Appendix

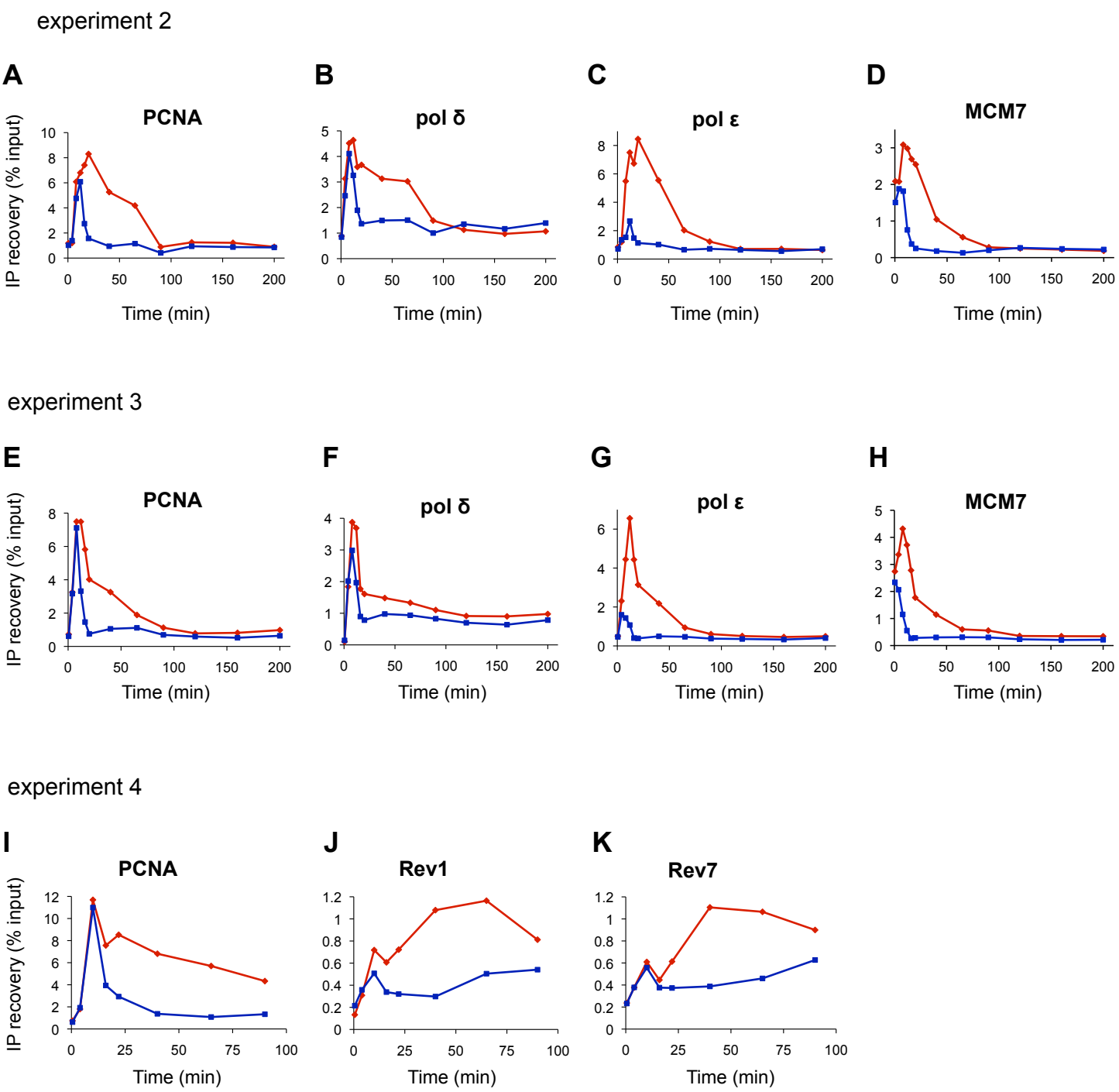
Table of contents:

Supplementary Figures 1-9

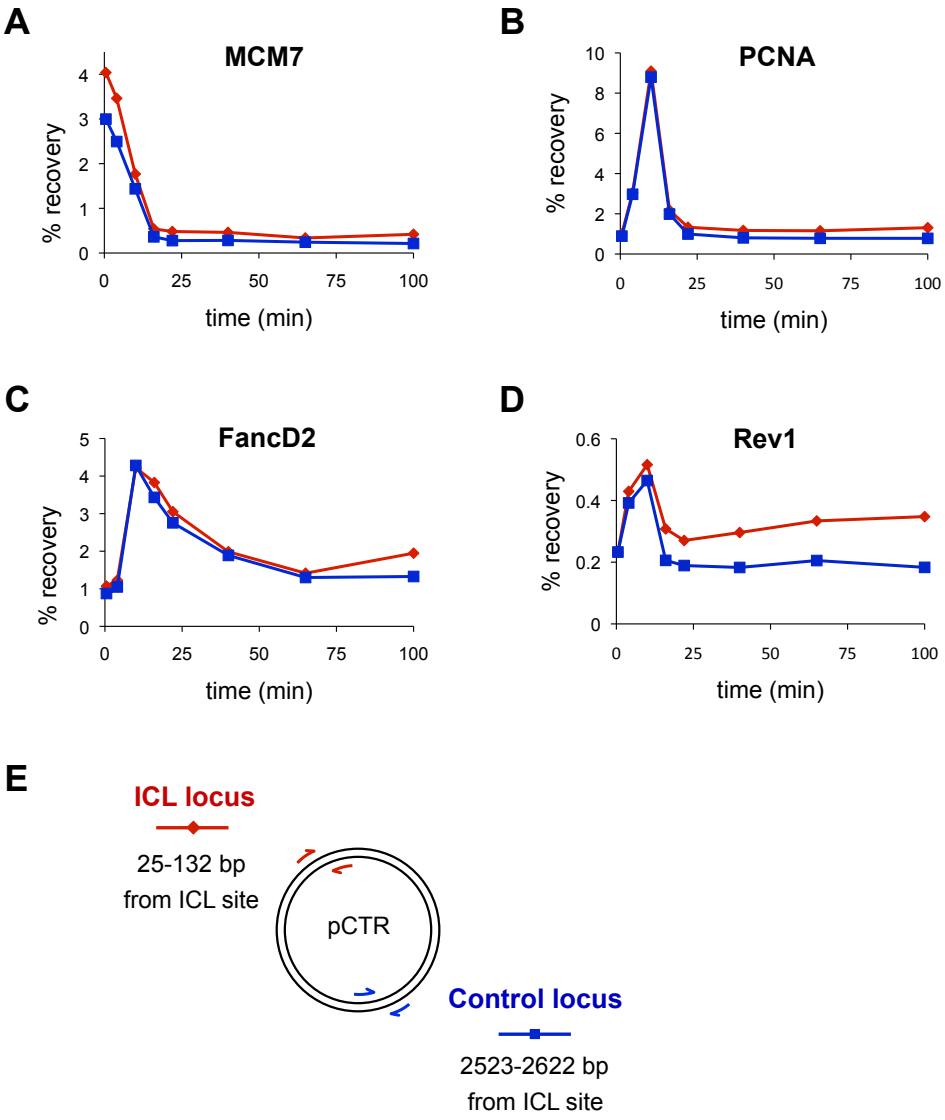
Supplementary Figure legends

Supplementary Table S1

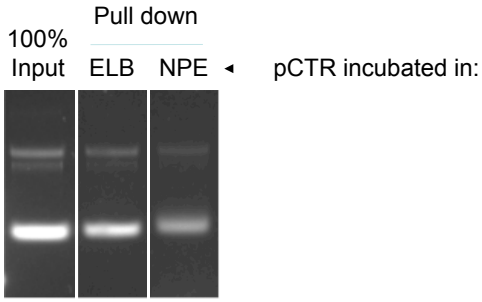
Supplementary Figure S1



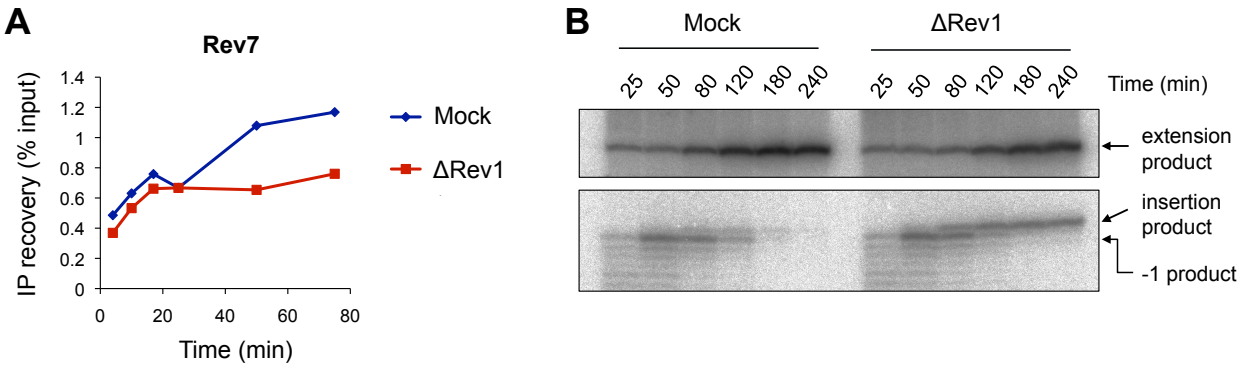
Supplementary Figure S2



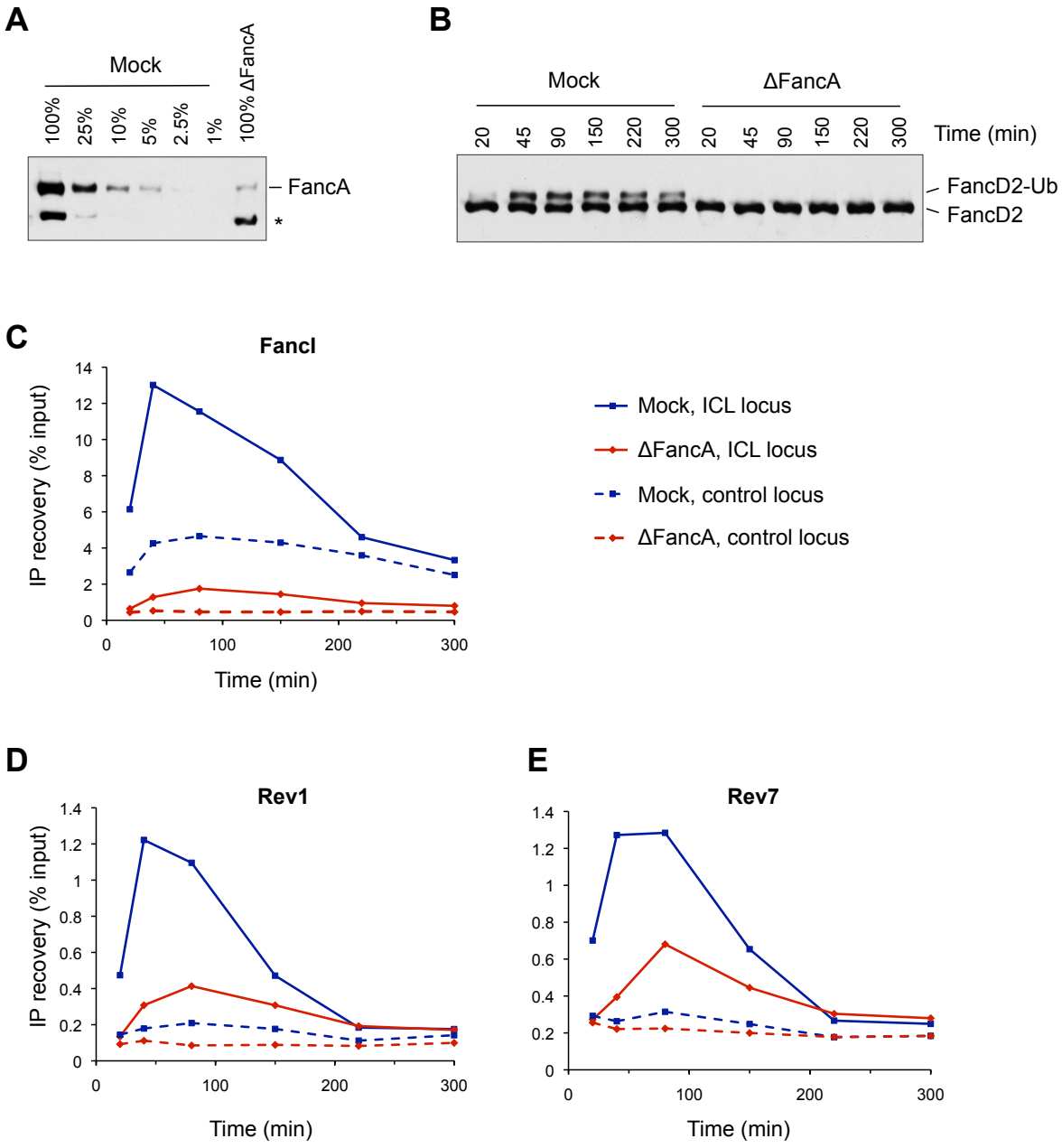
Supplementary Figure S3



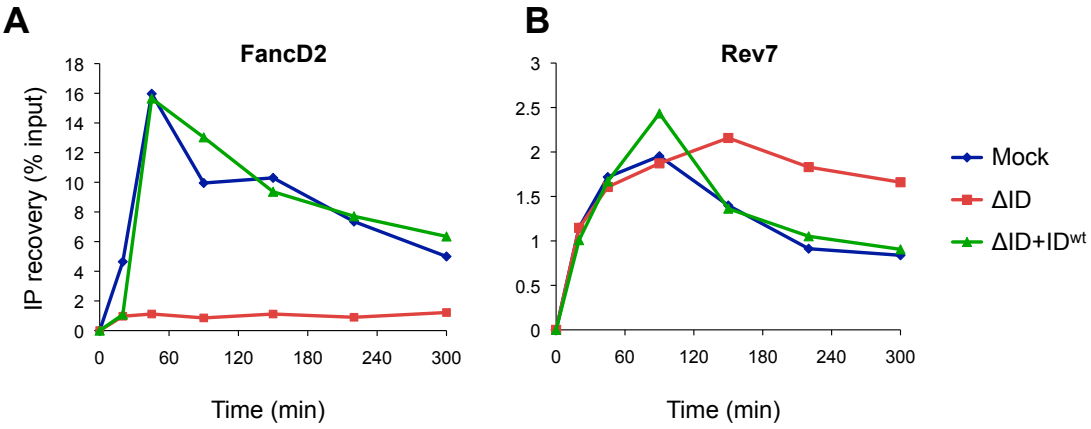
Supplementary Figure S4



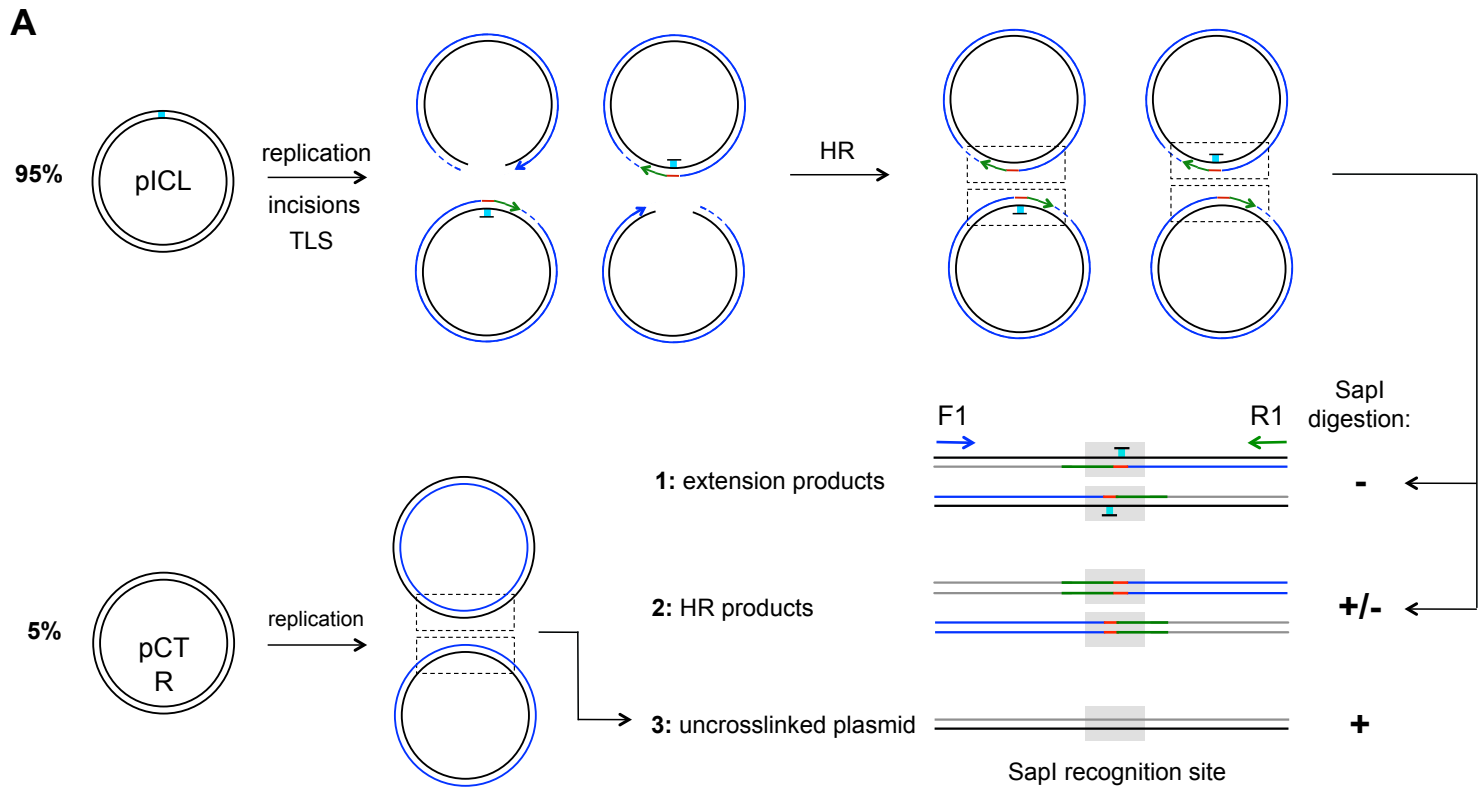
Supplementary Figure S5

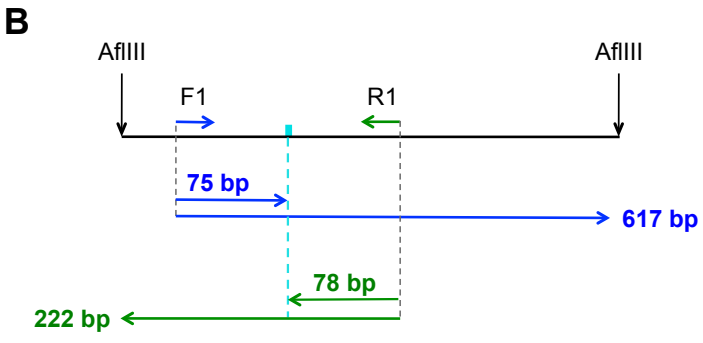


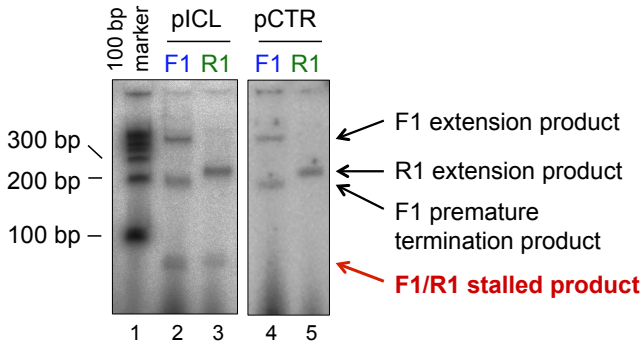
Supplementary Figure S6



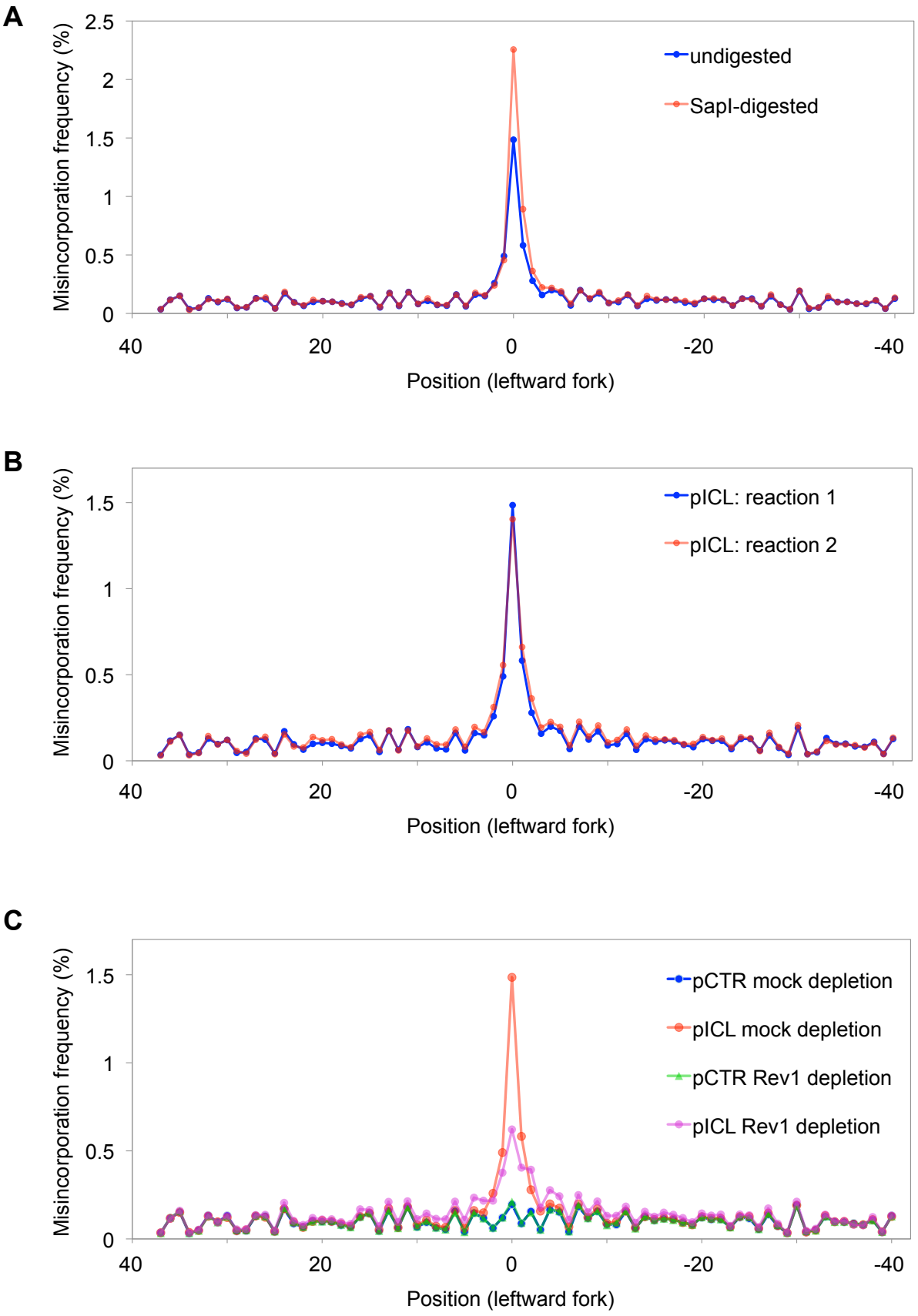
Supplementary Figure S7





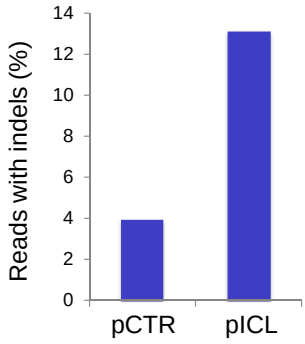


Supplementary Figure S8



Supplementary Figure S9

A



B

8 nt deletion: CAATGCATGCGGCCGCGAAGACAGCCCTCTTC TTTCGTGCGCGGCCGCGATCCGCTGCATTA

original sequence: CAATGCATGCGGCCGCGAAGACAGCCCTCTTCCGCTCTTCTTTCGTGCGCGGCCGCGATCCGCTGCATTA

SapI recognition site

Supplementary figure legends

Supplementary Figure S1. Replicates of experiments shown in primary Figure 2

A-K. Experimental replicates of the ChIP experiments shown in Figures 2D-K.

Supplementary Figure S2. Binding of replication and repair proteins to pCTR.

A-C. pCTR (undamaged plasmid of the same sequence as pICL) was replicated, and samples withdrawn at the indicated time points were analyzed by ChIP with antibodies against MCM7 **(A)**, PCNA **(B)**, FancD2 **(C),** and Rev1 **(D)**.

D. Location of primer pairs used for ChIP relative to the ICL site; the ICL is not present in pCTR.

Supplementary Figure S3. Isolation of chromatin using LacI-coated beads

pCTR was incubated in ELB buffer or in NPE and then mixed with an aliquot of LacI beads. The beads were recovered and washed. DNA was eluted, resolved on a native agarose gel, and stained with SYBR Gold.

Supplementary Figure S4. Replicate of experiments shown in primary Figure 4

A. Independent replicate of the experiment shown in Figure 4D.

B. Independent replicate of the experiment shown in Figure 4F using the same Rev1-depleted extract as in panel (A).

Supplementary Figure S5. Effect of FancA depletion on recruitment of TLS polymerases to the ICL locus

A. FancA immunodepletion. FancA-depleted NPE and a dilution series of mock-depleted NPE were analyzed by Western blotting using FancA antibodies. 100% corresponds to 0.25 μl of NPE. Asterisk, non-specific band.

B. pICL was replicated in mock- or FancA-depleted egg extract. At different times, protein samples were blotted for FancD2. The positions of unmodified (FancD2) and ubiquitylated FancD2 (FancD2-Ub) are indicated.

C-E. pICL was replicated in mock- or FancA-depleted egg extract and ChIP was performed at the ICL and control loci using antibodies to FancI (C), Rev1 (D), and Rev7 (E).

Supplementary Figure S6. Replicate of experiments shown in primary Figure 6

A-B. Independent replicate of the experiments shown in Figures 6C and 6E, except that in (A), ChIP was performed with antibodies against FancD2 instead

of FancI.

Supplementary Figure S7. Validation of deep sequencing strategy

A. Strategy for sequencing of DNA repair products. To define the mutation spectrum generated during pICL repair, pICL was allowed to undergo repair for four hours, which gives rise to adducted extension products (Species 1) and unadducted HR products (Species 2) (Räschle *et al*, 2008; Long *et al*, 2011). In addition, because all preparations of pICL contain ~5% undamaged DNA, undamaged replication product is also present (Species 3). The major species, and the one we were most interested in, is species 1 since it should contain the mutations generated during TLS. The HR products (Species 2) should copy any TLS-induced mutations from the un-adducted strand of species 1, and may also contain de novo TLS-induced mutations if D-loop extension uses the adducted strand as a DNA template. The sample containing these three species was subjected to PCR with primers F1 and R1 to amplify a 115 bp region surrounding the ICL. The resulting fragments were sequenced using a MiSeq sequencer.

B. Scheme to determine whether the PCR polymerase stalls at the unhooked ICL. We wanted to determine whether the *adducted* strand of species 1 in panel (A) could be amplified. If so, it might lead to mutations during PCR that mask the mutations generated during TLS in the extract. To address this question, we cut replicated pCTR or pICL with AflIII, performed primer-extension with either forward (F1) or reverse (R1) primers labeled at the 5' end with [γ^{32}P]ATP, and resolved the products on a denaturing gel. The length of the expected full length

and stalled products is indicated.

C. Results of the experiment described in (B). When pCTR was used as the template for primer extension with Phusion High-Fidelity DNA polymerase (NEB), we detected the expected full-length extension products for the forward F1 (617 nt) and reverse R1 (222 nt) primers (lanes 4 and 5). Unexpectedly, the F1 reaction with pCTR also generated a smaller product of ~200 nt (lane 4), which was due to premature termination. Importantly, when we used replicated pICL as the template, we observed the expected 75 and 78 nt stall products for the F1 and R1 primers (lanes 2 and 3), indicating that the polymerase was blocked by the adduct. We also observed the full-length extension products, as expected given that the major species (species 1 in panel A) contains one adducted strand and one undamaged strand, and that the adduct is present on either the top or the bottom strand (only bottom adduct shown in panel A). Additionally, the template should contain unadducted HR products (panel A), as well as a small amount of undamaged replication products, both of which would give rise to full-length extension products. The ratio of stalled to full length products suggests that stalling is efficient. This result suggests that the parental strand containing the adduct cannot be efficiently amplified, and that only the unadducted strand generated during repair will be amplified and sequenced. The same pattern and efficiency of stalling at the ICL adduct was observed with KAPA HiFi DNA polymerase (KAPA Biosystems), which was used for the experiments presented. Equivalent sequencing results were obtained with Phusion High-Fidelity DNA polymerase and KAPA HiFi DNA polymerase.

Supplementary Figure S8. Effect of SapI digestion and Rev1 depletion on the misincorporation frequency.

A. Effect of SapI digestion on the misincorporation frequency. Our pICL preparations contain on average 5% undamaged plasmid, which will replicate efficiently in egg extract (Räschle *et al*, 2008). As these molecules will be amplified during the preparation of samples for deep sequencing, we wanted to assess how their presence affects the mutation frequency. To eliminate these molecules, we cut the plasmid with SapI (whose recognition site overlaps the ICL). Thus, digesting the final repair products with SapI will eliminate uncrosslinked background molecules and faithfully repaired HR products. We then amplified and sequenced SapI-cut products as in Figure 7A. SapI digestion increased the misincorporation frequency from 1.5% to 2.3%, demonstrating that ~35% of the sequenced products contained an intact SapI site, while the majority were likely translesion synthesis products resistant to SapI digestion due to a persistent mono-adduct on the opposite strand (species 1 in panel A) (Räschle *et al*, 2008).

B. Experimental replicate of Figure 7A. The experiment shown in Figure 7A was repeated, and the results (red circles) were graphed with the data from Figure 7A (blue circles).

C. Effect of Rev1 depletion on misincorporation frequency. pCTR and pICL were replicated in mock- and Rev1-depleted extracts. Replication products were recovered after 60 minutes for pCTR and 240 minutes for pICL. A 115

nt long fragment surrounding the crosslink (present in pICL only) was deep-sequenced. For both plasmids, the misincorporation frequency in a 40 bp region surrounding the ICL is displayed. Nucleotide positions for the leftward fork are indicated.

Supplementary Figure S9. ICL repair involves formation of a specific deletion product

A. pICL replication products contain more indels than sequences of pCTR replication products. pCTR and pICL replication products were isolated at 60 and 240 minutes, respectively, and the DNA surrounding the ICL was amplified and sequenced as described in Fig 7 and Fig S7. The percentages of pCTR and pICL reads containing indels between positions e 20 and +20 relative to the ICL are graphed.

B. The sequence of the most common deletion found in pCTR and pICL (8 nt deletion) compared to the original sequence. The guanine that forms the ICL is indicated in red. Two tandem 6 nt repeats surrounding the ICL are indicated in cyan. SapI recognition site is indiated by a grey box. We speculate that when the 3' end of the leading strand stalls at the ICL, the 6 nt repeat becomes unpaired and anneals to the downstream repeat, looping out the intervening DNA in the template strand and resulting in an 8 nt deletion.

Supplemental Table S1: Acceptance and rejection statistics for sequencing reads.

	Accepted reads		Rejected reads		Total
	Correct length (78 bp)	Incorrect length	Type 1	Type 2	
pCTR	1228340 (72%)	50649 (3%)	80944 (5%)	335533 (20%)	1695466
pICL (reaction 1)	1563478 (61%)	236191 (9%)	140596 (5%)	632937 (25%)	2573202
pICL (reaction 2)	156168 (51%)	46582 (15%)	18959 (6%)	84512 (28%)	306221
pCTR (Rev1 depletion)	1034125 (72%)	46821 (3%)	64025 (4%)	281724 (20%)	1426695
pICL (Rev1 depletion)	639375 (33%)	468351 (25%)	144794 (8%)	658905 (34%)	1911425
pICL (reaction 1, Sap1 digestion)	674538 (54%)	161331 (13%)	73456 (6%)	333061 (27%)	1242386

Type 1 rejection: The first read could not be matched to two 6-nt reference sequences on either end of the PCR product.

Type 2 rejection: The first read could be aligned using the 6-nt reference sequences but did not perfectly match the second (paired-end) read