Title

Cryo-EM visualization of DNA-PKcs structural intermediates in NHEJ.

Journal

Science Advances, 9(22)

Authors

Vogt, Alex

Lee, Linda

Naila, Tasmin

et al.

Publication Date

2023-06-02

DOI

10.1126/sciadv.adg2838

STRUCTURAL BIOLOGY

Cryo-EM visualization of DNA-PKcs structural intermediates in NHEJ

Siyu Chen[1,2]*, Alex Vogt[1,2], Linda Lee[3]†, Tasmin Naila[4], Ryan McKeown[1,2], Alan E Tomkinson[4], Susan P Lees-Miller[3], Yuan He[1,2,5,6]*

DNA double-strand breaks (DSBs), one of the most cytotoxic forms of DNA damage, can be repaired by the tightly regulated nonhomologous end joining (NHEJ) machinery (Stinson and Loparo and Zhao *et al.*). Core NHEJ factors form an initial long-range (LR) synaptic complex that transitions into a DNA-PKcs (DNA-dependent protein kinase, catalytic subunit)–free, short-range state to align the DSB ends (Chen *et al.*). Using single-particle cryo–electron microscopy, we have visualized three additional key NHEJ complexes representing different transition states, with DNA-PKcs adopting distinct dimeric conformations within each of them. Upon DNA-PKcs autophosphorylation, the LR complex undergoes a substantial conformational change, with both Ku and DNA-PKcs rotating outward to promote DNA break exposure and DNA-PKcs dissociation. We also captured a dimeric state of catalytically inactive DNA-PKcs, which resembles structures of other PIKK (Phosphatidylinositol 3-kinase-related kinase) family kinases, revealing a model of the full regulatory cycle of DNA-PKcs during NHEJ.

INTRODUCTION

DNA double-strand breaks (DSBs) in nondividing cells are rapidly sensed and repaired by nonhomologous end joining (NHEJ) (*1*, *2*), thereby maintaining genome integrity and preventing cell death (*3*). In NHEJ, the DNA–protein kinase (PK) complex (DNA-PKcs and Ku70/80) assembled on DNA ends (*4*–*6*) recruits additional core NHEJ factors [DNA ligase 4 (Lig4)–x-ray repair cross-complementing protein 4 (XRCC4) and XRCC4-like factor (XLF) (*7*, *8*)] to form the long-range (LR) synaptic complex that holds two physically separated DNA ends together. Next, a transition to a DNA-PKcs–free short-range (SR) complex, in which the DNA ends are aligned, is triggered by in-trans DNA-PKcs autophosphorylation (*9*, *10*). How phosphorylated DNA-PKcs dissociates from the LR complex and facilitates SR complex formation remains unclear. Furthermore, the relationship of another NHEJ scaffolding factor, PAXX (paralog of XRCC4 and XLF) (*11*–*14*), with the dynamic core NHEJ machinery is poorly understood. To shed light on these open questions, we determined the cryo–electron microscopy (cryo-EM) structures of three intermediate states during NHEJ, an apo LR state that contains PAXX (PAXX-LR complex), a rearranged LR state induced by DNA-PKcs autophosphorylation [LR–adenosine 5′-triphosphate (ATP) complex], and a DNA-PKcs dimeric state that occurs in the absence of other core NHEJ factors.

[1]Department of Molecular Biosciences, Northwestern University, Evanston, IL, USA. [2]Interdisciplinary Biological Sciences Program, Northwestern University. Evanston, IL, USA. [3]Department of Biochemistry and Molecular Biology, Robson DNA Science Centre and Arnie Charbonneau Cancer Institute, University of Calgary, Calgary, Alberta T2N 4N1, Canada. [4]Department of Internal Medicine and Molecular Genetics and Microbiology and the University of New Mexico Comprehensive Cancer Center, University of New Mexico, Albuquerque, NM, USA. [5]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL, USA. [6]Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Northwestern University, Chicago, IL, USA.
*Corresponding author. Email: yuanhe@northwestern.edu
†Present address: Department of Molecular Biology, School of Biological Sciences, University of California San Diego, La Jolla, CA, USA.
‡Present address: Alberta Precision Laboratories, Calgary, Alberta, Canada.

RESULTS

The architecture of three NHEJ intermediate complexes

To obtain a stable LR complex during autophosphorylation-dependent activation and investigate the function of PAXX, we extended our previously established complex assembly protocol (*9*) by including PAXX throughout the entire process (see Materials and Methods). In the presence of PAXX, the apo LR complex forms with similar efficiency using the blunt-end and the overhang DNA substrate (*9*), whereas more individual DNA-PK complexes were detected in the absence of PAXX (fig. S1, A to F). In addition, PAXX promotes the formation of dimeric DNA-PKcs even without the XLF-XRCC4-LigIV scaffold, further confirming its ability to stabilize NHEJ complexes (fig. S1, G to I) (*15*).

When the eluted apo PAXX-LR complex assembled on the overhang DNA substrate was incubated with ATP for 15 min, approximately half of the particles adopted a previously unidentified conformation that we named the LR-ATP state (fig. S2, A to C). In similar reactions with blunt-end DNA substrates, all the apo PAXX-LR complexes transitioned into the LR-ATP state, suggesting that the homologous sequence in the middle of the overhang DNA design impedes and slows down the conformational transition (fig. S2, D to F). Unexpectedly, PAXX is still able to tether two copies of DNA-PK in a manner similar to that observed in the converted LR-ATP state without the XLF-XRCC4-LigIV scaffold and in the apo LR state (figs. S1, G to I, and S2, G to I). To test whether ATP hydrolysis is required to activate the transition into the LR-ATP state, ATP was replaced by slowly hydrolyzable adenosine 5′-*O*-(3-thiotriphosphate) (ATP-γ-S). A mixed population of PAXX-LR and LR-ATP states was in the presence of ATP-γ-S, resulting in the majority of class averages showing features resembling the apo PAXX-LR state (fig. S2, J to L).

We were able to obtain a high-resolution cryo-EM reconstruction of the apo PAXX-LR complex at an overall resolution of 4.4 and 3.63 Å for individual DNA-PK-LigIV regions (Fig. 1A and figs. S2, M and N, and S3), allowing us to visualize and model both an ATP ligand in the DNA-PKcs kinase active center and a previously unidentified Ku70 binding peptide, corresponding to

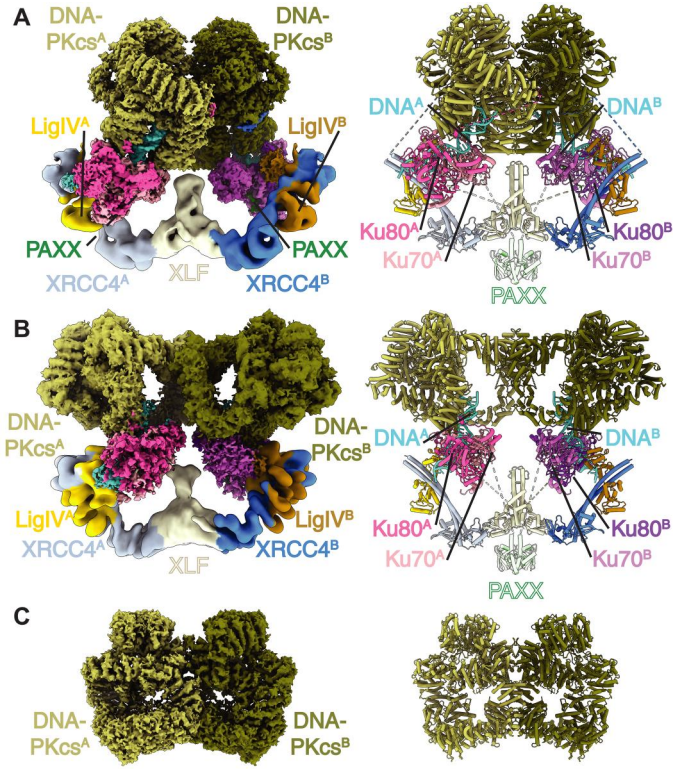

Fig. 1. Cryo-EM structures of three distinct NHEJ complexes that contain DNA-PKcs. (**A**) Cryo-EM composite map (left) (see Materials and Methods) and structural model (right) of the LR complex stabilized by PAXX. Subunits are colored here and in all subsequent figures, as in (A). (**B**) Cryo-EM composite map (left) and structural model (right) of the LR-ATP complex. (**C**) Cryo-EM composite map (left) and structural model (right) of a dimeric DNA-PKcs complex. The dimerization domain of PAXX [Protein Data Bank (PDB): 3WTF] (*12*) is not resolved in the reconstructions. It is placed in the middle of the complexes and colored as transparent to depict the complex architecture in (A) and (B) (right). For clarification, all subunits that are located on the left are named copy A, and subunits on the right are named copy B.

the conserved C terminus of PAXX Ku-binding motif (KBM). This apo PAXX-LR complex has a very similar overall architecture to the previously reported LR complex (*9*, *16*), with a root mean square deviation value of 1.745 Å between the two models.

The LR-ATP complex reconstructions appeared indistinguishable while using the overhang versus the blunt-end DNA substrates (fig. S2, M to P). We thus combined the two datasets and improved the resolution of this intermediate structure to a 9-Å overall resolution (fig. S4), with DNA-PKcs reaching 5.0-Å resolution and the DNA-PK complex reaching 5.9-Å resolution (Fig. 1B). The identities of both the apo PAXX-LR and LR-ATP complexes were verified by resolving the DNA-bound proteins on an SDS–polyacrylamide gel stained by silver nitrate (fig. S2Q). A clear and reproducible band shift of XRCC4 upon ATP addition was observed, consistent with its phosphorylation by DNA-PKcs during NHEJ (*17*).

When negative staining was initially used to optimize the assembly of LR complex, we obtained an alternative state of dimeric DNA-PKcs from a subset of the data (see Materials and Methods). We further optimized the protocol to enrich for dimeric DNA-PKcs using the blunt-end DNA substrate (Fig. 1C and figs. S5 and S6). The two copies of DNA-PKcs engage in a way that is different from any previously published NHEJ complexes containing DNA-PKcs (*9*, *16*, *18*). While it is expected that two activated DNA-PKcs dissociate from the LR complex, we were not able to obtain stably assembled SR complexes or dimeric DNA-PKcs complexes with prolonged ATP incubation, suggesting that additional regulatory mechanisms are involved in the transition from the LR to the SR state. Multiple dimerization states of DNA-PKcs presumably enable the repair complex to signal the successful detection of DSBs and the recruitment of additional accessory repair factors when end processing is necessary.

PAXX-Ku70 mimicry of the XLF-Ku80 interface

The local resolution of DNA-PK density map in the apo PAXX-LR state reaches 3.6 Å at the bottom crevice of Ku70 (fig. S3), revealing previously unidentified densities corresponding to the PAXX C-terminal KBM. We were able to unambiguously attribute residues 180 to 201 to this region (Fig. 2A). The short loop 184 to 188 forms a turn similar to the XLF KBM (*19*) and is inserted into a conserved pocket between the vWA (von Willebrand factor type A) domain and the core DNA binding region of Ku70 (Fig. 2B and fig. S7A). The location of the PAXX KBM binding at Ku70 is remarkably similar to the XLF KBM binding to a similar pocket of Ku80, consistent with the redundant and scaffolding roles of PAXX and XLF in the stabilization of the LR complex revealed by both biochemical and single-molecule studies (*12*, *15*, *20*). Within the apo PAXX-LR complex, Ku70 still adopts an open conformation similar to the published LR, SR, and DNA-PK complexes in the absence of PAXX, whereas the Ku80 vWA domain will not open unless XLF is bound (*9*, *6*, *18*, *19*, *21*). This subtle functional difference between PAXX and XLF binding may explain why PAXX was not initially characterized as one of the core NHEJ factors due to the partial overlapping role played by these two scaffolding factors (fig. S7) (*14*, *22*). This is further supported by the fact that the PAXX N-terminal dimerization domain (*12*) cannot be visualized and is likely only flexibly tethered to the two Ku70 molecules on the opposite ends of the LR complex through their C-terminal KBMs without stably interacting with the rest of the XLF-XRCC4-LigIV scaffold (Fig. 1A). The flexible linker between the PAXX dimerization domain and KBM motif contains 38 amino acids. With the general assumption of ~3.5 Å per residue, two copies of these linkers are long enough to cover the 110-Å distance between the two PAXX KBM motifs in the PAXX-LR model. Similarly, two 70–amino acid linkers flexibly tether the XLF dimerization domain and two corresponding XLF KBM motifs, which are 240 Å apart in the same model (Fig. 1A). The observation that a six–amino acid truncation of the XLF linker did not affect the NHEJ ligation efficiency (*23*) is consistent with the idea that the length of the linkers is more than sufficient to link the two XLF KBM motifs. To test whether PAXX is able to fully replace XLF′s role in facilitating synapsis and end joining, we assayed the activity of both LR and SR complexes in vitro (fig. S8). Ligation products are only observed in complexes containing XLF, indicating that while PAXX plays a redundant stabilizing role, it is not capable of supporting the assembly of an NHEJ complex that is competent for end joining.

To further confirm the assignment of PAXX KBM in the model, in silico structure prediction of Ku70/80 in complex with PAXX KBM was performed using AlphaFold2 (AF2) Multimer (*24*).

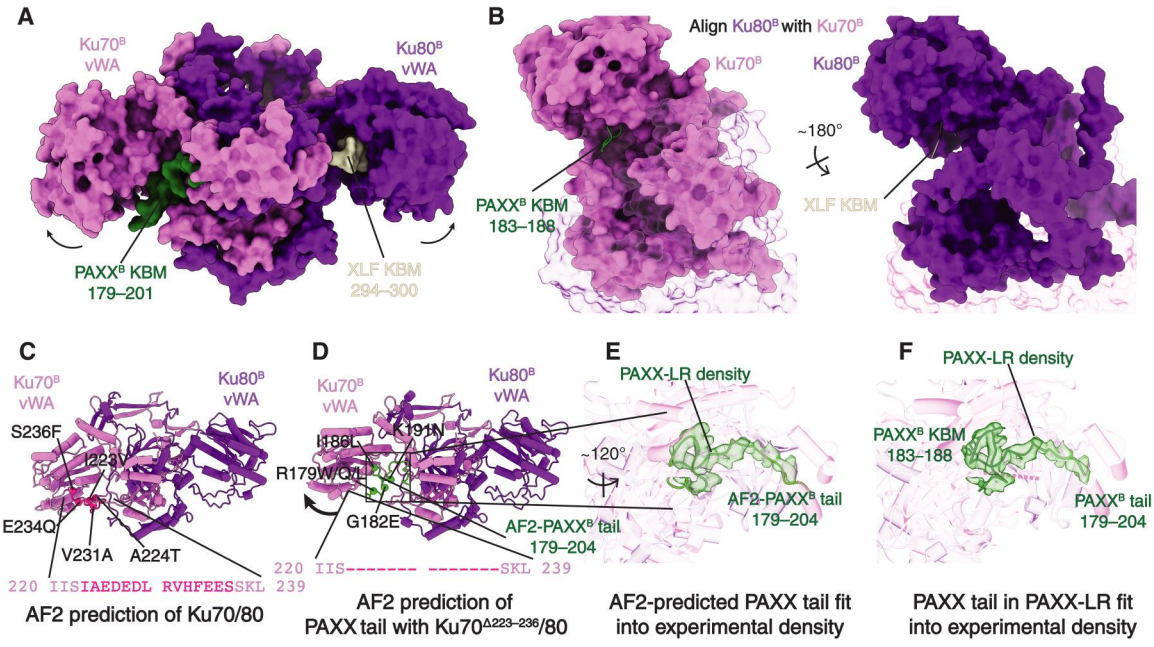

Fig. 2. PAXX stabilizes the LR complex by interacting with Ku70. (**A**) An overview of both PAXX KBM (179 to 201) and XLF KBM (294 to 300) bound to Ku70 and Ku80, respectively, in the LR complex. Both Ku70 and Ku80 vWA domains adopt open conformations, highlighted by curved arrows. (**B**) PAXX and XLF KBMs bind the corresponding Ku in a similar manner. Ku70 and Ku80 from the same LR complex protomer are aligned by a 180° rotation. The other copy of the Ku protein is transparent in each Ku70/80 heterodimer. (**C**) AlphaFold2 (AF2) (*24*) predicted Ku70/80 heterodimer adopts closed conformation for both Ku70 and Ku80 vWA domains. The flexible loop 223 to 236 in Ku70 involved in closing the Ku70 vWA domain is highlighted in deep pink. Spheres depict the locations of residues of missense mutations from patients with cancer identified from the COSMIC database (*26*). (**D**) AF2 prediction of a Ku70/80 heterodimer variant in complex with PAXX KBM. In silico truncation of Ku70 loop 223 to 236 is highlighted by the sequence. The open conformation of the Ku70 vWA domain is shown by the curved arrow. Missense mutations from patients with cancer identified in the PAXX KBM are highlighted. (**E**) Close-up view of AF2-predicted PAXX KBM fit into the experimental apo PAXX-LR complex density. Ku70$^{\Delta 223-236}$/80 ribbon model is shown as transparent background. The experimental density map corresponding to the PAXX KBM (179 to 201) is shown in green and transparency. (**F**) Close-up view of PAXX KBM model fit into the experimental density.

Naïve prediction using intact Ku70/80 core and PAXX KBM sequences resulted in a closed conformation for the Ku70 vWA domain, similar to that observed in the crystal structure (Fig. 2C) (*25*), avoiding a full engagement with the PAXX KBM. Since loop 220 to 239 of Ku70 is intrinsically flexible and has never been fully resolved in any crystallographic or EM structures before, we reason that it may play an autoinhibitory role by locking Ku70 into a closed conformation. Further, instead of inserting itself into the Ku70 pocket, the AF2-predicted loop 220 to 239 of Ku70 appears to be in proximity and blocking the pocket entry space (fig. S7, E to G). This prediction is consistent with the structural data, where PAXX KBM has stronger binding affinity to the Ku70 pocket compared with the Ku70 loop 220 to 239. When we supply a Ku70/80 core sequence with an internal truncation of Ku70 residues 220 to 239 for structure prediction, not only the Ku70 vWA domain adopts an open conformation in the resulting complex similar to that observed in the experimental model, but also the PAXX KBM binding to Ku70 is remarkably similar to the independently built cryo-EM model (Fig. 2, D to F). Our in silico AF2 prediction, in combination with experimental data, therefore, suggest that this conserved PAXX KBM plays an important role in NHEJ by directly tethering the two Ku molecules on opposite ends of the LR complex. We were able to map residues identified in the human cancer mutation database (*26*) onto both the Ku70 inhibitory loop 220 to 239 and the PAXX KBM (Fig. 2, C and D), suggesting that disruption of the regulatory elements in this region may contribute to human cancers.

ATP-induced rearrangement of the LR complex

The two copies of DNA-PKcs in the LR-ATP complex undergo substantial conformational change compared to the apo PAXX-LR state (movie S1). A ~60° rotation of the main body [without its N-terminal HEAT (N-HEAT) repeat regions] converts the dimeric interface from the forehead region by the ABCDE loop to the NUC194 domain that is close to the PQR loop (S2023, S2029, S2041, S2053, and S2056 phosphorylation sites) (Fig. 3A). The DNA end-blocking (DEB) helix becomes disordered and cannot be resolved in the LR-ATP state anymore. On the other hand, the ABCDE loop encompassing residues 2606 to 2649 undergoes disorder-to-order transition and shields the forehead region of DNA-PKcs, which constitutes a substantial portion of the dimeric surface in the apo PAXX-LR state (Fig. 3B) (*9*, *27*). This stabilized ABCDE loop has also been reported in a recently published cryo-EM study of the phosphorylated DNA-PK in complex with Artemis, a hairpin-opening nuclease involved in V(D)J recombination (*27*). During this apo PAXX-LR–to–LR-ATP state transition, both the disappearance of the DEB helix and the stabilization of the ABCDE loop coincide with the disruption of the YRPD/YRPD-I interaction observed in the LR state, resulting in the establishment of alternative interaction surfaces between these highly conserved motifs and DNA-PKcs (Fig. 3B and fig. S9A) (*28*). All of these must happen after both ABCDE clusters are phosphorylated because the kinase active center from the opposite side cannot reach the ABCDE sites in the apo PAXX-LR state unless the

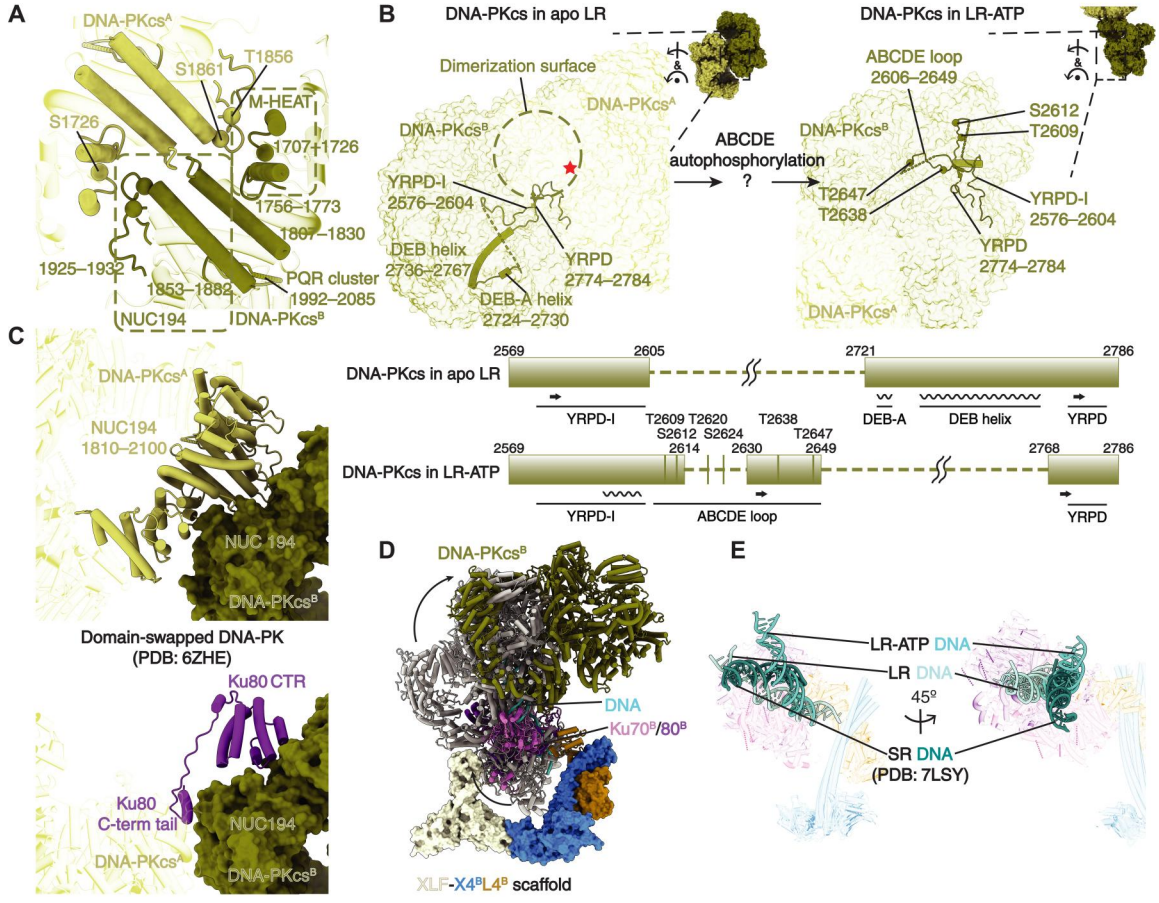

Fig. 3. DNA-PK undergoes substantial conformational change within the LR complex upon autophosphorylation. (**A**) Close-up view showing the previously unidentified interface between the two copies of DNA-PKcs in the LR-ATP state. The interacting loops and helices are highlighted and grouped accordingly. Nearby phosphorylation sites identified from the PhosphoSitePlus database (*59*) and the putative location of the PQR cluster are labeled as well. (**B**) Rearrangement of important regulatory elements at the surface of DNA-PKcs upon autophosphorylation of the ABCDE cluster. The DEB helix, the YRPD/YRPD-I motifs, and the previously unmodeled ABCDE loop are highlighted. The rest of DNA-PKcs are shown as surface representation in transparency. DNA-PKcs active center is highlighted by a red star. Locations of the phosphorylated residues are labeled and stacked on top of the forehead region. The presence (colored bar) and absence (dashed line) of the above regulatory elements in the structural models are highlighted in the domain diagram below. Secondary structures and the ABCDE phosphorylation sites are also labeled. (**C**) Side-by-side comparison of DNA-PKcs surface involved in either the dimerization within the LR-ATP state (left) or the Ku-CTR interaction within the previously reported domain-swapped DNA-PKcs dimeric state (right) (*18*). The NUC194 domain of one copy of DNA-PKcs is aligned and shown as surface representation. (**D**) Superimposition of the LR (gray) and LR-ATP (colored) complexes with XLF-XRCC4-LigIV BRCT scaffold aligned (shown as surface representation). Only half of the model is shown for clarity purposes. Movement of the Ku-DNA region and DNA-PKcs from the LR state to the LR-ATP state is highlighted by curved arrows. (**E**) Superimpositions of the LR, LR-ATP, and SR complexes aligned by the XLF-XRCC4-LigIV BRCT scaffold in two different views. DNA ends are shown as solid ribbons and colored with different cyan colors for highlighting and comparison. Other elements are shown in transparent representation.

YRPD-I loop is peeled off from the YRPD motif, consistent with their important regulatory role in coupling DNA-PKcs autophosphorylation with the state transition. It is likely that upon ATP addition, both copies of active DNA-PKcs autophosphorylate each other's ABCDE loops in trans, as suggested by both structural and biochemical observations (*9*, *10*). This, in turn, positions two negatively charged clusters right on top of both forehead regions near the DNA-PKcs dimerization surface, pushing the two large kinases away from each other and leading to the stabilization of LR-ATP state.

The previously unidentified DNA-PKcs interaction surface in the LR-ATP state also blocks the binding between the Ku80 C-terminal region (CTR) and DNA-PKcs (Fig. 3C) (*6*, *18*). Thus, the formation of such a cradle-cradle interaction can further destabilize Ku and DNA-PKcs, promoting the ultimate LR-to-SR transition (*9*). Expectedly, the weak interaction between DNA-PKcs and Ku-DNA results in medium-resolution reconstruction of Ku70/80 within the LR-ATP state. After focused refinement, we were able to observe a similar interaction between Ku and N-HEAT repeat of DNA-PKcs as in the apo PAXX-LR state, but the rest of DNA-PKcs clearly rotates out and forms an open cradle conformation (fig. S9, B to D). This opening of the cradle domain is more marked than the one captured in the phosphorylated DNA-PK-Artemis complex, possibly priming the recruitment of Artemis. It is also possible that the DNA ends can be exposed in this open space, enabling further processing by accessory NHEJ factors other than Artemis.

There are substantial rotations of the entire DNA-PK complex relative to the XLF-XRCC4-LigIV scaffold during the apo PAXX-LR–to–LR-ATP transition. Meanwhile, we did not observe any noticeable changes in the scaffold region during state transition, partially because of the resolution limitation in the designated region (Fig. 3D). With the scaffold aligned between the two states, Ku/DNA and DNA-PKcs are rotated relative to their original positions in the apo PAXX-LR state by ~30° and ~90°, respectively, with the anchoring point located somewhere between the tandem BRCT (BRCA1 C-terminal) domains of LigIV. The distance between the DSB ends within the LR-ATP state increases from ~115 to ~125 Å. While the LR-ATP state constitutes another version of the LR complex, it is distinct from the other versions in that the DNA ends are completely exposed in the LR-ATP state. To ultimately transition from the LR-ATP to the SR state, the two Ku need to rotate back to align the DSBs for the final ligation (Fig. 3E). The DSB ends observed in the LR-ATP complex do not reside on a linear trajectory between the apo PAXX-LR and the SR states. Instead, they form a triangle with similar distances separated among the three states (Fig. 3E). It remains to be tested whether these transitions are reversible or not, but it is clear that the LR-ATP complex is a rather stable NHEJ state due to the fact that we are able to completely convert apo PAXX-LR to the LR-ATP state by supplying ATP for 15 min in our sample preparation procedure (fig. S2).

A previously unidentified dimeric DNA-PKcs excludes Ku-DNA binding

The previously unidentified dimeric DNA-PKcs uses three distinct surfaces to form the alternative dimeric form, including the N-HEAT repeat (1 to 150), a segment of the M-HEAT within the central cradle domain (2300 to 2500), as well as the FAT domain in the head unit (2820 to 2850 and 3000 to 3100) (Fig. 4A). Among these interfaces, the HEAT repeat buries most of the interaction area (~900 $Å^2$) and is mutually exclusive with the Ku/DNA-DNA-PKcs interactions observed within both the apo PAXX-LR and the LR-ATP states (Fig. 4B). To our surprise, we cannot detect any DNA density within this dimeric DNA-PKcs complex, indicating that this form of DNA-PKcs can function as a standalone assembly and is most likely to be inactive without Ku-DNA binding. Therefore, we compared the active sites of DNA-PKcs from all three states determined in this study (Fig. 4C). The PIKK regulatory domain (PRD) is an evolutionarily conserved element that negatively regulates kinase activity by occupying the substrate-binding groove (*29*, *30*). The PRD is disordered in the apo PAXX-LR complex but structured in both the LR-ATP and the dimeric DNA-PKcs complexes.

To accommodate this dimer, both the N-HEAT and M-HEAT repeats of DNA-PKcs have to undergo conformational rearrangements, resulting in the tightest cradle conformation among all three states described in this study (fig. S10). These conformational changes that result in the closing of the cradle are in contrast to the opening of the domain observed in the transition from the apo PAXX-LR to the LR-ATP states. Without binding to Ku-DNA, the conformation of DNA-PKcs N-HEAT in the dimeric DNA-PKcs complex is more similar to the structure of standalone DNA-PKcs compared with the DNA-PK complex (fig. S10, F to H).

On the basis of the structural observations described above, it is possible that this DNA-PKcs–only dimeric state exists in vivo and functions as a reservoir of DNA-PKcs molecules held in an inactive state. Recruiting two copies instead of one DNA-PKcs by Ku-bound DSB could potentially increase the efficiency of forming the apo PAXX-LR complex. Alternatively, DNA-PKcs dimerization during the LR-ATP–to–SR state transition may contribute to the complete dissociation of DNA-PKcs from the two Ku-DNA assemblies. However, we were not able to capture this conversion in our assay even after an extended incubation with ATP, indicating that other unknown mechanism regulating this process remains to be found.

DISCUSSION

DNA-PKcs transition among multiple dimeric states during NHEJ

All three of the previously unidentified NHEJ intermediate complexes we describe in this study contain two copies of DNA-PKcs, with distinct regions being used for establishing the dimeric interactions (Figs. 1 and 4D and fig. S11). Complementary to the DNA-focused NHEJ reaction cycle we proposed previously (*9*), these three previously unidentified NHEJ states illustrate the substantial conformational plasticity of DNA-PKcs during NHEJ (Fig. 4D and movie S1). Most of the studies to date examine how autophosphorylation of DNA-PKcs regulates NHEJ progression (*31*). However, additional phosphatases, such as PP5 (protein phosphatase 5) (*32*), PP6 (*33*), or PP2A (*34*), are likely involved in resetting the DNA-PKcs enzymatic cycle (Fig. 4D). Notably, the relative orientation of the two DNA-PKcs within the inactive dimer closely resembles that of the self-inhibitory ATM and the ATR-ATRIP dimer structures (fig. S11) (*35*, *36*), implying that a similar regulatory mechanism may be shared by all PIKK family kinases involved in DSB repair pathways.

In summary, the previously unidentified protein complexes described address critical knowledge gaps in our understanding of how DSBs are repaired by NHEJ, a key repair pathway for maintaining genome integrity, preventing cancer formation, and determining the response to radiation therapy. Specifically, our study has revealed the remarkable conformational flexibility of DNA-PKcs and how its kinase activity is regulated in a number of additional steps that have not been observed previously (Fig. 4D). In addition, our studies have revealed critical insights as to how scaffolding NHEJ factor PAXX stabilizes different NHEJ states by mimicking its functional ortholog, XLF. Last, we describe a previously unidentified LR-ATP complex with the DNA ends fully exposed, which may serve as the platform for the recruitment of NHEJ end-processing factors. Together, our findings reveal previously unknown insights into the mechanism and regulation of NHEJ, providing opportunities for the development of therapeutic strategies that modulate responses to radiation therapy and chemotherapy.

MATERIALS AND METHODS

Protein purification

DNA-PKcs and Ku heterodimer were purified based on the previously described protocol (*37*). DNA-PKcs was purified endogenously from the nuclear pellet of HeLa cells, and Ku was purified from baculovirus-infected insect cells (*38*). The final concentration of DNA-PKcs was 0.5 mg/ml, and that of Ku70/80 was 1.53 mg/ml. The LigIV-XRCC4 complex was purified from baculovirus-infected

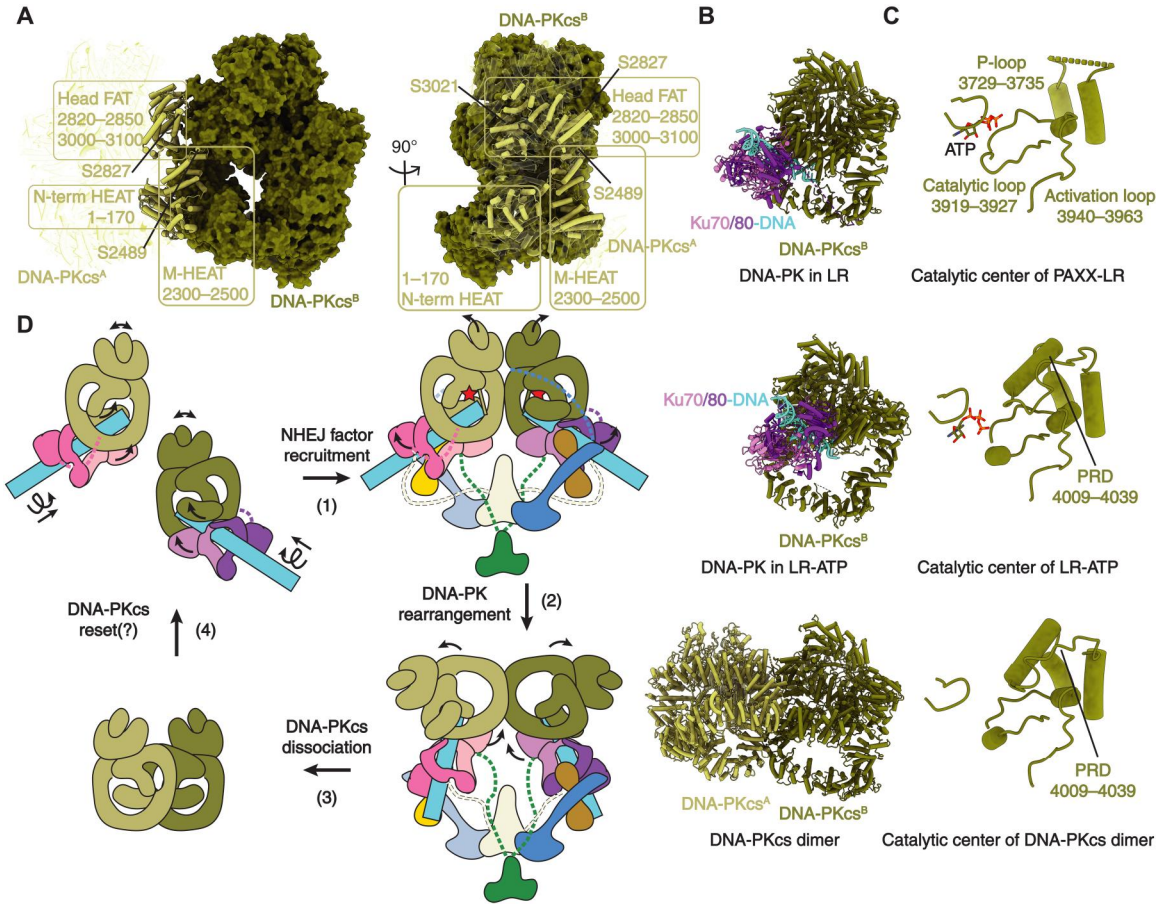

Fig. 4. The inactive dimeric state of DNA-PKcs potentially represents a missing link in the DNA-PKcs regulatory cycle during NHEJ. (**A**) Front (left) and side (right) overview of interaction surfaces between the two copies of DNA-PKcs in the dimeric state. Phosphorylation sites at the interaction surfaces are highlighted as spheres when visualizable. (**B**) The dimeric interface between two copies of DNA-PKcs is also used to bind Ku/DNA. Top to bottom: DNA-PK in the PAXX bound LR complex, DNA-PK in the LR-ATP state, and DNA-PKcs dimeric complex. The copy of DNA-PKcs on the right is aligned among the three states. (**C**) Comparison of the kinase active site between the apo PAXX-LR state (top), LR-ATP (middle), and dimeric DNA-PKcs (bottom). A flexible PRD represents the hallmark of the activated kinase, as shown in both the LR-ATP and the dimeric DNA-PKcs complexes. (**D**) A model of the full regulatory cycle of DNA-PKcs during NHEJ. When two DNA-PK are brought closer by the XLF-XRCC4-LigIV scaffold (1), PAXX further stabilizes the LR state. Upon DNA-PKcs activation and autophosphorylation in the presence of ATP, both Ku-DNA and DNA-PKcs rotate outward to expose the DSB ends and promote the dissociation of DNA-PKcs (2). Dimerization of DNA-PKcs using an evolutionarily conserved interface triggers the Ku-DNA dissociation (3), and this DNA-PKcs reservoir can be efficiently recruited by Ku in another round of DSB repair by NHEJ (4).

insect cells (*38*), and the final concentration was 7.9 mg/ml. Purification of XLF was performed as described previously (*39*), and the final concentration was 2 mg/ml. Purification of PAXX was also performed as described previously, with the final concentration of 1.3 mg/ml (*38*).

Assembly of the apo PAXX-LR, the LR-ATP, and the DNA-PKcs dimeric complexes

Sequences of oligonucleotides that were purchased from Integrated DNA Technology (IDT) are listed in table S1. Double-stranded DNA substrates with RNA-biotin tags were assembled as described previously (*9*).

Both the Y30-T40-c8 and Y30 DNA-RNA substrates were used for the assembly of both the apo PAXX-LR and the LR-ATP complexes. The overall protocol of complex assembly is similar as described previously (*9*) with the following modifications: First, when Y30 DNA-RNA substrate was used, the steps of combining half complexes and 37°C incubation were omitted since there are no half DNA-RNA substrates used as Y30-T40-c8. Second, PAXX was supplied together with LigIV-XRCC4 complex and XLF, followed by a 15-min incubation at room temperature. During the wash and elution steps, 50 nM PAXX was added throughout the process to achieve higher PAXX occupancy and further stabilize the complex. Third, to prepare the LR-ATP complex, 1 mM ATP (final concentration) was added to the sample after elution and incubated at room temperature for 15 min. Fourth, both the negative staining and cryo-EM samples of the LR-ATP complex were prepared with 0.01% NP-40 (final concentration). To protect the cryo-EM sample from the air-water interface, the NP-40 concentration was boosted up to 0.05% right before applying it to the graphene oxide (GO)–coated grids.

The DNA-PKcs dimeric complex was first observed in two independent experiments. The sample used for preliminary cryo-EM data collection was assembled using Y25 DNA-RNA substrate, DNA-PKcs, and Ku70/80 heterodimer in the absence of other NHEJ factors. The sample used for negative-staining data collection

was assembled with the same Y25 DNA-RNA substrate but with DNA-PKcs, Ku70/80, LigIV-XRCC4 complex, and XLF. The final factor concentration of each Y30-T40-c8 half complexes and Y30/Y25 complexes are the same as described previously (*9*). To better reproduce the result, the complex assembly protocol was optimized to include only DNA-PKcs and Y25 DNA-RNA substrate under low-salt conditions (20 mM KCl instead of 50 mM), while the other steps and the ratio between DNA-PKcs and substrate DNA remain the same.

Electron microscopy

The same procedure as described previously (*9*) was applied to prepare negative-staining and cryo-EM samples. Negative-staining samples were used for data collection with a JEOL 1400 transmission electron microscope with a Gatan 4000 × 4000 charge-coupled device camera operating at 120 kV at a nominal magnification of ×30,000 (3.71 Å per pixel) and a defocus range of −1.5 to −3 μm (20 e^- Å^{-2} exposures). Leginon (*40*) was used for data collection.

All of the datasets were collected on homemade GO grids (*41*) (GO applied on 2/1 copper/gold; Electron Microscopy Sciences). Solarus plasma cleaner (Gatan) was used for glow discharge, and Vitrobot IV (FEI) was used for plunge-freezing. The grid was blotted for 4 s at 25 force, plunge-frozen in liquid ethane, and stored in liquid nitrogen.

Two datasets were collected for the LR-ATP complex at the Pacific Northwest Center for cryo-EM (PNCC), one with Y30-T40-c8 substrate (super-resolution mode, pixel size of 0.5395 Å per pixel, and 11,006 movies) and the other with Y30 substrate (super-resolution mode, pixel size of 0.528 Å per pixel, and 19,442 movies). Titan Krios-3 transmission electron microscope (Thermo Fisher Scientific) operating at 300 kV was used with a K3 direct electron detector (Gatan) at ×30,000 magnification. For both datasets, the defocus range was set to −2 to −4 μm. Forty-five–frame exposures were taken during 4-s total exposure time, using a dose rate of 15 e^- per pixel per second, corresponding to a total dose of ~66 e^- Å^{-2} per movie series. Similar parameters were used for collecting the one dataset for the DNA-PKcs dimeric complex at PNCC, with slight differences in magnification (×105,000) and other parameters (super-resolution mode, pixel size of 0.413 Å per pixel, and 13,132 movies).

Image processing and three-dimensional reconstruction

Preprocessing of data obtained from negative stain samples was performed using the Appion processing environment (*42*). Particle picking [difference of Gaussians particle picker (*43*)], CTF (contrast transfer function) estimation [CTFFind4 (*44*)], two-dimensional (2D) classification [MSA-MRA (multivariate statistical analysis and multireference alignment) in IMAGIC (*45*)], and 3D initial model building [CryoSPARC (*46*) and EMAN2 (*47*)] were performed similarly with the previously described pipeline (*9*). For the initial 3D model of the LR-ATP complex, 82,622 particles were extracted and stacked with a box size of 144 pixels × 144 pixels (3.71 Å per pixel). The DNA-PKcs dimeric class appeared when the DNA-PK complex map was used as the initial model during preliminary data processing. Thus, the ab initio 3D model does not need to be generated ahead of time.

RELION 3.1.3 was then used for all the preprocessing, 3D classification, model refinement, postprocessing, and local-resolution estimation jobs (*48*). To preprocess the cryo-EM data, all the movie frames were aligned using RELION's own implementation.

To process the LR-ATP dataset with the Y30-T40-c8 substrate, 1,101,255 particles were picked using Gautomatch, binned by 2 (2.158 Å per pixel), and supplied to a 10-class 3D classification with the preprocessing reconstruction map (low pass–filtered to 30 Å) as an initial reference. Angular sampling interval (15°), 10 pixels offset search range, and 2 pixels of offset search step were used for the first 50 iterations and then followed by 120 iterations of default classification (7.5°, 5 pixels, and 1 pixel). While class 7 (134,748 particles) showed the expected LR-ATP state with sharp structural features, class 10 (138,252 particles) appeared to be very similar to the previously reported LR complex (*9*). Both classes were selected, autorefined with C2 symmetry, and recentered. However, class 10 was re-extracted without binning (1.079 Å per pixel, box size of 480), and all further processing on class 7 was done with the bin-2 stack. After that, both stacks were supplied to another round of 3D autorefinement with a soft mask applied around the whole complex and then subjected to per-particle CTF refinement (per-particle defocus and per-micrograph astigmatism estimation, as well as beam tilt estimation) and Bayesian particle polishing (*48*). Class 7 resulted in a reconstruction of 11.51-Å resolution, likely due to its intrinsic flexibility, and class 10 provided a 4.39-Å reconstruction. All reported resolutions correspond to the gold-standard Fourier shell correlation (FSC) using the 0.143 criteria (*49*), and the 3DFSC distribution plots are generated by the online 3DFSC calculation server (*50*). This map was used as the overall map for PAXX-LR model deposition.

To further improve the map resolution of class 10 (apo PAXX-LR complex), a similar strategy of symmetry expansion, signal subtraction, and classification methods as previously described (*9*) were used on both DNA-PKcs-Ku-LigIV BRCT domain and the XLF-XRCC4-LigIV scaffold. Specifically, the rigid region (DNA-PKcs, Ku70/80, DNA, and the first BRCT domain of LigIV) was masked, symmetry expanded with C2 (276,504 particles), and refined locally (initial angular sampling interval of 3.7° with local search) to 3.98 Å. Subsequent two-class 3D classification without alignment pooled out 144,755 good particles for autorefinement and postprocessing. The final resolution of this map reaches 3.65 Å. The local resolution of the map was also estimated within RELION 3.1.3 (*51*). The flexible body (XLF-XRCC4-LigIV) was binned by 2 (2.158 Å per pixel), recentered, and signal-subtracted without applying symmetry expansion. After a local refinement (3.7° interval with local search), the following three-class 3D classification without alignment resulted in a cleaned stack with 44,866 particles. The stack was further refined locally (3.7° interval with local search again) to 13.87 Å. The contour levels of the composite map shown in Fig. 1 are 0.02 (scaffold) and 0.015 (DNA-PK).

Similar initial 10-class classification was applied on the LR-ATP dataset with the Y30 substrate (2,684,531 total particles), resulting in a good class 1 with 361,071 particles. This class was combined with the class 7 from the Y30-T40-c8 dataset mentioned above. Combined refinement gives 8.9-Å resolution for the overall LR-ATP state model and was used for deposition. After that, a similar strategy was used to process the stack without un-binning (2.158 Å per pixel). After symmetry expansion, DNA-PKcs without N terminus was masked for signal subtraction (recentered), autorefinement, and three-class 3D no-alignment classification. The cleaned stack with 150,310 particles was refined locally and postprocessed to 5.0

Å. To visualize the Ku and N terminus of DNA-PKcs, the region corresponding to the entire DNA-PK complex was masked for signal subtraction (recentered) and autorefinement. Last, 176,398 particles give a DNA-PK complex map of 5.9-Å resolution. The stack corresponding to flexible scaffold gives 12.1-Å resolution after focused local refinement. The contour levels of the composite map shown in Fig. 1 are 0.0145 (scaffold) and 0.031 (DNA-PK).

A similar strategy was applied to the DNA-PKcs dimeric dataset. A total of 1,587,734 particles were picked up and binned by 2 (2.158 Å per pixel) for a six-class 3D classification. Class 3 (251,070 particles) was selected and refined to 5.67 Å using C2 symmetry. Further symmetry expansion and signal subtraction focused on one copy of DNA-PKcs. Another round of alignment-free two-class sorting resulted in a clean stack with 64,775 particles, and CTF-refinement and autorefinement provided an improved 4.27-Å map. DeepEMhancer (*52*) was used to denoise the final map to obtain better features for model building. The contour level of the composite map shown in Fig. 1 is 0.3.

UCSF Chimera (*53*) and ChimeraX (*54*) were used for volume segmentation, figure/movie making, and rigid-body docking. In general, the masks used for classifications, refinements and postprocessing jobs have direct extension of 3 pixels and cosine-shaped soft edge of 6 pixels. For all individual refinement jobs, the references are from previous steps of 3D classification or 3D refinement.

Model building

Sequence alignment of all the subunits (DNA-PKcs, Ku70, Ku80, XLF, XRCC4, and LigIV) was performed using CLC Sequence Viewer 7 (figs. S1 to S6). The published cryo-EM model of the LR complex (*9*) was referenced as a homology model to build the apo PAXX-LR complex. To build the PAXX KBM into the density, Coot (*55*) was used to generate the peptide, and then Coot and ISOLDE (*56*) in ChimeraX were used to flexibly fit the peptide into the density. The AF2 (*24*) predicted model of DNA-PKcs was also referenced during model building. The registers can be clearly identified as most of the densities and bulky sidechains aligned well between the map and the model. Because of the resolution limitation, only the secondary motifs of the flexible regions are flexibly fitted into density with minimal disruption of existing model constraints. Manual rebuilding of model with Coot and ISOLDE and the Phenix real-space refinement (*57*) were performed iteratively to further refine the model.

A similar strategy was used to build the model of the LR-ATP complex. The cryo-EM models of both the LR complex and the published DNA-PKcs-Artemis complex model [Protein Data Bank (PDB): 7SGL] (*27*) were referenced during model building. ISOLDE was used to fit in the flexible regions with minimal disruption. Phenix real-space refinement was performed as the final step to refine the model.

DNA-PKcs within the published cryo-EM model of the LR complex was referenced as a homology model to build the alternative DNA-PKcs dimeric complex. Rigid fitting of the model into the map results in clashed areas at the dimeric interaction surface. Because of the resolution limitation, ISOLDE with distance restraints was applied to flexibly fit the secondary structures into the density. To completely resolve the clashing contacts between two copies of DNA-PKcs, Coot and ISOLDE were also used to manually build the backbones at the dimerization surfaces. When one copy of the DNA-PKcs model was refined substantially, it was duplicated to rigidly fit into the other half of the cryo-EM map to ensure no clashing in between. This process was performed for several iterations; the refined model was supplied to Namdinator (*58*) and Phenix real-space refinement for further refinement. For clarification, all subunits located on the left in Fig. 1 are named copy A, and the subunits on the right are named copy B.

In vitro ligation assay

The sequences of Cy5-labeled DNA substrates, purchased from IDT, are listed in table S1. DNA substrates were annealed as described above. The substrate architecture and different experimental conditions are shown as in fig. S8.

The experimental procedures were similar as the previously published protocol (*9*). The LR complexes with Cy5-labeled substrate were assembled similarly as described for EM sample preparation above. After the intact complex was assembled and bound to the streptavidin-coated magnetic beads, samples were washed with buffer containing 5 mM $MgCl_2$ without EDTA. ATP (1 mM) was then added to the sample, and the reaction was carried out for 60 min at 37°C. For the control lanes, the complex assembly buffer remained unmodified, but various buffer conditions used after bead pull-down were modified accordingly. Next, the beads were boiled at 95°C for 5 min to elute off the DNA, and proteinase K was added to allow incubation at 37°C for 15 min. Formamide (50% volume) and 10 mM of EDTA were added to prepare the sample for denatured urea gel [TBE (tris-borate-EDTA) gel with 8% acrylamide (19:1)]. Gel electrophoresis was performed at 250 V for 40 min and scanned using a Sapphire Biomolecular Imager (Azure) at optimal absorbance.

Supplementary Materials

This PDF file includes:

Figs. S1 to S11
Tables S1 and S2
Data S1 to S8
Legend for movie S1

Other Supplementary Material for this manuscript includes the following:

Legend for movie S1

REFERENCES AND NOTES

1. B. M. Stinson, J. J. Loparo, Repair of DNA double-strand breaks by the nonhomologous end joining pathway. *Annu. Rev. Biochem.* **90**, 137–164 (2021).
2. B. Zhao, E. Rothenberg, D. A. Ramsden, M. R. Lieber, The molecular basis and disease relevance of non-homologous DNA end joining. *Nat. Rev. Mol. Cell Biol.* **21**, 765–781 (2020).
3. P. D. Aplan, Causes of oncogenic chromosomal translocation. *Trends Genet.* **22**, 46–55 (2006).
4. T. M. Gottlieb, S. P. Jackson, The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen. *Cell* **72**, 131–142 (1993).
5. S. P. Lees-Miller, Y. R. Chen, C. W. Anderson, Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen. *Mol. Cell. Biol.* **10**, 6472–6481 (1990).
6. X. Chen, X. Xu, Y. Chen, J. C. Cheung, H. Wang, J. Jiang, N. de Val, T. Fox, M. Gellert, W. Yang, Structure of an activated DNA-PK and its implications for NHEJ. *Mol. Cell* **81**, 801–810.e3 (2021).
7. P. O. Mari, B. I. Florea, S. P. Persengiev, N. S. Verkaik, H. T. Bruggenwirth, M. Modesti, G. Giglia-Mari, K. Bezstarosti, J. A. Demmers, T. M. Luider, A. B. Houtsmuller, D. C. van Gent, Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18597–18602 (2006).

8. K. Yano, D. J. Chen, Live cell imaging of XLF and XRCC4 reveals a novel view of protein assembly in the non-homologous end-joining pathway. *Cell Cycle* **7**, 1321–1325 (2008).
9. S. Chen, L. Lee, T. Naila, S. Fishbain, A. Wang, A. E. Tomkinson, S. P. Lees-Miller, Y. He, Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature* **593**, 294–298 (2021).
10. K. Meek, P. Douglas, X. Cui, Q. Ding, S. P. Lees-Miller, Trans autophosphorylation at DNA-dependent protein kinase's two major autophosphorylation site clusters facilitates end processing but not end joining. *Mol. Cell. Biol.* **27**, 3881–3890 (2007).
11. M. Xing, M. Yang, W. Huo, F. Feng, L. Wei, W. Jiang, S. Ning, Z. Yan, W. Li, Q. Wang, M. Hou, C. Dong, R. Guo, G. Gao, J. Ji, S. Zha, L. Lan, H. Liang, D. Xu, Interactome analysis identifies a new paralogue of XRCC4 in non-homologous end joining DNA repair pathway. *Nat. Commun.* **6**, 6233 (2015).
12. T. Ochi, A. N. Blackford, J. Coates, S. Jhujh, S. Mehmood, N. Tamura, J. Travers, Q. Wu, V. M. Draviam, C. V. Robinson, T. L. Blundell, S. P. Jackson, DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. *Science* **347**, 185–188 (2015).
13. A. Craxton, J. Somers, D. Munnur, R. Jukes-Jones, K. Cain, M. Malewicz, XLS (c9orf142) is a new component of mammalian DNA double-stranded break repair. *Cell Death Differ.* **22**, 890–897 (2015).
14. X. Liu, Z. Shao, W. Jiang, B. J. Lee, S. Zha, PAXX promotes KU accumulation at DNA breaks and is essential for end-joining in XLF-deficient mice. *Nat. Commun.* **8**, 13816 (2017).
15. J. L. Wang, C. Duboc, Q. Wu, T. Ochi, S. Liang, S. E. Tsutakawa, S. P. Lees-Miller, M. Nadal, J. A. Tainer, T. L. Blundell, T. R. Strick, Dissection of DNA double-strand-break repair using novel single-molecule forceps. *Nat. Struct. Mol. Biol.* **25**, 482–487 (2018).
16. A. K. Chaplin, S. W. Hardwick, A. K. Stavridi, C. J. Buehl, N. J. Goff, V. Ropars, S. Liang, T. M. De Oliveira, D. Y. Chirgadze, K. Meek, J. B. Charbonnier, T. L. Blundell, Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Mol. Cell* **81**, 3400–3409.e3 (2021).
17. Y. P. Yu, W. Wang, Q. Ding, R. Q. Ye, D. Chen, D. Merkle, D. Schriemer, K. Meek, S. P. Lees-Miller, DNA-PK phosphorylation sites in XRCC4 are not required for survival after radiation or for V(D)J recombination. *DNA Repair* **2**, 1239–1252 (2003).
18. A. K. Chaplin, S. W. Hardwick, S. Liang, A. Kefala Stavridi, A. Hnizda, L. R. Cooper, T. M. De Oliveira, D. Y. Chirgadze, T. L. Blundell, Dimers of DNA-PK create a stage for DNA double-strand break repair. *Nat. Struct. Mol. Biol.* **28**, 13–19 (2021).
19. C. Nemoz, V. Ropars, P. Frit, A. Gontier, P. Drevet, J. Yu, R. Guerois, A. Pitois, A. Comte, C. Delteil, N. Barboule, P. Legrand, S. Baconnais, Y. Yin, S. Tadi, E. Barbet-Massin, I. Berger, E. Le Cam, M. Modesti, E. Rothenberg, P. Calsou, J. B. Charbonnier, XLF and APLF bind Ku80 at two remote sites to ensure DNA repair by non-homologous end joining. *Nat. Struct. Mol. Biol.* **25**, 971–980 (2018).
20. S. K. Tadi, C. Tellier-Lebegue, C. Nemoz, P. Drevet, S. Audebert, S. Roy, K. Meek, J. B. Charbonnier, M. Modesti, PAXX is an accessory c-NHEJ factor that associates with Ku70 and has overlapping functions with XLF. *Cell Rep.* **17**, 541–555 (2016).
21. X. Yin, M. Liu, Y. Tian, J. Wang, Y. Xu, Cryo-EM structure of human DNA-PK holoenzyme. *Cell Res.* **27**, 1341–1350 (2017).
22. P. J. Hung, B. R. Chen, R. George, C. Liberman, A. J. Morales, P. Colon-Ortiz, J. K. Tyler, B. P. Sleckman, A. L. Bredemeyer, Deficiency of XLF and PAXX prevents DNA double-strand break repair by non-homologous end joining in lymphocytes. *Cell Cycle* **16**, 286–295 (2017).
23. S. M. Carney, A. T. Moreno, S. C. Piatt, M. Cisneros-Aguirre, F. W. Lopezcolorado, J. M. Stark, J. J. Loparo, XLF acts as a flexible connector during non-homologous end joining. *eLife* **9**, e61920 (2020).
24. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
25. J. R. Walker, R. A. Corpina, J. Goldberg, Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607–614 (2001).
26. J. G. Tate, S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L. Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang, S. Ward, P. J. Campbell, S. A. Forbes, COSMIC: The catalogue of somatic mutations in cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).
27. L. Liu, X. Chen, J. Li, H. Wang, C. J. Buehl, N. J. Goff, K. Meek, W. Yang, M. Gellert, Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Mol. Cell* **82**, 177–189.e4 (2022).
28. J. P. Lees-Miller, A. Cobban, P. Katsonis, A. Bacolla, S. E. Tsutakawa, M. Hammel, K. Meek, D. W. Anderson, O. Lichtarge, J. A. Tainer, S. P. Lees-Miller, Uncovering DNA-PKcs ancient phylogeny, unique sequence motifs and insights for human disease. *Prog. Biophys. Mol. Biol.* **163**, 87–108 (2021).
29. B. L. Sibanda, D. Y. Chirgadze, D. B. Ascher, T. L. Blundell, DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. *Science* **355**, 520–524 (2017).
30. H. Yang, D. G. Rudge, J. D. Koos, B. Vaidialingam, H. J. Yang, N. P. Pavletich, mTOR kinase structure, mechanism and regulation. *Nature* **497**, 217–223 (2013).
31. D. W. Chan, S. P. Lees-Miller, The DNA-dependent protein kinase is inactivated by autophosphorylation of the catalytic subunit. *J. Biol. Chem.* **271**, 8936–8941 (1996).
32. T. Wechsler, B. P. Chen, R. Harper, K. Morotomi-Yano, B. C. Huang, K. Meek, J. E. Cleaver, D. J. Chen, M. Wabl, DNA-PKcs function regulated specifically by protein phosphatase 5. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1247–1252 (2004).
33. P. Douglas, J. Zhong, R. Ye, G. B. Moorhead, X. Xu, S. P. Lees-Miller, Protein phosphatase 6 interacts with the DNA-dependent protein kinase catalytic subunit and dephosphorylates gamma-H2AX. *Mol. Cell. Biol.* **30**, 1368–1381 (2010).
34. P. Douglas, G. B. Moorhead, R. Ye, S. P. Lees-Miller, Protein phosphatases regulate DNA-dependent protein kinase activity. *J. Biol. Chem.* **276**, 18992–18998 (2001).
35. D. Baretic, H. K. Pollard, D. I. Fisher, C. M. Johnson, B. Santhanam, C. M. Truman, T. Kouba, A. R. Fersht, C. Phillips, R. L. Williams, Structures of closed and open conformations of dimeric human ATM. *Sci. Adv.* **3**, e1700933 (2017).
36. Q. Rao, M. Liu, Y. Tian, Z. Wu, Y. Hao, L. Song, Z. Qin, C. Ding, H. W. Wang, J. Wang, Y. Xu, Cryo-EM structure of human ATR-ATRIP complex. *Cell Res.* **28**, 143–156 (2018).
37. A. A. Goodarzi, S. P. Lees-Miller, Biochemical characterization of the ataxia-telangiectasia mutated (ATM) protein from human cells. *DNA Repair* **3**, 753–767 (2004).
38. M. Hammel, Y. Yu, S. K. Radhakrishnan, C. Chokshi, M. S. Tsai, Y. Matsumoto, M. Kuzdovich, S. G. Remesh, S. Fang, A. E. Tomkinson, S. P. Lees-Miller, J. A. Tainer, An intrinsically disordered APLF links Ku, DNA-PKcs, and XRCC4-DNA ligase IV in an extended flexible non-homologous end joining complex. *J. Biol. Chem.* **291**, 26987–27006 (2016).
39. Y. Yu, B. L. Mahaney, K. Yano, R. Ye, S. Fang, P. Douglas, D. J. Chen, S. P. Lees-Miller, DNA-PK and ATM phosphorylation sites in XLF/Cernunnos are not required for repair of DNA double strand breaks. *DNA Repair* **7**, 1680–1692 (2008).
40. C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter, B. Carragher, Automated molecular microscopy: The new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
41. A. Patel, D. Toso, A. Litvak, E. Nogales, Efficient graphene oxide coating improves cryo-EM sample preparation and data collection from tilted grids. biorxiv 2021.03.08.434344 [**Preprint**]. 8 March 2021. https://doi.org/10.1101/2021.03.08.434344.
42. G. C. Lander, S. M. Stagg, N. R. Voss, A. Cheng, D. Fellmann, J. Pulokas, C. Yoshioka, C. Irving, A. Mulder, P. W. Lau, D. Lyumkis, C. S. Potter, B. Carragher, Appion: An integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009).
43. N. R. Voss, C. K. Yoshioka, M. Radermacher, C. S. Potter, B. Carragher, DoG picker and TiltPicker: Software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166**, 205–213 (2009).
44. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
45. M. van Heel, G. Harauz, E. V. Orlova, R. Schmidt, M. Schatz, A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
46. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
47. G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, S. J. Ludtke, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
48. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl, S. H. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
49. S. H. Scheres, S. Chen, Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
50. Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher, D. Lyumkis, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
51. A. Kucukelbir, F. J. Sigworth, H. D. Tagare, Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
52. R. Sanchez-Garcia, J. Gomez-Blanco, A. Cuervo, J. M. Carazo, C. O. S. Sorzano, J. Vargas, DeepEMhancer: A deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
53. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

54. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
55. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
56. T. I. Croll, ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).
57. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
58. R. T. Kidmose, J. Juhl, P. Nissen, T. Boesen, J. L. Karlsen, B. P. Pedersen, Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526–531 (2019).
59. P. V. Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham, E. Skrzypek, PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015).

Acknowledgments: We thank J. Pattie for computer support; J. Meyers, R. M. Haynes, and H. Scott at the PNCC for data collection support; and S. Fang for purification of XLF. We also acknowledge the use of the Ametek K3 direct electron detector, which was provided by R. A. Lamb. **Funding:** This work was supported by American Cancer Society (IRG-15-173-21), H Foundation Core Facility Pilot Project Award, National Institutes of Health grant R01GM135651, and National Institutes of Health grant R01GM144559 to Y.H.; Molecular Biophysics Training Program from NIGMS/NIH (5 T32 GM008382) to S.C. and A.V.; PNCC at OHSU: EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research; National Institutes of Health grant U24GM129547; Northwestern University Structural Biology Facility, NCI CCSG P30 CA060553 (Robert H. Lurie Comprehensive Cancer Center); National Institutes of Health grant P01CA092584 to S.P.L.-M., A.E.T., and Y.H.; and National Institutes of Health grant R01GM047251, and NCI CCSG P30 CA118100 (University of New Mexico Comprehensive Cancer Center) to A.E.T. **Author contributions:** Conceptualization: Y.H., S.P.L.-M., and A.E.T. Methodology: S.C. and Y.H. Investigation: S.C., A.V., R.M., L.L., and T.N. Visualization: S.C. and Y.H. Funding acquisition: A.E.T., S.P.L.-M., and Y.H. Supervision: Y.H. Writing—original draft: S.C., A.V., and Y.H. Writing—review and editing: S.C., A.V., L.L., T.N., R.M., A.E.T., S.P.L.-M., and Y.H. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMD-28732 (overall apo PAXX-LR complex), EMD-28735 (DNA-PK-N-BRCT in the apo PAXX-LR complex), EMD-28736 (LigIV-XRCC4-XLF-XRCC4-LigIV in the apo PAXX-LR complex), EMD-28733 (overall LR-ATP complex), EMD-28737 (DNA-PKcs without N terminus in LR-ATP complex), EMD-28738 (DNA-PK in the LR-ATP complex), EMD-28739 (LigIV-XRCC4-XLF-XRCC4-LigIV in the LR-ATP complex), EMD-28731 (overall DNA-PKcs dimeric complex), and EMD-28734 (DNA-PKcs in the DNA-PKcs dimeric complex). Model coordinates have been deposited in the PDB under accession numbers 8EZA (the apo PAXX-LR complex), 8EZB (the LR-ATP complex), and 8EZ9 (the DNA-PKcs dimeric complex). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 13 December 2022
Accepted 28 April 2023
Published 31 May 2023
10.1126/sciadv.adg2838