SUPPLEMENTAL FIGURES

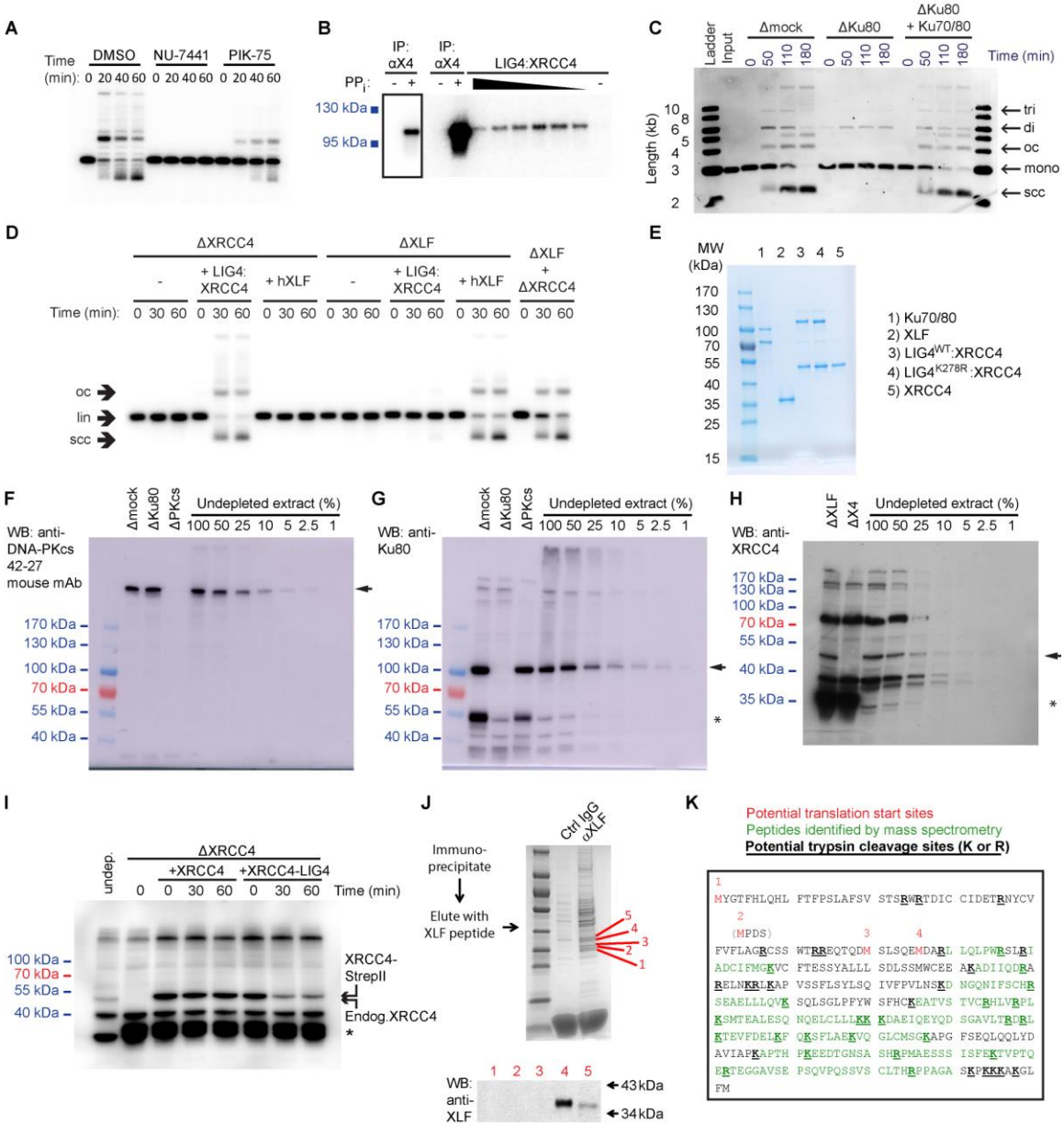

Figure S1. Bulk End Joining Assay Controls, Related to Figure 1

(A) Inhibition of end joining by the DNA-PKcs inhibitors NU-7441 (87 μM) and PIK-75 (11 μM).

(B) An adenylated protein with molecular mass similar to recombinant LIG4 is present in anti-XRCC4 immunoprecipitates. XRCC4 was immunoprecipitated from egg extract, and immunoprecipitates were incubated with [α^{32}P]-labeled ATP to detect autoadenylation activity. Appearance of a radiolabeled ~100 kDa band required pretreatment of the immunoprecipitate with 0.5 mM pyrophosphate (which drives LIG4 deadenylation), consistent with previous results indicating that the majority of LIG4 in cell lysates is adenylated (Grawunder et al., 1997). Recombinant LIG4:XRCC4 (which in this case had been purified from

Sf9 cell lysates using buffers without pyrophosphate) was also adenylated in vitro in a manner that required pre-treatment with pyrophosphate (concentration series: 0.45, 0.22, 0.11, 0.06, 0.03, 0.014, and 0 mM). High concentrations of pyrophosphate reduced the adenylation signal. (Note that in reactions with recombinant protein, pyrophosphate was not removed prior to addition of [α^{32}P]-ATP.) Boxed inset: image of the first two lanes with reduced contrast.

(C) SYBR gold-stained gel of an end-joining experiment using alternative conditions similar to those previously described (Labhart, 1999; Di Virgilio and Gautier, 2005). Under these conditions, residual multimer formation was seen in the absence of Ku. ladder, 2-log ladder (New England Biolabs). mono, monomeric linear DNA. oc, open-circular products. scc, supercoiled closed-circular products. di, linear dimers. tri, linear trimers.

(D) Orthogonality of XRCC4 and XLF immunodepletions. End joining was restored in XRCC4-immunodepleted (ΔXRCC4) extract by addition of LIG4:XRCC4 but not by addition of XLF (human XLF [hXLF] in this experiment). Conversely, end joining was restored in XLF-immunodepleted (ΔXLF) extract by human XLF but not by LIG4:XRCC4. End joining was also restored when ΔXRCC4 and ΔXLF extracts were mixed together (ΔXRCC4 + ΔXLF).

(E) SDS-PAGE gel of purified recombinant *X. laevis* NHEJ proteins stained with colloidal Coomassie blue (InstantBlue).

(F-H) Full-size versions of the western blots shown in Fig. 1A,B,D. The blot in (G) was obtained by re-probing the blot in (F) with rabbit anti-Ku80 antibody followed by HRP-conjugated anti-rabbit IgG secondary antibody. Arrows, bands likely to be the proteins of interest. Asterisks, contaminating IgG heavy chain (G) or light chain (H).

(I) To verify that XRCC4 is stable in egg extract independently of LIG4, recombinant XRCC4 or LIG4:XRCC4 was added to XRCC4-immunodepleted extract containing linear substrate DNA, and the extract was blotted for XRCC4 at the indicated times. The result shows that XRCC4 persists in the extract at 30 and 60 min, even in the absence of LIG4. *, contaminating IgG light chain. Endog. XRCC4, endogenous XRCC4.

(J-K) Analysis of the *X. laevis* XLF translation start site. (J) XLF was immunoprecipitated from egg extract as described in "Antibody generation and immunodepletion." Bands of approximately the correct molecular weight were excised from a Coomassie blue-stained gel (upper panel). Protein was eluted in SDS-PAGE sample buffer, separated again on an SDS-PAGE gel, and analyzed by western blotting with the same anti-XLF antibody used for immunoprecipitation (lower panel). The band that gave the strongest signal on the western blot was excised from a Coomassie blue-stained gel of a second immunoprecipitation and submitted for trypsin digestion and peptide identification by mass spectrometry. (K) Peptides identified by mass spectrometry are highlighted in green on a reference sequence of XLF with candidate start sites indicated in red. The reference sequence was stitched together from the *X. laevis* Xenbase/JGI 6.1 gene model for XLF (NHEJ1), the PHROG database (Wühr et al., 2014), and partial cDNA sequences in GenBank (accession numbers BC077538.1, BJ093763.1, CA988039.1, and CA982456.1). While the data do not definitively indicate the translation start site, the absence of any peptides or cDNA sequences corresponding to the possible N-terminal extension upstream of start site 3, and the fact that this extension is not conserved (see sequence alignment in Supplemental Text 3), suggest that start sites 3 or 4 are the most plausible. A construct beginning at start site 4 rescued end joining in XLF-immunodepleted extract but was poorly soluble (data not shown). A construct beginning at start site 3 was well expressed and soluble and was used for all experiments shown in the figures.

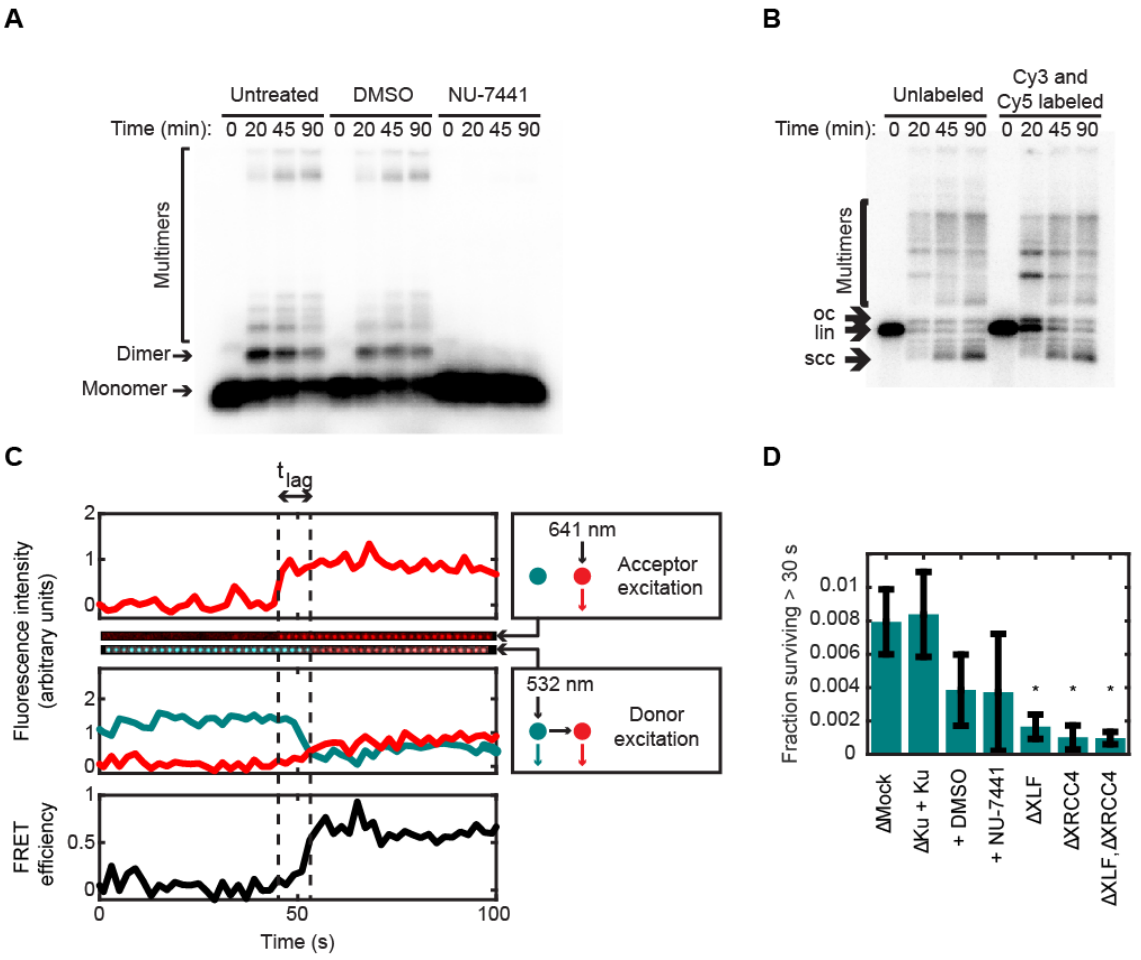

Figure S2. Bulk End Joining of Fluorophore-Labeled Substrates and Additional Tethering Assay Data, Related to Figure 2

(A) End-joining of the 100 bp Cy5-labeled DNA duplex that was used for tethering experiments. As with longer DNA fragments, end-joining was inhibited by a DNA-PK inhibitor, NU-7441, but not by solvent alone (DMSO).

(B) End joining of a 2 kilobase pair PCR product labeled 7 bp from one end with Cy3 and 7 bp from the other end with Cy5 (right) and of a control PCR product without Cy3 and Cy5 labels (left). oc, open-circular products. lin, linear DNA. scc, supercoiled closed-circular products.

(C) Single-molecule kymograph of Cy5-DNA tethering followed by a transition to high FRET in the intermolecular synapsis assay depicted in Fig. 2A. Upper plot, Cy5 intensity with direct (641 nm) excitation. Middle plot, Cy3 (cyan) and Cy5 (red) intensity with Cy3 (532 nm) excitation. Lower plot, calculated FRET efficiency. The kymograph between the upper and middle plots shows the image of the spot whose intensity is quantified. The upper row of the kymograph shows the image of the spot in the Cy5 channel with direct (641 nm) excitation, corresponding to the upper plot. The lower row of the kymograph shows an overlay of images in the Cy3 (cyan) and Cy5 (red) channels with Cy3 (532 nm) excitation, corresponding to the middle plot. A time delay between tethering and appearance of FRET (t_{lag}) is evident from both the raw images and quantified intensities. The two rows of the kymograph are staggered along the horizontal axis to accurately reflect alternating excitation with 532 nm and 641 nm light.

(D) Fraction of long-range complexes surviving greater than 30 s (*, $p < 0.01$ compared to ΔMock, one-way

ANOVA followed by Tukey's post-hoc test; $p > 0.05$ for other comparisons with ΔMock). The reduction in long-lived complexes in extract lacking XLF, XRCC4 or both may reflect a destabilization of the long-range synaptic complex in the absence of XLF or XRCC4-LIG4. Alternatively, it is possible that some fraction of (much longer-lived) short-range complexes were erroneously classified as long-range complexes. This could happen if the Cy5 label at the synapsed end was lost due to incomplete labeling, photobleaching, or limited resection while the Cy5 label at the opposite end persisted (see Fig. 2A and Supplemental Experimental Procedures).

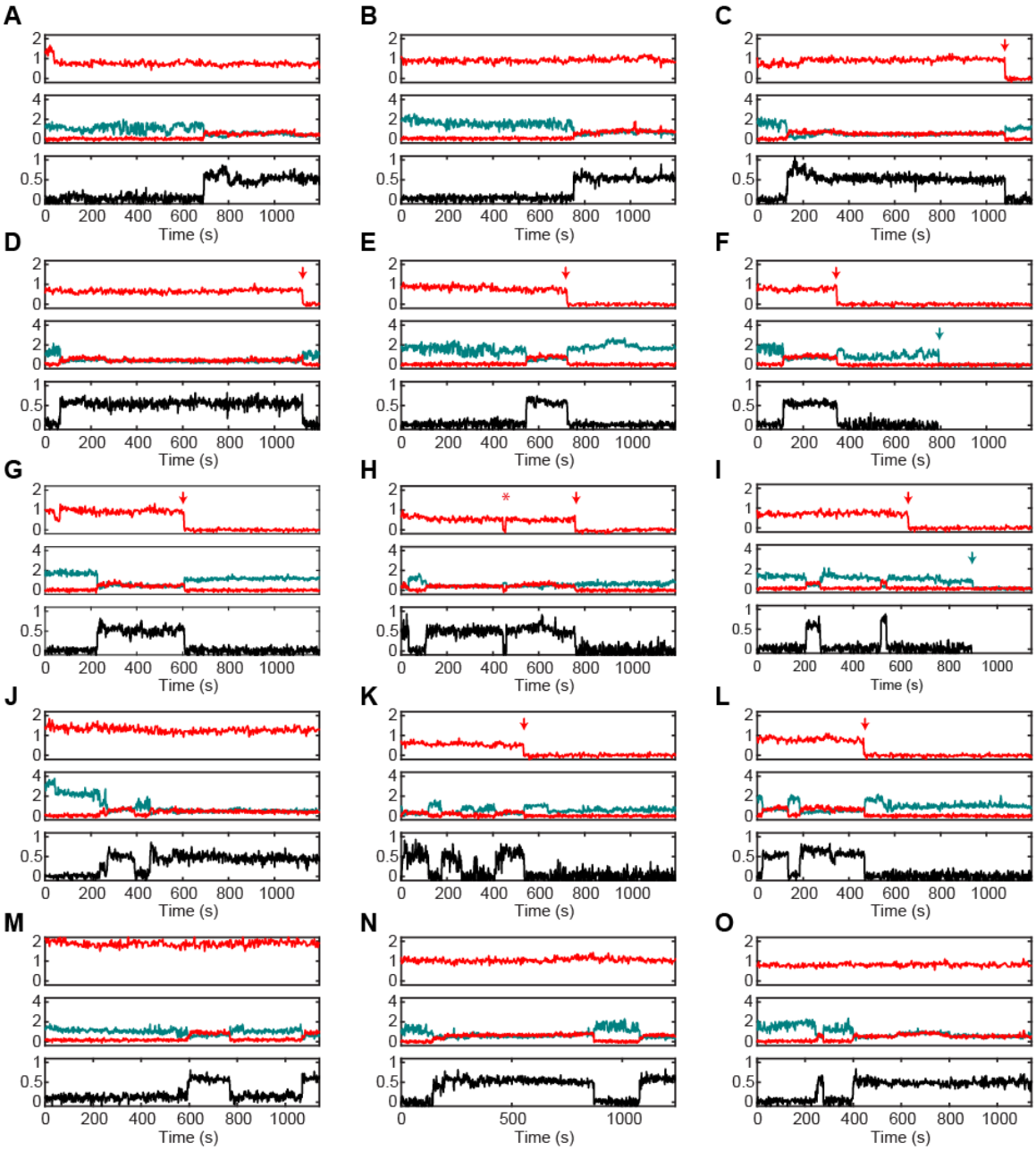

Figure S3. Sample Single-Molecule FRET Trajectories for the Intramolecular End Joining Substrate (shown in Fig. 3A), Related to Figure 3

The top panel shows Cy5 intensity with direct (641 nm) excitation (arbitrary units), the middle panel shows Cy3 (cyan) and Cy5 (red) intensities with Cy3 (532 nm) excitation (arbitrary units), and the bottom panel shows calculated FRET efficiency. Cy5 and Cy3 photobleaching events are indicated by red and cyan arrows,

respectively. Some substrates transitioned to and remained in a high-FRET state until the end of the observation interval (A,B) or until acceptor photobleaching (C-G). Other substrates exhibited multiple transitions between a low-FRET state and a high-FRET state (H-O). The apparent loss of FRET between 400 and 500 s in (H) was due to a brief acceptor blinking event (red *). The calculated FRET efficiency in (F) and (I) is shown only prior to donor photobleaching. The “0 s” timepoint in all trajectories corresponds to the start of the observation in a particular field of view, not the time of extract addition to the flowcell.

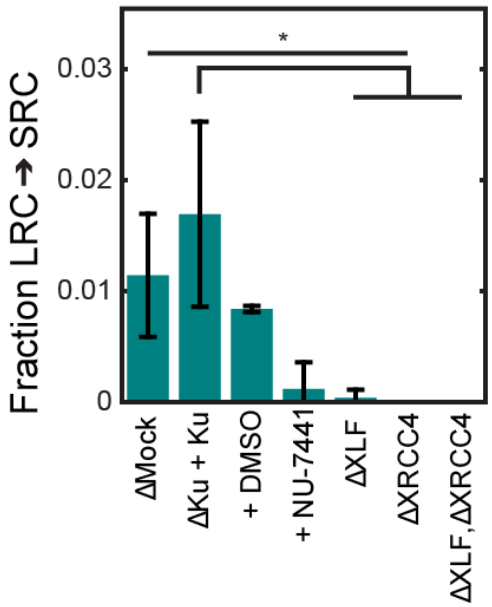

Figure S4. Long-Range Complex to Short-Range Complex Transition Probabilities in the Intermolecular Tethering Assay, Related to Figures 4 and 5

Number of transitions to high FRET as a fraction of long-range complexes surviving at least 4 acceptor excitation frames, the minimum survival time required for detection of FRET transitions using our algorithm (see Supplemental Experimental Procedures). *, $p < 0.05$, one-way ANOVA with Tukey's post-hoc test.

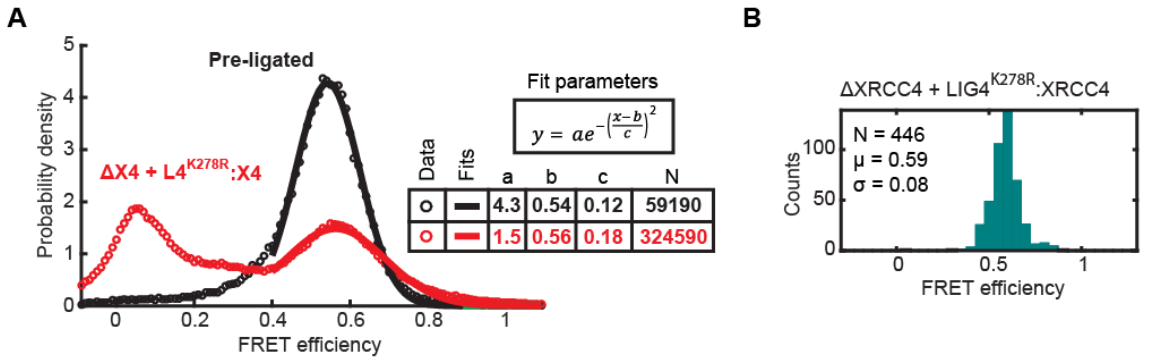

Figure S5. FRET Efficiency of the High-FRET State, Related to Figure 5

(A) FRET efficiency histogram of all frames from extended timecourse imaging experiments in XRCC4-depleted extract supplemented with catalytically inactive LIG4^{K278R}:XRCC4 (red). A FRET efficiency histogram of pre-ligated substrate imaged in extract (as in Fig. 3C) is shown as a standard (black). Gaussian fits to the high-FRET peak in each distribution are shown as solid lines, and fit parameters are displayed in the table to the right. Noncovalent short-range complexes formed in the presence of catalytically inactive LIG4 had a similar mean FRET efficiency as for pre-ligated substrate, indicating the close apposition of DNA ends in the short-range complex.

(B) Average FRET efficiency of high-FRET segments extracted from long timecourse trajectories in XRCC4-immunodepleted extract supplemented with catalytically inactive LIG4^{K278R}:XRCC4 (c.f. Fig. 3C).

SUPPLEMENTAL TABLES

Table S1, Related to Fig. 2.
Sample sizes for inter-molecular synapsis experiments.

Condition	**Number of replicates**	**Total number of Cy3-DNAs tracked (for ~3 min each)**	**Total number of tethering events detected**
ΔMock	13	63773	71794
ΔKu80	3	13697	618
ΔKu80 + Ku70/80	3	14228	22739
ΔDNA-PKcs	3	13550	398
ΔXLF	5	25901	28708
ΔXRCC4	5	23444	30642
ΔXLF,ΔXRCC4	4	16619	30506
DMSO	3	10593	13842
NU-7441	4	14523	19092

Table S2, Related to Figs. 4 and 5.
Sample sizes for intra-molecular synapsis experiments.

Condition	**Number of replicates**	**Total number of substrates tracked (for ~18 s each)**
ΔMock	5	124574
ΔKu80	2	43602
ΔKu80 + Ku70/80	2	48740
ΔDNA-PKcs	2	53159
DMSO	4	56797
NU-7441	3	47639
PIK-75	2	35448
ΔXRCC4	4	73169
ΔXRCC4 + $LIG4^{WT}$:XRCC4	3	47175
ΔXRCC4 + $LIG4^{K278R}$:XRCC4	4	60028
ΔXRCC4 + XRCC4	2*	47481
ΔXLF	3	68069
ΔXLF + XLF	2	47213

* One replicate of the ΔXRCC4 + XRCC4 condition was done with an elevated concentration of XRCC4, 0.8 pmol (monomer) per µl extract.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Protein expression and purification

X. laevis XRCC4 and LIG4 cDNA clones were purchased from Thermo Scientific (CloneIDs 6635763 and 6957895). *X. laevis* Ku70 and Ku80 cDNA clones were a kind gift of Hironori Funabiki (Rockefeller University). A synthetic codon-optimized *X. laevis* XLF coding sequence (Supplemental Text 1) was ordered from Life Technologies. C-terminally StrepII-tagged XRCC4, N-terminally Flag-tagged and C-terminally His6-tagged LIG4, C-terminally His6-tagged Ku70, and untagged Ku80 were subcloned into pFastBac1 (Invitrogen) as described in "Plasmid generation". Baculoviruses were generated following the manufacturer's instructions. XLF was cloned into a His10-SUMO expression vector with a 4-amino acid linker (GGGS) to improve cleavage efficiency by SUMO protease. A Coomassie blue-stained gel of the purified proteins described below is shown in Fig. S1E.

The translation start site of *X. laevis* XLF was ambiguous based on available database sequences (Fig. S1J-K). To determine the actual translation start site, we immunoprecipitated XLF from extract and sequenced tryptic peptides by mass spectrometry (Fig. S1J-K). XLF was immunoprecipitated from egg extract following the protocol described in "Antibody generation and immunodepletion". Anti-XLF and control IgG beads were washed extensively in egg lysis buffer containing sucrose (ELBS; 10 mM HEPES, pH 7.7, 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose), and bound proteins were eluted by incubation with 1 volume 5 mg/ml XLF C-terminal peptide in egg lysis buffer (ELB; 10 mM HEPES, pH 7.7, 50 mM KCl, 2.5 mM $MgCl_2$) for 20 min at room temperature. Immunoprecipitated proteins from 66 µl of extract were separated on a precast 10% SDS-PAGE gel (BioRad Mini-PROTEAN TGX), which was stained with InstantBlue (Expedeon) and destained with distilled water. To determine which band most likely corresponded to XLF, the five indicated bands of approximately the predicted molecular weight were excised with a scalpel, and proteins were eluted by shaking in 20 µl SDS-PAGE sample buffer for 1 h at 50°C. The eluates were separated on a second gel and western blotted with anti-XLF antibody (Fig. S1J). The immunoprecipitation procedure was repeated, and band four was excised and submitted for trypsin digestion and peptide identification by mass spectrometry at the Taplin Mass Spectrometry Core Facility at Harvard Medical School. As shown in Fig. S1J, the amino acids present in the identified peptides (green) suggested two candidate start sites (3 and 4). Recombinant proteins corresponding to both were purified, and both rescued end joining in XLF-immunodepleted extract (Figs. 1C and data not shown). The longer construct (start site #3 in Fig. S1J), which was much more soluble, was used for the experiments shown in the main text.

Ku70-H6:Ku80: Sf9 cells at a density of 4 x 10^6/ml in 200 ml of Sf900 III serum-free medium (Invitrogen) were infected with 1 ml each of *X. laevis* Ku70-H6 and Ku80 baculoviral stocks. Cells were harvested after 60 h of shaking at 27°C, pelleted by centrifugation for 5 min at 4°C in an SLA-3000 rotor at 1700 rpm, washed with phosphate-buffered saline (PBS; Teknova, 135 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4), and centrifuged again for 5 min at 1000 *g* at 4°C. Pellets were flash-frozen in liquid nitrogen and stored at -80°C. Cell pellets were thawed, resuspended in ~4 volumes lysis buffer (20 mM Tris-HCl, pH 8, 0.5 M NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, and one cOmplete protease inhibitor cocktail tablet (Sigma-Aldrich) per 50 ml), and lysed by sonication with a Branson Digital Sonifier at 40% amplitude for 30 s with pulses of 1 s separated by 1 s pauses. Lysates were clarified by centrifugation at 25,000 rpm in an SW-41 rotor for 1 h at 4°C, and supernatants were incubated with 1.5 ml NiNTA-agarose (Qiagen) for 1.5 h at 4°C. The resin was washed twice with 25 ml lysis buffer, poured into a disposable column, and washed with 25 ml additional lysis buffer followed by lysis buffer with 30 mM and 60 mM imidazole. Protein was eluted in 1.5 ml increments with lysis buffer (without protease

inhibitors) containing 300 mM imidazole. Peak fractions were pooled, dialyzed twice at 4°C against 1 L dialysis buffer (20 mM Tris-HCl, pH 8, 0.5 M NaCl, 10% glycerol, 5 mM BME), concentrated by spinning at 7000 *g* in a 5 kDa MWCO centrifugal concentrator (Vivaspin-6, Sartorius Stedim), flash-frozen in liquid nitrogen, and stored at -80°C. For additional purification, the protein was thawed, diluted with 9 volumes Q dilution buffer (100 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM β-mercaptoethanol), and bound to 1 ml bed volume Nuvia Q resin (BioRad) pre-equilibrated with the same buffer. The resin was washed with 5 ml Q dilution buffer followed by 3 ml Q wash buffer (100 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol), and eluted with Q elution buffer (100 mM Tris-HCl, pH 7.5, 250 mM NaCl, 5 mM β-mercaptoethanol). Peak fractions were pooled, concentrated with a 5 kDa MWCO centrifugal concentrator, aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C.

LIG4:XRCC4: X. laevis LIG4:XRCC4 complexes containing wild-type or K278R (catalytically inactive) mutant LIG4 were purified in parallel. Sf9 cells at a density of 3.5 x 10^6/ml in 200 ml were infected with 2 ml each xXRCC4-StrepII and either Flag-xLIG4-H6 or Flag-xLIG4^{K278R}-H6 baculoviral stocks and incubated with shaking at 27°C for ~60 h. Cells were pelleted by centrifugation for 5 min at 4°C at 500 *g* in an SLA-3000 rotor, and pellets were resuspended in 10 ml lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 10 mM imidazole, 0.5 mM sodium pyrophosphate, 5 mM BME) supplemented with one protease inhibitor cocktail tablet (cOmplete EDTA-free, Roche). Cells were sonicated on ice for 30 s at 40% intensity, and additional lysis buffer was added to a volume of 20 ml. Lysates were cleared by centrifuging 1.5 h at 25,000 rpm in an SW-41 rotor. The supernatant was incubated with 1 ml NiNTA agarose (Qiagen) pre-washed with lysis buffer for 1 h on a rotator at 4°C. The slurry was poured into a disposable polypropylene column, washed with 20 ml lysis buffer, and eluted with seven 1 ml volumes of NiNTA elution buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 250 mM imidazole, 5 mM BME). Peak fractions were identified by Bradford assay, pooled, and passed twice by gravity flow over 1.5 ml Strep-Tactin Sepharose (IBA) packed in a disposable column. The column was washed with 15 ml lysis buffer without protease inhibitors and eluted with six 1 ml volumes of ST elution buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 2.5 mM desthiobiotin, 5 mM BME). Peak fractions were pooled and concentrated ~5-fold in 5 kDa MWCO centrifugal concentrators (VivaSpin 6, Sartorius Stedim). Aliquots were flash-frozen in liquid nitrogen and stored at -80°C.

XRCC4: 100 ml Sf9 cells at a density of 4 x 10^6/ml were infected with 2 ml xXRCC4-StrepII baculoviral stock, incubated with shaking at 27°C for ~60 h, pelleted as above, washed with 1x PBS, flash-frozen in liquid nitrogen, and stored at -80°C. The cell pellet was thawed and resuspended in 10 ml lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 5 mM BME) supplemented with one protease inhibitor cocktail tablet (cOmplete EDTA-free, Roche) and avidin at a final concentration of 50 µg/ml. Following sonication on ice for 30 s at 40% intensity, the lysate was cleared by centrifuging 1 h at 25,000 rpm in an SW-41 rotor. The supernatant was passed over 1.5 ml Strep-Tactin Sepharose (IBA), which was then washed with 32.5 ml lysis buffer, and eluted with seven 1.5 ml volumes of ST elution buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 2.5 mM desthiobiotin, 5 mM BME). Peak fractions were pooled, concentrated ~6-fold in a 5 kDa MWCO centrifugal concentrator (VivaSpin 6, Sartorius Stedim), flash-frozen in liquid nitrogen, and stored at -80°C.

XLF: H10-SUMO-xXLF (pTG296) was transformed into Rosetta II (DE3) cells. Several colonies were used to inoculate a 500 ml culture of Luria broth (LB) containing 100 µg/ml ampicillin and 17 µg/ml chloramphenicol, which was grown with shaking at 37°C to an OD600 of 1.4, induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3 h at ~25-30°C, pelleted by centrifugation, washed with 1x PBS, flash-frozen in liquid nitrogen, and stored at -80°C. His-SUMO purification steps were performed as described previously for the bacterial protein Spo0J (Graham et al., 2014). Protein in His-SUMO dialysis

buffer (20 mM Tris-HCl, pH 8, 350 mM NaCl, 10 mM imidazole, 10% glycerol, 5 mM BME) was diluted with 1.36 volumes SP buffer (50 mM Na-MES, pH 6.5, 10% glycerol, 1 mM DTT) and passed over 300 µl SP Sepharose Fast Flow (Amersham), equilibrated with SP buffer containing 150 mM NaCl. The resin was washed with 1.5 ml SP buffer with 150 mM NaCl and eluted with SP buffer containing 350 mM NaCl in ten 300 µl volumes. Peak fractions were pooled, aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C. Human XLF (pTG259; used in Fig. S1D) was purified in the same manner but without the cation exchange step.

DNA-PKcs PIKK-FATC antigen (amino acids 3765-4146): BL21 (DE3) pLysS cells transformed with pTG314 were grown in 1 L of LB with 50 µg/ml kanamycin and 17 µg/ml chloramphenicol at 37°C. At an OD600 of 0.8, the culture was transferred to 30°C and induced with 1 mM IPTG for 6 h. Bacteria were pelleted by centrifugation, flash-frozen in liquid nitrogen, and stored at -80°C. The cell pellet was resuspended in 10 ml of high-salt lysis bufer (20 mM Tris-HCl pH 8, 1 M NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, and 1 mM PMSF), lysed by sonication, and treated for 10 min at room temperature with ~250 U benzonase. Insoluble material was pelleted in a fixed-angle benchtop centrifuge at 7000 *g* for 10 min at 4°C. The pellet was washed three times with lysis buffer and then resuspended in solubilization buffer (1x PBS supplemented with 8 M urea, 10 mM imidazole, and 5 mM β-mercaptoethanol). The resuspended mixture was homogenized by sonication and then clarified by centrifugation for 10 min at 7000 *g* followed by 10 min at 16,000 *g*. The clarified supernatant was diluted with 20 ml solubilization buffer and passed by gravity flow over 12 ml NiNTA-agarose (Qiagen). The column was washed with 5 column volumes of solubilization buffer and eluted with the same buffer containing 300 mM imidazole. Eluates were pooled, and the protein was further purified by SDS-PAGE on a horizontal gel rig. The gel was stained with InstantBlue (Expedeon), and the band of interest was excised. Protein was electroeluted in SDS-PAGE running buffer and concentrated to at least 1 mg/ml using a 5 kDa MWCO centrifugal concentrator (Amicon Ultra).

Antibody generation and immunodepletion

For immunodepletion, antibodies were bound to protein A sepharose beads at a ratio of 3 µg antibody per 1 µl beads. In the case of anti-DNA-PKcs antibody and the associated mock IgG control, 20 µg of antibody were bound per 1 µl beads. Beads were washed thoroughly with egg lysis buffer with sucrose (ELBS; 10 mM HEPES, pH 7.7, 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose), and the buffer was aspirated thoroughly with an ultrafine gel loading tip (Eppendorf GELoader 20 µl) before addition of extract. One volume of extract was incubated for 2 rounds (for XLF and XRCC4) or 3 rounds (for Ku80 and DNA-PKcs) with 1/5 volume (for DNA-PKcs) or 1/10 volume (for XLF, XRCC4, and Ku80) of antibody-bound beads for 20 min per round on a rotator at room temperature. Beads were pelleted by centrifugation at 2000 *g* after each round, and extract was collected with a standard P200 pipette tip followed by an ultrafine gel-loading tip. For rescue experiments, roughly 0.05 pmol LIG4:XRCC4 complex, 0.05-0.1 pmol XRCC4 (quantity of monomer), 0.3 pmol of Ku70/80, or 0.2 pmol (quantity of monomer) *X. laevis* XLF were added per µl of extract. Protein concentrations were estimated relative to a BSA standard on a colloidal Coomassie blue-stained SDS-PAGE gel. For the XRCC4 stability experiment in Fig. S1I, proteins were added at the usual concentration to end-joining reactions containing non-radiolabeled DNA. Samples were withdrawn at the indicated times, mixed with 4 volumes of SDS-PAGE loading dye, separated on a 12% SDS-PAGE gel, and analyzed by western blotting with anti-XRCC4 antibody as described below.

LIG4 autoadenylation assay

Bead-bound immunoprecipitates from the first round of an XRCC4 immunodepletion were incubated 15 min at room temperature in ELBS with or without 0.5 mM sodium pyrophosphate to reverse endogenous adenylation of LIG4, and pyrophosphate was removed by washing twice more with ELBS. As a control, recombinant LIG4:XRCC4 complex (from an early preparation, purified using buffers not containing pyrophosphate) was incubated with different amounts of sodium pyrophosphate for 15 min at room temperature in adenylation buffer (60 mM Tris-HCl, pH 8, 10 mM MgCl, 5 mM DTT, 50 µg/ml BSA) (Chen et al., 2006). Pyrophosphate was not removed from the recombinant protein prior to the adenylation reaction. Beads and recombinant protein were incubated in adenylation buffer with [α^{32}P]-labeled ATP at room temperature for 15 min, and reactions were stopped by adding one volume of SDS-PAGE sample buffer. Proteins were separated on a 7.5% SDS-polyacrylamide gel, which was dried on a gel dryer, exposed to a storage phosphorscreen, and imaged using a BioRad Personal Molecular Imager.

Western blotting

For the anti-DNA-PKcs western blot shown in Fig. 1B and Fig. S1F, samples were separated on a 4-15% precast SDS-PAGE gel (BioRad), transferred to a polyvinylidine fluoride membrane for 1 h at 100 V at 4°C, blocked with 5% powdered nonfat milk in PBST (PBS with 0.05% Tween 20), and probed overnight with a 1:5000 dilution of concentrated hybridoma supernatant containing αDNA-PKcs 42-27 mouse monoclonal antibody (a kind gift of Prof. Katheryn Meek, Michigan State University). The blot was washed with PBST, probed with 1:20,000 horseradish peroxidase (HRP)-conjugated rabbit anti-mouse IgG (H+L) secondary antibody (Jackson ImmunoResearch Cat. #315-035-003) in 5% milk/PBST for approximately 1 h at room temperature, washed again with PBST, incubated with chemiluminescent HRP substrate (HyGLO Quick Spray, Denville Scientific), and imaged on an Amersham Imager 600 (GE Healthcare Life Sciences). The same blot was reprobed with 0.1 µg/ml affinity-purified anti-Ku80 antibody followed by 1:20,000 horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (H+L) secondary antibody (Jackson ImmunoResearch Cat. #111-035-003) and imaged as before (data shown in Fig. 1A and Fig. S1G). For the anti-XRCC4 western blot shown in Fig. 1D and Fig. S1H, samples were separated on a 12% precast SDS-PAGE gel, 0.2 µg/ml affinity-purified anti-XRCC4 antibody was used for the primary step, and 1:5000 HRP-conjugated anti-rabbit IgG light chain-specific antibody (Jackson ImmunoResearch, Cat. #211-032-171) was used for the secondary step. For the anti-XLF western blot in Fig. S1J, eluates from XLF IP bands were separated on a precast 12% SDS-PAGE gel and transferred to a polyvinylidine fluoride (PVDF) membrane, 1 µg/ml affinity-purified anti-XLF antibody was used for the primary step, and 1:20,000 HRP-conjugated goat anti-rabbit IgG (H+L) was used for the secondary step. The blot was incubated with chemiluminescent HRP substrate and exposed to film.

Bulk end-joining assay

Substrate preparation: For the experiments shown in Fig. 1, a fragment corresponding to basepairs 21226-26104 (4.9 kb) in bacteriophage λ genomic DNA was excised from pTG313 by cleavage with EcoRI-HF (New England Biolabs), separated on a 1x TBE agarose gel containing ethidium bromide, and extracted by electroelution. Electroeluted DNA was extracted twice with one volume of 2-butanol to remove ethidium bromide and ethanol precipitated. DNA was labeled by fill-in of EcoRI overhangs with the DNA polymerase I Klenow fragment (New England Biolabs) in the presence of [α-^{32}P]dATP (EasyTides, Perkin Elmer) and

dTTP. To prevent irreversible resection of DNA ends beyond the EcoRI site by the 3'-5' exonuclease activity of the Klenow fragment, dCTP and dGTP were also included in the reaction. After a 15 min incubation at 25°C, labeled DNA was purified using a spin column PCR purification kit (Qiagen) and eluted with 5 mM Tris-HCl, pH 7.5. For the experiments in Figs. S1A,D, EcoRI-digested pBluescript II KS(-) was used as a substrate rather than the λ/EcoRI fragment.

For the short DNA duplex end-joining experiment in Fig. S2A, Cy5-labeled duplex DNA, generated as described in the following section, was end-labeled with ^{32}P phosphate using T4 polynucleotide kinase (New England Biolabs). Labeled DNA was purified using a spin column PCR purification kit (Qiagen) and eluted with 5 mM Tris, pH 7.5.

For comparing end-joining efficiencies of Cy3 and Cy5-labeled or unlabeled DNA fragments (Fig. S2B), PCR products were generated using pTG059 (see list of plasmids below) as a template and primer pairs oTG415-6 and oTG413-4 (see list of oligonucleotides below), respectively. DNAs were gel-purified using a spin column kit (Qiagen) and end-labeled with [γ-^{32}P]ATP (EasyTides, Perkin Elmer) and T4 polynucleotide kinase (New England Biolabs). Labeled DNA was purified using a spin column PCR purification kit (Qiagen) and eluted with 5 mM Tris, pH 7.5.

End joining reactions:

See Experimental Procedures for a description of bulk end joining experiments shown in the main text. For the experiment shown in Fig. S1C, a higher concentration of unlabeled substrate was used, as in previous work (Labhart, 1999; Di Virgilio and Gautier, 2005). 15 µl extract was combined on ice with 0.45 µl ATP regeneration system and 1.5 µl of 100 ng/µl EcoRI-digested pBluescript II KS(-) substrate, and reactions were transferred to room temperature. 3 µl samples were withdrawn at the indicated times, mixed with 5 µl stop solution/loading dye, and digested overnight with 1 µg of proteinase K per sample. Samples were separated on a 1x TBE 0.8% agarose gel, which was stained with SYBR gold while digesting egg extract RNA *in situ* with RNase A.

Single molecule microscope and chamber preparation

Single-molecule FRET measurements were made on a dual-color through-objective TIRF microscope, constructed using an inverted microscope body (Olympus IX71). 532 nm and 641 nm laser beams (Coherent, Sapphire 532 and Cube 641) were expanded and focused through the rear port of the microscope onto the back aperture of an oil immersion objective (Olympus UPlanSApo, 100x, NA 1.40) using a focusing lens on a vertical translation stage to adjust the TIRF angle. Cy3 and Cy5 emission wavelengths were separated using a set of dichroic mirrors and filters (Chroma) and imaged on two halves of an electron multiplying charge-coupled device (EM-CCD) camera (Hamamatsu, ImageEM 9100-13). The camera was operated at maximum EM gain for single-molecule imaging. Focus was maintained manually, and the horizontal position of the sample was controlled by an automated microstage (Mad City Labs).

Glass coverslips were functionalized with a mixture of methoxypolyethylene glycol (mPEG) and biotin-mPEG (Laysan Bio, Inc.), and flowcells were constructed as described previously (Tanner et al., 2008). To make closely spaced channels while avoiding leakage between channels, 20 mm long, 1 mm thick pieces of quartz (Technical Glass Products, Inc.) were cut into segments ~5 mm wide with a Dremel tool and a diamond-coated rotary blade. Holes were drilled 10 mm apart and fitted with PE20 and PE60 polyethylene tubing for the inlet and outlet, respectively. Flowcells were sealed with epoxy (Devcon #14250) and stored

under vacuum until needed. The biotin-mPEG functionalized surface was incubated for at least five minutes with streptavidin at a concentration of 0.7 mg/ml and washed with 200 µl of ELB prior to the addition of biotinylated DNA. For the majority of experiments, solutions were drawn into the channel with a P200 pipette by inserting a gel loading tip into the outlet tubing.

Single-molecule tethering assay and image analysis

Substrate preparation: Cy5-labeled 100 bp duplex (for the experiments in Fig. 2 and Fig. S2B): 100 pmol each of oTG538 and oTG539 were mixed with 200 pmol oTG416 in 25 µl 1x NEBuffer 3 (New England Biolabs), heated to 95°C for 1 min, and cooled at 0.1°C/s to 25°C. 1.3 µl of 20 mM ATP and 1 µl of T4 DNA ligase (NEB) were added, and the reaction was incubated at 37°C. After approximately 1 h, the reaction was separated on a 1x TBE, 5% polyacrylamide (19:1 bis-acrylamide:acrylamide) gel, alongside 40 pmol each of oTG538 and oTG416 to confirm clear separation of the final product from the reactants. The full-length product, which appeared as a light blue band, was excised, and the remainder of the gel was stained with ethidium bromide. The full-length product band was pulverized by centrifugation through a 27-gauge needle hole in the bottom of a silanized 0.65 ml microcentrifuge tube. The fragments were resuspended in 1.1 ml egg lysis buffer (ELB; 10 mM HEPES, pH 7.7, 50 mM KCl, 2.5 mM $MgCl_2$), and DNA was eluted by rotating overnight at room temperature. Gel fragments were removed by filtering the slurry through a 0.22 µm cellulose acetate centrifuge tube filter (Spin-X, Costar). DNA was isopropanol precipitated, washed with 70% ethanol, and resuspended in ELB. The concentration of DNA determined based on absorbance at 260 nm was 0.89 µM and the concentration of Cy5 determined based on absorbance at 650 nm was 1.8 µM, in excellent agreement with the expected stoichiometry of 2 Cy5 dyes per DNA duplex. Cy5-labeled duplex from the same preparation was used for all tethering experiments.

Cy3-labeled biotinylated duplex (for the experiments in Fig. 2): 20 pmol each of oTG048F, oTG415, oTG532, and oTG533 were combined in 20 µl 1x TE with 100 mM NaCl, heated to 95°C for 1 min, and cooled at 0.1°C/s to 25°C. Nicks were sealed by adding 2.3 µl 10x T4 DNA ligase buffer and 1 µl T4 DNA ligase (New England Biolabs) and incubating overnight at 37°C. Products were separated on a 0.5x TBE, 8% polyacrylamide (19:1 bis-acrylamide:acrylamide) gel for 30 min at 200 V. The band corresponding to the full-length product was excised after staining with ethidium bromide, and pulverized as above. The crumbled pieces were suspended in 100 µl 1x TE, and DNA was eluted by shaking at 1200 rpm for 1 h at 50°C. Gel fragments were removed by filtering the slurry through a 0.22 µm cellulose acetate centrifuge tube filter (Spin-X, Costar), and the filtrate was used without further purification.

For the Cy5 photostability control shown in Fig. 2D, a 1 kb biotinylated PCR product with the same Cy5-labeled end sequence as the 100 bp duplex used in the tethering assay was generated using primers oTG416 and oTG042F and template pTG064. The fragment was purified by agarose gel electrophoresis, electroeluted, extracted twice with isobutanol, and ethanol precipitated.

Data analysis: Using custom MATLAB scripts, movies containing multiple fields of view were automatically segmented in time based on the sudden drop in correlation in pixel intensities between successive images when the stage was moved to a different field of view. Integrated intensities derived from raw images were used for FRET efficiency calculations. For particle detection and selection of spots having both donor and acceptor fluorophores, images from each channel were spatially bandpass filtered (using the function bpass.m, available at http://site.physics.georgetown.edu/matlab/code.html) as follows: Image I_A was generated by convolving the original image with a Gaussian kernel of width (standard deviation) 1.4 pixels, and image I_B was generated by convolving the original image with a boxcar kernel of width 10 pixels. The

bandpass filtered image was obtained by subtracting I_B from I_A. Unlike in the original bpass.m script, pixels with negative values were not set equal to zero.

To align the Cy3 and Cy5 channels, calibration images were taken during each experiment, either of a trans-illuminated grid of 150 nm holes in a thin film of aluminum or of a sample of TetraSpeck™ fluorescent microspheres (Life Technologies). A linear translation was determined that maximized the Pearson correlation coefficient between pixel intensities in the two channels. Particles were identified in one channel, and the positions of the corresponding particles in the other channel were estimated by applying the linear translation. The particles in both channels were then localized more precisely by fitting to a 2-dimensional Gaussian function, and a nonlinear transformation between coordinates in the two channels was constructed using the MATLAB function cp2tform. Particles were located in the bandpass-filtered Cy3 channel using a script derived from pkfnd.m (http://site.physics.georgetown.edu/matlab/code.html) to find local intensity maxima greater than the 95th percentile of all pixel intensities. Particles within a radius of 5 pixels of other particles were excluded from the analysis to avoid crosstalk. The fluorescence intensity of each particle over time was determined by summing the pixels within a circular region of interest (ROI) of radius 4 pixels centered on the particle. The local background intensity was determined as the average of the intensities of pixels flanking the ROI. Integrated intensities were determined for both the raw image (for FRET efficiency calculation) and the bandpass-filtered image (for tethering and photobleaching event detection). The statistics of Cy5-labeled duplex binding and dissociation were determined as follows: The zero-Cy5 peak of the Cy5 intensity histogram was fit locally to a Gaussian distribution, and the detection threshold for Cy5 appearance was set at two standard deviations above the mean of the zero-Cy5 peak. Similarly, the single-Cy3 peak in the Cy3 intensity histogram was fit locally to a Gaussian distribution, and a threshold for Cy3 photobleaching was set at two standard deviations below the mean. Sets of successive acceptor excitation frames in which the Cy5 intensity exceeded the detection threshold were considered tethering events. Frames subsequent to Cy3 photobleaching were excluded, as Cy3 photobleaching made it impossible to determine if the complex was in a low-FRET (long-range complex) or high-FRET (short-range complex) state (see below). The tethering rate, *r*, was calculated using the formula

$$r = \frac{\sum_i e_i}{\sum_i t_i}$$

where e_i is the number of tethering events observed for the ith Biotin-100 bp-Cy3 duplex and t_i is the total amount of time that the ith Biotin-100 bp-Cy3 duplex was "available" for detection of binding events (i.e., Cy5 not already bound and Cy3 not photobleached). Total numbers of Cy3-DNAs tracked are given in Table S1. The nonspecific tethering rate was determined from movies without Biotin-100 bp-Cy3 by applying the same analysis (without the Cy3 photobleaching criterion) to randomly chosen spot centers in the Cy5 channel. Survival times were taken to be the number of frames in which Cy5 was present multiplied by the alternating laser excitation period (2 s).

Survival curves were constructed using the Kaplan-Meier procedure, regarding as right-censored those particles for which 1) the molecule transitioned to a high-FRET (short-range complex) state with FRET efficiency > 0.4, 2) Cy3 photobleached, making it impossible to determine if the molecule was in a high-FRET or low-FRET state, or 3) the molecule remained in a tethered state at the end of the observation interval. To estimate the fraction of long-range complexes that transitioned to a short-range complex (Fig. S4), transitions to high FRET were considered to occur when the FRET efficiency exceeded 0.4 for four consecutive donor excitation frames (~8 seconds). This criterion is reasonable, given that most short-range complexes survive much longer than this (Fig. 5B).

Statistical comparison of tethering rates and long-range complex lifetimes between different conditions was

performed using one-way ANOVA (MATLAB function anova1) followed by Tukey's post-hoc test (MATLAB function multcompare). The Kolmogorov-Smirnov test (MATLAB function kstest) was used to check for normality of the input data. Because of the order-of-magnitude difference in tethering rates observed between different conditions, statistical tests were performed on the *logarithm* of the tethering rate to better satisfy the assumption of homoscedasticity.

For Cy5 photostability measurements, photobleaching (or blinking) events were detected using an intensity threshold two standard deviations below the single-Cy5 intensity peak, and a survival curve was calculated using the Kaplan-Meier procedure, considering events to be right-censored if Cy5 had not bleached or blinked by the end of the movie.

Single-molecule circularization assay and image analysis

Substrate generation: A 2 kb long DNA fragment with a Cy3 label 7 bp from one end and a Cy5 label 7 bp from the other end was generated by PCR with template pTG059 and primers oTG415 and oTG416. This was gel purified and digested with the nicking restriction endonuclease Nb.BbvCI, which cuts at two tandem sites 25 bp apart in the center of the amplicon. The doubly nicked DNA was mixed with an approximate 10-fold molar excess of an internally biotinylated oligonucleotide (oTG441) with the same sequence as the DNA strand between the tandem nicks. The mixture was heated to 80°C for 20 min to inactivate the nicking enzyme and denature the short DNA segment between the nicks. It was then slowly cooled at 0.1°C/s to room temperature to permit annealing of oTG441. The DNA was treated with T4 DNA ligase to seal the nicks in the biotinylated strand, and the ligase was heat-inactivated by incubation at 65°C for 10 min.

Circularization assay: A detailed description of the circularization assay can be found in Experimental Procedures. For sodium dodecyl sulfate (SDS) wash experiments (Fig. 3h), substrates were pre-incubated for 30 min with XRCC4-depleted extract supplemented with either wild-type LIG4:XRCC4 or catalytically inactive $LIG4^{K278R}$:XRCC4. Following the 30 min incubation, substrates were imaged stroboscopically for 5 min with a 100 ms exposure every 500 ms, alternating between two frames of 532 nm excitation (surface power density 16 W/cm^2) and one frame of 641 nm excitation (surface power density 7 W/cm^2), and moving to a new field of view every 18 s. The flowcell was then washed sequentially with 200 µl each of ELB, TE, TE with 1% SDS, TE, and ELB at a flow rate of 100 µl/min using a syringe pump (Harvard Apparatus). Finally, the flowcell was filled with imaging solution (1x PBS with 45% glycerol, 0.2% n-propyl gallate, and the same concentrations of PCA, PCD, and Trolox used for imaging in extract), and substrates were imaged continuously with an exposure time of 500 ms/frame and the same illumination intensities as before. A longer exposure time was necessary due to reduced brightness of the fluorophores following substrate deproteinization.

Image analysis: Image analysis was performed essentially as described above for tethering experiments, with the following modifications: Particles were identified in the Cy5 channel with 641 nm excitation to avoid introducing a bias against high-FRET molecules. Particles within a radius of 7 pixels of other particles were excluded from the analysis to avoid crosstalk. A piecewise translation function between the two channels was generated by dividing the image into 8 sectors and determining the linear translation that gave the highest Pearson correlation coefficient between pixel intensities in the two channels in each sector, averaged over 5 different fields of view. Stage drift was estimated in long timecourse movies using image correlation, and ROI positions were translated to compensate. Transitions between low and high FRET and photobleaching events were annotated manually in long trajectories. During manual annotation, thumbnail movies of each ROI were reviewed, and trajectories were excluded if 1) the Cy3 and Cy5 spots were poorly centered in the

ROI, 2) more than one particle was visible in the ROI, or 3) more than one Cy5 photobleaching step was evident. Kaplan-Meier survival curves of the high-FRET state (Fig. 5B) were constructed from manual annotations, regarding as right-censored 1) substrates that remained in a high-FRET state at the end of the observation interval or 2) substrates for which Cy3 or Cy5 photobleached.

Calculation of corrected FRET efficiencies

FRET efficiencies were corrected for Cy3 bleedthrough into the Cy5 channel and Cy5 direct excitation by 532 nm light as previously described (Lee et al., 2005). Signal in the Cy5 channel due to Cy3 bleedthrough was determined to be ~6.5% of the Cy3 signal in the Cy3 channel. Cy5 signal due to direct excitation by 532 nm light was found to be ~4.6% of the Cy5 signal due to excitation by 641 nm laser light under illumination conditions used in circularization experiments and ~6.8% under illumination conditions used in tethering experiments and long timecourse circularization experiments. A correction factor for differences in Cy3 and Cy5 quantum yield and detection efficiency (the "γ" factor in (Lee et al., 2005)) was determined by monitoring the change in Cy3 and Cy5 signal upon photobleaching of Cy5. A γ factor of ~0.8 was obtained in extract, that is, (quantum yield)*(detection efficiency) for Cy5 was about 80% that of Cy3.

Plasmids used in this study

Plasmid number	Construct name	Description	Source	Associated figures
pTG059	pBS-2xBbvCI	pBluescript II KS(-)-derived plasmid containing tandem BbvCI sites for oligonucleotide insertion; used as PCR template to generate the single-molecule circularization substrate.	This work	Figs. 3-5, S3, and S5
pTG064	*parS*/ pBluescript II KS(-)	Used as a PCR template to generate a Cy5 photostability control substrate.	This work	Fig. 2D
pTG234	xKu70/pCMV-Sport6	*X. laevis* Ku70 cDNA	Hironori Funabiki (Rockefeller University)	
pTG235	xKu80/pCMV-Sport6	*X. laevis* Ku80 cDNA	Hironori Funabiki (Rockefeller University)	
pTG246	xKu80/ pFastBac1	Untagged *X. laevis* Ku80 in baculoviral expression vector	This work	Figs. 1A, 2E-F, 3, 4C, S1C, S2D, S4
pTG247	xKu70-Flag/ pFastBac1	C-terminally Flag-tagged *X. laevis* Ku70 in baculoviral expression vector	This work	
pTG257	H10-SUMO	SUMO fusion expression vector vector with 10xHis tag	This work	
pTG259	H6-SUMO-	Codon-optimized SUMO expression	This work	Fig. S1D

	hXLF	vector for *Homo sapiens* XLF		
pTG272	XGC xrcc4 cDNA	*X. laevis* XRCC4 cDNA (CloneID: 6635763)	Thermo Scientific	
pTG273	XGC lig4 cDNA	*X. laevis* LIG4 cDNA (CloneID: 6957895)	Thermo Scientific	
pTG275	Flag-xLIG4-H6/ pFastBac1	*X. laevis* LIG4 with N-terminal Flag tag and C-terminal H6 tag in pFastBac1	This work	Figs. 1D, 5, S1
pTG276	xXRCC4-PP-StrepII/ pFastBac1	*X. laevis* XRCC4 with C-terminal StrepII tag preceded by a PreScission Protease cleavage site	This work	Figs. 1D, 5, S1, S5
pTG281	Flag-xLIG4-K278R-H6/ pFastBac1	Catalytically inactive mutant of *X. laevis* LIG4 in baculoviral expression vector	This work	Figs. 1D, 5, S1, S5
pTG283	xDNA-PKcs/ pCMV-Sport6	Partial *X. laevis* DNA-PKcs cDNA clone (CloneId: 8549275)	Thermo Scientific	
pTG284	xKu70-H6/pFastBac1	C-terminally His6-tagged *X. laevis* Ku70 in baculoviral expression vector	This work	Figs. 1A, 2E-F, 3, 4C, S1C, S2D, S4
pTG288	xXLF/pMA-T	Codon-optimized *X. laevis* XLF coding sequence	This work (synthesized by Life Technologies)	
pTG292	H10-SUMO-xXLF (start site #1)	H10-SUMO-tagged *X. laevis* XLF starting from the first candidate translation start site (not used for protein expression)	This work	
pTG296	H10-SUMO-xXLF (start site #3)	H10-SUMO-tagged *X. laevis* XLF starting from the third candidate translation start site	This work	Fig. 1D, 4F
pTG313	λ-EcoRI/ pBluescript II KS(-)	Plasmid containing the EcoRI fragment (λ bacteriophage DNA bp 21226-26104) used for bulk end joining experiments	This work	Fig. 1
pTG314	H6-xDNA-PKcs-PIKK-FATC/ pET28b	Expression vector for C-terminal PIKK-FATC fragment of *X. laevis* DNA-PKcs.	This work	

Plasmid generation

The plasmids listed in the preceding section were generated as follows and verified by sequencing.

pTG059: Oligonucleotides oTG033F and oTG033R were annealed, and the resulting duplex was ligated into pBluescript II KS(-). The insert contains two tandem BbvCI sites that can be nicked by the nicking restriction endonuclease Nb.BbvCI to release a short oligonucleotide.

pTG064: Oligonucleotides oTG041F and oTG041R were annealed, phosphorylated, and blunt-end ligated into the EcoRV site of pBluescript II KS(-).

pTG246: The *X. laevis* Ku80 coding sequence was PCR amplified from pTG235 using primers oTG370 and oTG371 and inserted into pFastBac1 cut with BamHI by Gibson assembly using a homemade master mix (see https://wiki.med.harvard.edu/SysBio/Megason/IsothermalAssembly).

pTG247: The *X. laevis* Ku70 coding sequence was PCR amplified from pTG234 using primers oTG368 and oTG369, and the pFastBac1 vector backbone was amplified with oTG397 and oTG398. The two fragments were combined by Gibson assembly.

pTG257: pH6-SUMO (Graham et al., 2014) was PCR amplified with primers oTG410 and oTG432. The PCR product was phosphorylated with T4 polynucleotide kinase (PNK) and circularized with T4 DNA ligase.

pTG259: A codon-optimized human XLF coding sequence (Supplemental Text 2) was amplified with primers oTG429 and oTG430 and inserted into BamHI-digested pH6-SUMO by Gibson assembly.

pTG275: *X. laevis* LIG4 was PCR amplified from pTG273 using primers oTG484 and oTG485, and a pFastBac1 vector fragment was PCR amplified using primers oTG482 and oTG483. The two fragments were combined by Gibson assembly.

pTG276: *X. laevis* XRCC4 was PCR amplified from pTG272 using primers oTG480 and oTG481, and a pFastBac1 vector fragment was PCR amplified using primers oTG397 and oTG446. The two fragments were combined by Gibson assembly.

pTG281: A linear DNA fragment was generated by PCR with template pTG275 and primers oTG502 and oTG503, phosphorylated with T4 PNK, and circularized with T4 DNA ligase.

pTG284: A PCR product was generated using template pTG247 and primers oTG506 and oTG483, phosphorylated with T4 PNK, and circularized with T4 DNA ligase.

pTG292: A codon-optimized version of *X. laevis* XLF (largest putative coding sequence beginning with start codon #1 in Fig. S1K) was PCR amplified from pTG288 with primers oTG508 and oTG509 and subcloned by Gibson assembly into BamHI-digested pTG257.

pTG295: A PCR product was generated with template pTG292 and primers oTG410 and oTG541, phosphorylated with T4 PNK, and circularized with T4 DNA ligase.

pTG296: A PCR product was generated with template pTG292 and primers oTG410 and oTG540, phosphorylated with T4 PNK, and circularized with T4 DNA ligase.

pTG313: λ bacteriophage genomic DNA (New England Biolabs) was digested with EcoRI, and the fragment corresponding to basepairs 21226-26104 was ligated into EcoRI-digested pBluescript II KS(-).

pTG314: A fragment from the 3' end of the *X. laevis* DNA-PKcs cDNA sequence was amplified from pTG283 with primers oTG525 and oTG583 and subcloned by Gibson assembly into pET28b digested with NdeI.

Oligonucleotides used in this study

Oligonucleotides were obtained from Integrated DNA Technologies (IDT).

Manufacturer's abbreviations:

/5Biosg/	5' biotin
/iCy3N/	C6-amino-dT conjugated to Cy3 N-hydroxysuccinimidyl ester (off-catalog modification)
/iCy5N/	C6-amino-dT conjugated to Cy5 N-hydroxysuccinimidyl ester (off-catalog modification)
/5Phos/	5' phosphate
/iBiodUK/	Internal biotin (azide)-dT

Number	**Name**	**Sequence**	**Description**
oTG041F	parS1_for	CAGTTGAATCAGAATGTTCCACGTGAAACAAAGAAAAAA	Used to generate pTG064
oTG041R	parS1_rev	TTTTTTCTTTGTTTCACGTGGAACATTCTGATTCAACTG	Used to generate pTG064
oTG042F	Bio_parS1_for	/5Biosg/ CAGTTGAATCAGAATGTTCCACGTGAAACAAAGAAAAAA	Used to generate Cy5 photostability control substrate
oTG048F	Bio_parS5_for	/5Biosg/CAGTTGAATCTGACAAATGACTAACAATGAGAGCAAAAAGAACCTGTGAAAGATGCGGTTCTAA	For generating biotin-100 bp-Cy3 duplex.
oTG368	pFastBac1-	ACCATCGGGCGCGGATCCATGACTGAATGGGGGGAT	for subcloning *X. laevis*

	xKu70_F	CA	Ku70 into pFastBac1
oTG369	pFastBac1-xKu70-Flag_R	TCATCGTCGTCCTTGTAGTCGTTCTTCTTGAAATACTCTACAAGGGC	for subcloning *X. laevis* Ku70 into pFastBac1
oTG370	pFastBac1-xKu80_F	ACCATCGGGCGCGGATCCATGGCGCGAGCAGCCAAAAG	for subcloning *X. laevis* Ku80 into pFastBac1
oTG371	pFastBac1-xKu80_R	TTCCGCGCGCTTCGGACCGGGATTACATCATGTCAAGCAAGTCGT	for subcloning *X. laevis* Ku80 into pFastBac1
oTG397	pFastBac1_R	GGATCCGCGCCCGATGGT	for cloning various inserts into pFastBac1 by Gibson assembly
oTG398	pFastBac1-Flag_F	GACTACAAGGACGACGATGACAAGTAACGGTCCGAAGCGCGCGGAA	for cloning various inserts into pFastBac1 by Gibson assembly
oTG410	SUMO_R	TCCACCAATCTGTTCTCTGTGA	For generating pH10-SUMO
oTG413	pBS_2671_F	GGATCTTACCGCTGTTGAGATC	Unlabeled version of the Cy3-labeled primer used to generate single-molecule substrates.
oTG414	pBS_1709_R	AACTCTTTTTCCGAAGGTAACTGG	Unlabeled version of the Cy5-labeled primer used to generate single-molecule substrates.
oTG415	pBS_2671_Cy3_F	GGATCT/iCy3N/ACCGCTGTTGAGATC	Internally Cy3-labeled primer used to generate single-molecule substrates.
oTG416	pBS_1709_Cy5_R	AACTCT/iCy5N/TTTCCGAAGGTAACTGG	Internally Cy5-labeled primer used to generate single-molecule substrates.
oTG429	H6-SUMO-hXLFco-ITA_F	GGCTCACAGAGAACAGATTGGTGGAGGCGGAGGTTCAATGGAAGAACTGGAACAGGG	for subcloning codon-optimized human XLF into pH6-SUMO
oTG430	H6-SUMO-	ACAAGCTTATTACTCGAGGAGCTCGGATTAGCTGAAC	for subcloning codon-optimized human XLF into

	hXLFco-ITA_R	AGACCGCGCGG	pH6-SUMO
oTG432	pH10-SUMO_F2	CACCATCACCATCACCACCACCACCACCATATGGCTAGCG	for generating pH10-SUMO
oTG441	IntBio-TOM-insert	/5Phos/TGAGGGATATCGAA/iBiodUK/TCCTGCAGGC	for inserting internal biotin tether in NHEJ circularization substrate
oTG446	pFastBac1-PP-StrepII_F	TTTCAAGGTCCTTGGAGCCACCCTCAGTTCGAAAAGTAACGGTCCGAAGCGCGCGGAA	to subclone constructs into pFastBac-1 with a PreScission protease-cleavable C-terminal StrepII tag
oTG480	pFB1-xXRCC4_F	ACCATCGGGCGCGGATCCATGGAGAAAAAAATAAGAAGTATCTGCATATC	For cloning *X. laevis* XRCC4 into pFastBac1
oTG481	pFB1-xXRCC4-PP-StrepII_R	AGGGTGGCTCCAAGGACCTTGAAAAAGAACTTCAAGAATGTCACTGAAGAGGTCATCTGG	For cloning *X. laevis* XRCC4 into pFastBac1
oTG482	pFastBac1_Flag_R	CTTGTCATCGTCGTCCTTGTAGTCCATGGATCCGCGCCCGATGGT	for cloning into pFastBac1 by Gibson assembly with an N-terminal Flag tag
oTG483	pFastBac1-H6_F	CATCACCACCACCACCATTAACGGTCCGAAGCGCGCGGAA	for cloning into pFastBac1 by Gibson assembly with a C-terminal H6 tag
oTG484	pFastBac1-Flag-xLIG4_F	ACAAGGACGACGATGACAAGTCGACGCCGAAAGCCTCAGA	for cloning xLIG4 into pFastBac1 with a Flag tag at the N terminus
oTG485	pFastBac1-xLIG4-H6_R	TTAATGGTGGTGGTGGTGATGTAGCAAATAGGCATTTTCCATCTGA	for cloning xLIG4 into pFastBac1 with a H6 tag at the C terminus
oTG502	xLIG4_K278R_F	AGACTGGACGGAGAACGGATG	for making K278R mutant *X. laevis* LIG4
oTG503	xLIG4_831_R	CGTTTCAATGAAAAAACTTTGGTGGTTC	for making K278R mutant *X. laevis* LIG4
oTG506	xKu70_R	GTTCTTCTTGAAATACTCTACAAGGGC	xKu70 3’ reverse primer

oTG508	SUMO-xXLF_F	GGCTCACAGAGAACAGATTGGTGGAATGTATGGCACCTTTCATCTGC	For cloning xXLF into SUMO vectors
oTG509	SUMO-xXLF_R	AAGCTTATTACTCGAGGAGCTCGGATTACATAAACAGGCCTTTTGCTT	For cloning xXLF into SUMO vectors
oTG525	H6-xDNA-PKcs-PIKK_F	GCCTGGTGCCGCGCGGCAGCCATATGTACCCCTTTCTGGTAAAAGGC	For subcloning a fragment of xDNA-PKcs into pET28b
oTG532	Adapter1	/5Phos/GTTACATCGAACTGGATCTCAACAGCGGTAAGATCC	For generating Biotin-100 bp-Cy3 duplex
oTG533	Adapter2	/5Phos/CAGTTCGATGTAACTTAGAACCGCATCTTTCACAGGTTCTTTTTGCTCTCATTGTTAGTCATTTGTCAGATTCAACTG	For generating Biotin-100 bp-Cy3 duplex
oTG538	Cy5-Cy5-adapter1	/5Phos/ACATTTACTCTCTAACATCACGCCTAGATAGAAACAGATAGCTTGAACAGATCCAGTTACCTTCGGAAAAAGAGTT	for generating Cy5-labeled 100 bp duplex
oTG539	Cy5-Cy5-adapter2	/5Phos/ATCTGTTCAAGCTATCTGTTTCTATCTAGGCGTGATGTTAGAGAGTAAATGTCCAGTTACCTTCGGAAAAAGAGTT	for generating Cy5-labeled 100 bp duplex
oTG540	SUMO-GGGS-xXLF_start1_F	GGCGGAGGTTCAATGAGCCTGAGCCAAGAAATG	for generating H10-SUMO-xXLF beginning with candidate start site 3
oTG541	SUMO-GGGS-xXLF_start2_F	GGCGGAGGTTCAATGGATGCACGTCTGCTGCAG	for generating H10-SUMO-xXLF beginning with candidate start site 4
oTG583	H6-xDNA-PKcs-PIKK-FATC_R	TGTCCACCAGTCATGCTAGCCATTTATATCCATGGCTCCCACCC	For subcloning a fragment of xDNA-PKcs into pET28b

Supplemental Text

Supplemental Text 1: Codon-optimized *X. laevis* XLF sequence

ATGAGCCTGAGCCAAGAAATGGATGCACGTCTGCTGCAGCTGCCGTGGCGTAGCCTGCGTATTGCAGATTG
CATTTTTATGGGTAAAGTGTGCTTTACCGAAAGCAGCTATGCACTGCTGCTGAGCGATCTGAGCAGCATGT
GGTGTGAAGAAGCAAAAGCAGATATCATTCAGGATCGTGCACGTGAACTGAATAAACGTCTGAAAGCACCG
GTTAGCAGCTTTCTGAGCTATCTGTCACAGATTGTTTTTCCGGTTCTGAACAGCAAAGATAACGGCCAGAA
CATTTTTAGCTGTCATCGTAGCGAAGCAGAGCTGCTGCTGCAGGTTAAAAGCCAGCTGAGCGGTCTGCCGT
TTTATTGGAGCTTTCATTGTAAAGAAGCAACCGTTAGTACCGTTTGTCGTCATCTGGTTCGTCCGCTGAAA
AGCATGACCGAAGCACTGGAAAGCCAGAATCAAGAACTGTGTCTGCTGCTGAAAAAAAAAGATGCCGAAAT
CCAAGAGTATCAGGATAGCGGTGCAGTTCTGACCCGTGATCGTCTGAAAACCGAAGTTTTTGATGAACTGA
AATTCCAGAAAAGCTTCCTGGCAGAAAAAGTTCAGGGTCTGTGTATGAGCGGTAAAGCTCCGGGTTTTAGC
GAACAGCTGCAACAGCTGTATGATGCAGTTATTGCACCTAAAGCACCGACCCATCCGAAAGAAGAAGATAC
CGGTAATAGCGCAAGCCATCGTCCGATGGCAGAAAGCAGCAGCATTAGCTTTGAAAAAACCGTTCCGACCC
AAGAACGTACCGAAGGTGGTGCAGTTAGCGAACCGAGCCAGGTTCCGCAGAGCAGCGTTAGCTGTCTGACC
CATCGTCCTCCGGCAGGCGCAAGCAAACCTAAAAAAAAAGCAAAAGGCCTGTTTATGTAA

Supplemental Text 2: Codon-optimized *H. sapiens* XLF sequence

ATGGAAGAACTGGAACAGGGTCTGCTGATGCAACCGTGGGCCTGGCTGCAGCTGGCCGAAAATAGCCTGCT
GGCCAAGGTGTTTATCACCAAACAGGGCTACGCCCTGCTGGTTAGTGATCTGCAGCAGGTGTGGCATGAGC
AGGTGGATACAAGCGTGGTGAGCCAGCGTGCCAAAGAACTGAATAAACGCCTGACCGCCCCGCCTGCCGCA
TTTCTGTGTCACCTGGATAACCTGCTGCGCCCTCTGCTGAAAGATGCCGCACATCCTAGCGAAGCCACATT
CAGTTGCGACTGCGTGGCCGATGCACTGATCCTGCGTGTGCGCAGCGAACTGAGCGGTCTGCCGTTCTACT
GGAACTTCCATTGCATGCTGGCCAGCCCGAGCCTGGTTAGCCAGCATCTGATCCGCCCGCTGATGGGTATG
AGTCTGGCACTGCAGTGTCAGGTGCGTGAGCTGGCCACCCTGCTGCACATGAAGGACCTGGAGATTCAGGA
TTACCAGGAGAGTGGCGCAACACTGATTCGCGATCGCCTGAAAACCGAGCCGTTCGAAGAAAATAGCTTCC
TGGAGCAGTTCATGATCGAGAAACTGCCGGAAGCCTGCAGTATCGGCGACGGCAAGCCTTTCGTGATGAAT
CTGCAGGACCTGTATATGGCCGTGACAACACAGGAAGTGCAGGTGGGTCAAAAACACCAGGGTGCCGGTGA
TCCGCATACCAGCAATAGCGCCAGCCTGCAGGGCATTGACAGTCAGTGCGTTAATCAGCCGGAACAGCTGG
TTAGTAGCGCCCCGACACTGAGCGCACCGGAAAAAGAAAGCACCGGTACAAGCGGTCCGCTGCAGCGTCCT
CAGCTGAGCAAAGTGAAGCGTAAAAAACCGCGCGGTCTGTTCAGC

Supplemental Text 3: XLF sequence alignment

```
Xl_XLF          -----MYGTFHLQHLFTFPSLAFSVSTSRWRT-DICCID---ETRNYCVFVFLAGRCSSW 51
Xt_XLF          MSYNEVYGAIKQTGDHKKPLEGHIHNSPAHDTAALECTERFSETRNYRVFVFLAGRCSSW 60
Hs_XLF          ------------------------------------------------------------

                        3     4
Xl_XLF          TRREQTQDMSLSQEMDARLLQLPWRSLRIADCIFMGKVCFTESSYALLLSDLSSMWCEEA 111
Xt_XLF          TRRQ--QDMSFSQGTDAQLLQLPWRTLRIGDCTFLGKVCFAESGYALLLSDLSSVWCEEA 118
Hs_XLF          -----------MEELEQGLLMQPWAWLQLAENSLLAKVFITKQGYALLVSDLQQVWHEQV 49
                           :  :  **  **  *::.:  ::.** ::::..****:***..:* *:.

Xl_XLF          KADIIQDRARELNKRLKAPVSSFLSYLSQIVFPVLNSKDN-GQNIFSCHRSEAELLLQVK 170
Xt_XLF          EADAIQERARKLNKRLKAPVSSFLSYLSQIVFPALDSKDN-GQNVFSCHRTEAELVLQVK 177
Hs_XLF          DTSVVSQRAKELNKRLTAPPAAFLCHLDNLLRPLLKDAAHPSEATFSCDCVADALILRVR 109
                .:. :.:**::*****.** ::**.:*.::: * *..  : .:  ***.     *:*:*:

Xl_XLF          SQLSGLPFYWSFHCKEATVSTVCRHLVRPLKSMTEALESQNQELCLLLKKKDAEIQEYQD 230
Xt_XLF          SQLSGLPFYWNFHCKEAKVSTVCRHLVNPLMSMAEALERQNQELYLLLGKKDAEIQEYQD 237
Hs_XLF          SELSGLPFYWNFHCMLASPSLVSQHLIRPLMGMSLALQCQVRELATLLHMKDLEIQDYQE 169
                *:********.***  *. * *.:**:.** .*: **: * :**  **  ** ***:**:

Xl_XLF          SGAVLTRDRLKTEVFDELKFQKSFLAEKVQGLCMSGKAPGFSEQLQQLYDAVIAPKAPTH 290
```

```
Xt_XLF          SGAVLTRGRLKTEVFDELKFQESFMAEKVQVLSKSGETHGFSEQLQRLYHALTAP-APVN 296
Hs_XLF          SGATLIRDRLKTEPFEENSFLEQFMIEKLPEACSIGDGKPFVMNLQDLYMAVTTQEVQVG 229
                ***.* *.***** *:* .* :.*: **:   .  *.   *  :** ** *: :  . .

Xl_XLF          PKEEDTGNSASHRPMAESSSISFEKTVPTQERTEGGAVSEPSQVPQSSVSCLTHRPPAGA 350
Xt_XLF          PKGEDVGDSDNLPPVAESNRENTNIPFGKTEPTQEGAGPETSQVPQCSLPCLTHRPPTAA 356
Hs_XLF          QKHQGAGDPHTSNSASLQGIDSQCVNQPEQLVSSAPTLSAPEKES-TGTSGPLQRP--QL 286
                 * :..*:. .  . : ..  .          :.  : . ..: .  . .   :**

Xl_XLF          SKPKKK-AKGLFM- 362
Xt_XLF          SKPKKKKAKGLFT- 369
Hs_XLF          SKVKRKKPRGLFS- 299
                ** *:* .:***
```

Xl_XLF: Longest candidate *Xenopus laevis* XLF sequence (see Fig. S1K).

Xt_XLF: Predicted *Xenopus tropicalis* protein XP_002933989.2

Hs_XLF: Human XLF, NP_079058.1

M #3 and M #4 – possible translation start sites based on sequence alignment and mass spectrometry data (Fig. S1K); recombinant XLF proteins beginning at either start site rescued the end joining defect of XLF-immunodepleted extract (Fig. 1D and data not shown)

SUPPLEMENTAL REFERENCES:

Chen, X., Pascal, J., Vijayakumar, S., Wilson, G.M., Ellenberger, T., and Tomkinson, A.E. (2006). Human DNA Ligases I, III, and IV-Purification and New Specific Assays for These Enzymes. Methods Enzymol. 409, 39–52.

Graham, T.G.W., Wang, X., Song, D., Etson, C.M., van Oijen, A.M., Rudner, D.Z., and Loparo, J.J. (2014). ParB spreading requires DNA bridging. Genes Dev. 28, 1228–1238.

Lee, N.K., Kapanidis, A.N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R.H., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. Biophys. J. 88, 2939–2953.

Wühr, M., Freeman, R.M., Presler, M., Horb, M.E., Peshkin, L., Gygi, S.P., and Kirschner, M.W. (2014). Deep proteomics of the xenopus laevis egg using an mRNA-derived reference database. Curr. Biol. 24, 1467–1475.