This is a repository copy of *Understanding the structure and role of DNA-PK in NHEJ: How X-ray diffraction and cryo-EM contribute in complementary ways*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/162233/

Version: Accepted Version

Understanding the structure and role of DNA-PK in NHEJ: How X-ray diffraction and cryo-EM contribute in complementary ways.

Qian Wu[1,2], Shikang Liang[1], Takashi Ochi[1,2], Dimitri Y Chirgadze[1], Juha T Huiskonen [3], Tom L Blundell[1*]

[1] Department of Biochemistry, University of Cambridge,
Tennis Court Road, Cambridge CB2 1GA
[2] Present Address: Astbury Centre for Structural Molecular Biology
School of Molecular and Cellular Biology, Faculty of Biological Sciences
University of Leeds, Leeds, LS2 9JT, UK
[3] Division of Structural Biology, University of Oxford,
Roosevelt Drive, Oxford OX3 7BN
[*] Corresponding Author

Abstract

DNA double-strand breaks (DSBs), generated by ionizing radiation, reactive oxygen species and DNA replication across nicks, are the most severe DNA damage in eukaryotic cells. Non-Homologous End Joining repairs DNA double-strand breaks directly without a template and so can take place at any point in the cell cycle. Ku70/80 heterodimers rapidly assemble around broken DNA ends, allowing DNA-PKcs, the catalytic subunit of DNA-dependent protein kinase, to be recruited and facilitating synapsis of broken DNA ends. This then provides a stage for end-processing and ligation. Here we review progress leading in 2017 to the medium resolution X-ray structure of DNA-PKcs, a single polypeptide chain of 4128 amino acids. This was followed quickly by chain tracing of cryo-EM structures of DNA-PKcs in complex with Ku and DNA. We discuss how combination of structural information from X-ray and cryo-EM studies can produce a working model for complex multicomponent molecular assemblies such as those found in DNA-double-strand-break repair.

Keywords

DNA-PK, DNA-PKcs, Ku70/80, NHEJ, DSBs, X-ray crystallography, cryo-EM, DNA repair

Abbreviations

AFM: atomic force microscopy

DSBs: DNA double-strand breaks

EM: electron microscopy

HEAT: Huntingtin, Elongation Factor 3, PP2 A, and TOR1

Ku80CTR: Ku80 C-terminal region

LigIV: DNA ligase IV

NHEJ: Non-Homologous End Joining

1. Introduction: DNA-PK in NHEJ

DNA double-strand breaks (DSBs) are the most severe DNA damage in eukaryotic cells. They are generated by ionizing radiation, reactive oxygen species and DNA replication across nicks. Physiological DSBs are also present in many cellular activities such as V(D)J and class-switch recombination for generating the diversity in immunoglobulin genes (Chaudhuri and Alt, 2004; Gellert, 2002). DSBs are repaired by two major pathways: Non-Homologous End Joining (NHEJ) and Homologous Recombination (HR). The phase in the cell cycle is a major factor in the choice of DSBs repair pathway. HR reaches its peak activity in mid-S phase and requires a sister chromatid as a template, whereas NHEJ repair of DSBs directly without a template can take place at any point in the cell cycle and is the dominant repair pathway in both G1 and G2 phases (Beucher et al., 2009; Her and Bunting, 2018; Karanam et al., 2012). Recent studies have shown that modulating the DSB repair pathway choice can also influence the outcome of precise genome editing by CRISPR-Cas9, which generates site-specific DSBs in the genome (Pawelczak et al., 2018). Inhibiting the NHEJ pathway was found to enhance the efficiency of HR repair and CRISPR-Cas9-induced precise genome editing (Chu et al., 2015).

Compared to HR, NHEJ functions as a faster and simpler repairing mode through rapid recruitment of various NHEJ proteins at the DNA ends. These proteins function in three general stages of NHEJ. During the first stage, synapsis, Ku70/80 (Ku) heterodimers recognize the DSBs and assemble around broken DNA ends (Walker et al., 2001). They recruit the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs), a phosphatidylinositol 3-kinase-related serine/threonine kinase, through interaction with Ku80 C-terminus (Gell and Jackson, 1999; Singleton et al., 1999) to form DNA-PK. Interactions between DNA-PKcs molecules likely contribute to the synapsis of broken DNA ends together (Spagnolo et al., 2006). This step is further stabilised by PAXX (Wang et all, 2018). The second, end-processing stage involves nucleases such as Artemis, which exhibits endonuclease activity after activation by DNA-PKcs. Finally,

in the third stage, ligation is mediated by DNA ligase IV (LigIV) in a stable complex with dimeric XRCC4 (Bryans et al., 1999; Grawunder et al., 1998). XLF, which interacts with XRCC4, further enhances the ligation step (Ahnesorg et al., 2006; Buck et al., 2006).

It has been evident for two decades that understanding how the components of NHEJ are coordinated in space and time during synapsis, end processing and ligation must be a major priority. Many research groups have addressed this by carrying out structural and functional studies in NHEJ as reviewed in Ochi *et al.,* 2014 (Ochi et al., 2014). DNA-PKcs, a single polypeptide chain of 4128 amino acids, interacts with Ku forming DNA-PK at DNA broken ends. X-ray analysis of the core ring structure of the heterodimeric Ku70/80 encircles the DNA providing a strong interaction with DNA ends (Walker et al., 2001) (**Figure 1A, B**). The C-terminal regions of both Ku70 and Ku80 contain a flexible long linker and a globular structural domain. The globular domain (residues 595-704) in Ku80 C-terminal region (Ku80CTR) (residues 543-732) comprises six helices including two "helical hairpin" structures, determined in solution by NMR methods (Harris et al., 2004; Zhang et al., 2004) (**Figure 1B**). The extreme 12 residues of the Ku80 C-terminus are crucial for mediating interaction with DNA-PKcs for recruitment of DNA-PKcs to the DNA damage site (Gell and Jackson, 1999) (**Figure 1A**).

In 2004 Oscar Llorca and Laurence Pearl reviewed electron microscopy (EM) studies on DNA recognition by DNA-PK, suggesting how electron microscopists could collaborate with X-ray crystallographers and NMR spectroscopists to study the structure and dynamics of the molecular assemblies and machines, such as DNA-PK, gaining insights into the structural basis of recognition of damaged DNA and the activation of DNA-PK kinase (Llorca and Pearl, 2004). Here we briefly review progress over two decades in understanding the structure of DNA-PK but focus on recent developments describing higher order complexes in NHEJ studied using X-ray and cryo-EM. The publication of chain tracing at medium resolution of the cryo-EM structures of the DNA-PKcs in complex with Ku and DNA in (Sharif et al., 2017; Yin et al., 2017), following the publication of the X-ray crystal structure of DNA-PK in 2017 (Sibanda et al., 2017), shows how quickly cryo-EM can build on and significantly extend our understanding of the structure and function of large molecules that

contribute to multicomponent complexes. By combining structural information from both X-ray and cryo-EM studies, we can produce a working model for the complex formation of DNA-PK with DNA.

2. Early EM studies of the structure of DNA-PK

Structural studies on DNA-PK were initiated using EM and atomic force microscopy (AFM) over 20 years ago (Cary et al., 1997; Yaneva et al., 1997). 2D electron crystallography, EM of negative-stain samples and cryo-EM have all subsequently contributed to the determination of the general architecture of DNA-PK.

In 1998 (Chiu et al., 1998) reported cryo-EM imaging of the catalytic subunit of the DNA-PKcs at 22 Å resolution. This was followed in 1999 by Leuther, Hammarsten, Kornberg, & Chu (K. Leuther et al., 1999), who produced a map defined by electron crystallography at a similar resolution. The structures showed an open channel "similar to those seen in other double-stranded DNA-binding proteins", and an "enclosed cavity with three openings large enough to accommodate single-stranded DNA". The authors demonstrated that efficient kinase activation required DNA longer than 12 bp, the minimal length accommodated in the open channel.

DeFazio et al. (2002), in an analysis of DNA-PKcs, showed that kinase activity is cooperative with respect to DNA concentration and demonstrated using electron microscopy that two DNA-PKcs molecules form complexes, possibly a mechanism for bringing together broken DNA ends (DeFazio et al., 2002). This was elaborated by Spagnolo, Rivera-Calzada, Pearl & Llorca (2006), who described the 3D-structure of the DNA-PK assembled on DNA and discussed its implications for DNA DSB repair. These studies showed that DNA-PKcs has a large circular structure with a head/crown structure including the kinase (Boskovic et al., 2003; Chiu et al., 1998; K. K. Leuther et al., 1999; Rivera-Calzada et al., 2007, 2005; Williams et al., 2008). DNA-PKcs interacts with DNA through the cavity in the circular structure with the Ku interaction site located close by (Boskovic et al., 2003; Rivera-Calzada et al., 2005; Spagnolo et al., 2006). Furthermore the Pearl and Llorca groups (Boskovic et al., 2003) demonstrated by electron microscopy that DNA-PKcs undergoes large conformational changes in the presence of double-stranded DNA that correlate with the activation of

the kinase. In 2005 this collaboration (Rivera-Calzada et al., 2005) used single-particle cryo-EM of DNA-PKcs to extend the resolution to 13 Å, suggesting how conformational changes in FAT and FATC domains affect activity of the kinase domain.

Better prediction methods for the HEAT (Huntingtin, Elongation Factor 3, PP2 A, and TOR1) repeats and structural homology modelling allowed docking of different parts of DNA-PKcs into the EM densities (Brewerton et al., 2004; Rivera-Calzada et al., 2005). However, due to the low resolution of these EM structures and difficulty of modelling unambiguously long regions of HEAT repeat structures with the correct curvature and amino acid registration, the chain trace remained unclear.

3. Understanding molecular structures and interactions using X-ray crystallography

In parallel over the first decade of the 21st century other groups, including ours, defined structures using X-ray crystallography of individual smaller components of NHEJ, including Ku (Walker et al., 2001), LigIV (Ochi et al., 2013, 2012), XLF (Andres et al., 2007; Li et al., 2008), XRCC4 (Junop et al., 2000) and PAXX (Ochi et al., 2015; Xing et al., 2015) as well as binary complexes of Lig IV/XRCC4, XLF/XRCC4, LigIV/Artemis (Ioanned et al., 2012; Ochi et al., 2012) and Ku/XLF & Ku/APLF (Nemoz et al., 2018).

X-ray analysis of the much larger NHEJ protein DNA-PKcs, complexed with the Ku80CTR, was reported first as an incomplete molecular trace, omitting the poorly defined N-terminal region, which at very low resolution was unclear (Sibanda et al., 2010). A more complete X-ray analysis at 4.3 Å resolution was published in 2017 (Chirgadze et al., 2017; Sibanda et al., 2017) (**Figure 1A,C**). It defined the structure of DNA-PKcs complex with part of the Ku80CTR, assigning ~90% of the 4128 amino acid residues of DNA-PKcs. The structural units of DNA-PKcs can be described as: 1) N-terminal arm comprising residues 1 to 892 (38 α-helices assembled as heat repeats in four supersecondary structures: N1-N4); 2) Circular cradle of residues 893-2801 (85 α-helices in five supersecondary structures, CC1-CC5, predominantly of HEAT repeats); 3) C-terminal Head comprising residues 2802-4128 (64 α-helices in FAT,

FRB, kinase and FATC domains). As the regular repeat structures have similar sequence patterns, the amino acid sequence register was checked by positions of the introduced Se-Met groups defined by anomalous scattering of selenium atoms but some ambiguities remain due to poor density.
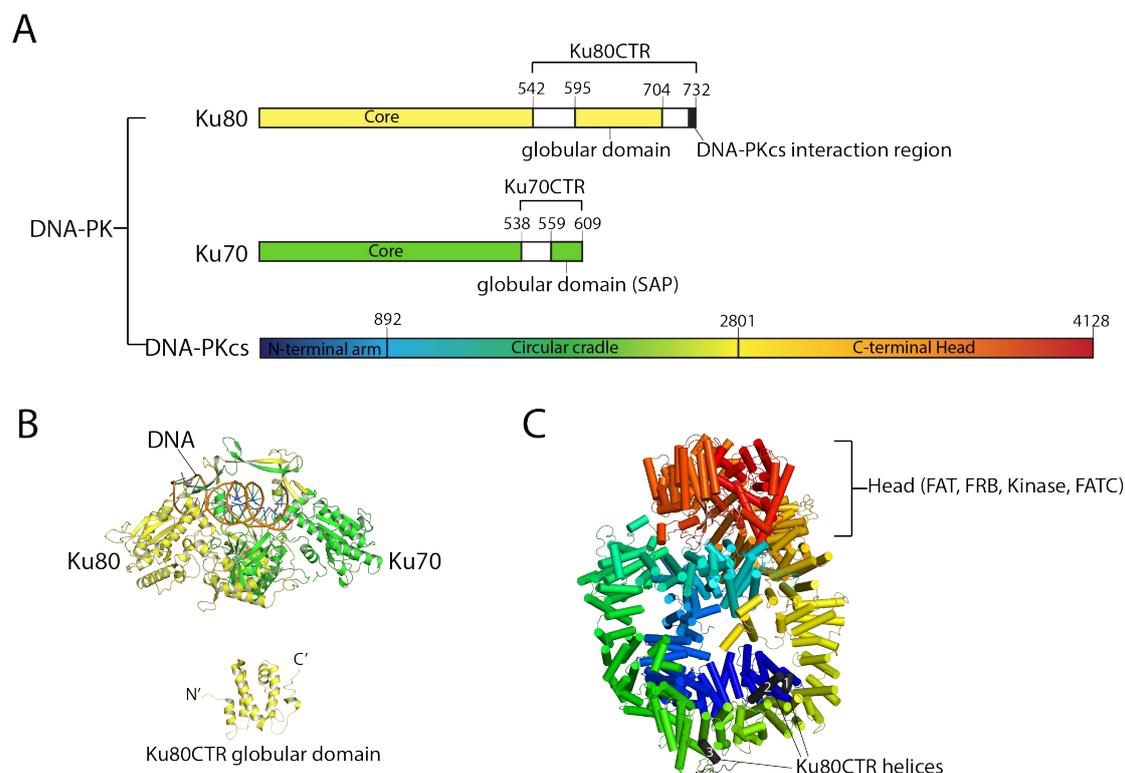
Figure 1. X-ray and NMR structures of Ku70/80 and DNA-PKcs. **A)** Domain organisation of Ku70/80 and DNA-PKcs. Ku80 is in yellow, Ku70 is in green and DNA-PKcs is in rainbow colour. Coloured regions indicate structures determined. The core is the region of Ku70 and Ku80 that forms the Ku heterodimer bound with DNA as shown in B. Ku80CTR and Ku70CTR represent the C-terminal regions of the two proteins. The C-terminal region interacting with DNA-PKcs is coloured black; **B)** Crystal structure of Ku70/80 bound with DNA (PDB code: 1JEY) (Walker et al., 2001). Lower panel is the NMR structure of the globular domain of Ku80CTR (PDB code: 1Q2Z) (Harris et al., 2004). **C)** Crystal structure of DNA-PKcs (PDB code: 5LUQ) (Sibanda et al., 2017) in rainbow colours corresponding to the colouring scheme and domain regions shown in A. Head includes FAT (FRAP, ATM, TRRAP), FRB (FKBP12-rapamycin-binding), kinase and FATC (FAT C-terminal) domains. Three helices from Ku80CTR are numbered and coloured in black.

4. The impact of the cryo-EM resolution revolution on the structure determination of DNA-PK

In recent years, single-particle cryo-EM studies of DNA-PKcs have benefited hugely from the development and improvement of EM instruments, detectors and image

processing software (Bai et al., 2015). These advances not only reflect the improved resolution of cryo-EM density maps, but also provide the possibility for the structure determination of large protein complexes. Two recent studies have determined the cryo-EM structures of the apo form of DNA-PKcs at 4.4 Å, DNA-PKcs in complex with Ku70/80 at 5.8 Å (Sharif et al., 2017) and DNA-PK holoenzyme bound with DNA at 6.6 Å resolution (Yin et al., 2017). The Blundell group has also determined the cryo-EM structure of the apo form of DNA-PKcs at 5.7 Å resolution (see Supplementary Data; EMDataBank (EMDB) ID code EMD-4425). While two apo cryo-EM structures of DNA-PKcs from (Sharif et al., 2017) and our study are very similar (**Figure 2A**), significant structural changes occur when DNA-PKcs is bound with different constructs of Ku70/80 and DNA. By combining the structural information obtained so far from X-ray crystallography and cryo-EM, we can compare the mode of the interaction between DNA-PKcs and Ku70/80 with and without DNA.

5. Ku interaction with DNA-PKcs

Although all structural studies of DNA-PK contain the same full-length DNA-PKcs, different constructs of Ku protein are used in the complexes defined by X-ray and cryo-EM studies. The X-ray structure of DNA-PK contains Ku80CTR (Sibanda et al., 2017), while two cryo-EM studies contain full length Ku70/80 (Sharif et al., 2017; Yin et al., 2017). The globular domain of Ku80 (residues 595 to 704) was not observed in the X-ray structure. The remaining Ku80 residues including the extreme 12 C-terminal residues of Ku80CTD were observed in the X-ray map as three helices (**Figure 1C**). One helix of Ku (helix 3) was identified as that of the C-terminus based on selenomethionine labelling (Sibanda et al., 2017). These three helices are absent in the apo-structures of DNA-PKcs defined by cryo-EM (**Figure 2B**). At the same time, a region built as polyalanine for residues 2577-2773 in the crystal structure of DNA-PKcs with Ku80CTD is not observed in the apo cryo-EM structures (**Figure 2B**). This could be due to the flexibility of this region, which is under more constraints in the crystal structure.
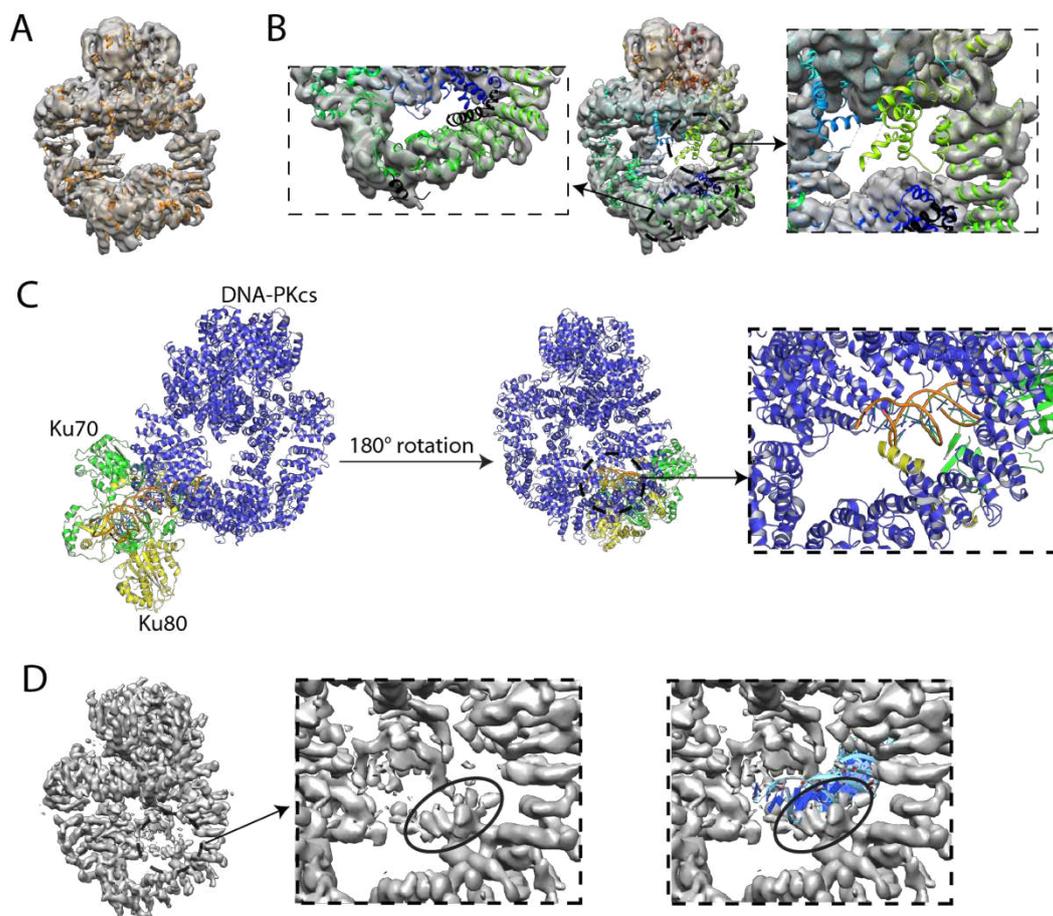
Figure 2. EM structures of DNA-PK. **A)** Structure superimposition of the EM density (grey colour) of DNA-PKcs from this study EMDataBank (EMDB ID code EMD-4425, see Supplementary Data) with EM structure model (orange colour) of DNA-PKcs (PDB code: 5W1R) from (Sharif et al., 2017) showing similarity. **B)** Crystal structure of DNA-PKcs (rainbow colour, PDB code: 5LUQ) superimposed with EM density (grey colour, from this study). Left panel: zoomed-in view shows that three helices from Ku80CTR (black colour) are not present in the apo form of DNA-PKcs. Right panel: zoomed-in view shows the region does not contain the density in the EM structure but is shown in the crystal structure of DNA-PKcs – see text; **C)** Structure of DNA-PK including DNA (PDB code: 5Y3R) (Yin et al., 2017) in two different views. DNA-PKcs is in blue, Ku70 is in yellow, Ku80 is in green and DNA is in orange. **D)** EM map of DNA-PK in (Sharif et al., 2017). The extra density of the globular domain of Ku80CTR is inside the black circle (in both middle and right panels). The EM structure model of DNA-PK with DNA (PDB code: 5Y3R) in (Yin et al., 2017) was docked into the EM density of DNA-PK in Sharif et al. (Sharif et al., 2017). The position of extra density overlaps with the DNA (blue colour in right-hand side).

In the EM structure of DNA-PK without DNA (Sharif et al., 2017), extra density between the N-terminus of DNA-PKcs and the circular cradle HEAT-repeat structure was observed in some classes of cryo-EM maps and this density was assigned as the globular domain of Ku80CTR on the basis of its size (**Figure 2D**). This implies that the

interaction between DNA-PKcs and the globular domain of Ku80CTR is weak and its occupancy at this site is low. This might be why we did not observe the corresponding electron density in our crystal structure of DNA-PKcs/Ku80CTR. Interestingly, in the cryo-EM structure of DNA-PK bound with DNA (Yin et al., 2017), this globular domain was not observed in any classes. Instead, DNA occupied the same binding space. These observations imply that the globular domain of Ku80CTR might transiently interact with DNA-PKcs and regulate the DNA binding of DNA-PKcs through the same region. One helix of Ku80 C-terminal region (described as helix 2 in the crystal structure) was observed close to the DNA end in the cryo-EM structure in (Yin et al., 2017). The authors predicted that the function of this helix may be to prevent inward sliding of DNA-PKcs along the DNA and also offer protection of the DNA end. This structure also demonstrates that when both DNA-PKcs and Ku bind to the DNA to form the holoenyzme, there are extra interactions between Ku70/80 heterodimer core structure with the N-terminal and middle HEAT repeats regions of DNA-PKcs. Thus, the interaction between DNA-PKcs and Ku70/80 is not only a simple C-terminal region single-site interaction but also involves multiple interactions sites among these proteins with DNA.

The functions of the globular domains of both Ku70CTR and Ku80CTR are still not very clear. A computational study has indicated that the SAP domain can form DNA-bound complexes with relatively high affinities (Hu et al., 2012). As the location of the SAP domain changes after DNA binds to Ku, it has been suggested that the SAP domain of Ku70 regulates the DNA loading to Ku (Walker et al., 2001), although there is still no structural evidence for this. For Ku80CTR, the globular domain contained exposed and conserved residues, which could potentially mediate protein-protein interaction (Zhang et al., 2004; Harris et al., 2004). One potential interaction partner is DNA-PKcs as observed in the EM structure of DNA-PK without DNA (Sharif et al., 2017). Taking all the structural information together, a likely series of events for DNA-PK binding to DNA is as follows. First, the Ku70/80 heterodimer binds to DNA ends through Both Ku70 and Ku80 core domains of the Ku-heterodimer in the presence of DNA DSBs. Ku80CTR (including the globular domain and the DNA-PKcs interaction region) then recruits DNA-PKcs to DSB sites. While a stable anchoring of DNA-PKcs occurs through binding the C-terminal 12 residues of Ku, the C-terminal globular domain of Ku80 mediates further dynamic and weak protein-protein interactions with DNA-PKcs

N-terminal HEAT repeats and circular cradle region to bring Ku and DNA-PKcs closer together. Once DNA-PKcs is recruited by Ku70/80 and is close to the DNA end and while the Ku80 C-terminal helices are still bound with DNA-PKcs, the position once occupied by the Ku80CTR globular domain is rapidly replaced by stronger interaction from DNA, which has already bound with the core structure of Ku70/80 (**Figure 3B**). In this working model, the globular domain of Ku80CTR functions to regulate the Ku loading to DNA-PKcs. The N-terminal region of DNA-PKcs thus functions as a gate opening to mediate and accommodate the interaction with Ku and DNA. The interaction between the core structure of Ku70/80 with DNA-PKcs then further stabilizes the formation of DNA-PK holoenzyme at the DNA end.
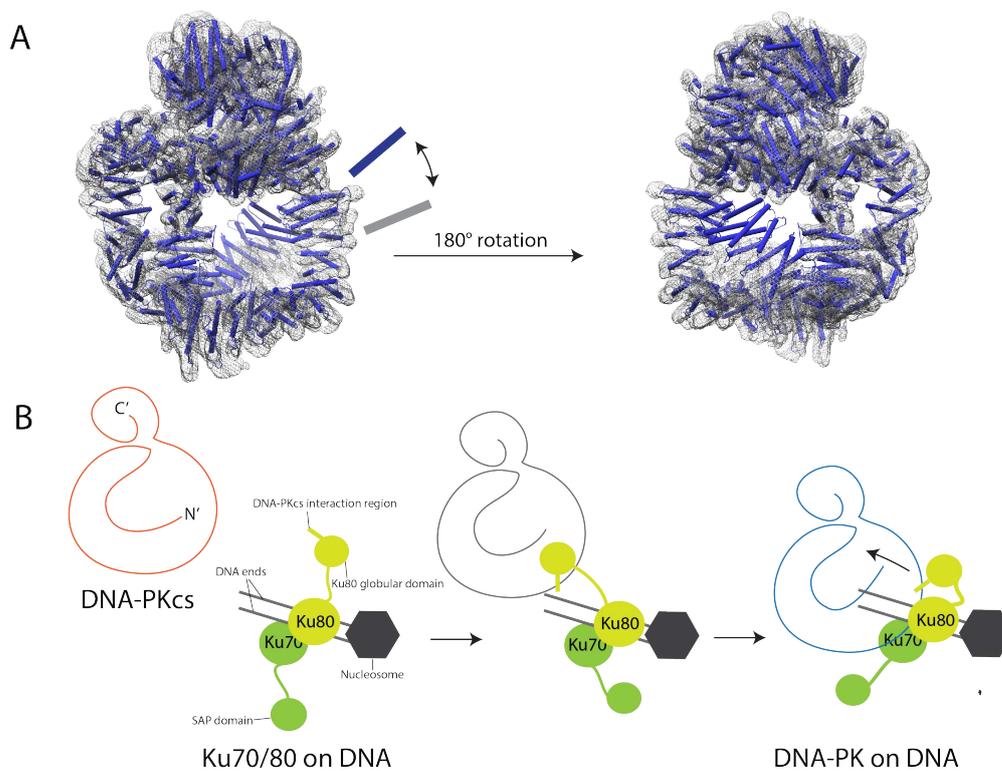
Figure 3: Structural differences of the DNA-PK N-terminal region correlated with interactions between Ku and DNA-PKcs. **A)** The position of the N-terminal region differs between that seen in the apo DNA-PKcs EM density (from this study, grey mesh) and that of the DNA-PKcs bound with Ku70/80 and DNA (only DNA-PKcs is shown in blue colour for clearer illustration) (PDB code: 5Y3R) (Yin et al., 2017). **B)** Cartoon of the mode of interaction between DNA-PKcs with Ku at the DNA ends. Ku80 is in yellow colour, while Ku70 is in green. Big yellow/green circles represent the core structures in Ku70/80 heterodimer. Smaller circles represent the globular domain of the C-terminal region (CTR). The DNA-PKcs interaction region in the C-terminal of Ku80 is labelled. The apo form of DNA-PKcs is shown in orange,

and the DNA-PKcs structure interacting with the globular domain of Ku80CTR is in grey. The DNA-PKcs structure on the DNA is in blue. Two sided black arrows indicate the movement of the N-terminal domain of DNA-PKcs.

6. The allosteric activation of the kinase domain of DNA-PKcs

In order to understand the allosteric activation of the kinase domain of DNA-PKcs, careful observation of the conformational changes of DNA-PKcs is critical. So far, four atomic models of DNA-PKcs have been built: two from the asymmetric unit of the crystal structure (Sibanda et al., 2017), one from the cryo-EM structure of apo DNA-PKcs (Sharif et al., 2017), and one from the cryo-EM structure of DNA-PK holoenzyme (Yin et al., 2017).

Conformational changes in the two models of the crystal structure (top panel of **Figure 4**) show that the whole molecule of DNA-PKcs is twisting, involving movement of the N terminal region, consistent with the proposal that the interaction between E3933 and K881, next to the highly conserved region I (HCRI) of 979 to 1121, is important for the activation of the activation of kinase activity as it opens up the active site (Sibanda et al., 2017). This is of course a transition between two states, presumably influenced by the crystal packing, but nevertheless showing a correlation between changes in the N-terminal region and those of the kinase active site.

The two cryo-EM structures, corresponding to apo DNA-PKcs and DNA-PK holoenzyme, show two well-defined functional states of DNA-PKcs, namely inactive and active. Conformational changes (bottom panel of **Figure 4**) between these are more relevant to understanding the activation of the kinase. Comparison of the structures show that DNA-PKcs becomes more compacted when activated by DNA-Ku70/80 complex and there is a significant uplift of the N-terminus (10-382) compared to the apo-structure, including HEAT1-HEAT8 repeats. The interaction in DNA-PKcs between N terminus and the FAT domain has been proposed to be the main driving factor of kinase activation (Yin et al., 2017), but other changes in the overall conformation undoubtedly affect the conformation and activation of the kinase.
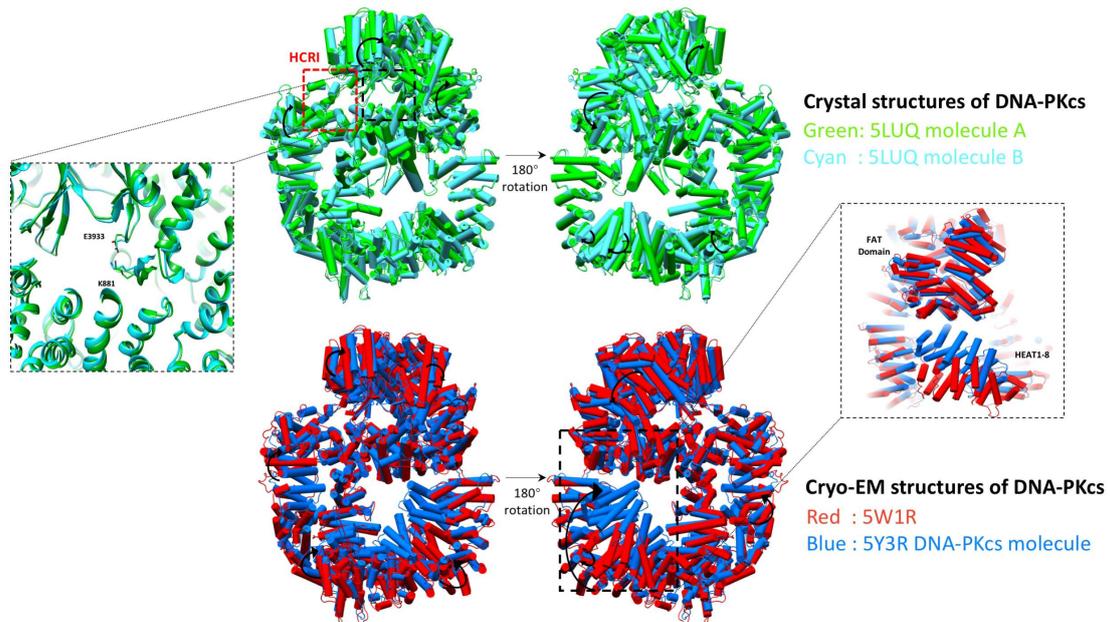
Figure 4: Allosteric activation of DNA-PKcs. Four DNA-PKcs secondary structural models are depicted. In the top panel, two apo DNA-PKcs molecules (in green and cyan) from crystal structure (PDB code: 5LUQ) are superposed and show a relatively small conformational difference (Sibanda et al., 2017). The Highly Conserved Region I (HCRI), involving interaction between E3933 and K881, is shown in the amplified window on the left. In the bottom panel the cryo-EM structure of the apo DNA-PKcs (PDB code: 5W1R) (Sharif et al., 2017) is shown in red, superposed on the structure of DNA-PKcs in blue from the cryo-EM structure DNA-PK (PDB code: 5Y3R) (Yin et al., 2017) in which DNA-PKcs is bound with Ku70/80 and DNA. The uplift of the N terminus leads to interaction between HEAT 1-8 and the FAT domain is shown in the right amplified window. For clarity the Ku C-terminal helices are omitted from (Sibanda et al., 2017) and the Ku and DNA molecules are omitted from (Yin et al., 2017) structures. The arrows indicate the conformational changes between the compared models.

7. Future structural studies of DNA-PK and NHEJ

Further high-resolution structures of different states of DNA-PK will be required to understand the interaction mode of DNA-PKcs with Ku70/80, and the detailed conformational changes that modulate the DNA-binding and allosteric kinase active site. The DNA-PKcs structure bound with a small-molecule kinase inhibitor would be very useful for study of the structural changes in the kinase domain. It may be beneficial for future structural work of DNA-PKcs to produce the protein recombinantly in mammalian cells (Yin et al., 2017) as opposed to purifying the protein endogenously expressed in HeLa cells. This will enable mutagenesis of DNA-PKcs protein to test hypotheses about kinase function and interactions with other NHEJ proteins and DNA.

Research over the past two decades on DNA DSB repair through NHEJ, particularly that on DNA-PKcs, demonstrates the complementarity of X-ray and cryo-EM

approaches. The future will no doubt bring new advances in cryo-EM that will assist study of larger multicomponent assemblies and will also likely become increasingly powerful with smaller proteins. Furthermore a recent study of NHEJ using a single-molecule approach has indicated a requirement for all core NHEJ proteins for stable interaction (Wang et al., 2018). Over the short term, with many structures of individual proteins and small protein assemblies already available from X-ray studies, cryo-EM can take advantage of this knowledge to build up a picture of individual steps in NHEJ and possibly even explore the dynamics of this process. In conclusion, cryo-EM study of DNA-PK is just the beginning of the new era of structural biology for understanding the large multi-component assembly of NHEJ at the DNA damage sites and the structural relationship with upstream DNA damage response proteins and chromatin environment.

References

Ahnesorg, P., Smith, P., Jackson, S.P., 2006. XLF Interacts with the XRCC4-DNA Ligase IV Complex to Promote DNA Nonhomologous End-Joining. Cell 124, 301–313. https://doi.org/10.1016/j.cell.2005.12.031

Andres, S.N., Modesti, M., Tsai, C.J., Chu, G., Junop, M.S., 2007. Crystal Structure of Human XLF: A Twist in Nonhomologous DNA End-Joining. Mol. Cell 28, 1093–1101. https://doi.org/10.1016/j.molcel.2007.10.024

Bai, X., McMullan, G., Scheres, S.H.W., 2015. How cryo-EM is revolutionizing structural biology. Trends Biochem. Sci. 40, 49–57. https://doi.org/https://doi.org/10.1016/j.tibs.2014.10.005

Beucher, A., Birraux, J., Tchouandong, L., Barton, O., Shibata, A., Conrad, S., Goodarzi, A.A., Krempler, A., Jeggo, P.A., Löbrich, M., 2009. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. EMBO J. 28, 3413–3427. https://doi.org/10.1038/emboj.2009.276

Boskovic, J., Rivera-Calzada, A., Maman, J.D., Chacón, P., Willison, K.R., Pearl, L.H., Llorca, O., 2003. Visualization of DNA-induced conformational changes

in the DNA repair kinase DNA-PKcs. EMBO J. 22, 5875–5882. https://doi.org/10.1093/emboj/cdg555

Brewerton, S.C., Doré, A.S., Drake, A.C., Leuther, K.K., Blundell, T.L., 2004. Structural analysis of DNA-PKcs: modelling of the repeat units and insights into the detailed molecular architecture. J. Struct. Biol. 145, 295–306. https://doi.org/10.1016/j.jsb.2003.11.024

Bryans, M., Valenzano, M.C., Stamato, T.D., 1999. Absence of DNA ligase IV protein in XR-1 cells: evidence for stabilization by XRCC4. Mutat. Res. Repair 433, 53–58. https://doi.org/10.1016/S0921-8777(98)00063-9

Buck, D., Malivert, L., de Chasseval, R., Barraud, A., Fondanèche, M.-C., Sanal, O., Plebani, A., Stéphan, J.-L., Hufnagel, M., le Deist, F., Fischer, A., Durandy, A., de Villartay, J.-P., Revy, P., 2006. Cernunnos, a Novel Nonhomologous End-Joining Factor, Is Mutated in Human Immunodeficiency with Microcephaly. Cell 124, 287–299. https://doi.org/10.1016/j.cell.2005.12.030

Cary, R.B., Peterson, S.R., Wang, J., Bear, D.G., Bradbury, E.M., Chen, D.J., 1997. DNA looping by Ku and the DNA-dependent protein kinase. Proc. Natl. Acad. Sci. 94, 4267–4272.

Chaudhuri, J., Alt, F.W., 2004. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat. Rev. Immunol. 4, 541–552. https://doi.org/10.1038/nri1395

Chirgadze, D.Y., Ascher, D.B., Blundell, T.L., Sibanda, B.L., 2017. Chapter Seven - DNA-PKcs, Allostery, and DNA Double-Strand Break Repair: Defining the Structure and Setting the Stage, in: Eichman, B.F.B.T.-M. in E. (Ed.), DNA Repair Enzymes: Structure, Biophysics, and Mechanism. Academic Press, pp. 145–157. https://doi.org/https://doi.org/10.1016/bs.mie.2017.04.001

Chiu, C.Y., Cary, R.B., Chen, D.J., Peterson, S.R., Stewart, P.L., 1998. Cryo-EM imaging of the catalytic subunit of the DNA-dependent protein kinase. J. Mol. Biol. 284, 1075–81. https://doi.org/10.1006/jmbi.1998.2212

Chu, V., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., Kuhn, R., 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat. Biotechnol. 33, 543–548. https://doi.org/doi: 10.1038/nbt.3198

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

https://doi.org/10.1093/emboj/cdf299

Gell, D., Jackson, S.P., 1999. Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. Nucleic Acids Res. 27, 3494–3502. https://doi.org/10.1093/nar/27.17.3494

Gellert, M., 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. Annu. Rev. Biochem. 71, 101–132. https://doi.org/10.1146/annurev.biochem.71.090501.150203

Grawunder, U., Zimmer, D., Lieber, M.R., 1998. DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains. Curr. Biol. 8, 873–879. https://doi.org/10.1016/S0960-9822(07)00349-1

Harris, R., Esposito, D., Sankar, A., Maman, J.D., Hinks, J.A., Pearl, L.H., Driscoll, P.C., 2004. The 3D solution structure of the C-terminal region of Ku86 (Ku86CTR). J. Mol. Biol. 335, 573–582.

Her, J., Bunting, S.F., 2018. How cells ensure correct repair of DNA double-strand breaks. J. Biol. Chem. 293, 10502–10511. https://doi.org/10.1074/jbc.TM118.000371

Hu, S., Pluth, J.M., Cucinotta, F.A., 2012. Putative binding modes of Ku70-SAP domain with double strand DNA: a molecular modeling study. J. Mol. Model. 18, 2163–2174. https://doi.org/10.1007/s00894-011-1234-x

Junop, M., Modesti, M., Guarné, A., Ghirlando, R., Gellert, M., Yang, W., 2000. Crystal structure of the Xrcc4 DNA repair protein and implications for end joining. EMBO J. 19, 5962–5970. https://doi.org/10.1093/emboj/19.22.5962

Karanam, K., Kafri, R., Loewer, A., Lahav, G., 2012. Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. Mol. Cell 47, 320–329. https://doi.org/10.1016/j.molcel.2012.05.052

Leuther, K., Hammarsten, O., Kornberg, R., Chu, G., 1999. Structure of DNA-dependent protein kinase: implications for its regulation by DNA. EMBO J. 18, 1114–1123.

Leuther, K.K., Hammarsten, O., Kornberg, R.D., Chu, G., 1999. Structure of DNA-dependent protein kinase: implications for its regulation by DNA. EMBO J. 18, 1114–1123. https://doi.org/10.1093/emboj/18.5.1114

Li, Y., Chirgadze, D.Y., Bolanos-Garcia, V.M., Sibanda, B.L., Davies, O.R., Ahnesorg, P., Jackson, S.P., Blundell, T.L., 2008. Crystal structure of human

XLF/Cernunnos reveals unexpected differences from XRCC4 with implications
for NHEJ. EMBO J. 27, 290–300. https://doi.org/10.1038/sj.emboj.7601942

Llorca, O., Pearl, L.H., 2004. Electron microscopy studies on DNA recognition by
DNA-PK. Micron 35, 625–633.
https://doi.org/https://doi.org/10.1016/j.micron.2004.05.004

Ochi, T., Blackford, A., Coates, J., Jhujh, S., Mehmood, S., Tamura, N., Travers, J.,
Wu, Q., Draviam, V., Robinson, C., Blundell, T., Jackson, S., 2015. DNA repair.
PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA
double-strand break repair. Science 347, 185–188. https://doi.org/doi:
10.1126/science.1261971

Ochi, T., Gu, X., Blundell, T., 2013. Structure of the Catalytic Region of DNA Ligase
IV in Complex with an Artemis Fragment Sheds Light on Double-Strand Break
Repair. Structure 21, 672–679. https://doi.org/doi: 10.1016/j.str.2013.02.014

Ochi, T., Wu, Q., Blundell, T.L., 2014. The spatial organization of non-homologous
end joining: From bridging to end joining. DNA Repair (Amst).
https://doi.org/10.1016/j.dnarep.2014.02.010

Ochi, T., Wu, Q., Chirgadze, D., Grossmann, G., Bolanos-Garcia, V., Blundell, T.,
2012. Structural insights into the role of domain flexibility in human DNA ligase
IV. Structure 20, 1212–1222. https://doi.org/doi: 10.1016/j.str.2012.04.012

Pawelczak, K.S., Gavande, N.S., VanderVere-Carozza, P.S., Turchi, J.J., 2018.
Modulating DNA Repair Pathways to Improve Precision Genome Engineering.
ACS Chem. Biol. 13, 389–396. https://doi.org/10.1021/acschembio.7b00777

Rivera-Calzada, A., Maman, J.P., Spagnolo, L., Pearl, L.H., Llorca, O., 2005. Three-
Dimensional Structure and Regulation of the DNA-Dependent Protein Kinase
Catalytic Subunit (DNA-PKcs). Structure 13, 243–255.
https://doi.org/10.1016/j.str.2004.12.006

Rivera-Calzada, A., Spagnolo, L., Pearl, L.H., Llorca, O., 2007. Structural model of
full-length human Ku70-Ku80 heterodimer and its recognition of DNA and
DNA-PKcs. EMBO Rep. 8, 56–62. https://doi.org/10.1038/sj.embor.7400847

Sharif, H., Li, Y., Dong, Y., Dong, L., Wang, W., Mao, Y., Wu, H., 2017. Cryo-EM
structure of the DNA-PK holoenzyme. Proc. Natl. Acad. Sci. 114, 7367–7372.
https://doi.org/doi: 10.1073/pnas.1707386114

Sibanda, B., Chirgadze, D., Ascher, D., Blundell, T., 2017. DNA-PKcs structure
suggests an allosteric mechanism modulating DNA double-strand break repair.

Science (80-.). 355, 520–524. https://doi.org/doi: 10.1126/science.aak9654

Sibanda, B., Chirgadze, D., Blundell, T., 2010. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. Nature 463, 118–121.

Singleton, B., Torres-Arzayus, M., Rottinghaus, S., Taccioli, G., Jeggo, P., 1999. The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit. Mol. Cell. Biol. 19, 3267–3277.

Spagnolo, L., Rivera-Calzada, A., Pearl, L.H., Llorca, O., 2006. 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. https://doi.org/10.1016/j.molcel.2006.04.013

Walker, J.R., Corpina, R.A., Goldberg, J., 2001. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412, 607–614. https://doi.org/10.1038/35088000

Wang, J.L., Duboc, C., Wu, Q., Ochi, T., Liang, S., Tsutakawa, S.E., Lees-Miller, S.P., Nadal, M., Tainer, J.A., Blundell, T.L., Strick, T.R., 2018. Dissection of DNA double-strand-break repair using novel single-molecule forceps. Nat. Struct. Mol. Biol. 25, 482–487. https://doi.org/10.1038/s41594-018-0065-1

Williams, D.R., Lee, K.-J.J., Shi, J., Chen, D.J., Stewart, P.L., 2008. Cryo-EM structure of the DNA-dependent protein kinase catalytic subunit at subnanometer resolution reveals alpha helices and insight into DNA binding. Structure 16, 468–477. https://doi.org/10.1016/j.str.2007.12.014

Xing, M., Yang, M., Huo, W., Feng, F., Wei, L., Jiang, W., Ning, S., Yan, Z., Li, W., Wang, Q., Hou, M., Dong, C., Guo, R., Gao, G., Ji, J., Zha, S., Lan, L., Liang, H., Xu, D., 2015. Interactome analysis identifies a new paralogue of XRCC4 in non-homologous end joining DNA repair pathway. Nat. Commun. 6, 6233. https://doi.org/10.1038/ncomms7233

Yaneva, M., Kowalewski, T., Lieber, M.R., 1997. Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. EMBO J. 16, 5098–5112. https://doi.org/10.1093/emboj/16.16.5098

Yin, X., Liu, M., Tian, Y., Wang, J., Xu, Y., 2017. Cryo-EM structure of human DNA-PK holoenzyme. Cell Res. 27, 1341–1350.

Zhang, Z., Hu, W., Cano, L., Lee, T.D., Chen, D.J., Chen, Y., 2004. Solution structure

of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. Structure 12, 495–502. https://doi.org/10.1016/j.str.2004.02.007

DNA-PKcs Review Supplementary Material

The apo form of DNA-PKcs at 5.7 Å resolution (EMDataBank ID code EMD-4425)

Purification of DNA-PKcs

DNA-PKcs was purified from Hela nuclear extract as described previously (Sibanda et al., 2010). The final protein buffer of DNA-PKcs is 20 mM Hepes pH 7.6, 200 mM NaCl, 0.5 mM EDTA, 2 mM $MgCl_2$, 5 mM DTT. Cryo-EM grids of DNA-PKcs were prepared using a FEI Vitrobot Mark IV operated at 4°C and 100% relative humidity. Aliquots of 3 μl purified DNA-PKcs sample at a concentration of 0.2 mg/ml were applied to glow-discharged holey carbon grid (Quantifoil Cu R1.2/1.3, 300 mesh), blotted with filter paper for 2.5 s to remove excess sample and then plunge-frozen in liquid ethane.

Cryo-EM structure of DNA-PKcs

Cryo-EM data for DNA-PKcs were collected on a FEI Titan Krios electron microscope operated at 300 kV and movie image stacks were recorded on FEI Falcon 2 detector (integration mode). Each stack contained 40 movie frames and collected over 3 s using an exposure rate of 50 electrons per pixel per second at a calibrated magnification of 95,000× (yielding a pixel size of 1.43 Å at the sample level). The final data set is composed of 1,557 image stacks with defocus values ranging from 2.0 to 3.8 μm.

The beam-induced motion correction was performed using the MotionCor2 (Zheng et al., 2017). The contrast transfer function parameters were estimated by CTFFIND3 (Mindell and Grigorieff, 2003). The remaining image processing was done using RELION 2.0 (Scheres, 2012). The flow-chart of data processing for DNA-PKcs is shown in **Figure S1A**. 1557 micrographs were collected (**Figure S1B**) and 1683 particles were manually picked for initial 2D classification. 371484 particles were later auto-picked for 2D classification (**Figure S1C**). There were 289445 particles selected for 3D classification and there were two classes. Finally, class 1 with 129117 particles were used in 'gold standard' 3D refinement to yield the final map. The resolution of the map was estimated by Fourier shell correlation (0.143 threshold) to be 5.7 Å (**Figure S1D**) and estimated local resolution is shown in **Figure S1E**. The electron density map was deposited in the EMDataBank with ID code: EMD-4425.
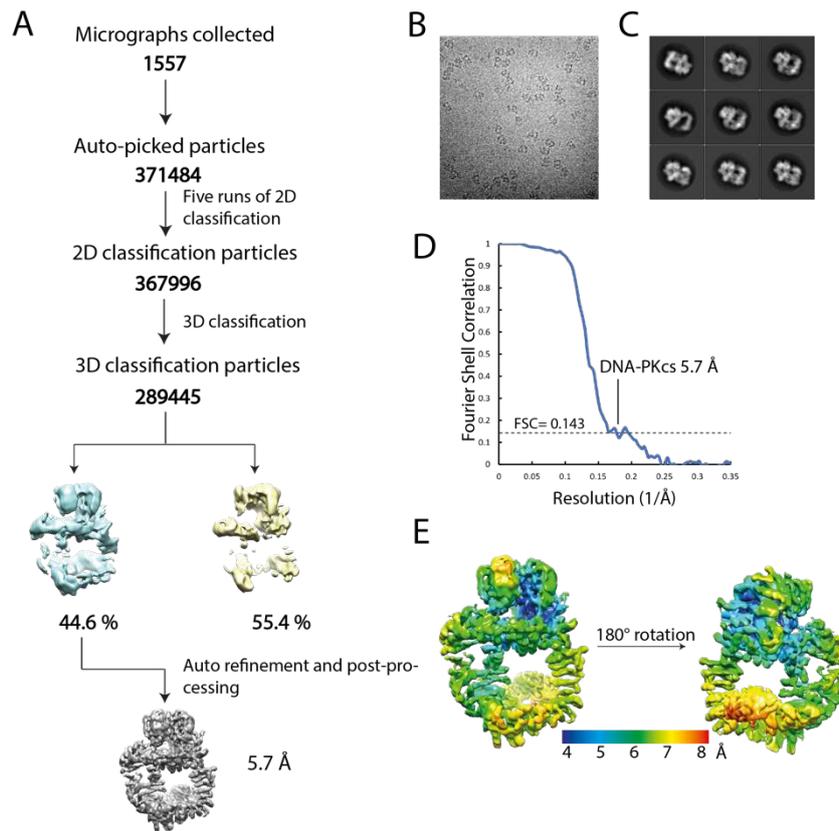
Figure S1: The determination of the cryo-EM structure of DNA-PKcs. **A)** Flow-chart of data processing. **B)** Representative images of cryo-EM micrography of DNA-PKcs. **C)** Representative images of 2D class averages for DNA-PKcs. **D)** FSC curve for the final refined DNA-PKcs structure. **E)** Overall resolution distribution of DNA-PKcs cryo-EM map. The N-terminal region of DNA-PKcs is the most flexible region (red) while the kinase domain contains the highest resolution (blue).

References

Mindell, J., Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. 142, 334–347.

Scheres, S., 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530.

Sibanda, B., Chirgadze, D., Blundell, T., 2010. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. Nature 463, 118–121.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., Agard, D.A., 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-

electron microscopy. Nat. Methods 14, 331–332.

https://doi.org/10.1038/nmeth.4193