A single XLF dimer bridges DNA ends during nonhomologous end joining

Thomas G. W. Graham, Sean M. Carney, Johannes C. Walter, and Joseph J. Loparo

Nonhomologous end joining (NHEJ) is the primary pathway of DNA double-strand-break repair in vertebrate cells, yet how NHEJ factors assemble a synaptic complex that bridges DNA ends remains unclear. To address the role of XRCC4-like factor (XLF) in synaptic-complex assembly, we used single-molecule fluorescence imaging in Xenopus laevis egg extract, a system that efficiently joins DNA ends. We found that a single XLF dimer binds DNA substrates just before the formation of a ligation-competent synaptic complex between DNA ends. The interaction of both globular head domains of the XLF dimer with XRCC4 is required for efficient formation of this synaptic complex. Our results indicate that, in contrast to a model in which filaments of XLF and XRCC4 bridge DNA ends, binding of a single XLF dimer facilitates the assembly of a stoichiometrically well-defined synaptic complex.

Canonical NHEJ is the major pathway of DNA double-strand break (DSB) repair in vertebrate cells. Successful rejoining of broken DNA ends requires that they be held together in a manner that permits their processing and ligation. DNA ends are first bound by the Ku heterodimer, composed of Ku70 and Ku80, which fits over the DNA end like a ring. Ku recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), whose kinase activity is required for efficient end joining. Ends are ultimately ligated by DNA ligase 4 (LIG4), which resides in a complex with XRCC4, referred to as the SR-complex. A homolog of XRCC4, named XLF or Cernunnos, was identified as an XRCC4-interacting protein whose disruption in humans results in immunodeficiency, developmental defects, and cellular radiosensitivity. Despite the identification of these core components of the NHEJ machinery, the nature of the synaptic complex that aligns broken DNA ends, in particular the roles of XLF and XRCC4–LIG4, remains poorly defined.

Structure characterization of XLF and XRCC4 has revealed that both proteins form symmetric homodimers that interact through their globular head domains, and mutations in XLF or XRCC4 that disrupt this interaction impair NHEJ in cell culture-based assays. Two possible roles of the XLF–XRCC4 interaction in NHEJ have been proposed. First, XLF may directly promote catalysis by XRCC4–LIG4. XLF stimulates XRCC4–LIG4 activity in vitro, especially on noncohesive and mismatched ends. XLF likewise increases the rate of LIG4 autoadenylation, the initial step in the LIG4 catalytic cycle. The second proposal is that alternating filaments of XLF and XRCC4 bridge broken DNA ends. Filaments of purified XLF and XRCC4 are seen in crystal structures and electron micrographs, and a mixture of XLF and XRCC4 bridges DNA molecules in bulk and single-molecule assays. Additionally, super-resolution imaging of fixed cells stained with anti-XLF or anti-XRCC4 antibodies has revealed elongated nuclear foci proposed to be XLF–XRCC4 filaments. Although these results are suggestive, XLF–XRCC4 filaments have not been demonstrated to be necessary for physiological end joining.

To study the role of the XLF–XRCC4 interaction in NHEJ, we used Xenopus laevis egg extract, which performs efficient NHEJ dependent on the core pathway components Ku, DNA-PKcs, XLF, XRCC4, and LIG4. We previously combined this cell-free system with a single-molecule Förster resonance energy transfer (smFRET) assay that allowed us to visualize synopsis of single pairs of DNA ends during NHEJ. These experiments revealed a stepwise transition from an initial ‘long-range’ synaptic complex (LR complex), in which DNA ends are physically tethered but not held closely together, to a ‘short-range’ synaptic complex (SR complex), in which DNA ends are closely aligned for ligation. The SR complex either ligates DNA ends or dissociates, requiring another round of SR-complex formation before eventual ligation. Formation of the LR complex requires Ku and DNA-PKcs, in agreement with the previously reported dimerization and DNA end-bridging activities of the DNA-PK holoenzyme. Transition from the LR complex to the SR complex requires DNA-PK catalytic activity, XLF, and XRCC4–LIG4, but not LIG4 catalytic activity. A noncatalytic role of LIG4 in the stable bridging of DNA ends is consistent with previous results from pulldown experiments in mammalian cell-free extracts. A subsequent study in a reconstituted mixture of human NHEJ proteins has indicated a transition from a transient synaptic complex dependent on Ku and DNA-PKcs to a more stable complex dependent on XLF and XRCC4–LIG4. These observations closely parallel our results in Xenopus egg extract, thus suggesting that the mechanisms of synaptic-complex assembly are conserved between frogs and humans.

Here, we investigated the role of the XLF–XRCC4 interaction in synaptic-complex formation. Using X. laevis egg extract, we show that the XLF–XRCC4 interaction is required for formation of the SR complex. In three-color single-molecule imaging experiments, we visualized binding of fluorescently labeled XLF protein to individual DNA substrates undergoing synopsis. Strikingly, formation of the SR complex involves association of only a single XLF dimer with DNA ends. Experiments with a synthetic XLF tandem dimer imply that XLF nonetheless interacts with XRCC4 through both globular head domains of the XLF dimer. These results support a model in which XLF binds with defined stoichiometry to the synaptic complex, thereby facilitating a structural rearrangement that closely aligns DNA ends.
Results

XLF–XRCC4 interaction is required for end joining in Xenopus egg extract. To explore the function of the XLF–XRCC4 interaction during physiological end joining, we first tested whether Xenopus laevis XLF and XRCC4 interact. To this end, we used biolayer interferometry (BLI) to measure the association of X. laevis XLF with immobilized X. laevis XRCC4 (unless otherwise specified, XLF and XRCC4 refer to the X. laevis proteins herein). Full-length XLF bound XRCC4 and only partially dissociated when washed with buffer, thus suggesting irreversible aggregation on the surface (Supplementary Fig. 1a). Similar aggregation was previously observed in surface plasmon resonance experiments with human XLF and XRCC4 (ref. 10). However, a truncated version of XLF lacking the unstructured C-terminal region (XLF1–226, corresponding to human XLF21–224, ref. 15) bound XRCC4 and completely dissociated when washed with buffer (Fig. 1a and Supplementary Fig. 1b). X. laevis XLFL117D and XLFL68D did not rescue end joining in XLF-depleted extract (Fig. 1c) or by addition of recombinant XLF to XLF-depleted extract (Fig. 1c) or by addition of recombinant XRCC4–LIG4 to XRCC4-depleted extract (Supplementary Fig. 1d). Thus, XLF–XRCC4 interaction is required for short-range synapsis of DNA ends.

The XRCC4 K104E mutation weakened but did not eliminate interaction with XLF (Fig. 1a), thus reflecting a higher dissociation rate (Supplementary Fig. 1e,f). The denaturation temperatures of all mutants, measured by differential scanning fluorimetry, were within 1.45 °C of the wild-type (WT) protein, thus suggesting that the defect is not due to large-scale misfolding (Supplementary Fig. 1g). Mutations that disrupted the XLF–XRCC4 interaction also disrupted formation of large XLF–XRCC4 complexes in DLS experiments and abolished DNA bridging in pulldown experiments (Supplementary Fig. 1c,d).

We next tested whether interaction-deficient XLF and XRCC4 mutants might support cell-free NHEJ. Egg extract was immunodepleted of XLF or XRCC4, thus abolishing end joining of linear DNA substrates (Fig. 1c,d and ref. 38). This effect was fully rescued by addition of recombinant XLF to XLF-depleted extract (Fig. 1c) or by addition of recombinant XRCC4–LIG4 to XRCC4-depleted extract (Fig. 1d; note, depletion of XRCC4 codepletes LIG4 (ref. 38)). In contrast, XLFL117D and XLFL68D did not rescue end joining in XLF-depleted egg extract (Fig. 1c), and XRCC4 F111E–L4 did not rescue end joining in XRCC4-depleted extract (Fig. 1d), in agreement with the residual low-affinity interaction between XRCC4 F111E and XLF (Fig. 1a and Supplementary Fig. 1e,f). Together, these results suggest that in frog egg extract, as in mammals, the interaction between XLF and XRCC4 is required for end joining.

XLF–XRCC4 interaction is required for short-range synopsis of DNA ends. We previously described two related single-molecule fluorescence assays for monitoring synaptic-complex formation in egg extract38. In the ‘intermolecular’ assay, a surface-tethered DNA duplex labeled with Cy3 is exposed to extract containing a DNA duplex labeled with Cy5. Colocalization of Cy3 and Cy5 without FRET reports on the formation of the LR complex, whereas the appearance of FRET indicates the formation of the SR complex, responding mutations in the interaction between XRCC4 K104E and XLF (Fig. 1a and Supplementary Fig. 1e,f). The denaturation temperatures of all mutants, measured by differential scanning fluorimetry, were within 1.45 °C of the wild-type (WT) protein, thus suggesting that the defect is not due to large-scale misfolding (Supplementary Fig. 1g). Mutations that disrupted the XLF–XRCC4 interaction also disrupted formation of large XLF–XRCC4 complexes in DLS experiments and abolished DNA bridging in pulldown experiments (Supplementary Fig. 1c,d).

The XRCC4 K104E mutation weakened but did not eliminate interaction with XLF (Fig. 1a), thus reflecting a higher dissociation rate (Supplementary Fig. 1e,f). The denaturation temperatures of all mutants, measured by differential scanning fluorimetry, were within 1.45 °C of the wild-type (WT) protein, thus suggesting that the defect is not due to large-scale misfolding (Supplementary Fig. 1g). Mutations that disrupted the XLF–XRCC4 interaction also disrupted formation of large XLF–XRCC4 complexes in DLS experiments and abolished DNA bridging in pulldown experiments (Supplementary Fig. 1c,d).

We next tested whether interaction-deficient XLF and XRCC4 mutants might support cell-free NHEJ. Egg extract was immunodepleted of XLF or XRCC4, thus abolishing end joining of linear DNA substrates (Fig. 1c,d and ref. 38). This effect was fully rescued by addition of recombinant XLF to XLF-depleted extract (Fig. 1c) or by addition of recombinant XRCC4–LIG4 to XRCC4-depleted extract (Fig. 1d; note, depletion of XRCC4 codepletes LIG4 (ref. 38)). In contrast, XLFL117D and XLFL68D did not rescue end joining in XLF-depleted egg extract (Fig. 1c), and XRCC4 F111E–L4 did not rescue end joining in XRCC4-depleted extract (Fig. 1d), in agreement with the residual low-affinity interaction between XRCC4 F111E and XLF (Fig. 1a and Supplementary Fig. 1e,f). Together, these results suggest that in frog egg extract, as in mammals, the interaction between XLF and XRCC4 is required for end joining.
in which DNA ends are closely aligned and poised for ligation. To increase the efficiency of SR-complex formation, we also used an intramolecular assay in which a 2-kb DNA substrate labeled 7 nt from one end with Cy3 and 7 nt from the other end with Cy5 is tethered to a glass coverslip via an internal biotin (Fig. 2a). FRET between Cy3 and Cy5, detected by TIRF microscopy, indicates assembly of the SR complex. Because the Cy3 and Cy5 labels on the intramolecular substrate always colocalize, this assay cannot detect LR-complex assembly.

Given our previous demonstration that XLF and XRCC4–LIG4 are dispensable for LR-complex formation36, we used the intramolecular assay to address whether the XLF–XRCC4 interaction is required for SR-complex formation. For these experiments, many fields of view were imaged sequentially, and the fraction of DNA substrates with FRET efficiency >0.25 was plotted as a function of time after extract addition (Fig. 2b,c). As shown previously36, depletion of XLF or XRCC4 abolished SR-complex formation, and these defects were rescued by purified recombinant XLF or XRCC4–LIG4, respectively (Fig. 2b,c). In contrast, interaction-deficient XLF<sup>K104E</sup>, XRCC4<sup>K104E</sup>, and XRCC4<sup>F111E</sup>–LIG4 did not support SR-complex formation (Fig. 2b,c), whereas XRCC4<sup>K104E</sup>–LIG4 exhibited low activity, in agreement with its residual binding to XLF (Fig. 2c). Thus, a physical interaction between XLF and XRCC4 is required for the close alignment of DNA ends within the SR synaptic complex.

The above results do not distinguish whether XLF–XRCC4 interaction is required for the formation or maintenance of the SR complex. To address this question, we performed long time-course imaging experiments to compare the kinetics of SR-complex formation and dissociation in the presence of WT XRCC4 and the hypomorphic K104E mutant. To prevent ligation of DNA substrates in these experiments, we supplemented extract immunodepleted of XRCC4 with XRCC4 WT or K104E in complex with catalytically inactive LIG4<sup>K278R</sup>. The SR complex formed at a fivefold-slower rate in the presence of XRCC4 K104E compared with XRCC4<sup>WT</sup> (2.7 ± 0.3 s<sup>−1</sup>, s.e.m.), thus suggesting that the XLF–XRCC4 interaction is required for efficient establishment of the SR complex. Once formed, however, the SR complex had a similar lifetime in the presence of WT XRCC4<sup>WT</sup> and hypomorphic XRCC4<sup>K104E</sup> (153 ± 23 s) (Fig. 2d,e), thus suggesting that the stability of the assembled SR complex becomes independent of XRCC4–XLF interaction. In both cases, the lifetime distribution of the SR complex is significantly nonexponential (P<0.001, Lilliefors test<sup>48</sup>), thus indicating that SR-complex dissociation is not governed by a single rate constant.

The SR synaptic complex contains a single XLF dimer. A model in which end synopsis involves the formation of XLF–XRCC4

---

**Fig. 2** | XLF–XRCC4 interaction is required for SR-complex formation. a. Schematic of intramolecular smFRET reporter for monitoring SR-complex formation<sup>36</sup>. DNA is tethered internally via a biotin–streptavidin interaction and labeled 7 bp from each end with Cy3 (donor) and Cy5 (acceptor) fluorophores. No FRET is detected when DNA ends are untethered or in the LR complex. Close alignment of DNA ends within the SR complex is indicated by energy transfer between Cy3 and Cy5. b,c. Kinetics of SR-complex formation in extract depleted of XLF or XRCC4 and supplemented with purified recombinant protein, as in Fig. 1c,d. The mean fraction of FRET-positive (SR complex or ligated) substrates is plotted as a function of time after extract addition. Bars indicate the minimum and maximum values obtained from multiple experimental replicates (n values in Supplementary Table 2). Source data for b and c are available with the paper online. d. Sample smFRET trajectories showing SR-complex formation and dissociation in the presence of WT XRCC4 and the hypomorphic K104E mutant. To prevent ligation of DNA ends, catalytically inactive LIG4<sup>K278R</sup> was used here and in e to prevent ligation of DNA ends. e. Cumulative distribution functions of SR-complex lifetimes. The difference between the two conditions was not statistically significant (P=0.45, log-rank test; sample sizes in Supplementary Table 3). The data shown in b and c were collected with a 0.5-s exposure time, alternating between two frames of Cy3 excitation and one frame of Cy5 excitation. The data shown in d and e were collected with a 1-s exposure with a 1-s delay between exposures, alternating between four frames of Cy3 excitation and one frame of Cy5 excitation.

---

**NATURE STRUCTURAL & MOLECULAR BIOLOGY**
Filaments predicts binding of numerous XLF dimers to each pair of DNA ends. To determine the stoichiometry of XLF within the synaptic complex, we used three-color imaging of Cy3/Cy5-labeled DNA ends and Alexa Fluor 488–labeled, Halo-tagged XLF (AF488-XLF). AF488-XLF rescued end joining in XLF-immunodepleted extracts with Cy5 (Cy5-XLF; Fig. 3c, red curve, and Supplementary Fig. 3a–c). AF488-XLF rescued end joining in XLF-immunodepleted extracts without AF488-XLF protein (Fig. 3c, black curve), thus ruling out a photophysical artifact. Similar results were obtained when Halo-XLF was labeled with Cy5 (Cy5-XLF; Fig. 3c, red curve, and Supplementary Fig. 3a–c). Remarkably, despite the ability of Halo-XLF to support NHEJ with similar efficiency as untagged XLF (Supplementary Fig. 2a), a mixture of Halo-XLF and XRCC4 did not bridge DNA in bulk pulldown experiments (Supplementary Fig. 1h), thus indicating that the

The number of AF488 fluorophores present within 10s of SR-complex formation, determined by comparison with the internal single-AF488 standard, was mostly between 0 and 2 (Fig. 3b, red curve), in agreement with binding of a single incompletely labeled dimer of XLF. To test this interpretation, we fit the distribution of fluorophore number in the AF488 frame just before high FRET (Methods). The decay of the black curve below zero results from photobleaching of the Cy5 label on DNA. 

At any instant, most DNA substrates exhibited no AF488 signal (Fig. 3b, black curve). Given that AF488-XLF monomer was 68% labeled, on the basis of UV–visible absorbance measurements (Methods), this result implies low average occupancy of XLF on DNA substrates, and it is inconsistent with widespread formation of XLF–XRCC4 filaments under these conditions. We next focused on substrates that undergo a transition to the SR complex, as indicated by the appearance of FRET between Cy3 and Cy5. Strikingly, SR-complex formation was typically preceded by step-like AF488-XLF binding events (Fig. 3a and Supplementary Fig. 2f).
DNA-bridging activity detected by this pulldown assay is not essential for NHEJ. To ensure that the observed low stoichiometry of XLF in single-molecule experiments was not due to the inability of Halo-XLF to form DNA-bridging filaments with XRCC4, we labeled untagged XLF on endogenous cysteines with Cy5-maleimide. Cy5-maleimide-labeled XLF (Cy5-mal-XLF) rescued end joining in XLF-depleted extract (Supplementary Fig. 3d), and, unlike Halo-XLF, it assembled DNA-bridging complexes with XRCC4 in bulk pulldown experiments (Supplementary Fig. 3e). Nonetheless, Cy5-mal-XLF still bound with an average stoichiometry consistent with one dimer per DNA substrate at the time of SR-complex formation (Fig. 3d). These results suggest that SR-complex formation involves binding of a single XLF dimer.

**End synopsis requires interaction of both XLF head domains with XRCC4.** Although our data indicate that extensive XLF–XRCC4 filaments are not required for SR-complex formation, the possibility remained that both subunits of the XLF dimer might be required to interact with XRCC4. To address this possibility, we generated an XLF tandem dimer construct (tdXLF; Fig. 4a) in which the number of XRCC4 interaction sites could be precisely defined. tdXLF had the same molecular weight as the WT XLF dimer, as measured by size-exclusion chromatography with multiangle light scattering (Supplementary Fig. 4a,b). End joining in XLF-immunodepleted egg extract was rescued by tdXLF, thus demonstrating that this construct functioned in NHEJ (Fig. 4b,c).

To test whether interaction with XRCC4 on both sides of the XLF dimer is required for end joining, we supplemented XLF-depleted extract with tdXLFWT/WT or tdXLFWT/L68D L117D, in which the second XLF sequence in the tandem dimer was mutated to prevent interaction with XRCC4 (Fig. 4a). Whereas tdXLFWT/L68D L117D partially rescued end joining, its activity was impaired ~8- to 16-fold relative to tdXLFWT/WT—an impairment greater than would be expected from a 50% decrease in the effective concentration of WT XLF head domains (Fig. 4b,c). Equivalent results were obtained for the tdXLFWT/L117D single mutant (data not shown). Importantly, tdXLFWT/L68D L117D exhibited a 15-fold decrease in the SR-complex formation rate (1.9 ± 0.51 × 10^{-4}s^{-1}) relative to tdXLFWT/WT (2.8 ± 0.2 × 10^{-3}s^{-1}) in intramolecular smFRET circularization assays (Fig. 4d). However, a low level of SR-complex formation and end joining still occurred (Fig. 4b–d). Given the presence of two interaction-blocking mutations, this residual activity of tdXLFWT/L68D L117D is unlikely to reflect binding of the mutant head domain to XRCC4. Thus, binding of XRCC4 to both head domains of XLF is required to form the SR complex with normal efficiency, whereas binding of a single XLF head domain to XRCC4 permits SR-complex formation at a substantially slower rate.

Fig. 4 | End joining in the presence of a synthetic tandem dimer of XLF. **a,** Schematic of the XLF tandem dimer (tdXLF) construct. Two XLF coding sequences are concatenated by a flexible 15-aa linker. Asterisks denote point mutations. **b,** Dose dependence of end joining as a function of tdXLF concentration for WT/WT and WT/L68D L117D constructs. Black wedges represent a twofold serial dilution series from 910 nM to 0.44 nM. Control reactions not supplemented with tdXLF are labeled ‘–’. Lin, linear DNA substrate; oc, open-circular product; mult, dimeric and multimeric products. Uncropped image in Supplementary Dataset 1. **c,** Quantification of product formation in the gel in **b.** Fraction joined was quantified as background-subtracted intensity of product bands divided by total background-subtracted intensity of substrate and product bands. Source data for **c** are available online. **d,** Histogram of FRET efficiency from tdXLF rescue of smFRET intramolecular circularization experiments, accumulated over 20 min.

Fig. 4 | a, Schematic of the XLF tandem dimer (tdXLF) construct. Two XLF coding sequences are concatenated by a flexible 15-aa linker. Asterisks denote point mutations. b, Dose dependence of end joining as a function of tdXLF concentration for WT/WT and WT/L68D L117D constructs. Black wedges represent a twofold serial dilution series from 910 nM to 0.44 nM. Control reactions not supplemented with tdXLF are labeled ‘–’. Lin, linear DNA substrate; oc, open-circular product; mult, dimeric and multimeric products. Uncropped image in Supplementary Dataset 1. c, Quantification of product formation in the gel in b. Fraction joined was quantified as background-subtracted intensity of product bands divided by total background-subtracted intensity of substrate and product bands. Source data for c are available online. d, Histogram of FRET efficiency from tdXLF rescue of smFRET intramolecular circularization experiments, accumulated over 20 min.
As an independent means of examining the requirement for XRCC4–XLF interaction through both XLF head domains, we generated synthetic XLF heterodimers by coexpression of XLF monomers with different affinity tags, then performed tandem affinity purification (Methods). The exchange of monomers between purified XLF dimers was negligible (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d). In agreement with results for tdXLF, XLF heterodimers with one WT subunit and one L117D mutant subunit were severely deficient in end joining, as compared with WT–WT heterodimers (Supplementary Fig. 4c,d).
One possibility is that DNA-binding motifs from XLF, XRCC4, and LIG4 directly contact DNA ends and bring them together (Fig. 5d, i). Alternatively, assembly of the XLF–XRCC4–LIG4 subcomplex within the DNA-PK holoenzyme may induce a conformational transition of DNA-PK that aligns DNA ends (Fig. 5d, ii). In support of the latter model, the SR complex remains stable in the presence of XRCC4K104E, which dissociates much more rapidly from XLF than does WT XRCC4, thus suggesting that XLF–XRCC4 interactions are required only for the initial formation of the SR complex but not for its maintenance (Fig. 2e). Given that the lifetime of the SR complex is nonexponentially distributed, characterizing the multiple kinetic steps involved in SR-complex dissociation and determining which factors influence these steps should be interesting.

There are several conceptual problems associated with the XLF–XRCC4 filament model. First, end-bridging filaments of XLF–XRCC4 should block LIG4 and other end-processing enzymes from gaining access to the DNA ends46. Second, how the inner channel of these filaments could accommodate the large size of DNA-PKcs is unclear. Third, some mechanism would be needed to promote stable assembly of filaments near DNA ends while preventing their assembly on the much larger quantity of unbroken DNA in the cell. Fourth, phosphorylation of XLF and XRCC4 by DNA-PK inhibits filament formation, an effect that would be expected to suppress rather than promote filament assembly near DNA ends46,47. In contrast, none of these difficulties apply to a model in which XLF and XRCC4 form a short oligomer within a stoichiometrically defined synaptic complex.

Our experiments cannot rule out the possibility that XLF and XRCC4 might form filaments in cells. For instance, other protein–protein interactions might contribute XLF and XRCC4 within chromosomal DNA-damage-response foci in a manner that is not recapitulated on the relatively short DNA substrates in our assay. However, our observation of efficient NHEJ in the absence of XLF–XRCC4 filaments shows that these structures are not required and suggests that a careful reappraisal of the filament model is needed. In vitro DNA-bridging experiments with purified XLF and XRCC4 have generally used buffers with low ionic strength (25–75 mM KCl). In contrast, we found that DNA pulldown by XLF–XRCC4 mixtures, as well as formation of large complexes measured by dynamic light scattering, was abolished at a more physiological salt concentration of 150 mM (Supplementary Fig. 1h,i). A more direct line of evidence supporting XLF–XRCC4 filament formation in cells has come from super-resolution imaging of fixed cells stained with antibodies against XLF and XRCC4 (ref. 10). XLF and XRCC4 appeared punctate in these experiments, and a subset of puncta exhibited an elongated morphology suggestive of filaments. However, further characterization of these structures is needed, because colocalization of points in super-resolution imaging does not prove the existence of physically continuous filaments. Moreover, whether XLF–XRCC4 interaction is required for focus formation remains to be tested (e.g., by analysis of mutants).

Previously, human XLFL115A was shown to rescue the DNA-damage sensitivity of most XLF-deficient mammalian cell lines despite being unable to bridge DNA in vitro, thus calling into question whether XLF–XRCC4 filaments are generally required for NHEJ14,47. Likewise, we observed that X. laevis XLFL117A (corresponding to hXLFK115A) and N-terminally Halo-tagged XLF supported NHEJ in egg extract despite exhibiting no DNA-bridging activity in a pulldown assay (Figs. 3a–c and 5a,b and Supplementary Figs. 1h and 2a). We found that, contrary to previous qualitative observations14, hXLFK115A is not deficient in hXRCC4 binding (Fig. 5c and Supplementary Table 1). These results support the idea that, as in our cell-free extract, end joining in cells requires XLF–XRCC4 interaction but not DNA bridging by XLF–XRCC4 filaments.

In summary, a single XLF dimer, interacting through both head domains with XRCC4, promotes the close alignment of DNA ends. This result suggests that rather than being tethered by protein filaments, DNA ends are precisely aligned within a stoichiometrically defined protein complex. Reconstitution and structural characterization of this complex should be an important direction for future work.

Online content
Any methods, additional references, Nature Research Reporting Summary, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41594-018-0120-y.

Received: 7 March 2018; Accepted: 5 July 2018; Published online: 3 September 2018

References
1. Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412, 607–614 (2001).
2. Dobbs, T. A., Tainer, J. A. & Lees-Miller, S. P. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. DNA Repair (Amst.) 9, 1307–1314 (2010).
3. Jette, N. & Lees-Miller, S. P. The DNA-dependent protein kinase: a multifunctional protein kinase with roles in DNA double strand break repair and mitosis. Prog. Biophys. Mol. Biol. 117, 194–205 (2015).
4. Jiang, W. et al. Differential phosphorylation of DNA-PKcs regulates the interplay between end-processing and end-ligation during nonhomologous end-joining. Mol. Cell. 58, 172–185 (2015).
5. Critchlow, S. E., Bowater, R. P. & Jackson, S. P. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligation IV. Curr. Biol. 7, 588–598 (1997).
6. Grawunder, U. et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature 388, 492–495 (1997).
7. Buck, D. et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. Cell 124, 287–299 (2006).
8. Ahnesorg, P., Smith, P. & Jackson, S. P. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 124, 301–313 (2006).
9. Junop, M. S. et al. Crystal structure of the Xrcc4 DNA repair protein and implications for end joining. EMBO J. 19, 5962–5970 (2000).
10. Li, Y. et al. Crystal structure of human XLF/Cernunnos reveals unexpected differences from XRCC4 with implications for NHEJ. EMBO J. 27, 290–300 (2008).
11. Andres, S. N., Modesti, M., Tsai, C. J., Chu, G. & Junop, M. S. Crystal structure of human XLF: a twist in nonhomologous DNA end-joining. Mol. Cell. 28, 1093–1101 (2007).
12. Hammel, M., Yu, Y., Fang, S., Lees-Miller, S. P. & Tainer, J. A. XLF regulates filament architecture of the XRCC4-ligase IV complex. Structure 18, 1431–1442 (2010).
13. Ropars, V. et al. Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. Proc. Natl Acad. Sci. USA 108, 12663–12668 (2011).
14. Hammel, M. et al. XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended groove scaffold for DNA ligation and double strand break repair. J. Biol. Chem. 286, 32638–32650 (2011).
15. Malwert, L. et al. Delination of the Xcc4-interacting region in the globular head domain of cernunnos/XLF J. Biol. Chem. 285, 26475–26483 (2010).
16. Roy, S. et al. SRCC4 protein interaction with XLF is required for coding (but not signal) end joining. Nucleic Acids Res. 40, 1684–1694 (2012).
17. Roy, S. et al. XRC4/XLF interaction is variably required for DNA repair, and is not required for Ligase IV stimulation. Mol. Cell. Biol. 35, 3017–3028 (2015).
18. Lu, H., Pannicke, U., Schwarz, K. & Lieber, M. R. Length-dependent binding of human XLF to DNA and stimulation of XRCC4-DNA ligase IV activity. J. Biol. Chem. 282, 11155–11162 (2007).
19. Guo, J., Lu, H., Tsai, A. G., Schwarz, K. & Lieber, M. R. Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence. Nucleic Acids Res. 35, 5755–5762 (2007).
20. Tsai, C., Kim, S. A. & Chu, G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. Proc. Natl Acad. Sci. USA 104, 7851–7856 (2007).
21. Tadi, S. K. et al. PAXX is an accessory c-NHEJ factor that associates with Ku70 and has overlapping functions with XLF. Cell Rep. 17, 541–555 (2016).
ARTICLES

22. Riballo, E. et al. XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. *Nucleic Acids Res.* **37**, 482–492 (2009).

23. Mahaney, B. L., Hammel, M., Meek, K., Tainer, J. A. & Lees-Miller, S. P. XRCC4 and XLF form long helical protein filaments suitable for DNA end protection and alignment to facilitate DNA double strand break repair. *Biochem. Cell Biol.* **91**, 31–41 (2013).

24. Wu, Q. et al. Non-homologous end-joining partners in a helical dance: structural studies of XLF-XRCC4 interactions. *Biochem. Soc. Trans.* **39**, 1387–1392 (2011).

25. Andres, S. N. et al. A human XRCC4-XLF complex bridges DNA. *Nucleic Acids Res.* **40**, 1868–1878 (2012).

26. Brouwer, I. et al. Sliding sleeves of XRCC4-XLF bridge DNA and connect fragments of broken DNA. *Nature* **535**, 566–569 (2016).

27. Reid, D. A. et al. Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proc. Natl Acad. Sci. USA* **112**, E2575–E2584 (2015).

28. Di Virgilio, M. & Gautier, J. Repair of double-strand breaks by nonhomologous end joining in the absence of Mre11. *J. Cell Biol.* **171**, 765–771 (2005).

29. Postow, L. et al. Ku80 removal from DNA through double strand break-induced ubiquitination. *J. Cell Biol.* **182**, 467–479 (2008).

30. Labhart, P. Nonhomologous DNA end joining in cell-free systems. *Eur. J. Biochem.* **265**, 849–861 (1999).

31. Labhart, P. Ku-dependent nonhomologous DNA end joining in *Xenopus* egg extracts. *Mol. Cell. Biol.* **19**, 2585–2593 (1999).

32. Sandoval, A. & Labhart, P. Joining of DNA ends bearing non-matching 3′-overhangs. *DNA Repair (Amst.)* **1**, 397–410 (2002).

33. Taylor, E. M. et al. The Mre11/Rad50/Nbs1 complex functions in nonhomologous end joining complex. *J. Biol. Chem.* **285**, 397–410 (2010).

34. Zhu, S. & Peng, A. Non-homologous end joining repair in *Xenopus* egg extract. *Sci. Rep.* **6**, 27797 (2016).

35. Thode, S., Schäfer, A., Pfeiffer, P. & Vielmetter, W. A novel pathway of DNA end-to-end joining. *Cell 60*, 921–928 (1990).

36. Pfeiffer, P. & Vielmetter, W. Joining of nonhomologous DNA double strand breaks in vitro. *Nucleic Acids Res.* **16**, 907–924 (1988).

37. Graham, T. G. W., Walter, J. C. & Loparo, J. J. Ensemble and single-molecule analysis of non-homologous end joining in frog egg extracts. *Methods Enzymol.* **591**, 233–270 (2017).

38. Graham, T. G. W., Walter, J. C. & Loparo, J. J. Two-stage synopsis of DNA ends during non-homologous end joining. *Mol. Cell* **61**, 850–858 (2016).

39. DeFazio, L. G., Stansel, R. M., Griffith, J. D. & Chu, G. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J.* **21**, 3192–3200 (2002).

40. Hammel, M. et al. Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J. Biol. Chem.* **285**, 1414–1423 (2010).

41. Cary, R. B. et al. DNA looping by Ku and the DNA-dependent protein kinase. *Proc. Natl Acad. Sci. USA* **94**, 4267–4272 (1997).

42. Spagnolo, L., Rivera-Calzada, A., Pearl, L. H. & Llorca, O. Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. *Mol. Cell** 22**, 511–519 (2006).

43. Cottaret, J. et al. A non-catalytic function of the ligation complex during nonhomologous end joining. *J. Cell Biol.* **200**, 173–186 (2013).

44. Wang, J. L. et al. Dissection of DNA double-strand-break repair using novel single-molecule forceps. *Nat. Struct. Mol. Biol.* **25**, 482–487 (2018).

45. Normanno, D. et al. Mutational phospho-mimicry reveals a regulatory role for the XRCC4 and XLF C-terminal tails in modulating DNA bridging during classical non-homologous end joining. *Life* **6**, e22900 (2017).

46. Lilellors, H. W. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *J. Am. Stat. Assoc.* **64**, 387 (1969).

47. Fattah, E. J. et al. A role for XLF in DNA repair and recombination in human somatic cells. *DNA Repair (Amst.)* **15**, 39–53 (2014).

48. Ochi, T. et al. Structural insights into the role of domain flexibility in human DNA ligase IV. *Structure* **20**, 1212–1222 (2012).

49. Menon, V. & Povirk, L. F. XLF/Cernunnos: an important but puzzling participant in the nonhomologous end joining DNA repair pathway: *DNA Repair (Amst.)* **58**, 29–37 (2017).

Acknowledgements

We thank E. van Arnam (Harvard Medical School) for assistance with purification of Cy5 Halo ligand, B. Sinson (Harvard Medical School) for providing improved FRET analysis code and generating the pETDuet vector containing XRCC4 and LIG4, K. Schmitz (Massachusetts Institute of Technology) for advice on generating an ‘orthogonal’ codon-optimized XLF sequence, K. Arnett (HMS Center for Macromolecular Interactions) for assistance with dynamic light scattering. We would also like to thank members of the laboratories of J.J.L. and J.C.W. for helpful discussions. This work was funded by a National Institutes of Health grant R01GM115487 (to J.J.L.) and the Howard Hughes Medical Institute (J.C.W.).

Author contributions

All authors designed experiments and wrote the manuscript. T.G.W.G. and S.M.C. performed experiments and data analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41594-018-0120-y.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.J.L.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
**Methods**

**Plasmid construction.** XLF and XRCC4–LIG4 expression constructs. WT *X. laevis* XLF, XRCC4, and LIG4 plasmids were as previously described. A plasmid encoding human XRCC4 was obtained from Addgene (#13331) and was used to contain a single-nucleotide substitution resulting in the absence of the N-terminus of the carboxy acid chain I134T. This was restored to the reference sequence by site-directed mutagenesis, and the XRCC4 coding sequence was subcloned into a H10-SUMO expression vector. For human XLF, a bacterial codon-optimized gene block (Integrated DNA Technologies) was cloned into a H10-SUMO expression vector. Point mutants, epitope-tagged proteins, and C-terminally truncated XLF, were generated by standard round the horn site-directed mutagenesis. An H10-SUMO-Halo-XLF expression plasmid was created by isothermal assembly of a Haloencoding fragment and a PCR product of the original H10-SUMO-XLF plasmid. A dual vector expression for XRCC4 and H6-SUMO-LIG4 was constructed by amplifying XRCC4, H6-SUMO, and LIG4 from previously described plasmids, and inserting the fragments in a pETDuet vector by isothermal assembly. All constructs were verified by Sanger sequencing.

**Tandem XLF dimers.** A custom Python script was used to generate a second codon-optimized XLF sequence (XLF*) as dissimilar as possible from our existing codon-optimized XLF sequence, to facilitate cloning and to prevent recombination between the two XLF sequences. Additional sequence encoding the zinc finger GGGGGGSSGGGGGGGG was appended at the 3′ end of the XLF sequence, and this combined fragment was ordered as a block from Integrated DNA Technologies, inserted into a cloning vector, and sequence verified. The XLF tandem dimer was then inserted into the isothermal assembly immediately upstream of the initial XLF methionine in H10-SUMO-XLF, thus yielding the tandem dimer H10-SUMO-XLF–‘linker’–XLF.

**XLF heterodimers.** A plasmid for expression of H10-tagged XLF (pTG355) was created by using round-the-horn mutagenesis to delete the SUMO tag in H10-SUMO-XLF (pTG296). The alternative codon-optimized XLF sequence described above was inserted together with a T7 promoter and an N-terminal Flag-Avitag into the H10-SUMO-XLF plasmid, thereby generating a vector for dual expression of two versions of XLF with different affinity tags.

**Protein purification.** XLF. Plasmids encoding H10-SUMO-tagged WT XLF and XLF point mutants were transformed into *Escherichia coli* BL21(DE3)pLysS cells. Cultures were grown in LB medium at 37 °C to an *OD* of ~0.6, IPTG was added to a final concentration of 1 mM to induce expression, and cultures were grown an additional 3 h at 37 °C. Cells were collected by centrifugation, and cell pellets were resuspended in 1 ml of lysis buffer and stored at –80 °C. H10-SUMO-tagged proteins were purified essentially as described previously. Briefly, cells were lysed by sonication in His-SUMO lysis buffer (20 mM Tris-HCl, pH 8, 1 M NaCl, 30 mM imidazole, 5 mM BME, and 1 mM PMSE), and lysates were clarified by centrifugation for 1 h at 25,000 r.p.m. in an SW-41 rotor at 4 °C. Clarified lysates were incubated with Ni–NTA–agarose (Qiagen) for 1 h at 4 °C to bind histidine-tagged proteins. The resin was washed with 25 CV of the same buffer and bound to a HiTrap SP HP column equilibrated with the His-SUMO dialysis buffer (20 mM Tris-HCl, pH 8, 350 mM NaCl, 10% glycerol, and 5 mM BME). Human XLF1–224 was purified with the same protocol.

**Catalytic domain proteins.** Halo-XLF was expressed as a H10-SUMO fusion and subjected to His-SUMO purification steps used for XLF (described above). Biotinylated protein was eluted with 20 CV of the same buffer containing 300 mM imidazole. Next, eluates were passed by gravity flow over SoftLink Avidin resin to captureAvitagged Halo-XLF–Flag-Avitag XLF heterodimers. The resin was washed with 25 CV of 20 mM Tris-HCl, pH 8, 1 M NaCl, and 5 mM BME, followed by 20 CV of 20 mM Tris-HCl, pH 8, 350 mM NaCl, and 5 mM BME. We suspect that the use of high-salt buffers at this stage is important to remove nonspecifically bound H10-XLF–H10-XLF homodimer, because early WT–mutant heterodimer preparations purified under less stringent conditions exhibited variable and elevated activity (data not shown). Bound protein was eluted in 1–CV increments with 20 mM Tris-HCl, pH 8, 350 mM NaCl, 5 mM BME, and 5 mM biotin, with a pause of at least 5 min between elution steps to facilitate dissociation of bound protein. Protein was flash-frozen and stored at –80 °C in elution buffer.

**XRCC4.** N-terminal H10-SUMO fusion constructs of *X. laevis* XRCC4 were transformed into BL21(DE3)pLysS *E. coli* cells. The cells were grown at 37 °C to an *OD* of 0.6 and then induced with 1 mM IPTG for 3 h at 37 °C. Cells were pelleted by centrifugation, flash frozen, and stored at –80°C. Cell pellets were then thawed and subjected to the His-SUMO purification procedure previously.

**Biotinylated XRCC4–Avitag.** *X. laevis* XRCC4 with an N-terminal H10-SUMO tag and a C-terminal Flag-Avitag was expressed in BL21(DE3) E. coli cells cotransformed with the BirA biotin ligase expression plasmid pBirAc and a dual-expression plasmid encoding H10-XLF as well as Flag-Avatig XRCC4–Avitag (pTG448) or Flag-Avitag XRCC4–L117D (pTG449). After the cells had grown to an *OD* of 0.4, expression was induced by addition of IPTG to 1 mM, and the growth medium was supplemented with biotin to a final concentration of 25 μM. Cells were collected by centrifugation after 3 h of induction at 25–30°C, flash-frozen, and stored at –80°C. Frozen cell pellets were resuspended in 20 mM Tris-HCl, pH 8, 1 M NaCl, 30 mM imidazole, and 5 mM BME and lysed by three 30-s cycles of sonication (1 on, 1 off). Lysates were clarified by centrifugation for 1 h at 25,000 r.p.m. in an SW-41 rotor. Supernatants were incubated for 1 h at 4°C with 250 μl of Ni–NTA agarose (Qiagen) per liter of culture. Resin was washed with 40 CV of lysis buffer, and H10-tagged complexes were eluted with 20 CV of the same buffer containing 300 mM imidazole. Next, eluates were passed by gravity flow over SoftLink Avidin resin to capture Avitagged XRCC4–Avitag XLF heterodimers.

**XRCC4 N-terminal H10-SUMO fusion constructs.** A plasmid encoding the catalytic fragment of XRCC4 was constructed as described above. The diluted sample was passed over a Mono Q column equilibrated with 20 mM Tris-HCl, pH 8, 100 mM NaCl, 10% glycerol, and 5 mM BME. The flow through, containing the protein of interest, was concentrated with a Amicon 10-kDa MWCO centrifugal concentrator and dialyzed overnight against 20 mM Tris-HCl, pH 8, 350 mM NaCl, 10% glycerol, and 5 mM BME. Human XRCC4 was purified through the same protocol. However, it bound the MonoQ column and was eluted with a NaCl gradient of 100–1,000 mM.

**XL fertility.** XL fertility constructs were expressed as fusions to H10-SUMO in BL21(DE3) pLysS cells and subjected to the His-SUMO purification described above for WT XL fertility. Flow through from the second Ni–NTA step was eluted with 12.36 volumes of buffer A (500 mM Na-2-ME, pH 6.5, 10% glycerol, and 5 mM BME) and bound to a Hitrap SP column equilibrated with the same buffer containing 150 mM NaCl. The protein was eluted with a gradient of 150 mM to 1 M NaCl over 20 column volumes (CV). Peak fractions were pooled, concentrated on a 10-kDa MWCO Amicon centrifugal concentrator, and subjected to size-exclusion purification on a Superdex 200 Increase 10/300 column equilibrated with 50 mM Na-2-ME, pH 6.5, 350 mM NaCl, 10% glycerol, and 5 mM BME. Peak fractions were pooled and concentrated as described above.
fractions with lysis buffer containing 250 mM imidazole. Fractions containing the protein of interest were pooled, supplemented with Ulp1 protease, and dialyzed into cleavage buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DTT; and 10% glycerol). The dialyzed sample was then incubated with 250 μl of Ni-NTA resin per liter of culture for 1.5 h at 4°C. The flow through was collected, passed through a 0.22 μm filter, and diluted in cleavage buffer without NaCl to bring the final NaCl concentration to 50 mM. Protein was bound to a MonoQ 5/50 GL column (GE Healthcare Life Sciences) equilibrated in buffer A (cleavage buffer with 50 mM NaCl). Protein was eluted with a 20-ml gradient of 0–60% buffer B (cleavage buffer with 1 M NaCl). Peak fractions were pooled and further separated on a S200 16/600 pg sizing column (GE Healthcare Life Sciences) equilibrated with cleavage buffer. Protein-containing fractions were pooled and concentrated with an Amicon 10-kDa MWCO centrifugal concentrator.

Biotin-Avitag-Halo. N-terminally SUMO-tagged Avitag-Halo protein was expressed in BL21(DE3) cells containing the pBirAcm plasmid, which encodes the BirA enzyme that covalently attaches biotin to the N-terminus of proteins. The BL21(DE3) cells were grown in LB medium containing 100 μg/ml of ampicillin and 30 μg/ml of kanamycin to a final OD600 of 0.4, induced with 1 mM IPTG for 3 h at 25–30°C, and collected by centrifugation. Cells were lysed in 20 mM Tris-HCl, pH 7.5, and 150 mM NaCl, without reducing agent. The lysate was clarified by centrifugation at 14,000 × g for 30 min at 4°C. Protein containing fractions were pooled and concentrated with an Amicon 10-kDa MWCO centrifugal concentrator.

Fluorescence labeling of Halo-XLF. A twofold molar excess of Cy5- or Alexa Fluor 488-Halo ligand in DMSO was added to Halo-XLF protein in its storage buffer. The mixture was allowed to react for 10 min while being centrifuged at 16,000 × g. Proteins containing 10% of total protein were applied to a Superdex 200 Increase 10/300 column equilibrated with 20 mM Tris-HCl, pH 8, 350 mM NaCl, 10% glycerol, and 5% β-mercaptoethanol. Peak fractions were pooled and concentrated with an Amicon 10-kDa MWCO centrifugal concentrator. Protein concentrations were measured on the basis of absorbance at 280 nm, assuming an extinction coefficient of 82,390 M⁻¹ cm⁻¹. Dye concentrations were measured according to absorbance at 650 nm for Cy5 (assuming an extinction coefficient of 250,000 M⁻¹ cm⁻¹) or 495 nm for AF488 (assuming an extinction coefficient of 73,000 M⁻¹ cm⁻¹). For AF488-labeled protein, the absorbance at 280 nm due to protein was determined by subtraction of the AF488 absorbance at 280 nm from the total absorbance, with the correction factor (AF488 absorbance at 280 nm = 0.11 × AF488 absorbance at 495 nm) subtracted. To test the dependence of the dye-to-protein ratio of the labeled fraction on the concentration of dye, the molar concentration of dye was determined by division of the molar concentration of dye by the molar concentration of protein. Measurements of the labeled fraction obtained with a NanoDrop spectrophotometer agreed with those calculated from dual-wavelength absorbance measurements of labeled protein peaks during size-exclusion chromatography.

Fluorescence labeling of XLF with Cy5-maleimide. 130 μl of 53 μM WT X. laevis XLF was incubated for 1 h at room temperature with 50 μM sulfo-Cy5-maleimide (Lumiprobe) in 20 mM Tris-HCl, pH 7.5, 350 mM NaCl, and 10% glycerol. The labeled sample was mixed with an AF488-labeled XLF control and incubated with DTT for 1 h at 4°C. Labeled protein was separated from free dye on an S200 Increase 10/300 GL sizing column (GE Healthcare Life Sciences). Protein-containing fractions were collected, and the Cy5/XLF ratio was determined according to the absorbance of each, measured with a NanoDrop spectrophotometer. With an extinction coefficient of 250,000 M⁻¹ cm⁻¹ for Cy5 and an extinction coefficient of 73,000 M⁻¹ cm⁻¹ for AF488, the dye-to-protein ratio was measured to be 0.77. This same procedure was performed in parallel for an AF488-labeled XLF sample and a mock-labeled XLF sample.

Preparation of Xenopus egg extracts. Xenopus egg extract was prepared as previously described[8,22]. The female frogs used to produce oocytes were cared for by the Center for Animal Resources and Comparative Medicine at Harvard Medical School (AAALAC accredited). Work performed for this study was in accordance with the rules and regulations set by AAALAC. The Institutional Animal Care and Use Committee (IACUC) of Harvard Medical School approved the work.

Antibodies and immunodepletion. XLF and XRCC4 antibodies were the same as described previously[23,30]. For rescue experiments in Figs. 1c and 1b, immunodepleted extract was supplemented with 500 nM XLF or 50 nM XRCC4-LIG4 complex.

Bulk end-joining assays. Bulk end-joining assays were performed as described previously[24]. The final concentration of radiolabeled linear substrate DNA in bulk end-joining assays was approximately 1 ng/μl (~0.5 nM). The integrated intensity of gel bands (Fig. 4c and Supplementary Fig. 4f) was quantified in Fiji.

dXLF and XLF heterodimer titration experiments. To test the dependence of the protein concentration on the protein concentration of XLF tandem dimer (dXLF; Fig. 1b–d) and heterodimer (Supplementary Fig. 4e,f) constructs, we depleted extract of endogenous XLF, then supplemented the depleted extract with an ATP-regeneration system and closed-circular carrier DNA, as described previously[24]. Radiolabeled linear substrate DNA was added to extract on ice to a final concentration of ~1 ng/μl, and a portion of this extract mixture was supplemented with XLF at the indicated maximum concentration. The same fractional volume of XLF storage buffer was added to the remaining extract mixture. Extract supplemented with XLF was then serially diluted with extract mixture, while tubes were placed in a 37°C block to permit melting of the TIRF angle. Emission light was set from excitation light with a multipass dichroic mirror, and laser lines were further attenuated with a StopLine 488/532/635 notch filter. A home-built beam splitter (ref. 31) was used to separate Alexa Fluor 488 and Cy3 emission from Cy5 emission; these two channels were imaged on separate halves of an electron-multiplying charge-coupled device camera.
Articles

Nature Structural & Molecular Biology | www.nature.com/rismb

Single-molecule imaging and analysis. Single-molecule imaging and analysis is described in the Supplementary Note.

Biolayer interferometry binding assays. BLI experiments were performed on an Octet RED384 (ForteBio) with streptavidin-coated Dip and Read Biosensors (ForteBio) and 384-well plates. 100 nM of biotinylated C-terminally Avi-tagged XRC4C4 was bound for 5 min in a binding buffer consisting of 20 mM Tris-HCl, pH 7.5, 125 mM KCl, 10 mg/mL bovine serum albumin (Calbiochem), and 5 mM BME. To test for nonspecific binding of XLF<sup>232</sup>, reference tips were incubated in buffer alone. The tips were washed with buffer for 60 s to obtain a baseline reading and then transferred to wells containing XLF<sup>232</sup> or point mutants thereof (2 μM of each protein unless otherwise noted) for 120 s. After measurement of association, the tips were moved to wells containing buffer, and XLF dissociation was measured for 180 s. The response for each experimental condition shown in Fig. 1a is the average of three replicates, and each replicate is the average of three sequential rounds of association and dissociation. The s.d. is represented as shading around each trace. The data shown in Supplementary Fig. 1a are representative results from a single experiment. For the titration experiments shown in Supplementary Fig. 1b,c,f, the background subtracted association dissociation curve for each concentration was fit to a 1:1 model with GraphPad Prism software and the built-in fitting tools. Data generated from fits for which R<sup>2</sup> < 0.98 were not included in subsequent analysis. The resulting K<sub>d</sub>, k<sub>e</sub>, and apparent K<sub>e</sub> values were averaged. This procedure was performed for two replicates for each condition. The reported values are averages from the two replicates, and the reported error represents the minimum and maximum values. The same fitting and analysis procedure was used to generate the curves shown in Fig. 5c and the values in Supplementary Table 1. Those data represent two experimental replicates, with the error representing the minimum and maximum values.

Dynamic light scattering. DLS experiments were performed on a Wyatt/Protein Solutions DynaPro-99-E-50 Dynamic Light Scattering Module with a TC100-830 Temperature Controlled Microsampler. The data shown in Supplementary Fig. 1c,i were collected at 20 °C with an acquisition time of 10 s and a read interval of 1 s. Sample measurement began ~1 min after assembly and mixing of the specified proteins in a filtered buffer containing 20 mM HEPES, pH 8, 75 mM KCl, 0.5 mM EDTA, 1 mM DTT, and 5% glycerol. Dynamics V6 software (Wyatt/Protein Solutions) was used to calculate hydrodynamic radii with the built-in parameters for a 5% glycerol solution. For each condition, three replicates were averaged to generate the data shown in Supplementary Fig. 1c.i. The shaded region around each trace represents the s.e.m.

Bulk protein–DNA pulldown assays. DNA pulldown assays (Supplementary Figs. 1d,h and 3e) were performed as previously described<sup>14</sup>. A biotinylated 1,000-bp DNA fragment was generated by PCR (template pTG064; primers pTG414 and oTG042F) with Pfu polymerase. The 500-bp DNA was generated by PvuII digestion of pTG024. These DNA substrates were then gel-purified by AdvanceBio 300 column was preequilibrated overnight with running buffer (50 mM Na-MES, pH 6.5, 350 mM NaCl, 10% glycerol, and 5 mM BME) at a 0.2 ml/min flow rate and then allowed to equilibrate to a flow rate of 0.4 ml/min for 1 h. 30 μl of a 50 μM solution of tdXLFWT was separated on the column. As a control, 100 μl of 2mg/ml BSA fraction V (Omnipur) in the same buffer without BME was separated on the column, thus yielding monomer, dimer, and trimer peaks of the expected sizes.

Statistics and reproducibility. A two-tailed unpaired t test with unequal variance was performed to determine whether the melting temperatures of the mutants were significantly different from those of the WT proteins.

Size-exclusion chromatography with multilangle light scattering. SEC–MALS was performed on an Agilent 1260 Infinity Isocratic Liquid Chromatography System equipped with a Wyatt Dawn Helesos II Multi-Angle Light Scattering detector and a Wyatt Optilab T-reak Refractive Index Detector. An Agilent AdvanceBio 300 column was preequilibrated overnight with running buffer (50 mM Na-MES, pH 6.5, 350 mM NaCl, 10% glycerol, and 5 mM BME) at a 0.2 ml/min flow rate and then allowed to equilibrate to a flow rate of 0.4 ml/min for 1 h. 30 μl of a 50 μM solution of tdXLFWT was separated on the column. As a control, 100 μl of 2mg/ml BSA fraction V (Omnipur) in the same buffer without BME was separated on the column, thus yielding monomer, dimer, and trimer peaks of the expected sizes.

Differential scanning fluorimetry. Differential scanning fluorimetry protein thermal-shift assays were performed with a QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). 20-μl reactions containing 10 μg of the specified protein were assembled in MicroAmp FAST optical 96-well reaction plates (Life Technologies) and covered with MicroAmp Optical Adhesive Film (Life Technologies). SYPRO Orange dye was included in the reactions with a Protein Thermal Shift Dye Kit (Applied Biosystems), according to the manufacturer’s instructions. Protein samples were diluted in storage buffer (20 mM Tris-HCl, pH 8, 10 mM imidazole, 350 mM NaCl, and 10% glycerol). The temperature was raised from 25 to 99 °C at 3 °C per minute, and the fluorescent dye was excited and measured at 470 nm and 587 nm, respectively. To determine the melting temperatures, we plotted emission signals as a function of temperature, and these thermal denaturation curves were fit with the Boltzmann equation in Protein Thermal Shift software (Life Technologies). The average of three experimental replicates is shown for each condition. The error bars represent the s.e.m. A two-tailed, unpaired t test with unequal variance was performed to determine whether the melting temperatures of the mutants were significantly different from those of the WT proteins.

Supplementary information. Source data for Figs. 2b,c, 4c, and 5c are available with the paper online. Used to analyze them are available from the corresponding author upon reasonable request. Source data for Figs. 2c, 4e, and 5c are available with the paper online.

References

50. Hemsley, A., Arnhem, N., Toney, M. D., Cortopassi, G. & Galas, D. J. A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Res. 17, 6545–6551 (1989).
51. Graham, T. G. W. et al. ParB spreading requires DNA bridging. Genes Dev. 28, 1228–1238 (2014).
52. Lebofsky, R., Takahashi, T. & Walter, J. C. DNA replication in nucleus-free Xenopus egg extracts. Methods Mol. Biol. 521, 229–252 (2009).
53. Tanner, N. A. et al. Single-molecule studies of fork dynamics in Escherichia coli DNA replication. Nat. Struct. Mol. Biol. 15, 998 (2008).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| Item                                                                 | Confirmed |
|----------------------------------------------------------------------|-----------|
| The exact sample size \( (n) \) for each experimental group/condition, given as a discrete number and unit of measurement | x         |
| An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly | x         |
| The statistical test(s) used AND whether they are one- or two-sided | x         |
| Only common tests should be described solely by name; describe more complex techniques in the Methods section. | x         |
| A description of all covariates tested | x         |
| A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons | x         |
| A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) | x         |
| For null hypothesis testing, the test statistic (e.g. \( F, t, r \)) with confidence intervals, effect sizes, degrees of freedom and \( P \) value noted Give \( P \) values as exact values whenever suitable. | x         |
| For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings | x         |
| For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes | x         |
| Estimates of effect sizes (e.g. Cohen’s \( d \), Pearson’s \( r \)), indicating how they were calculated | x         |
| Clearly defined error bars State explicitly what error bars represent (e.g. SD, SE, CI) | x         |

Software and code

Policy information about availability of computer code

Data collection

The collection of all single-molecule data is described in the Methods section under ‘Single-molecule imaging and analysis’. ForteBio Octet 384 Data Acquisition 9.0 software was used to collect all bio-layer interferometry data. Wyatt/Protein Solutions Dynamics V6 software was used to collect dynamic light scattering data. Life Technologies Protein Thermal Shift software was used to collect differential scanning fluorimetry data.

Data analysis

The software used and analysis of all single-molecule data is described in the Methods section under ‘Single-molecule imaging and analysis’. ForteBio Octet 384 Data Analysis 8.0 software and GraphPad Prism software were used to analyze all bio-layer interferometry data as described in the Methods section under ‘Bio-layer interferometry (BLI) binding assays’. Wyatt/Protein Solutions Dynamics V6 software was used to analyze dynamic light scattering data as described in the Methods section under ‘Dynamic light scattering’. Life Technologies Protein Thermal Shift software was used to analyze differential scanning fluorimetry data as described in the Methods section under ‘Differential scanning fluorimetry’. Results from gel-based assays were quantified using Fiji.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study and the custom-written computer code used to analyze it are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample sizes for all datasets are listed in the figure captions, the associated Methods section, or in Supplementary tables. Generally, samples sizes were chosen to reliably reproduce experimental observations within reasonable margins of error.

Data exclusions

BLI response curves for which fitting to a 1:1 model as described in the Methods section, 'Bio-layer interferometry (BLI) binding assays’ generated R-square values less than 0.98 were not included in subsequent analysis. Poor fitting to this model only occurred at very low concentrations of the protein being titrated where the measurable response was minimal.

For single-molecule data analysis, spots were excluded if they showed evidence of containing more than one DNA substrate. Details are provided in the Methods.

Replication

All data have been successfully reproduced. The number of replicates and associated error are noted in the Methods section, figures, figure captions, main text, and/or Supplementary tables.

Randomization

Randomization was not relevant to this study.

Blinding

Blinding was not relevant to this study.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a  Involved in the study

☐  Unique biological materials
☐  Antibodies
☒  Eukaryotic cell lines
☒  Palaeontology
☐  Animals and other organisms
☒  Human research participants

Methods

n/a  Involved in the study

☒  ChIP-seq
☐  Flow cytometry
☒  MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials

All plasmids listed in Supplementary table 4 are readily available upon reasonable request.
### Antibodies

| Antibodies used     | Antibodies against Xenopus XLF, XRCC4, and IgG were used in this study. |
|---------------------|------------------------------------------------------------------------|
| Validation          | These antibodies were previously validated as described in Ref 38.     |

### Animals and other organisms

Policy information about studies involving animals: ARRIVE guidelines recommended for reporting animal research

| Laboratory animals | Xenopus laevis females older than 2 years were obtained from Nasco for use in this study. |
|--------------------|--------------------------------------------------------------------------------------------|
| Wild animals       | Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals. |
| Field-collected samples | For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field. |