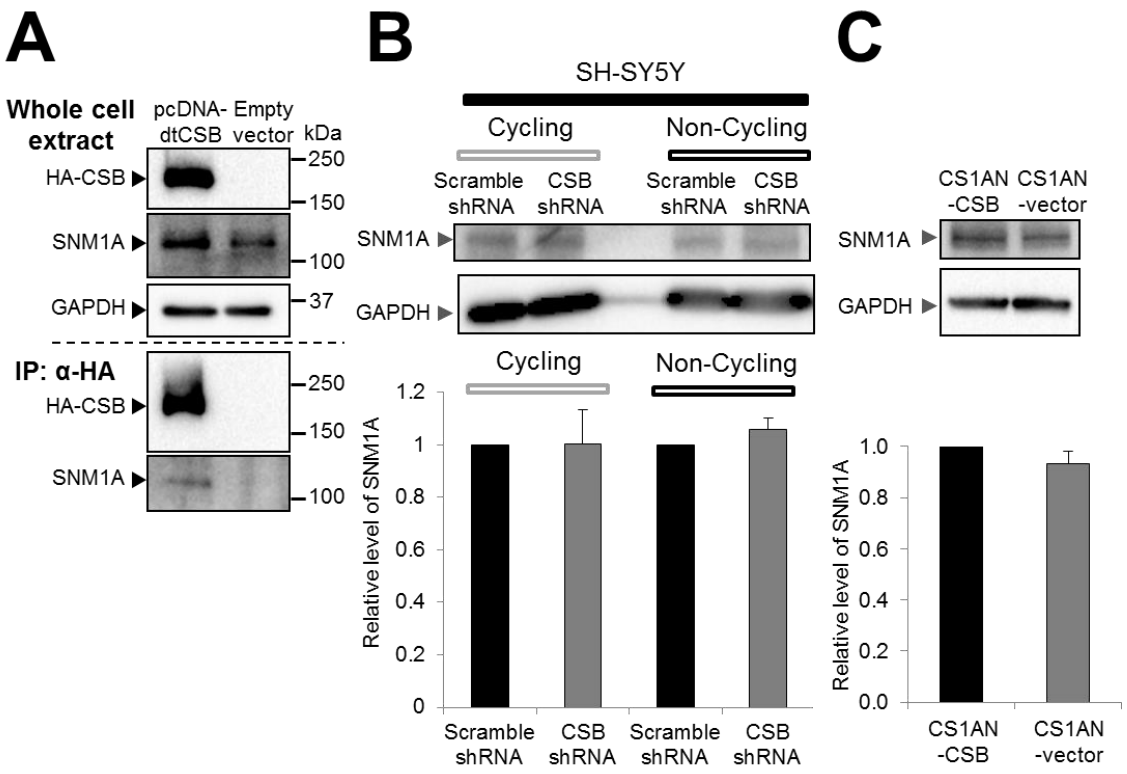

Figure S2. SNM1A protein co-purifies with CSB from human whole cell extracts, but its steady-state level is not affected by CSB deficiency. (A) Co-immunoprecipitation of double tagged-CSB (dtCSB) and SNM1A. The pcDNA-dtCSB expression vector, which produces a N-terminal HA-tagged and C-terminal His-tagged protein, was created by amplifying the *CSB* coding region using primers dtCSB BamHI-Fw and dtCSB XhoI-Rv (Table S1), and subcloning the digested product into the BamHI and XhoI restriction sites of pcDNA3.1(+) (Life Technologies). Approximately 24 hr after transfection of either the pcDNA-dtCSB or the pcDNA empty vector into HeLa cells, whole cell extracts were prepared as described in Materials and Methods. Prior to α-HA immunoprecipitation, the soluble whole cell extract was pre-treated with protein A/G magnetic beads at 4°C for 2 hr to remove nonspecific protein binders. Proteins were subsequently captured using α-HA magnetic beads at 4°C for 2 hr. The bead-bound material was washed five times, suspended in 2 x SDS-polyacrylamide gel electrophoresis loading dye, and incubated at 95°C for 5 min. Proteins were resolved on a polyacrylamide gel and detected by western blotting. The positions of the dtCSB, SNM1A and GAPDH proteins are shown, as are the protein standards in kDa. Expression of SNM1A protein in (B) cycling and non-cycling scramble shRNA or CSB shRNA SH-SY5Y cells or in (C) CS1AN-CSB or CS1AN-vector cells. SNM1A protein level was determined in whole cell extracts by western blotting, in comparison to GAPDH. The relative level of SNM1A to GAPDH is shown (average and standard deviation of three independent experiments), with the scramble control cell line or CS1AN-CSB being set as 1.

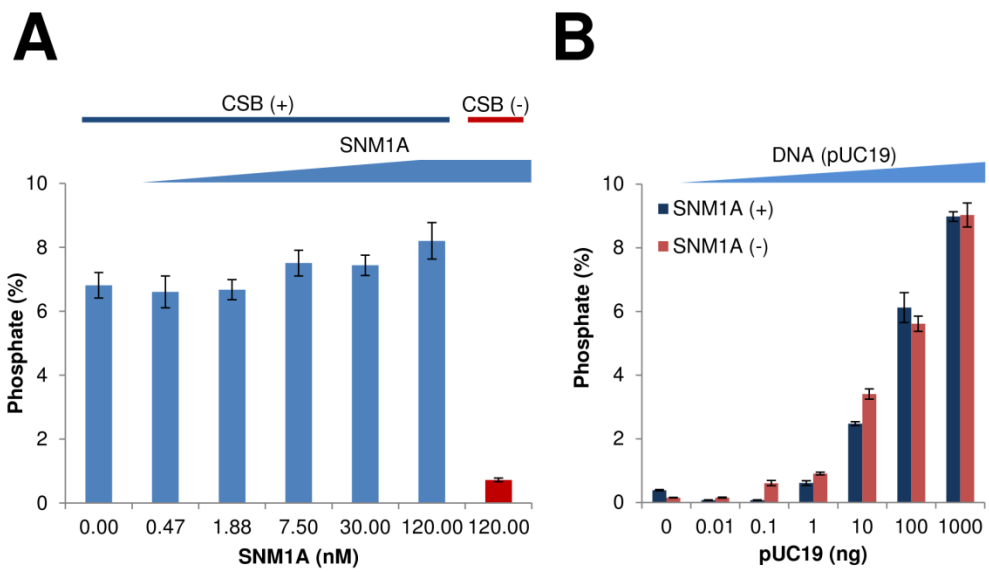

Figure S3. SNM1A does not affect the DNA-dependent ATPase activity of CSB. (A) ATP hydrolysis by CSB with the indicated amount of ΔN-SNM1A. ATPase reactions (10 µL) consisted of 20 mM HEPES pH 8.0, 4 mM $MgCl_2$, 2 mM ATP, 40 µg/mL BSA, 1 mM DTT, 300 ng pUC19, 50 ng (~30 nM) CSB and the indicated amount of ΔN-SNM1A (0-120 nM). Each reaction was performed for 1 hr at 30°C, and stopped by the addition of 5 µL of 0.5 M EDTA. CSB ATPase activity was measured using an inorganic phosphate (Pi) release assay with the Biomol Green phosphate assay reagent (Enzo Life Sciences, Loerrach, Germany). The percentage of free phosphate released from ATP was quantified by the addition of the Biomol Green reagent and reading of the absorbance at 620 nm after a 30 min incubation at 25°C. (B) ATP hydrolysis by CSB with indicated amount of DNA substrate. ATPase activity with various amounts of pUC19 was determined as above in the presence of 15 ng (~30 nM) ΔN-SNM1A. Each data point represents mean ± SD from triplicate experiments.

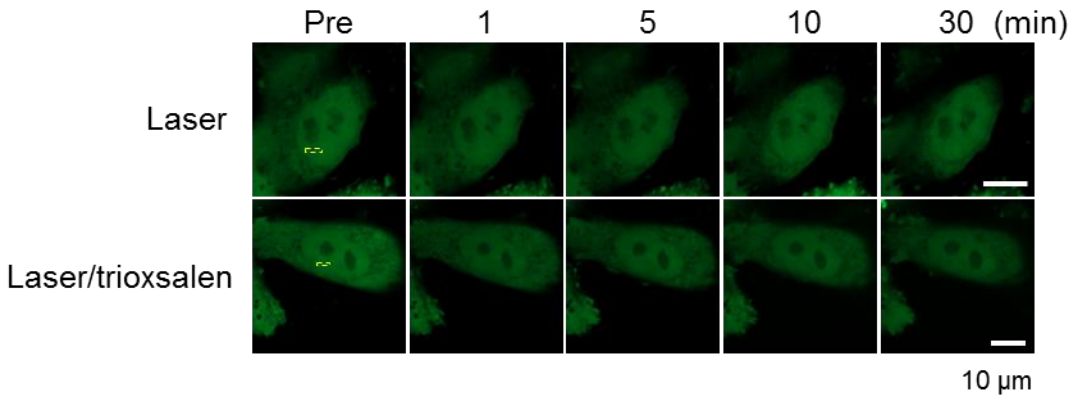

Figure S4. GFP alone does not respond to localized DNA damage. HeLa cells were transfected with control vector (pGFP), and after a 20-24 hr incubation, the cells were exposed (or not) to 6 μM trioxsalen for 30 min prior to targeted microirradiation (area designated by yellow box). Representative images of GFP at 0, 1, 5, 10 and 30 min after targeted laser irradiation are shown. Bar; 10 μm.

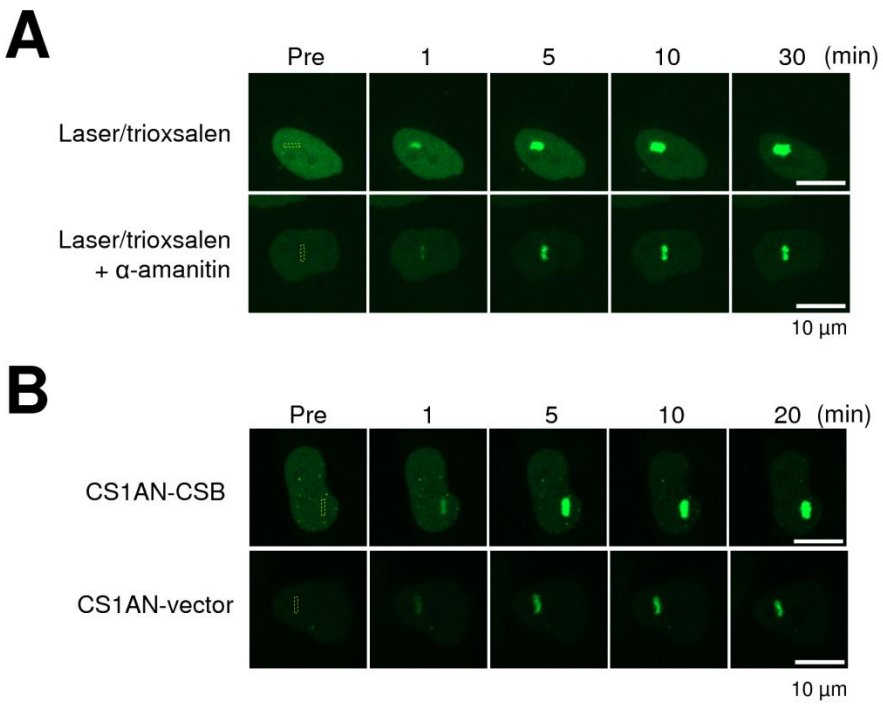

Figure S5. Representative images of SNM1A recruitment to site-specific ICL damage. (A) HeLa cells were transfected with pSNM1A-GFP, and after a 20-24 hr incubation, the cells were exposed to 6 µM trioxsalen for 30 min prior to targeted microirradiation (area designated by yellow box). Transfected cells were simultaneously treated with 20 µM α-amanitin during the 30 min incubation where indicated. Representative images show SNM1A-GFP accumulation at 0, 1, 5, 10 and 30 min after targeted laser irradiation. (B) CS1AN fibroblasts complemented with either the control vector (CS1AN-vector) or a CSB-expression plasmid (CS1AN-CSB) were transfected with pSNM1A-GFP and incubated with 6 µM trioxsalen for 30 min prior to targeted microirradiation as above (area designated by yellow box). Representative images show SNM1A-GFP accumulation at 0, 1, 5, 10 and 20 min after targeted laser irradiation. Bar; 10 µm.

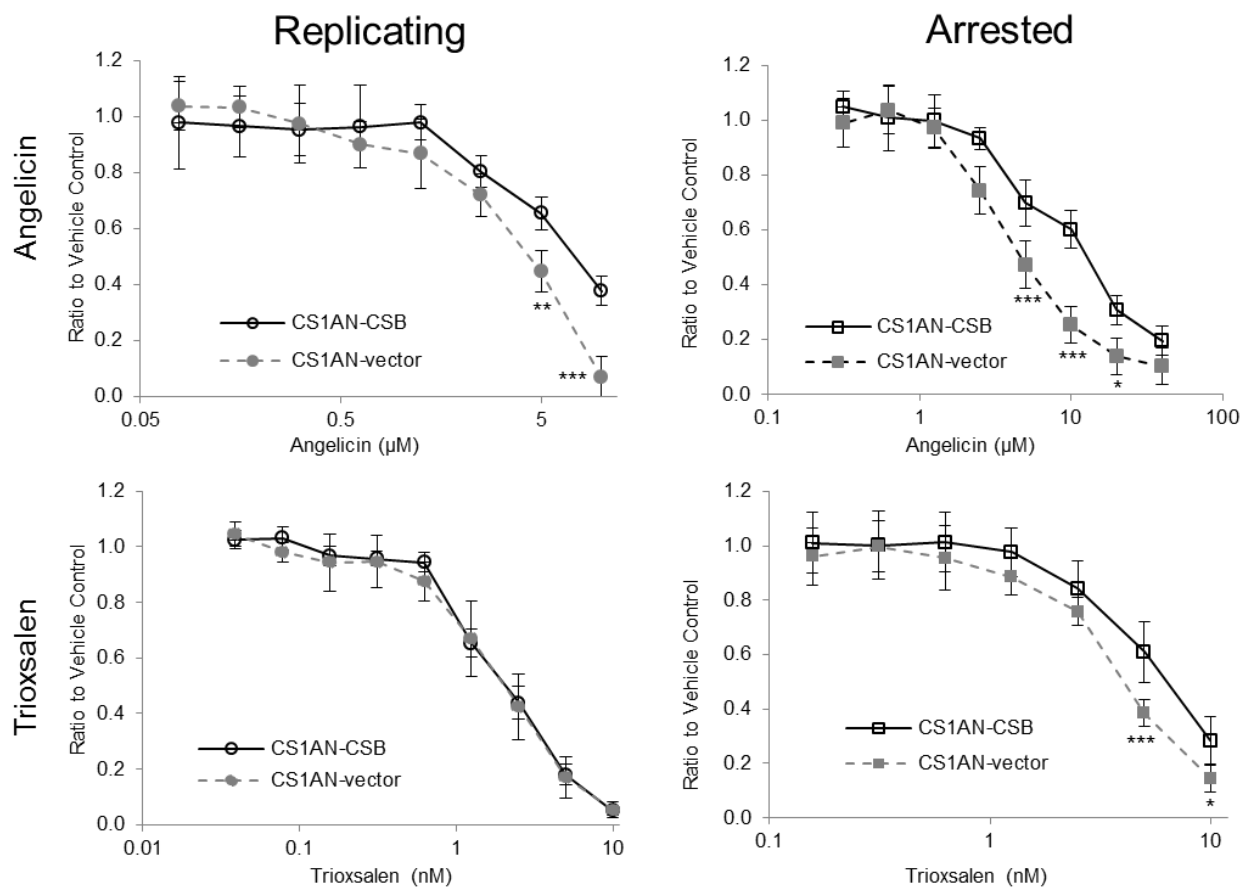

Figure S6. Sensitivity to angelicin and trioxsalen of replicating and arrested CS1AN mutant and complemented cells. Viability of replicating or serum starvation-arrested CS1AN-vector and CS1AN-CSB cells treated with either UVA/angelicin (replicating cells, 0.08-10 μM; arrested cells, 0.3-40 μM) or UVA/trioxsalen (replicating cells, 0.04-10 nM; arrested cells, 0.15-10 nM). Viability was measured using a WST-8 assay 72 hr after treatment. Results are presented as the ratio of the value obtained at the indicated concentration of angelicin or trioxsalen relative to that of the untreated (UVA only) cells. The data represent the mean ± SD from three independent experiments. The statistical analysis of each graph was performed by two way ANOVA followed by the Bonferroni post-hoc test (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0001$).

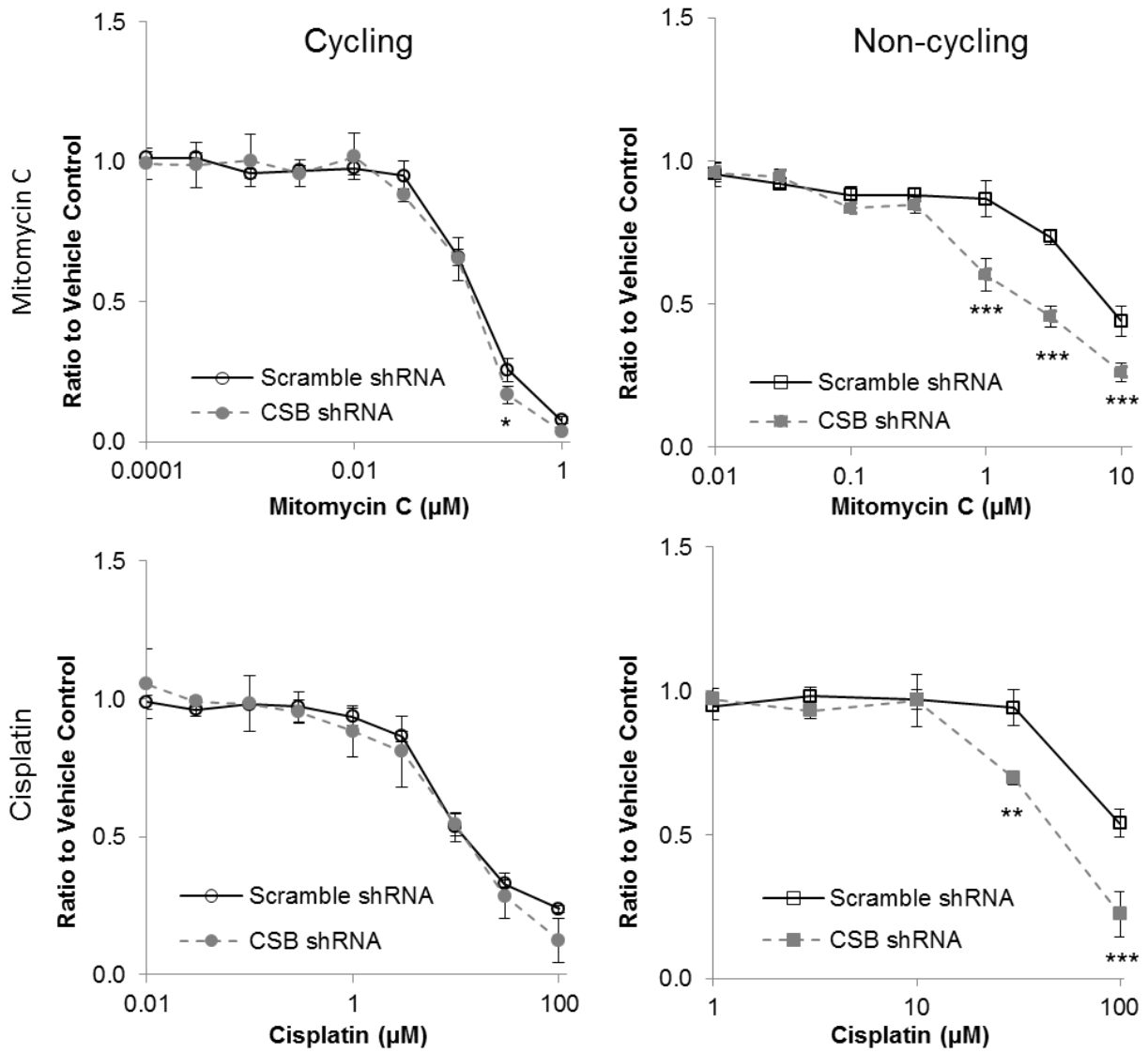

Figure S7. Sensitivity of cycling and non-cycling scramble shRNA and CSB shRNA SH-SY5Y cells to MMC and cisplatin. Viability of cycling and non-cycling scramble shRNA and CSB shRNA SH-SY5Y cells treated with MMC (cycling cells; 0.0001-1 μM , non-cycling cells; 0.01-10 μM) or cisplatin (cycling cells; 0.01-100 μM , non-cycling cells; 1-100 μM). Viability was measured using a WST-8 assay at 72 hr after treatment. Results are presented as the ratio of the value obtained at the indicated concentration of MMC or cisplatin relative to that of the untreated cells. The data in each graph represent the mean ± SD from three independent experiments. Statistical analysis was performed by two-way ANOVA followed by the Bonferroni post-hoc test (*$p < 0.05$, **$p < 0.001$, ***$p < 0.0001$).

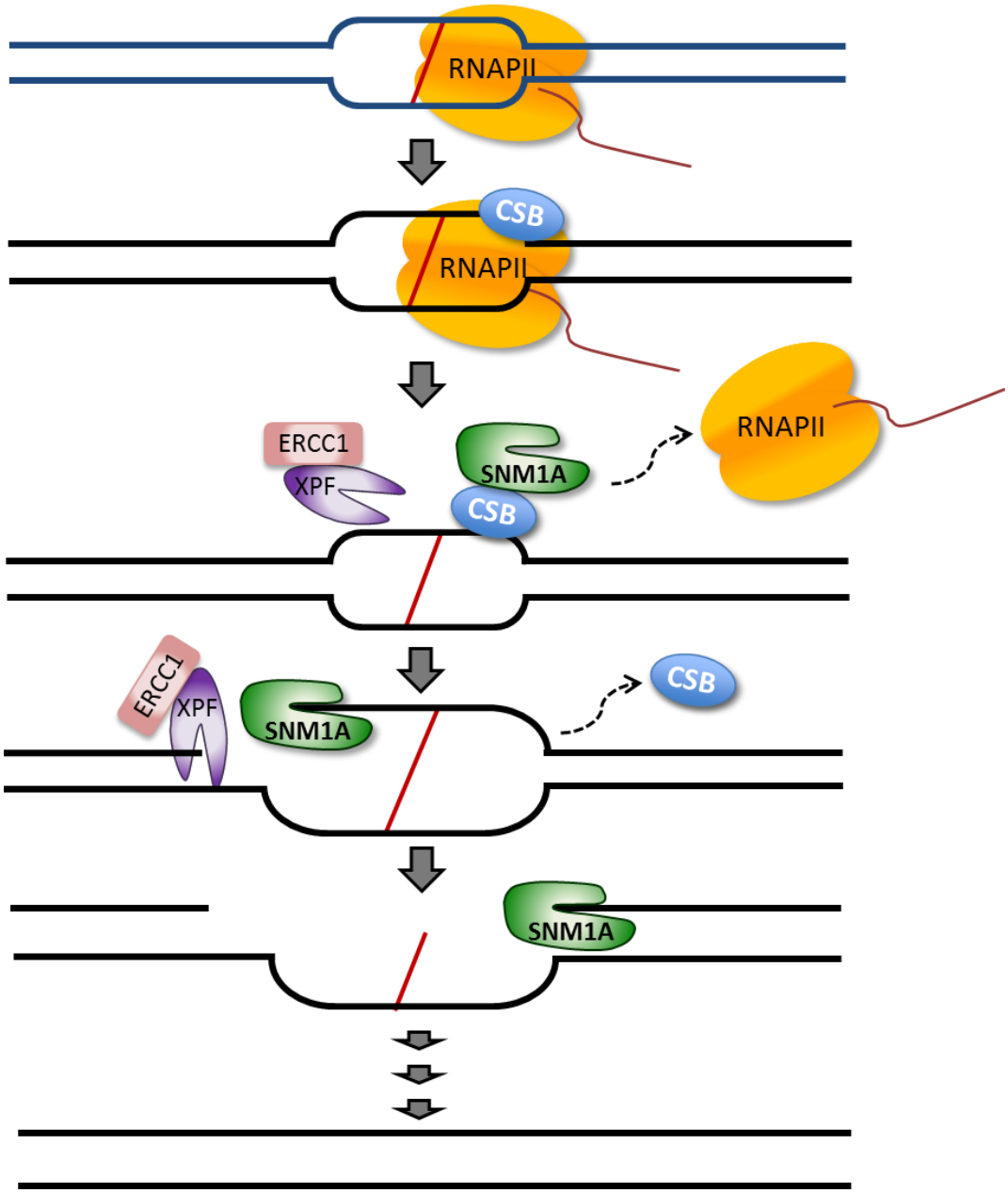

Figure S8. Proposed model for transcription-associated ICL repair involving CSB and SNM1A. Once RNAP II encounters a DNA ICL, a response is initiated to recruit various transcription-associated repair factors, such as CSB. CSB reorganizes RNAP II at the site of the lesion and facilitates assembly of repair nucleases, such as SNM1A. ERCC1/XPF participates in ICL removal by incising 5' of the damage. This nick serves as a substrate for the 5' to 3' exonuclease activity of SNM1A, which unhooks the ICL by degrading past the covalent linkage. The gap is filled by a translesion DNA polymerase and sealed by a ligase, prior to a second round of repair, likely classic nucleotide excision repair, to remove the crosslink remnant.