SUPPLEMENTARY INFORMATION

Polymerase μ in Non-Homologous DNA End Joining: Importance of the Order of Arrival at a Double-Strand Break in a Purified System

Bailin Zhao[1], Go Watanabe[1], and Michael R. Lieber[1,*]

[1] Department of Pathology, Department of Biochemistry & Molecular Biology, Department of Molecular Microbiology & Immunology, and Section of Computational & Molecular Biology, USC Norris Comprehensive Cancer Center, University of Southern California Keck School of Medicine, 1441 Eastlake Ave, Rm. 5428, Los Angeles, CA 90089, USA.

Supplementary Result

The 0.3 FRET state corresponds to a synaptic complex containing two downstream duplexes.

The total intensity of donor and acceptor of the 0.3 FRET synaptic complex is almost twice that of the 0.6 FRET state, while the two states have almost the same level of Cy5 intensity (Supplementary Figure S1A). The result suggests that the synaptic complex with 0.3 FRET may contain two Cy3 labeled downstream duplexes. That is, three duplexes appear to be present in the 0.3 FRET complex. To further test the possibility of a three-duplex complex, the two dyes were switched on the two duplexes with the immobilized upstream one labeled with Cy3 and the incoming downstream one with Cy5. If the 0.3 FRET state corresponds to a real intermediate structure of the synaptic complex which contains only one incoming downstream duplex, the 0.3 FRET state should be also observed for this new set of DNA duplexes. But if the 0.3 FRET peak corresponds to the complex having two incoming duplexes, the 0.3 FRET state would not be observed for the new set of duplexes. Because of the difference from the complex containing two donor-labeled duplexes, the complex containing two acceptor-labeled duplexes has the same FRET value as that containing only one acceptor-labeled duplex. The FRET distribution shows that only a single FRET peak of 0.64 is observed for the new set of duplexes (Supplementary Figure S1C). The result indicates that a small number of synaptic complexes indeed have three duplexes present, one upstream duplex and two downstream duplexes.

Supplementary Discussion

DNA-PKcs is not essential for NHEJ synapsis.

Implicit in the results here and in our previous work (see Introduction), DNA-PKcs is not essential for NHEJ synapsis. As we have shown previously, smFRET studies claiming a role for DNA-PKcs use oxygen scavenger conditions which inactivate DNA-PKcs kinase activity and presumably denature it. We have used conditions where DNA-PKcs activity is retained.

Supplementary Figures

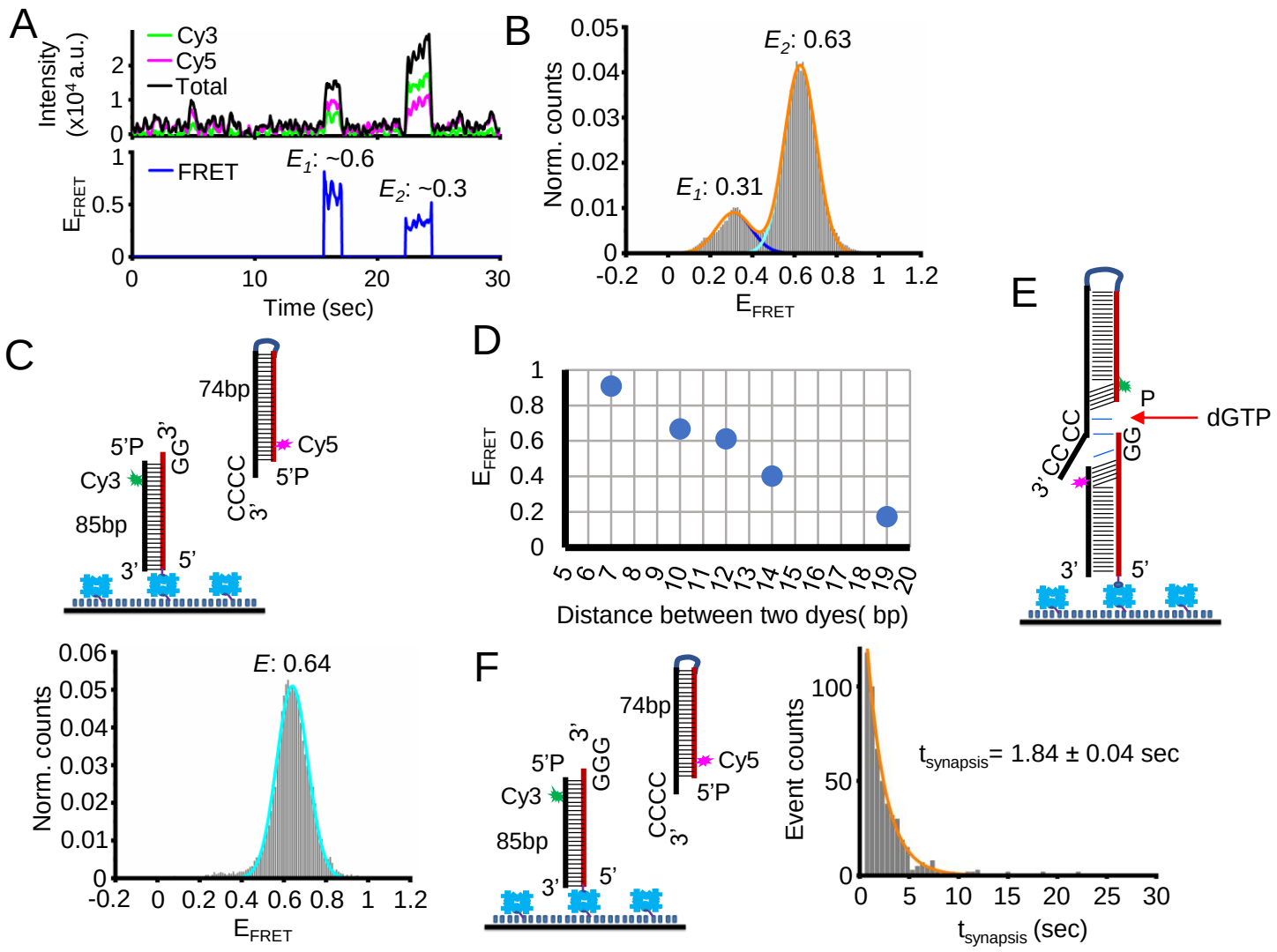

Figure S1. The 0.3 FRET state corresponds to a synaptic complex containing two downstream duplexes.

A, One of the single molecule time traces of donor (green) intensity, acceptor (magenta) intensity, total (black, donor + acceptor) intensity, and corresponding E_{FRET} values (blue) for synapsis mediated by 50 nM pol μ. The same DNA duplexes as shown in Figure 1**A** were used.

B, Histogram of E_{FRET} values of all synapsis events mediated by 50 nM pol μ. The same DNA duplexes as shown in Figure 1**A** were used. The *E* values shown on the graph were obtained from a Gaussian fit of the corresponding peaks. More than 400 molecules were included.

C, Top panel: Switching of the two dyes on the two duplexes. The immobilized duplex was labeled with Cy3 and the incoming duplex with Cy5 dye. Bottom panel: Histogram of E_{FRET} values of all synapsis events obtained using the new set of duplexes shown on the Top panel. The *E* value shown on the graph was obtained from a Gaussian fit of the corresponding peak. More than 500 molecules were included.

D, Distance dependent E_{FRET} measurement using our system. The distance between two dyes are 7, 10, 12, 14, and 19 bp. 10 – 12 bp distance between two dyes results in a FRET of ~0.6.

E, DNA configuration within the formed NHEJ synaptic complexes.

F, Histogram and corresponding single exponential fit of synapsis time stimulated by pol μ for the 3' 3Gs and 3' 4Cs overhangs. More than 500 synapsis events were included. Synapsis time shown on graph is represented as the mean ± SD of three independent replicates.

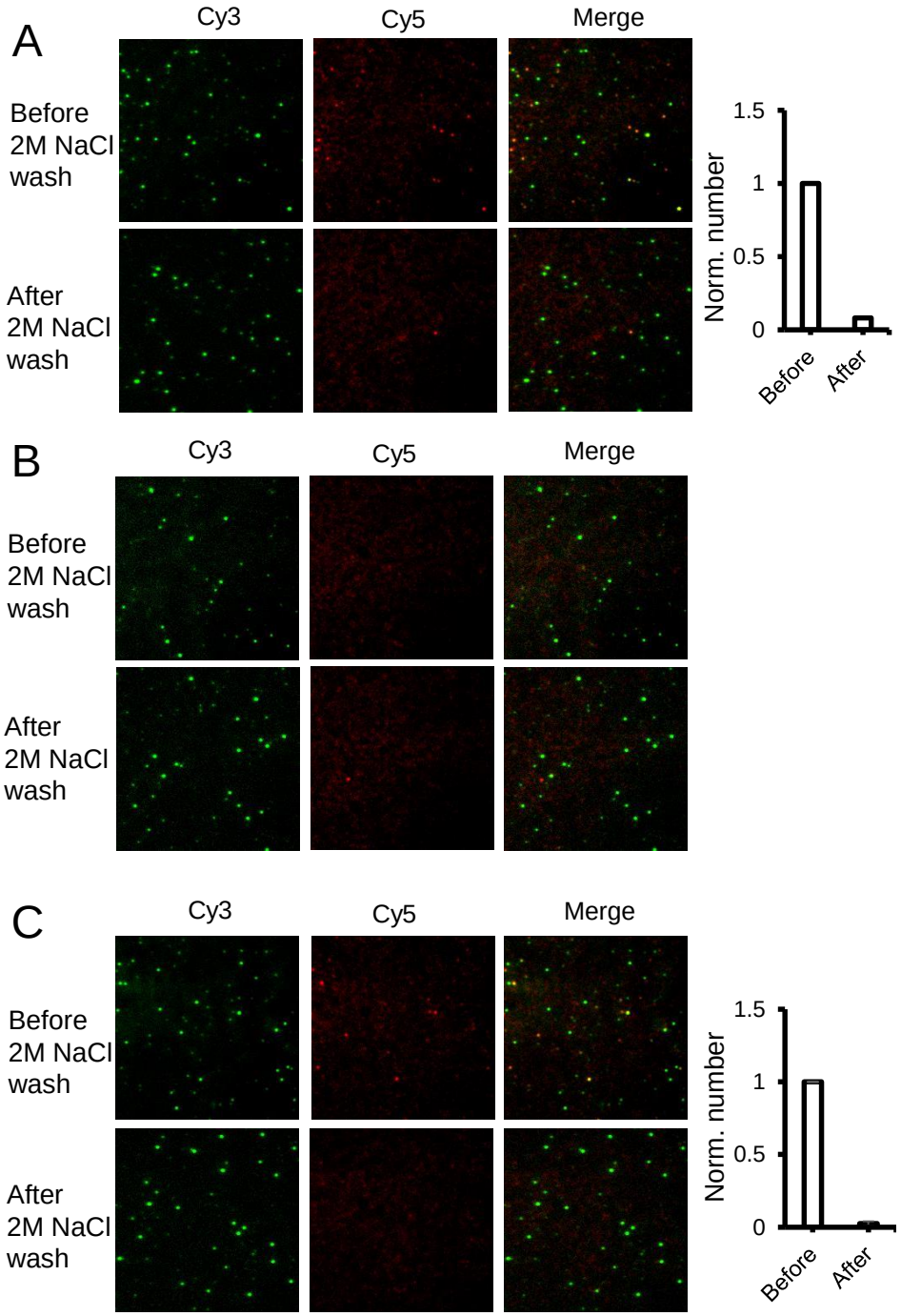

Figure S2. Covalent ligation of the two ends within the pol μ synaptic complex is polymerization-dependent.

A,**B**,**C**, Representative images of Cy3 channel, Cy5 channel, and overlay of two channels for NHEJ synaptic molecules before/after 2 M NaCl wash. The bar graph represents the corresponding quantified results of the reactions. The reaction contains 50 nM pol μ plus 100 μM dGTP (**A**) or 50 nM X4L4 plus 100 μM dGTP (**B**) or 50 nM pol μ plus 50 nM X4L4 (**C**). No nucleotide is present in the reaction of (**C**). Two independent repeats were performed for reaction shown in **C**. We used 2 M NaCl to remove any noncovalent associations of the overhangs, which leaves only the covalently ligated molecule remaining. Results in **A** confirm that 2 M NaCl can wash away the noncovalent association, because no ligase is present in the reaction. Results in **B** are used as a control for DNA substrate ligation merely by X4L4. Results in **C** confirm that pol μ polymerization is required for the two ends ligation within the synaptic complex.

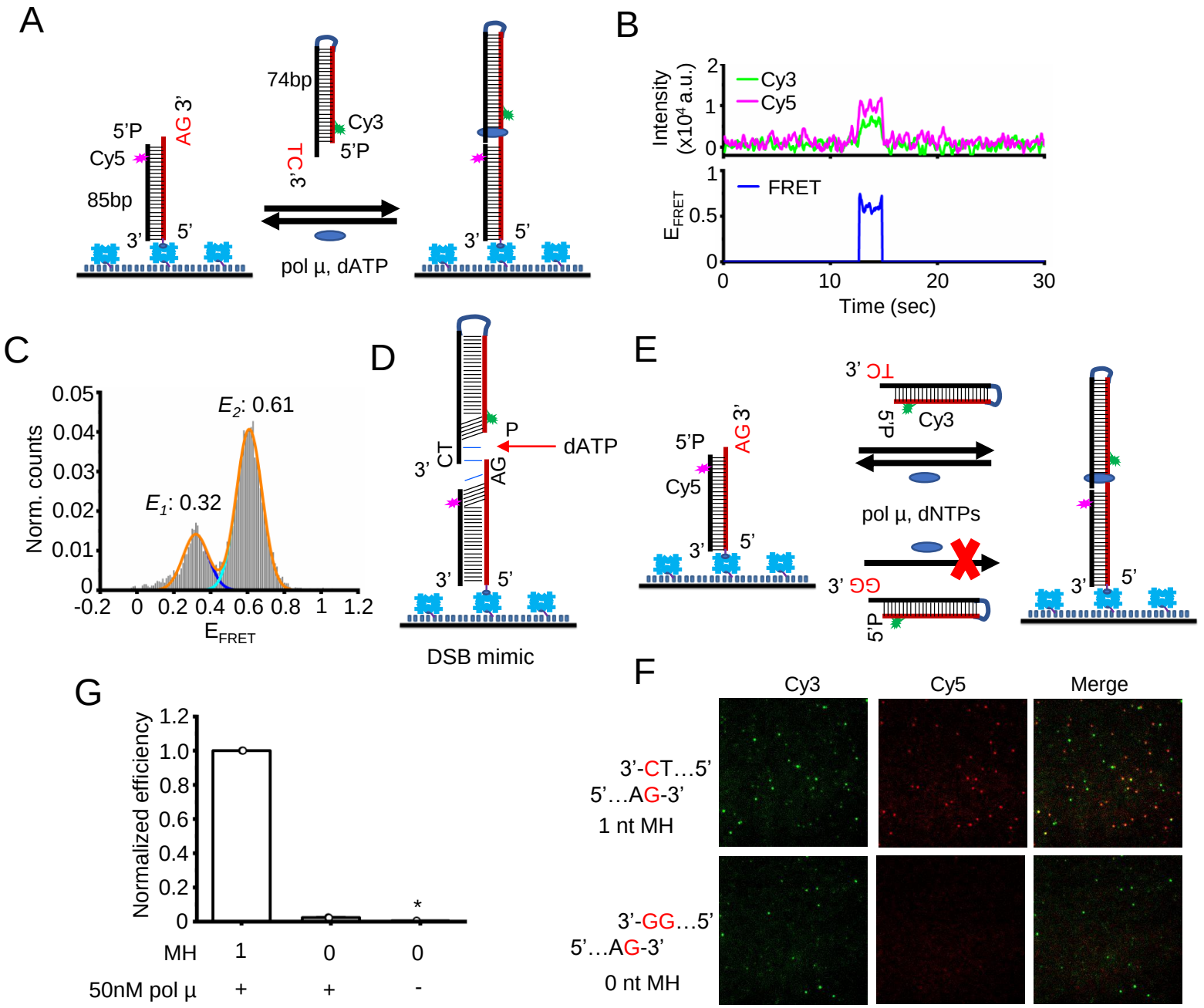

Figure S3. Microhomology is required for efficient NHEJ synapsis.

A, Schematic of smFRET assay for NHEJ synapsis of overhangs with 1 nt microhomology.

B, One of the single molecule time traces of donor (green) intensity, acceptor (magenta) intensity and corresponding E_{FRET} values (blue) for NHEJ synapsis of overhangs with 1 nt microhomology.

C, Histogram of E_{FRET} values of all synapsis events of the overhangs with 1 nt microhomology. The *E* values shown on the graph were obtained from a Gaussian fit of the corresponding peaks. More than 800 molecules were included.

D, DNA configuration within the formed NHEJ synaptic complex for duplexes with overhangs of 1 nt microhomology.

E, Schematic of smFRET assay for NHEJ synapsis of overhangs with or without microhomology (MH). The immobilized duplex has a 3' AG overhang, the solution duplex has either a TC (1 nt MH) or GG (0 nt MH) overhang. 50 nM pol μ and dNTPs (100 μM each) were included in the reactions.

F, Representative images of Cy3 channel, Cy5 channel, and overlay of the two channels for NHEJ synapsis of different sets of duplexes.

G, Normalized NHEJ synapsis efficiency for different sets of DNA duplexes. Data are represented as the mean ± SD of two independent replicates. MH represents the number of nucleotides of microhomology. * : only one sample was done.

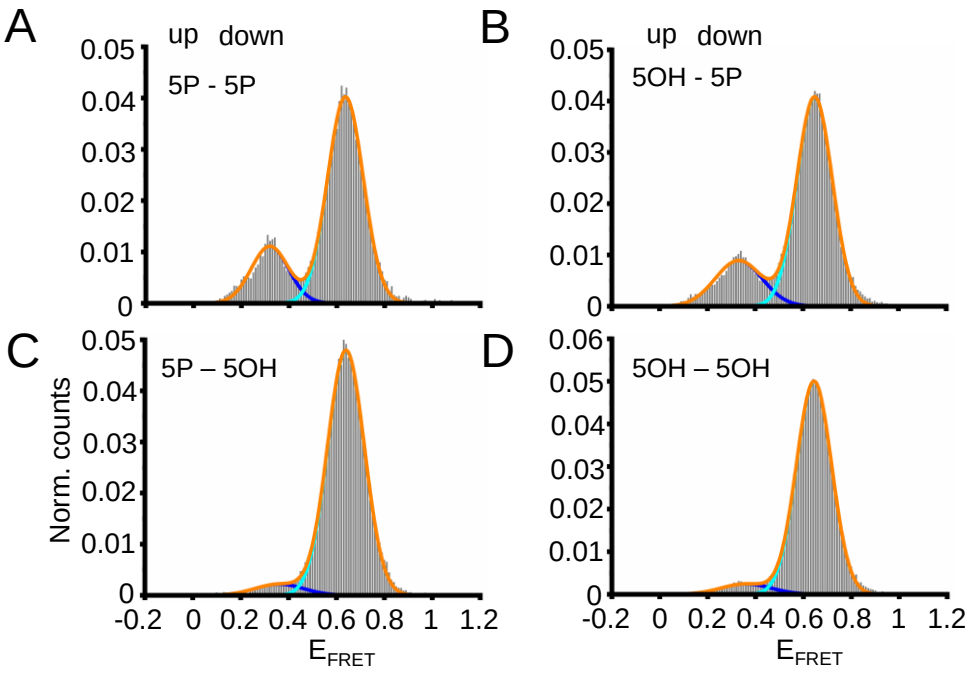

Figure S4. FRET distributions of NHEJ synaptic complexes formed with duplexes having different end chemistry.

The reactions included 50 nM pol μ and 100 μM dGTP. More than 1000 molecules were included for each FRET histogram.

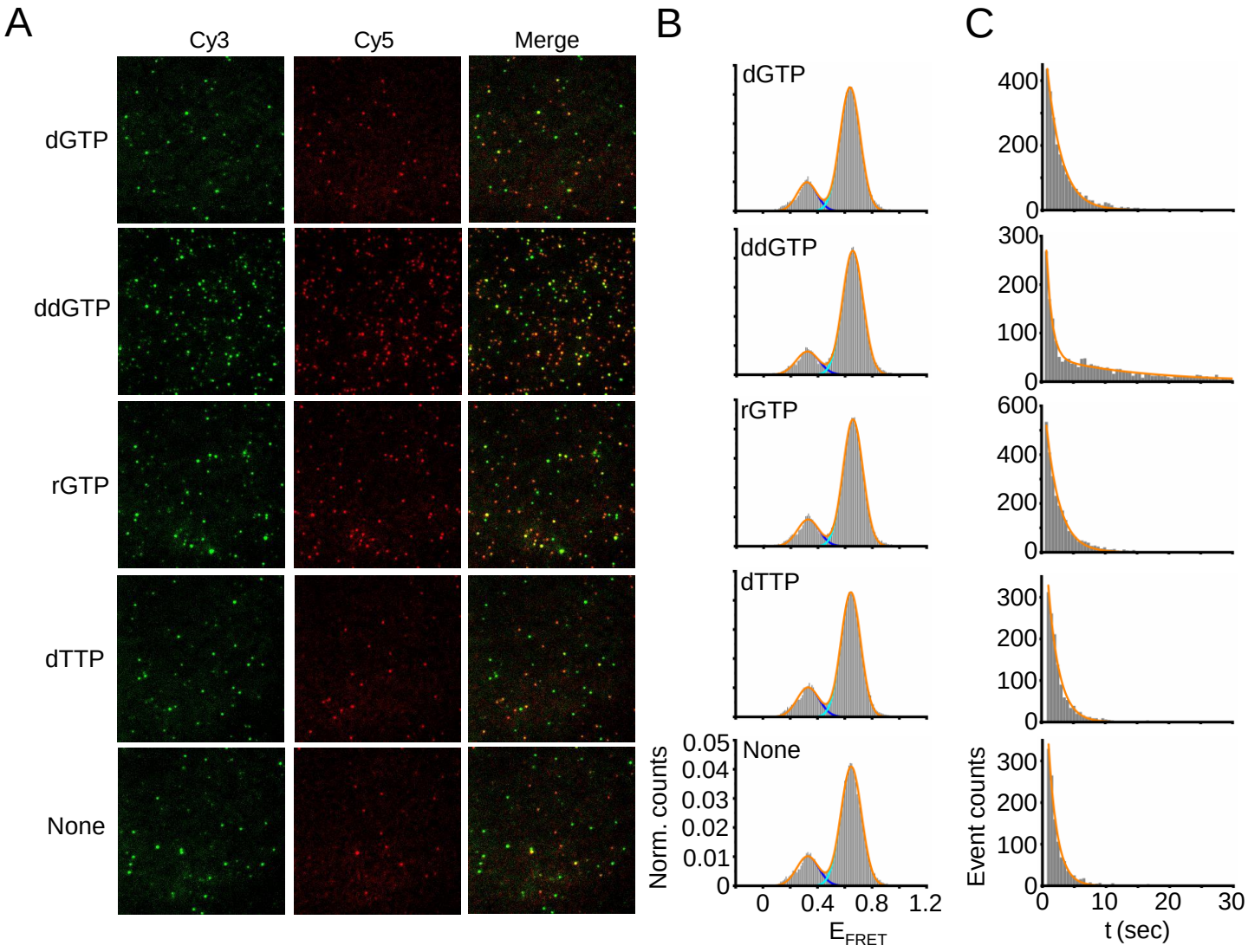

Figure S5. Effect of nucleotide on pol μ-mediated NHEJ synapsis.

A, Representative images of Cy3 channel, Cy5 channel, and overlay of the two channels for NHEJ synapsis stimulated by different nucleotide factors. The quantified efficiency is shown in Figure 4**A**.

B, Histogram of E_{FRET} values of all synapsis events stimulated by different nucleotides. Over 1000 molecules were included for each histogram.

C, Histogram and corresponding exponential fit of synapsis time stimulated by different nucleotide factors. At least 1000 synapsis events were included for each histogram. Histograms except that for ddGTP were fitted by a single exponential function. The histogram for ddGTP was fitted by a bi-exponential function,

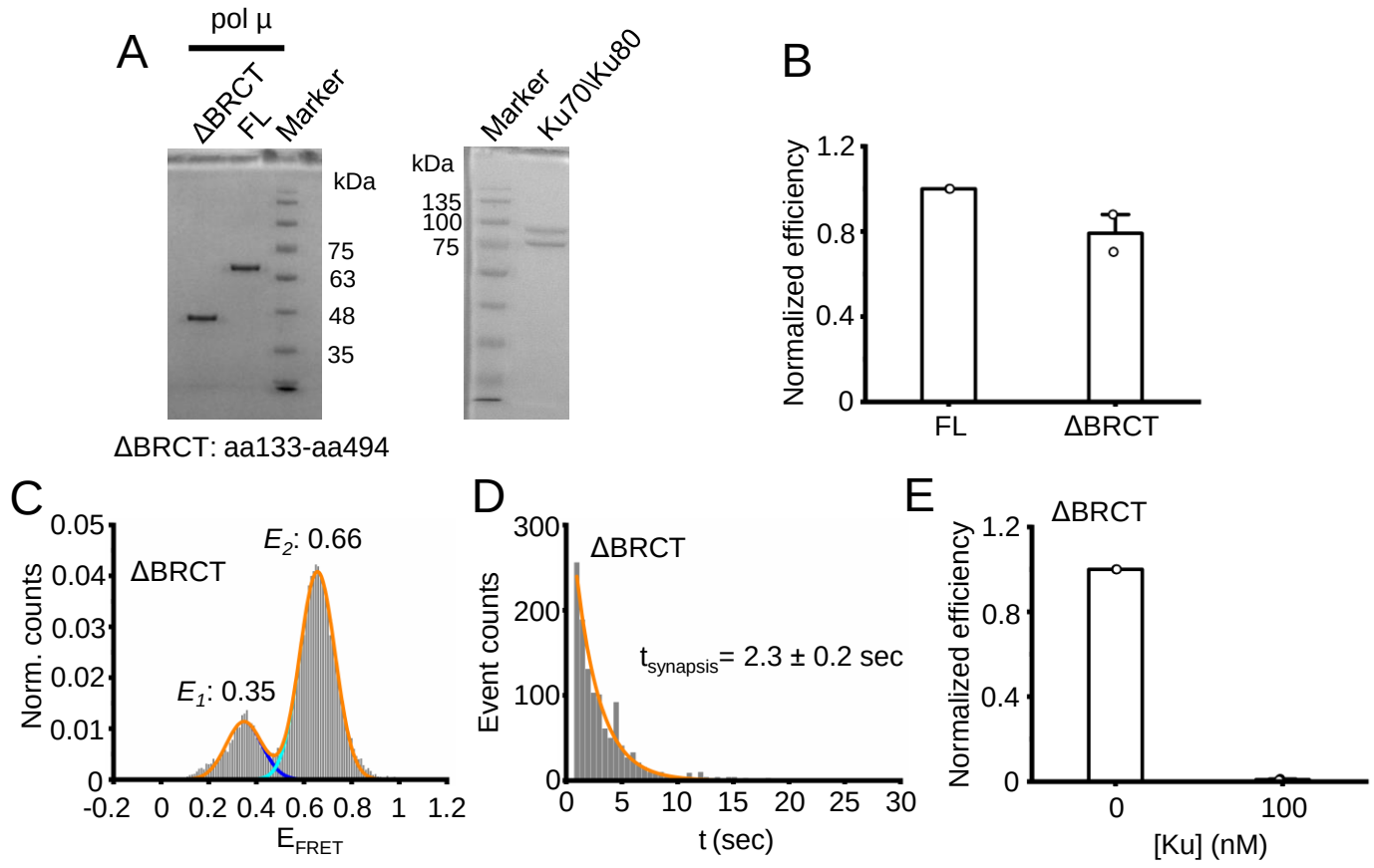

Figure S6. Ku inhibits synapsis that is mediated by pol μ lacking the BRCT domain (pol μ - ΔBRCT).

A, Full-length (FL) and BRCT domain deleted (ΔBRCT) pol μ and recombinant Ku70/Ku80 used in this study. pol μ - ΔBRCT corresponds to aa133 – aa494.

B, pol μ - ΔBRCT has equivalent capability as FL to promote synapsis of the 3' GG and 3' CCCC overhangs. 50 nM pol μ and 100 μM dGTP were included in the reactions. Both duplexes have terminal 5'-P groups. Data are represented as the mean ± SD of two independent replicates.

C, Histogram of E_{FRET} values of all synapsis events stimulated by 50 nM pol μ - ΔBRCT. The *E* values shown on the graph were obtained from a Gaussian fit of the corresponding peaks. More than 1000 molecules were included.

D, Histogram and corresponding single exponential fit of synapsis time stimulated by pol μ - ΔBRCT. More than 1000 synapsis events were included. Synapsis time shown on graph is represented as the mean ± SD of two replicates.

E, Ku inhibits the synapsis mediated by pol μ - ΔBRCT. The DNA duplexes have 3' GG and 3' CCCC overhangs, and have terminal 5'-OH groups. 50 nM pol μ - ΔBRCT, 100 μM ddGTP, and 100 nM Ku were used in the reactions. Data are represented as the mean ± SD of two independent replicates.

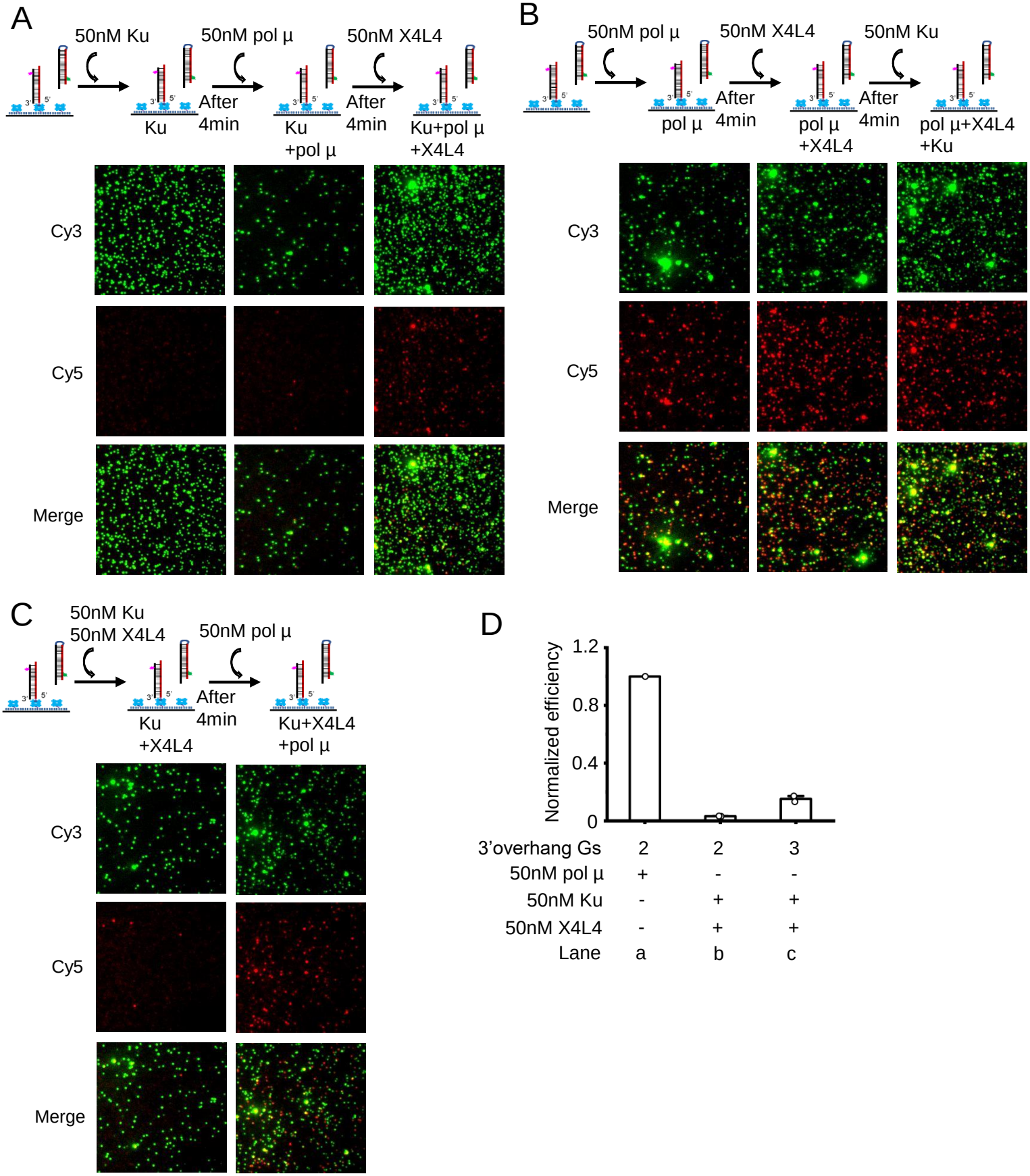

Figure S7. X4L4 can largely reverse Ku inhibition of pol μ-mediated synapsis.

A,**B**,**C**, Representative images of Cy3 channel, Cy5 channel, and overlay of the two channels for NHEJ synapsis corresponding to Figure 6**C**.

D, Normalized efficiency of NHEJ synapsis mediated by different combinations of 50 nM pol μ, 50 nM Ku, and 50 nM X4L4. The incoming duplex, proteins, and 100 μM ddGTP were first mixed well, then injected into the reaction chamber. The DNA duplexes have 3' GG or 3' GGG and 3' CCCC overhangs and have terminal 5'-OH groups. The numbers 2 and 3 on the graph represent two Gs and three Gs, respectively at the immobilized duplex 3' overhang. Data are represented as the mean ± SD of two independent replicates.

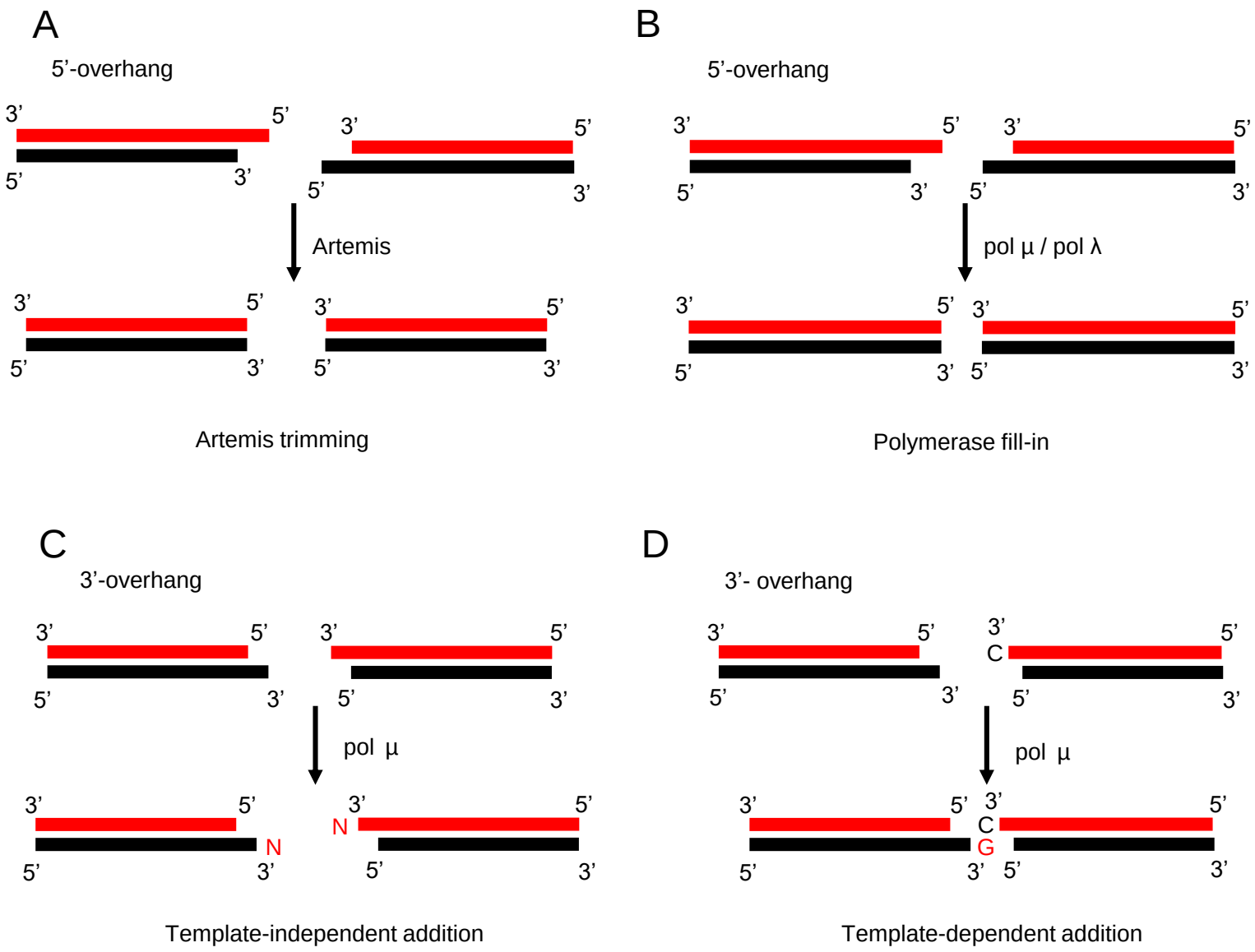

Figure S8. Processing and bridging of incompatible ends.
Broken dsDNAs with incompatible overhangs cannot be directly ligated, but rather they need to be processed by either a nuclease to remove incapatibility or a polymerase to add nucleotides that match existing compatibility with the other DNA duplex or create novel compatibility between the two DNA ends. Specifically, 5’ incompatible overhangs can be either trimmed by Artemis (**A**) or filled-in by pol μ and/or pol λ (**B**) to create blunt ends, which can be bridged and ligated by key NHEJ proteins, Ku, X4L4, and XLF. For 3’ incompatible overhangs, pol μ can randomly add nucleotides (N addition) in a template-independent way (**C**), which may then provide short microhomology for pol μ mediated synapsis (Figure 7). Or pol μ can participate in bridging the two overhangs aided by other NHEJ proteins and conduct template-dependent addition across a discontinuous template (**D**). Red letters represent newly added nucleotides.

Supplementary Table

Supplementary Table 1. DNA Oligos used in this study

Name	Sequence (5'->3')	comments
HC116	/5BiosG/CGA TAG TGG GTT CAG CAG GCA TTG TGC TAT GAT CAA CCG AAT CTG TAC ATA TAT CAG TGT CTG CAT CGT CGA CCT TGG AGG CAT CGG GG	
HC122	/5BiosG/CGA TAG TGG GTT CAG CAG GCA TTG TGC TAT GAT CAA CCG AAT CTG TAC ATA TAT CAG TGT CTG CAT CGT CGA CCT TGG AGG CAT CGG G	
HC123	/5BiosG/CGA TAG TGG GTT CAG CAG GCA TTG TGC TAT GAT CAA CCG AAT CTG TAC ATA TAT CAG TGT CTG CAT CGT CGA CCT TGG AGG CAT CGG	
BZ14	/5Phos/GAT G /iCy3/CC TCC AAG GTC GAC GAT GCA GAC ACT GAT ATA TGT ACA GAT TCG GTT GAT CAT AGC ACA ATG CCT GCT GAA CCC ACT ATC G -3'	
BZ15	/5Phos/GAT G /iCy5/CC TCC AAG GTC GAC GAT GCA GAC ACT GAT ATA TGT ACA GAT TCG GTT GAT CAT AGC ACA ATG CCT GCT GAA CCC ACT ATC G -3'	5'OH-BZ15: BZ15 was treated with SAP to remove 5'-PO_4
BZ24	5'-GATCC TC CAT CAA GTA AGA TGC AGA TAC TTA ACG TTTTTT CGT TAA GTA TCT GCA TCT TAC TTG ATG GAG-3'	Treated with PNK to add 5'-PO_4
BZ35	/5Phos/GG TT/ iCy3/C TCG ATG CGC TTG ACCA GTA GTC TAG CAC GTG ACA G-3'	5'OH-BZ35: BZ35 was treated with SAP to remove 5'-PO_4
BZ36	5'-GAT CCT GTC ACG TGC TAG ACT ACT GGT CAA GCG CAT CGA GAA CC CCCC -3'	Treated with PNK to add 5'-PO_4
BZ82	/5BiosG/CGA TAG TGG GTT CAG CAG GCA TTG TGC TAT GAT CAA CCG AAT CTG TAC ATA TAT CAG TGT CTG CAT CGT CGA CCT TGG AGG CAT C AG -3'	
BZ83	5'-GAT CCT GTC ACG TGC TAG ACT ACT GGT CAA GCG CAT CGA GAA CC TC -3'	Treated with PNK to add 5'-PO_4
BZ84	5'-GAT CCT GTC ACG TGC TAG ACT ACT GGT CAA GCG CAT CGA GAA CC GG -3'	Treated with PNK to add 5'-PO_4
BZ85	5'-GG TT/ iCy5/C TCG ATG CGC TTG ACCA GTA GTC TAG CAC GTG ACA G-3'	Treated with PNK to add 5'-PO_4
BZ42	5'-Cy5-GGA CTG CCG CCT GGG GAG CCG CAC GAC GAC ACG ACA AAG-Biotin-3'	For FRET standard

BZ43	5’-CGT GTC GTC GTG CGG CTC CCC AGG CG/Cy3/G CAG TCC-3’	For FRET standard
BZ44	5’-CGT GTC GTC GTG CGG CTC CCC AG /Cy3/ G CGG CAG TCC-3’	For FRET standard
BZ45	5’-CGT GTC GTC GTG CGG CTC C /Cy3/ CC AGG CGG CAG TCC-3’	For FRET standard
BZ46	5’-CGT GTC GTC GTG CG /Cy3/ G CTC CCC AGG CGG CAG TCC-3’	For FRET standard
BZ91	5’-GAT GCC TCC AAG GTC GAC GAT GCA G-3’	Ku binding assay
BZ107	5’-CTG CAT CGT CGA CCT TGG AGG CAT C-3’	Ku binding assay
BZ57	5’-GAT G /iCy3/CC TCC AAG GTC GAC GAT GCA G-3’	Ku binding assay
BZ109	5’-CTG C /iCy3/ AT CGT CGA CCT TGG AGG CAT C-3’	Ku binding assay