



Structural insights into the role of DNA-PK as a master regulator in NHEJ

Siyu Chen^{1,2} · James P. Lees-Miller³ · Yuan He^{1,2} · Susan P. Lees-Miller^{3,4}

Received: 27 May 2021 / Revised: 6 July 2021 / Accepted: 12 July 2021 / Published online: 23 July 2021
© The Author(s) 2021

Abstract

DNA-dependent protein kinase catalytic subunit DNA-PKcs/*PRKDC* is the largest serine/threonine protein kinase of the phosphatidylinositol 3-kinase-like protein kinase (PIKK) family and is the most highly expressed PIKK in human cells. With its DNA-binding partner Ku70/80, DNA-PKcs is required for regulated and efficient repair of ionizing radiation-induced DNA double-strand breaks via the non-homologous end joining (NHEJ) pathway. Loss of DNA-PKcs or other NHEJ factors leads to radiation sensitivity and unrepaired DNA double-strand breaks (DSBs), as well as defects in V(D)J recombination and immune defects. In this review, we highlight the contributions of the late Dr. Carl W. Anderson to the discovery and early characterization of DNA-PK. We furthermore build upon his foundational work to provide recent insights into the structure of NHEJ synaptic complexes, an evolutionarily conserved and functionally important YRPD motif, and the role of DNA-PKcs and its phosphorylation in NHEJ. The combined results identify DNA-PKcs as a master regulator that is activated by its detection of two double-strand DNA ends for a cascade of phosphorylation events that provide specificity and efficiency in assembling the synaptic complex for NHEJ.

Keywords DNA-PKcs · Non-homologous end joining · Cryo-EM · Structural biology · Phosphorylation

Introduction

The phosphatidylinositol 3-kinase-like protein kinases (PIKKs) are a family of serine/threonine protein kinases with similarity in their kinase domains to the lipid kinase *PIK3CA*, the catalytic subunit of phosphatidylinositol 3-kinase, PI3K. However, the PIKKs are serine/threonine protein kinases and three of them, DNA-dependent protein kinase catalytic subunit (DNA-PKcs, gene name *PRKDC*), Ataxia telangiectasia mutated (ATM) and ATM and Rad3

related (ATR), play important roles in the cellular response to DNA damage (Blackford & Jackson, 2017). In this review, we note and pay tribute to the scientist who first discovered the DNA-activated protein kinase, Dr. Carl W. Anderson, who sadly passed away October 20, 2020, aged 76. We will then place Anderson's discoveries into the context of breakthrough advances in understanding the structure and function of the largest member of the PIKK family, DNA-PKcs.

Carl spent a large part of his career at the Biology Department at Brookhaven National Laboratory, Upton, NY, where he was Chair from 1999 to his retirement in 2011. In 1985, Carl carried out a sabbatical in the lab of Dr. Tim Hunt at Cambridge University, UK where he planned to work on the RNA-dependent protein kinase. However, the RNA they purchased for their experiments was contaminated with double-stranded (ds)DNA, leading to the serendipitous discovery that human cells contain a DNA-activated protein kinase that phosphorylates the 90-kDa heat shock protein (Hsp90) and other proteins in HeLa cell extracts (Walker et al., 1985). One of the authors of this review, S. P. Lees-Miller, was fortunate to join Carl's lab as a post-doctoral fellow to attempt to purify and characterize this putative DNA-activated protein kinase. In his lab, we showed that the DNA-activated

✉ Susan P. Lees-Miller
leesmill@ucalgary.ca

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, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

⁴ Robson DNA Science Centre, Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

protein kinase phosphorylates the α isoform of Hsp90 on two N-terminal threonine residues that are each followed by a glutamine, in the sequence PEETQTQDQPME (Lees-Miller & Anderson, 1989). Thus, Hsp90- α was the first PIKK substrate shown to be phosphorylated on an SQ/TQ motif, now recognized as a characteristic signature of many PIKK substrates (O'Neill et al., 2000; Schlam-Babayov et al., 2021).

Using Hsp90 as our substrate, we attempted to purify the DNA-activated protein kinase and showed that its activity was associated with a large polypeptide we called p350 (now known as DNA-PKcs) and the dsDNA binding Ku70/80 heterodimer (Lees-Miller et al., 1990). p350 was independently identified as a DNA-activated protein kinase by Dr. Tim Carter of St. John's University, Jamaica, NY (Carter et al., 1990), and the groups of Drs. Steve Jackson (Gottlieb & Jackson, 1993; Jackson et al., 1990) and Bill Dynan (Dvir et al., 1992, 1993). As part of these efforts, Carl made important contributions to the cloning of the DNA-PKcs cDNA (Hartley et al., 1995), as well as its genomic location on 8q11 (Sipley et al., 1995) and reported that DNA-PKcs shares a bi-directional promoter with replication factor MCM2 (Connelly et al., 1998). Carl also had a long-standing interest in p53 (Nguyen et al., 2014, 2018) and showed that DNA-PK phosphorylates p53 on serine 15 *in vitro*, another SQ site, establishing the first quantitative assay for DNA-PK kinase activity (Lees-Miller et al., 1992).

In sum, Carl made many important contributions to the discovery and function of DNA-PK and the regulation of p53. He was a kind and generous man. He was a great mentor. He will be sorely missed by his many friends and colleagues. Fortunately, his many contributions to science and in particular to our understanding of DNA-PKcs will live on. In the subsequent sections, we build upon Carl's foundational results to provide the integrated structural and functional characterization of DNA-PK and its role in NHEJ.

Repair and non-repair functions of DNA-PKcs

DNA-PKcs is a 4128 amino acid polypeptide composed of a large N-terminal HEAT (Huntingtin, Elongation Factor 3, protein phosphatase 2 A subunit, TOR) repeat containing domain, a region of conservation between FRAP, ATM and TRRAP (the FAT domain) followed by the kinase domain and a short C-terminal region termed the FATC domain (Fig. 1). The N-terminal region is subdivided into an N-terminal HEAT repeat domain (N-HEAT), which is unique to DNA-PKcs, and the middle (M-HEAT) domain which forms an α -solenoid, also called the circular cradle (Baretic et al., 2019; Sibanda et al., 2017) (Fig. 1). DNA-PKcs interacts with Ku70/80 in the presence of dsDNA to form the active DNA-dependent protein kinase holoenzyme

DNA-PK, and is required for non-homologous end joining (NHEJ), the major pathway for the repair of ionizing radiation (IR)-induced DNA double-strand breaks (DSBs) in human cells (Yue et al., 2020; Zhao et al., 2020a). In NHEJ, the Ku heterodimer binds directly to DSB ends (Fig. 2a, b), then translocates inwards, allowing DNA-PKcs to interact with the DSB end (Yoo & Dynan, 1999) (Fig. 2a, c). This complex, composed of two distinct DNA-PK complexes assembled on the two adjacent DSB ends, is often termed a synaptic complex (Fig. 2c) and serves to tether and protect the DSB ends from inappropriate degradation or rejoining. IR-induced DSBs frequently contain non-ligatable end groups (indicated by stars in Fig. 2a) and other forms of DNA damage that must be removed or repaired prior to ligation (Menon & Povirk, 2016). This is thought to be carried out by processing enzymes such as the nuclease Artemis, the 3'-DNA phosphatase/5'-DNA kinase polynucleotide kinase phosphatase (PNKP) and other factors (Ghosh & Raghavan, 2021) (Fig. 2g). Ligatable ends are then rejoined by the XLF-XRCC4-DNA ligase IV complex (Menon & Povirk, 2016; Wang & Lees-Miller, 2013; Zhao et al., 2020a) (Fig. 2h–j). Thus, NHEJ can be considered to proceed in four distinct phases. (1) Detection of the DSB ends by the Ku70/80 heterodimer; (2) recruitment of DNA-PKcs and other proteins to tether the DSB ends together and protect them from nucleolytic digestion (Fig. 2c–e); (3) removal of non-ligatable end groups (Fig. 2g) and (4) ligation of the DSB ends (Fig. 2h–j). However, the precise order of recruitment of NHEJ factors to the DSB is still unclear. Indeed, rather than taking place sequentially, it has been proposed that the NHEJ proteins form a multi-protein assembly at the DSB (Wang & Lees-Miller, 2013; Williams et al., 2014; Yano & Chen, 2008). Assembly of such a complex could have advantages over a sequential pathway in that the DSB ends would be held within the protein complex, protected from inappropriate end joining or nuclease activity, that hand over from one component to another could occur more efficiently and that the complex might maintain some function even if an individual component was missing or mutated, as has been proposed in other DNA repair systems (Hitomi et al., 2007).

Within this NHEJ complex, DNA-PKcs is emerging as a critical master regulator. Interaction of DNA-PKcs with Ku activates its kinase activity and helps tether DSB ends in the initial synaptic complex. As discussed in detail below, DNA-PKcs autophosphorylation results in dissociation of DNA-PKcs from the DSB ends allowing processing enzymes such as Artemis and PNKP (Fig. 2e–g) and ultimately DNA ligase IV to access the DSB ends and complete repair (Fig. 2h–j). DNA-PKcs and Artemis are also required for opening DNA hairpin coding ends during V(D)J recombination and cells lacking DNA-PKcs, Ku, XRCC4 or DNA ligase IV are radiation sensitive and defective in NHEJ and V(D)J

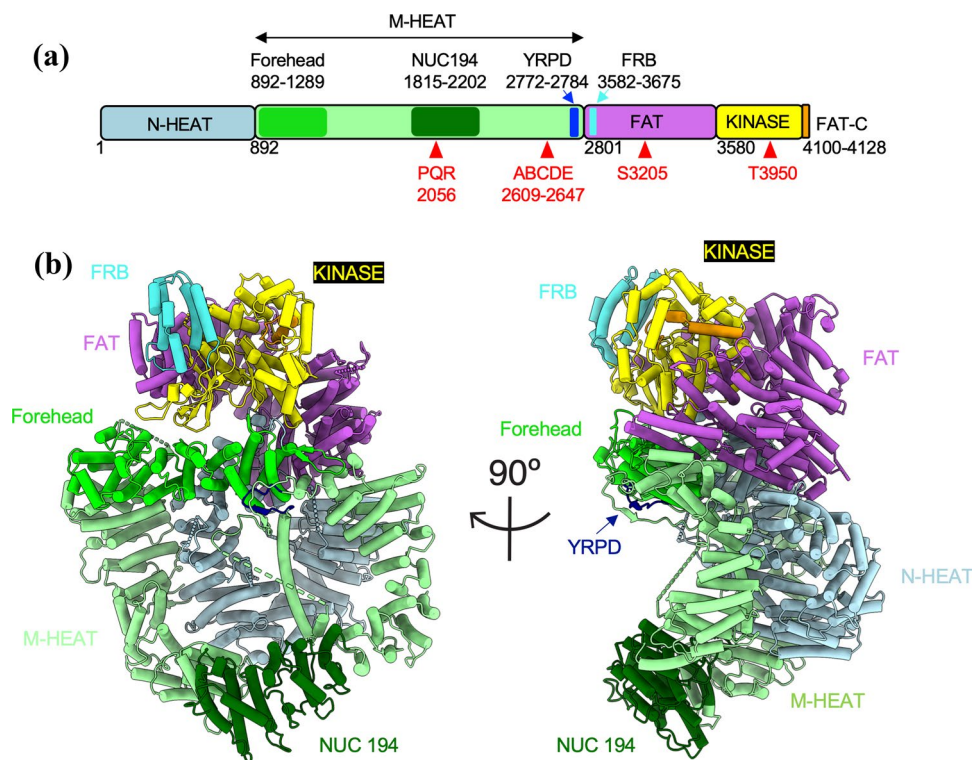


Fig. 1 Structure of DNA-PKcs. **A** Schematic of DNA-PKcs showing the N-HEAT domain (residues 1–891, light blue), the M-HEAT domain (residues 892–2800, light green), the FAT domain (residues 2801–3579, purple), the kinase domain (3580–4099, yellow) and the FATC domain (4100–4128, orange). Also shown is the position of the conserved forehead domain (892–1289, bright green) and the NUC194 domain (residues 1815–2202, dark green), the YRPD motif (residues 2772–2784, dark blue), the FRB domain (residues 3582–

3675, bright blue) and the positions of the PQR, ABCDE, S3205 and T3950 phosphorylation sites (red triangles). **B** Two views of DNA-PKcs (from PDB 7LT3) (Chen et al., 2021a), rotated by 90°, colored as in **A**. The YRPD motif is shown in dark blue, most clearly visible in the side view, indicated by an arrow. See Lees-Miller et al., (2020) for conservation of amino acids in the forehead and NUC194 domains in DNA-PKcs from eukaryotes

recombination (Pannunzio et al., 2018). Recently, additional roles for DNA-PKcs, outside DSB repair have been reported (Yue et al., 2020), including mitosis (Douglas et al., 2014, 2020; Jette & Lees-Miller, 2015; Lee et al., 2011), transcription (Goodwin et al., 2015; Ju et al., 2006), RNA processing (Shao et al., 2020), metastasis (Goodwin et al., 2015), metabolism (Chung, 2018; Park et al., 2017), the innate immune response (Yang et al., 2020) and HIV replication (Anisenko et al., 2020), but its most well-characterized role is in NHEJ.

DNA-PKcs autophosphorylation in NHEJ

DNA-PKcs phosphorylates many NHEJ proteins in vitro including Ku70 and Ku80 (Chan et al., 1999; Douglas et al., 2005), XRCC4 (Yu et al., 2003), XLF (Yu et al., 2008), DNA Ligase IV (Wang et al., 2004), PNKP (Zolner et al., 2011), APLF (Hammel et al., 2016) and Artemis (Goodarzi et al., 2006) but the functional impact of these phosphorylation events is, in many cases, unclear as ablation of

DNA-PKcs phosphorylation sites had little effect on cell survival after IR or on V(D)J recombination (Douglas et al., 2005; Yu et al., 2003, 2008). Moreover, phosphorylation of PNKP (Segal-Raz et al., 2011; Zolner et al., 2011), Artemis (Chen et al., 2005b) and APLF (Macrae et al., 2008) at the same sites is largely ATM-dependent in cells, indicating extensive crosstalk between the DNA-damage-induced PIKKs (Schlam-Babayov et al., 2021).

One of the most well-characterized in vitro and in vivo substrates of DNA-PKcs is itself. DNA-PK undergoes autophosphorylation in vitro and autophosphorylation correlates with loss of DNA-PK kinase activity and dissociation of DNA-PKcs from Ku (Chan & Lees-Miller, 1996). DNA-PKcs autophosphorylation sites include T2609, S2612, T2620, S2624, T2638 and T2647 in the M-HEAT domain (together called the ABCDE cluster), S3205 in the FAT domain, and T3950 in the kinase domain (Chan et al., 2002; Douglas et al., 2002, 2007) (Fig. 1A). Serine 2056 in the HEAT repeat region, part of the PQR cluster, was also identified as an in vitro and in vivo phosphorylation site (Chen et al., 2005a; Cui et al., 2005) and both ABCDE and PQR

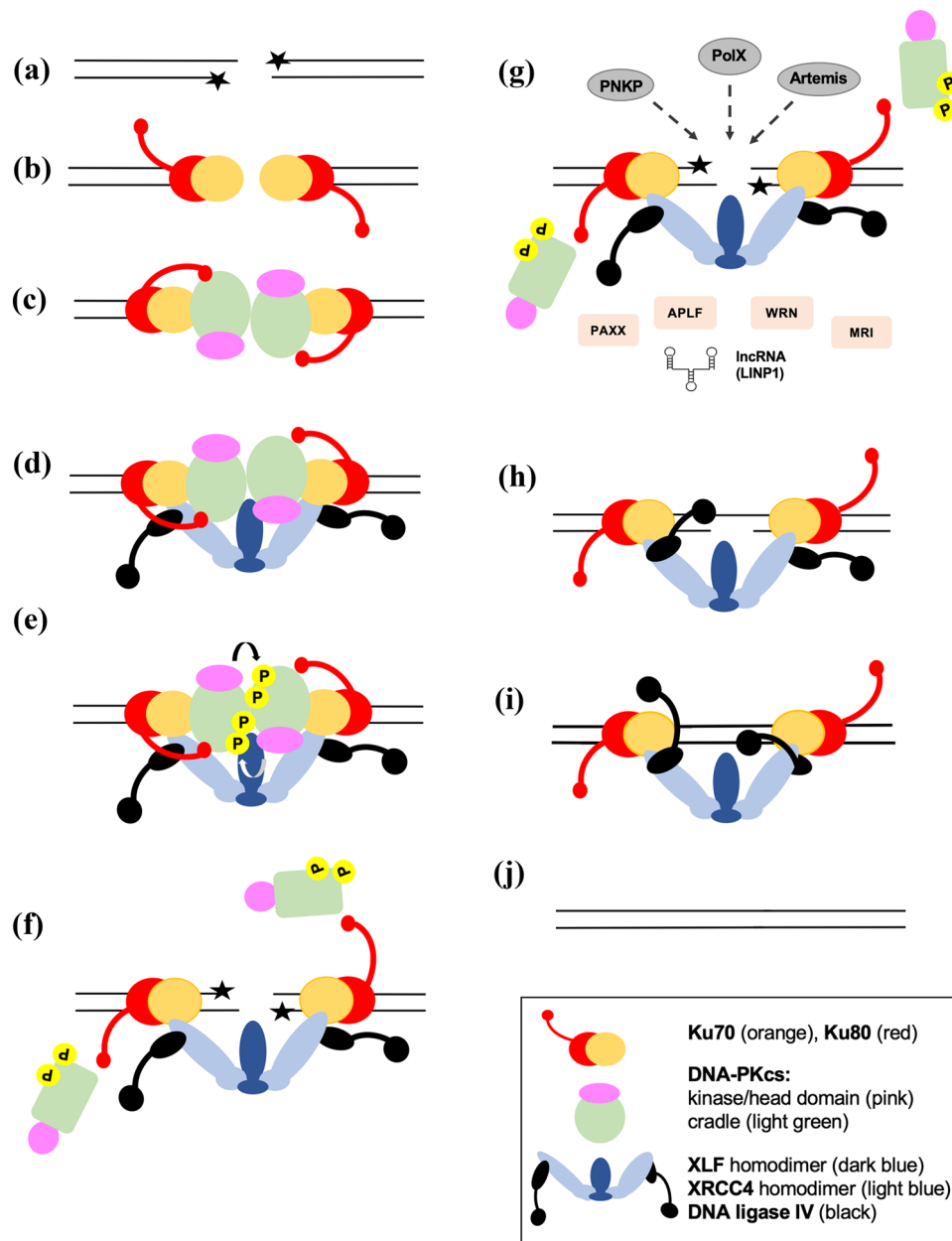


Fig. 2 Model for NHEJ. **A** IR induces DSBs, often with damaged ends that contain non-ligatable end groups (stars). **B** A Ku heterodimer (composed of Ku70, orange; Ku80, red) binds to each end of the DSB and **C** recruits DNA-PKcs (shown with head/kinase-FATC domain in pink and circular cradle in green) to form a synaptic complex. **D** The Ku complex also recruits a single homodimer of XLF (dark blue) that interacts via its head domain with two XRCC4 homodimers (light blue). The coiled-coil domain of XRCC4 interacts with tandem BRCT domains in the C-terminal region of DNA ligase IV (black). This forms a long-range synaptic complex as reported in Chen et al. (2021a). **E** Autophosphorylation of DNA-PKcs at ABCDE and possibly other sites, likely in trans leads to a conformational change that causes release of the DSB ends by DNA-PKcs. **F** autophosphorylated DNA-PKcs dissociates from the complex provid-

ing access to processing enzymes such as PNKP, Artemis and DNA polymerases mu and lambda which remove non-ligatable end groups and fill in gaps (**G**). **H** and **I** each DNA ligase IV is tethered to an XRCC4 homodimer through its C-terminal BRCT domains while the N-terminal catalytic domain is attached via a flexible linker region. This allows the catalytic domains to access the DSB ends while Ligase IV remains tethered to the synaptic complex. Each DNA Ligase IV repairs a single-strand break so that the two breaks are repaired sequentially. **J** The break is repaired restoring genome integrity. **G** shows additional NHEJ factors such as PAXX, APLF, WRN, Cyren/MRI and long non-coding RNA which may also be involved. The precise order of recruitment and dissociation of the NHEJ proteins is unknown

phosphorylation clusters are important for NHEJ and V(D)J recombination (Convery et al., 2005; Cui et al., 2005; Ding et al., 2003; Jiang et al., 2015; Neal et al., 2011). Biochemical assays using biotinylated DNA pull-down assays with purified proteins confirmed that phosphorylation disrupts the interaction of wt-DNA-PKcs with DNA-bound Ku, while DNA-PKcs in which the ABCDE sites have been mutated to alanine was less effective at phosphorylation-induced dissociation (Hammel et al., 2010; Jette & Lees-Miller, 2015). Together, these studies suggested that DNA-PKcs is recruited to DSBs through its interaction with Ku and that it undergoes autophosphorylation and subsequent release, thus allowing other NHEJ proteins to access the DSB ends to facilitate processing and ligation.

The first suggestion that ABCDE phosphorylation facilitates release of DNA-PKcs from Ku bound DNA in vivo was the observation that DNA-PKcs with all six sites blocked by alanine substitution imparts a more dramatic radiosensitization than cells completely lacking DNA-PKcs (Ding et al., 2003). This was explained by the observation that DNA-PKcs with ABCDE sites blocked remains associated with chromatin for much longer after IR than wild-type DNA-PKcs (Cui et al., 2005). Further support for this model came from laser micro-irradiation experiments showing that whereas wild-type DNA-PKcs was released from sites of DNA damage with a half-life of about 1 h, over 60% of DNA-PKcs in which the S2056, T2609, S2612, T2620, S2624, T2638 and T2647 were mutated to alanine as well as kinase-dead DNA-PKcs were retained at DSB sites for at least 2 h (Uematsu et al., 2007). Given the accumulating evidence of DNA-PKcs' importance in NHEJ, we re-examined DNA-PKcs evolution. We discovered it is more widespread and ancient than previously supposed and uncovered a high conserved motif that informs cryo-EM structures (Lees-Miller et al., 2020).

Evolutionary conservation of DNA-PKcs

An impediment to studying the cellular function of DNA-PKcs is its size, 4128 amino acids, making it difficult to express by recombinant means. In addition, unlike ATM and ATR, DNA-PKcs is absent from many model organisms including *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Arabidopsis thaliana*. This has led to the assumption that DNA-PKcs is a vertebrate-specific PIKK. However, the literature contains reports of putative DNA-PKcs homologues in Arthropods (Dore et al., 2004), *Dictyostelium discoideum* (Block & Lees-Miller, 2005; Pears & Lakin, 2014) and a variety of invertebrates and plants (Elias-Villalobos et al., 2019). To investigate these claims in more detail, we searched NCBI databases for putative DNA-PKcs/*PRKDC*

homologues and used multiple sequence analysis to identify conserved amino acids (Lees-Miller et al., 2020). We showed that DNA-PKcs is present in a wide diversity of eukaryotes including Arthropods, Echinoderms, Molluscs, Annelids, Nematodes, Cnidaria, oomycetes, plants, fungi, protozoa and amoeba (Lees-Miller et al., 2020). Yet, DNA-PKcs was not present uniformly within all phyla. For example, although it was present in fungi of phyla Mucoromycota and Chytridiomycota, it was absent from the Dikarya, a subkingdom that includes the yeasts *S. cerevisiae*, *S. pombe* and mushrooms. Similarly, although DNA-PKcs is present in Enoplean nematodes such as *Trichinella* and *Trichuris*, it is absent in Chromadorea, which includes *C. elegans*. Furthermore, although DNA-PKcs is present in many plants including green algae, liverworts, mosses, ferns, cycads, and conifers such as pine, fir and spruce, it is absent from flowering plants, including *Arabidopsis thaliana*. Within the insect order Diptera, DNA-PKcs is present in the suborder Nematocera including mosquitoes and midges, but absent in fruit flies from the suborder Brachycera (such as *Drosophila melanogaster*). Why so many model organisms lack DNA-PKcs is unclear, however, these data reveal that DNA-PKcs/*PRKDC* has a far more widespread and ancient lineage than previously supposed (Lees-Miller et al., 2020).

Analysis of DNA-PKcs sequence across species and phyla allowed us to identify regions of high amino acid conservation that likely maintained important functional roles. These included a sequence in the forehead domain at the beginning of the M-HEAT domain (residues 893–1289) and a region between residues 2772 and 2784 in the M-HEAT domain, immediately following the disordered loop containing the ABCDE phosphorylation sites and immediately prior to the FAT domain. Remarkably, in this region, which we termed the YRPD motif, the residues YRxGxxPD were invariant in DNA-PKcs through vertebrates, invertebrates, oomycetes (mildew), amoeba and plants (Lees-Miller et al., 2020), suggesting that they are critical for DNA-PKcs function.

Structures of NHEJ synaptic complexes

Over the past decade, elegant structural studies have revealed the structure of DNA-PKcs alone and in complex with Ku70/80 and dsDNA (Chaplin et al., 2021; Chen et al., 2021b; Sharif et al., 2017; Sibanda et al., 2010, 2017; Yin et al., 2017). The FAT-kinase and FATC domains form a head or crown at the top of the molecule (represented in pink in Fig. 2), while the N and M-HEAT domains form 2 α -solenoid rings, often referred to as the circular cradle (shown in green in Fig. 2). Double-stranded-DNA, bound within the pre-formed ring of Ku70/80 enters from the back and the base of DNA-PKcs, between the N and M-HEAT repeat rings such that the DSB end is positioned within its

central cavity (Chen et al., 2021b; Yin et al., 2017). While these structures were extremely informative, they were unable to provide information on the location of the flexible loop containing the ABCDE phosphorylation sites. In addition, none of the available structures of DNA-PKcs alone or DNA-PK holoenzyme addressed the structure of the proposed NHEJ synaptic complexes, where one molecule of Ku and one molecule of DNA-PKcs is positioned on each side of the DSB (Fig. 2c).

Clues to the structure of NHEJ synaptic complexes came from single molecule (sm) FRET studies in *Xenopus* egg extracts which showed that synapsis occurs in two stages. The first resulted in complex in which the two DSB ends face each other but are offset by more than 100 Å and a second in which the DSBs were more closely aligned (Graham et al., 2016). These were referred to as long-range and short-range complexes, respectively (Graham et al., 2016). Further support for the existence of a long-range complex came from structural mass spectrometry using hydrogen–deuterium exchange and cross-linking where two DNA-PK holoenzymes were shown to form a synaptic complex with the DSB ends staggered, approximately 100 Å apart (Hepburn et al., 2020). Hydrogen–deuterium exchange and cross-linking mass spectrometry studies also indicated that the path of dsDNA within DNA-PKcs was blocked by a flexible “plug domain”, corresponding to residues 2577–2773, i.e., encompassing the ABCDE phosphorylation cluster (Hepburn et al., 2020), however, at 13 Å, precisely how this plug domain functioned was unclear.

To better understand the nature of NHEJ synaptic complexes and the mechanism of NHEJ, we assembled Ku70/80 and DNA-PKcs on dsDNA, then added XLF and X4-L4 and determined the structures of synaptic complexes using cryo-EM (Chen et al., 2021a). We characterized a complex composed of DNA-PKcs, Ku70/80 and dsDNA held by a scaffold containing two XRCC4 homodimers, each bound to one molecule of Lig4 and linked via their head domains by a single XLF homodimer (Fig. 3). Since the DSB ends in this complex were staggered and ~115 Å apart (Chen et al., 2021a), we referred to this as the long-range synaptic complex due to its similarities with complexes inferred from FRET (Graham et al., 2016) and mass spectrometry (Hepburn et al., 2020). The stoichiometry of the complex was also consistent with quantitative smFRET data, indicating a single XLF homodimer linked to two XRCC4-Ligase IV molecules (Graham et al., 2018). The DSB ends were located within the central cavity of each DNA-PKcs, moreover, their path was blocked by a DNA end binding (DEB) helix motif composed of DEB-appended helix (DEB-A, residues 2724–2730), a 6 amino acid linker and a 37 amino acid DEB helix (residues 2736–2767) that directly interact with the DSB ends. Moreover, the highly conserved YRPD motif (residues 2772–2784) was shown to form a beta-sheet that interacts with an anti-parallel beta-sheet formed by the YRPD-interacting region (YRPD-I, residues 2586–2604) (Fig. 4).

From the structures, we suggest that two DNA-PKcs molecules interact within a long-range synaptic complex such

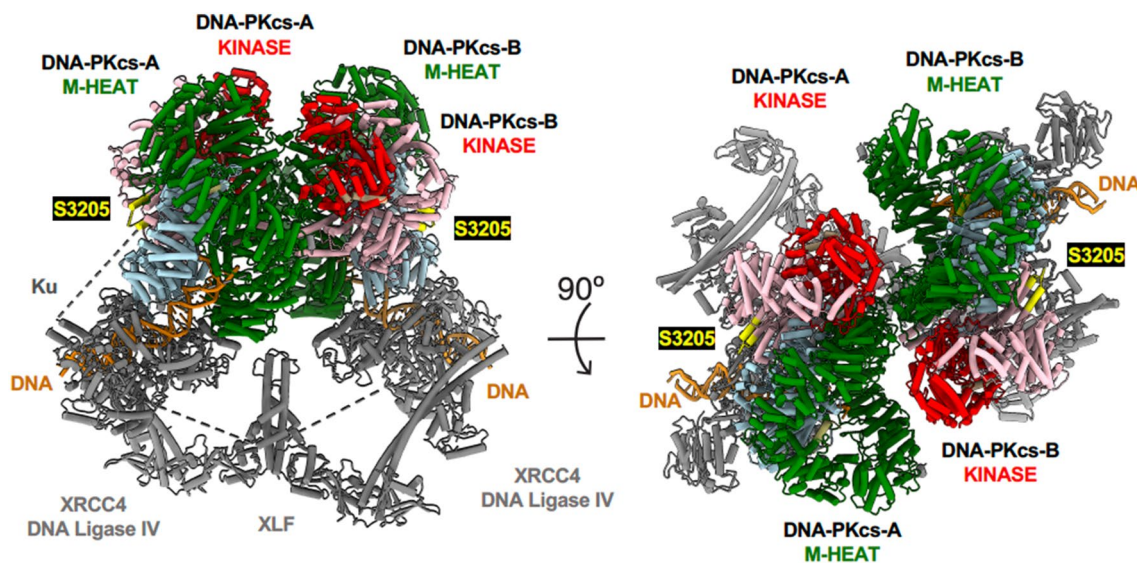


Fig. 3 Model for long-range synaptic complex. The long-range synaptic complex (PDB: 7LT3) (Chen et al., 2021a) was generated using ChimeraX 1.1.1 (Pettersen et al., 2021). The N-HEAT domain is shown in light blue, the M-HEAT in green and the kinase domain in red. The FAT domain is in pink and the FATC in tan. dsDNA is shown in dark orange. The Ku heterodimer, XLF and XRCC4-Lig4

are shown in grey. The structure on the right is rotated by 90° to reveal the relative positions of the kinase domains (red) and circular cradles (green) in the opposing DNA-PKcs molecules. The location of the ATM-dependent phosphorylation site S3205 is shown in yellow

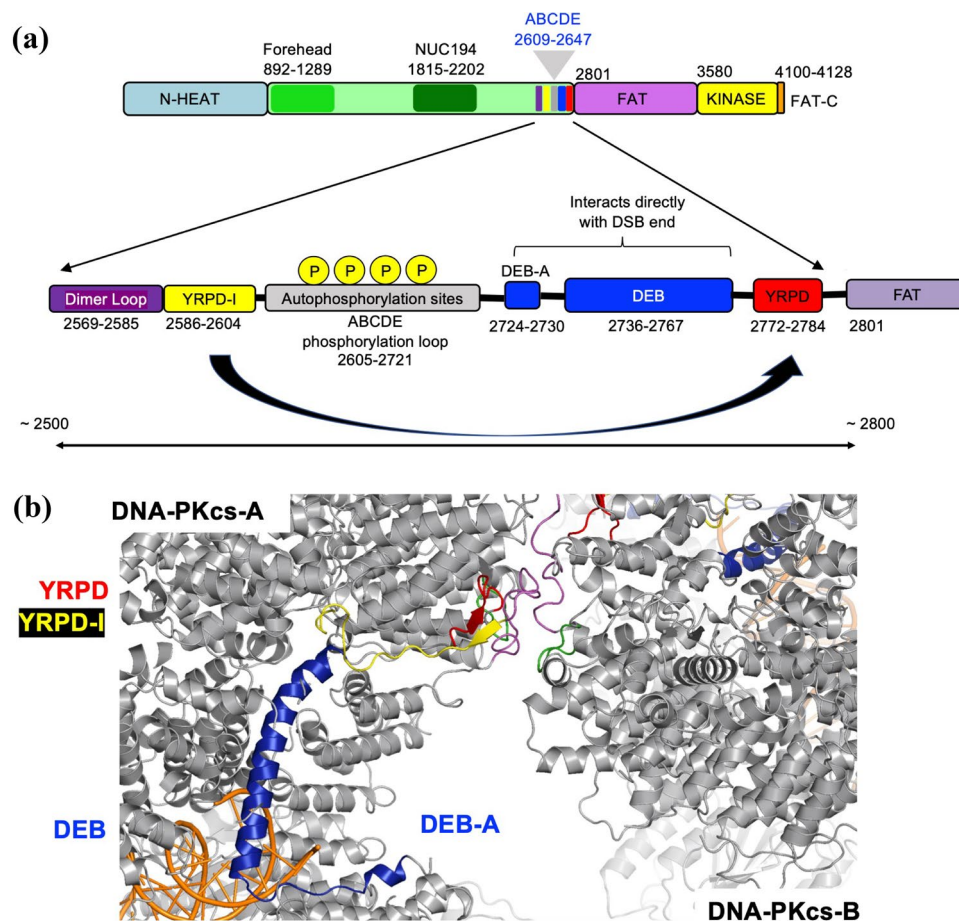


Fig. 4 Cartoon showing regulatory sequences located in DNA-PKcs between 2500 and 2800. **A** Schematic showing the region of DNA-PKcs between amino acids 2500 and 2800 that regulates the synaptic interface. The dimerization loop (residues 2569–2585), the YRPD-interacting motif (2586–2604), the DNA End Binding (DEB) helix (2736–2767) with DEB-appended (DEB-A, residues 2724–2730) and a short linker (2731–2735), and the YRPD motif (2722–2784) are from Chen et al. (2021a). DEB-A, the linker and 37 amino acid DEB form a basic surface that interacts directly with the ends of the DSB (Chen et al., 2021a). **B** Close up view of interaction interface between

two DNA-PKcs molecules, labeled DNA-PKcs-A and DNA-PKcs-B, showing the forehead loop (residues 896–903, green), the dimerization loop (2569–2585, purple), the YRPD-I (residues 2586–2604, yellow), the YRPD (2774–2784, red) and the DEB-A, linker DEB (2724–2767, blue). The position of the DEB is maintained through interaction of the YRPD and the YRPD-I, shown as yellow and red beta sheets respectively, the dimerization loop and the forehead loop. The disordered loop containing the ABCDE autophosphorylation sites (residues 2605–2721) is absent from the structures and is not shown

that autophosphorylation of the ABCDE and PQR phosphorylation sites occurs in trans (Meek et al., 2007), leading to the release of the DSB ends from the DEB/DEB-A motif by charge repulsion, and dissociation of DNA-PKcs. Basic amino acids in the DEB-A and DEB helices are highly conserved (Lees-Miller et al., 2020), indicating the importance of this DNA end binding mechanism. In line with these observations, DNA-PKcs in the long-range complex is captured in its activated state. Thus, we propose that autophosphorylation results in dissociation of DNA-PKcs from the DSB ends, while the Ku-XLF-XRCC4-Ligase IV scaffold retains and aligns the DNA ends for subsequent processing and ligation. However, we cannot exclude the possibility that autophosphorylation can also occur in cis, and that other

phosphorylation events are involved. Equally, the coordination of autophosphorylation, end-processing and ligation needs to be highly regulated, and the precise mechanism is yet to be resolved.

Biochemical studies revealed that DNA-PKcs interacts with Ku via a conserved motif within the sequence EEGGD-VDDLDMI (residues 720–732) at the extreme C-terminus of Ku80 (Falck et al., 2005; Gell & Jackson, 1999). This DNA-PKcs interaction motif is attached to the body of Ku80 via a flexible linker (Hammel et al., 2010; Harris et al., 2004) and may be important not only for recruiting DNA-PKcs to the DSB but also for tethering autophosphorylated DNA-PKcs to the synaptic complex after its dissociation from the DSB ends (Fig. 2f). Although the linker was not visible

in our structure, Ku80 residues 725–731 formed a discrete alpha-helix that interacted with the NUC194 domain of DNA-PKcs (residues 1815–2202) (Fig. 5). Low-resolution in this area of DNA-PKcs prevented assignment of interacting residues with certainty, but, from our structure it is likely that the DNA-PKcs helix between residues 1912–1923 as well as amino acids 1955–1970 are involved in the interaction with Ku80 CTR (Chen et al., 2021a). This assignment differs slightly from the Ku80-CTR-interacting regions identified in the recent structures of DNA-PKcs-Ku-DNA complexes (Chaplin et al., 2021; Chen et al., 2021b) perhaps indicating subtle differences in conformation of DNA-PKcs and Ku alone versus in the XLF-XRCC4-DNA Ligase IV complex (see also Hammel & Tainer, 2021; Hammel et al., 2020). Nevertheless, each of the structures reveals the importance of the NUC194 domain in the DNA-PKcs–Ku interaction. The NUC194 domain is conserved through evolution and is a signature of DNA-PKcs (Lees-Miller et al., 2020). Similarly, the Ku80 C-terminal motif DVDDLLDM is conserved in vertebrates, invertebrates, ciliates, molds and non-flowering plants that contain DNA-PKcs, but is absent from Dikarya yeast, flowering plants and organisms that lack DNA-PKcs (Figs. 5 and 6), suggesting that these regions co-evolved and that the function of DNA-PKcs in NHEJ is conserved in most eukaryotes. Interestingly, although *Drosophila* does not have DNA-PKcs/PRKDC, it retains a partial DNA-PKcs-interacting motif (DMEDVEM). This may relate to relatively recent loss, in evolutionary terms, of DNA-PKcs/PRKDC from *Drosophila* as it is present along with a conserved Ku80 COOH-terminus in other flies from the infra order Muscomorpha including the stable fly and robber

fly. This is compared to loss in yeast at the sub-Kingdom level and flowering plants and Chromadorean nematodes at the class level.

Our cryo-EM structures revealed that in the absence of DNA-PKcs, Ku, dsDNA, XLF and X4-L4 formed a complex in which the two Ku heterodimers face each other, positioned by the XRCC4-DNA Ligase IV-XLF-XRCC4-DNA Ligase IV scaffold, such that the DSB ends are horizontal and aligned (Chen et al., 2021a). We termed this a short-range synaptic complex, in reference to earlier FRET-based studies in *Xenopus* egg extracts (Graham et al., 2016, 2018). Significantly, in this complex, the 16 bp of DNA previously held within the cradle domain of DNA-PKcs is now free, providing an ideal substrate for efficient nick sealing by Ligase IV (S. Chen et al., 2021a). We propose that DNA-PKcs acts as a “molecular ruler”, initially drawing ends of the DSB into its central cavity until further passage is blocked by interaction with the DEB. In this conformation, the ends of the DSB are slightly opened, or melted, which may be important for subsequent end-processing and ligation. Based on these structures and the in vitro and cell-based data described above, we propose that DNA-PKcs autophosphorylation triggers transition from the long-range to the short-range synaptic complex (Chen et al., 2021a). Moreover, our structures further show that Ligase IV ligates one nick within the DSB at a time in a sequential manner, suggesting a unique catalytic mechanism for rejoining (Chen et al., 2021a) (Fig. 2h–j). This finding is also able to explain how the unusual single turn-over DNA ligase,

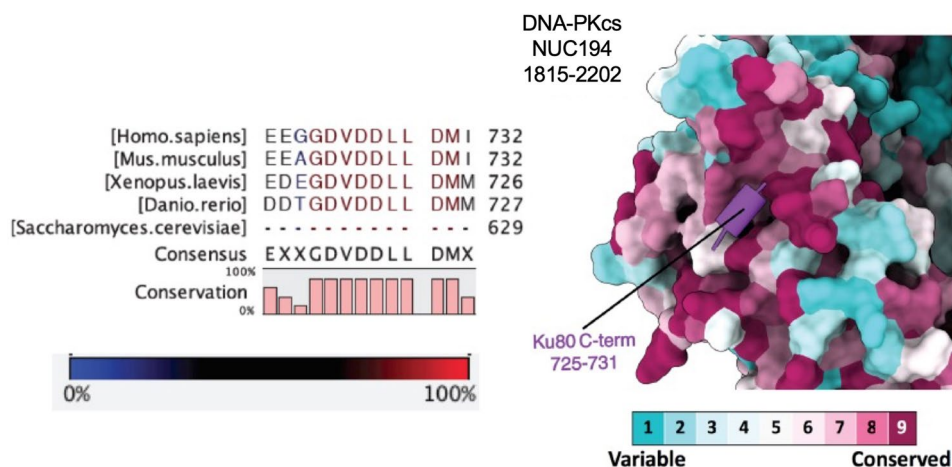


Fig. 5 Interaction of the C-terminal region of Ku80 (residues 725–731) with the NUC194 domain of DNA-PKcs. The amino acid sequence VDDLLDM at the extreme C-terminus of Ku80 interacts directly with DNA-PKcs (Falck et al., 2005). In the long-range complex (Chen et al., 2021a), this sequence forms an alpha-helix (pur-

ple) that interacts with the NUC194 domain (residues 1815–2202) of DNA-PKcs. Conserved residues within the NUC194 domain of vertebrate model organisms are shown in magenta and variable amino acids are shown in teal

DNA Ligase IV, is able to completely ligate two nicks in DSBs via the NHEJ pathway (Chen et al., 2009).

Future considerations

The structures of these NHEJ synaptic complexes provide unprecedented insight into the mechanism of NHEJ and also raise new questions. The location of the adjacent DNA-PKcs molecules such that the kinase domain of one DNA-PKcs molecule is adjacent to the forehead and concave ring of the opposite and partially inverted DNA-PKcs, places the one DNA-PKcs active site in position to phosphorylate the disordered ABCDE loop of the opposite molecule in trans (Fig. 3). This observation supports our hypothesis that autophosphorylation of the ABCDE cluster plays an important role in disassembly of the complex. For example, phosphorylation of the ABCDE cluster would introduce negative charges that could disrupt the interaction of the positively charged DEB helix with the ends of dsDNA (Fig. 4). However, DNA-PKcs autophosphorylation is more complex than just ABCDE phosphorylation, involving phosphorylation of the PQR cluster, and perhaps other sites, and we cannot exclude the possibility that phosphorylation at some sites may occur in cis as well as in trans. Ablation of the ABCDE and PQR phosphorylation sites has opposing effects on end-processing, with ABCDE promoting end-processing and PQR limiting it (Cui et al., 2005; Ding et al., 2003).

A recent study suggests that autophosphorylation occurs in a two-step process where different types of DNA promote autophosphorylation at different sites and affects the ability to phosphorylate exogenous substrates. For example, dsDNA with closed DNA hairpin ends (mimicking DNA produced by RAG endonucleases at coding ends in V(D)J recombination) supports robust phosphorylation of T2609 in the ABCDE cluster with minimal phosphorylation of either S2056 in the PQR cluster or exogenous substrates such as p53, Hsp90 and XRCC4. In contrast, dsDNA with open ends supports phosphorylation of T2609, S2056 and exogenous substrates (Meek, 2020). These findings have important implications not only for the mechanism of activation of DNA-PKcs and thus regulation of NHEJ, but also how DSB ends are processed during V(D)J recombination. The data suggests that in the first step of phosphorylation, RAG-induced DNA hairpins on coding ends support ABCDE phosphorylation, resulting release of DSB ends by DNA-PKcs, permitting activation of the hairpin opening activity of Artemis. Once the hairpins have been opened, the free DNA ends support phosphorylation of S2056 and full activation of DNA-PKcs (Meek, 2020).

It has also been reported that the related PIKKs ATM and ATR can phosphorylate DNA-PKcs on the ABCDE sites (Chen et al., 2007; Jiang et al., 2015; Yajima et al., 2006). How this would occur in the context of the synaptic complex described here is unclear, perhaps indicating that ATM-dependent phosphorylation of DNA-PKcs occurs when DNA-PKcs and Ku assemble on DSB ends but are unable to form a synaptic complex. In addition, from our model we suggest that once DNA-PKcs has undergone autophosphorylation it dissociates from the DSB ends. For it to re-form a synaptic complex and/or to be phosphorylated by other PIKKs, it seems likely that the ABCDE/PQR sites would need to be dephosphorylated. Although we have shown that protein phosphatase 6 (PP6) interacts with DNA-PKcs, depletion of PP6 did not result in increased ABCDE phosphorylation suggesting that PP6 may not be responsible for dephosphorylation of DNA-PKcs on the ABCDE sites (Douglas et al., 2010); therefore, DNA-PKcs dephosphorylation could involve other protein phosphatases. Alternatively, autophosphorylation could act as a signal for degradation and turnover of DNA-PKcs at the DSB.

One site that is clearly phosphorylated by ATM in response to DNA damage is serine 3205 in the FAT domain (Neal et al., 2011; Schlam-Babayov et al., 2021). Although the precise region surrounding S3205 was missing from our structure, the immediately surrounding amino acids are solvent facing, consistent with 3205 being located on an exposed flexible loop suitable to be phosphorylated directly by ATM (Fig. 3). S3205 is also an *in vitro* autophosphorylation site (Douglas et al., 2002). Phosphorylation is unlikely to occur in cis, therefore perhaps is phosphorylated by a neighboring DNA-PKcs molecule *in vitro*. In addition, DNA-PKcs phosphorylates ATM in response to DNA damage, inhibiting its activity (Zhou et al., 2017). Again, it is difficult to imagine how this could occur within the closed configuration of the synaptic complex, suggesting that DNA-PKcs can act in other conformations or complexes, independent of the synaptic complex.

Critical to NHEJ is the processing of the DSB ends to remove non-ligatable end groups. Potential processing enzymes include the nuclease Artemis, which interacts with both DNA-PKcs (Ma et al., 2002) and DNA Ligase IV (Ochi et al., 2013) and may play a role in removing blocked DNA ends; PNKP that interacts with CK2-phosphorylated XRCC4 and removes 3'-phosphates or phosphoglycolates and adds 5'-phosphates to create ligatable end groups (Weinfeld et al., 2011), as well as the Pol X family of DNA polymerases that can fill in gaps at the DNA termini (Kaminski et al., 2020; Ramsden & Asagoshi, 2012). Other NHEJ factors whose roles are less defined include PAXX (Xing & Oksenysh, 2019), APLF (Hammel et al., 2016), CYREN/MRI (Arnoult et al., 2017; Hung et al., 2018) and long-non-coding RNA

HUMAN	QQLNHFWIIVVQD---GILITIKKEAS-GSSVVAEAKKFLAPKDKPSGDTAA-----VFEEGGVDDLLDMI--732
PLASMIDIOPHORA	RVPFAPFLDKMASR---GLVPIISSECK-GSPVAPERLRFPTQPPPTAIA-----DTT-G---TGNDDESSITWDEWGD669
HONDEA	SCEPSEFKRVAIA---NLFLISSDDAVAVSGVSAKDAADFIKENVSRNDQAGAL---MAPAPAVPAADAFDDFD-739
APHANOMYCES_A	PSNGPVWQLLVEH---HVTLLSTRDDP-SQDTSIRAESEFLLPDAQVKDEPTGG---AE-VVDAGQDDSDLFADLE-734
STYLONCHIA	GTNKEFFQSIKHK---KISLITKDESSISVITADPAKRFLEIDKDDSGQMAQQ-----KKIKKEDFMDIE-734
NAEGLERIA	KKREDFWCSFIVQN---NISLITQDECD-ESEVDEEAKKFLTEAVIAAVIQE---E---SEDEKICNDLFDLLE-773
BARLEY	DDVVKFFQLLSK---NVLILKKEEAP-DSDVTEEMARSEFLK-TEIS-----SE-----701
ARABIDOPSIS	RNLSHLLEHFMK---KITLIPKSEAA-DSDIVDEAGDFIVKQESM-----LES-----680
SOYBEAN	KNLPSFCEYLASK---ELFLISKTEAI-DSNTDDPARKGFLVKSEPK-----VD-----691
KLEBSORMIDIUM	KRQHFWRERLVKN---QLFLITRDEEP-DSVLPPEEAQKFLTEAAEKPPS---AVEEPEKDEVDDEMAELLDMAE-722
LIVERWORT	KRQDFWELVTAK---NLFLISKDETP-DSDVTEEMARSEFLSRVQNSG---HVSEGVPEEVDDEIFALLDVI-714
DICTYOSTELIUM	KKKDDFWKYITTNQPIAFLISEDECD-SSSISKDEANEFLNRKQSN-----QTPQINTNTDSDVDDLFDLIE-769
S_CEREVISIAE	-----629
S_POMBE	PKENELINKIRSS---GLIILHDLT-RHDFNFI-----695
NEUROSPORA	GDRREMFWRIVGG---RLGLITQDESE-VSEVSEEAQKFLK-----725
NEOCALLIMASTIX	YIHRDFWQVMIDN---EMILIKS-EEG-KGVVEEYAKFLEIPSKAVENS-----IKESQALITGIDDLFDLLE-720
SPIZELLOMYCES	YRHAPFRRVVDQ---QISLITKDEAG-DSNVIREADAFRAEDPEPP-A-----AISKGDEVEEEDLLDMM-797
MONOSIGA	GKAQHFATTALLQH---NLGLITVETN-TQGVQDASSFLSQPEPSHEAMDD---APPEDNADDEDDLDL-784
CAPSAPORA	PQHYEFWQLIVKA---KITLISDEAY-DSEVTPPEAQKFLTEAVDVS---AQVAVAT-PKDDQVDDLLDMM-731
SPONGE	TRYSSFWDEMVK---SLSLITSECD-KSNVQDESKFLSESDVDTVAPVD---N---KVEEEDLDDLLDMM-746
HYDRA	KLKDFWQLLIKE---NLFLISGNEAK-DSITIEDVHSEFLNNEVEE-----KQ-IIE-LLNDVDDLLDMM-714
SOFT_CORAL	CNKASFWARIKKE---KVTLIKRNECG-DVNIIEEAQKFLLEDVSTP---EET-P---EPEVDEIDDLLEMM-735
FLATWORM	FLTQHLATEGRR---AMGLISKDETA-DSDLTAECAELLASSSGGADSGAAGK-----EDGAEEDDDLDL-759
C_ELEGANS	EDFEPFAE-VLSEEK-SCNPISSEVPS-MSEVVAEAEFWEED-----713
TRICHINELLA	GKQNFWLLKKNKD-SLGLITVDECC-SISIQEAVSYSEVSEDEILKN-----NPLDSDTDLIDLE-1005
ROTIFER	KYVENFWRNFMKD---KLGLISNLECS-DSNVSEDAEAKFLSCLINANNDA---NIVEKEVEEDLLDMM-766
SILKWORM	EKNTKLEVINEN---ELNVLKSESS-LSYINETS-QLYENDTIPNTLP-----VVIDPEVDALFDL-704
THRIP	LEKSFWEVLVRS---QAGLITSSFT-QSTVITERAKFLEMIVTKDEILQ-----PKTDFEEDLLDMM-736
DROSOPHILA	KNRSEFWQDVMVDK---QLGPLVLGEP-LDDELAL---KAYYTIENTWAESGAN-----DMEDV-EM-699
ROBBERFLY	RKXDFWQDQVVDKGLIPIQSGLI---SISIDLEAKFVYINVGPTTG---DDGPAEVDLDFEM-706
STABLEFLY	RNMNVFWDHVIVKH---ELGLCEIGEPS-LDEIKK---KDFYSLEFADAVKLTLP-----EEMEMDPENDEMLANINET31
MOSQUITO	RNKQEFWQDQVVRE---RFLIGASESD-MSITISADVEEYFNVNTAPKTGDTGA-----DADVVDADLDFEM-722
TERMITE	RGRKVFWDTIVEA---KLGLISEESA-KSVTDDPARKFLQYEEPIEQEVA-----EELVMDVDDLLDMM-715
DAPHNIA	NYFTDFWQNYVAKS---RLGLITKEESP-LSAVEKRAEQFLRLSENVNEI-----ESVDEDEELLNLE-718
TICK	IKREELWELLIKE---AVLPVKAEC-SKALDNDVVKREYVVEKEAEKKE-----DEPKDEDDLLDMM-704
PRIAPULA	KSRAWFWEAMVNE---RLFLISAVECT-PSDVQDEADAPLVKDTKEVDTI---P-----AVAVEEADLDDMM-725
SYMPHYLANS	SCPPALWQILVTK---NISLITFDESDDSHVTKQDAEFLKAGAGDSVQP-----TSIDEAADLDDMM-682
MILLIPEDE	SGKLFWDQIVEK---KIGLIGEDAP-HSGSLEADAFKIEIKTSASA-P---SPSPAEDADLDDMM-715
CENTIPEDE	KEQKFWRMIREE---KISVITKDNP-LSITIVS-ARLEEEVEVNETPD-----EIIDDDADLDDMM-715
CAPITELLA	KGRSDFWDDLLKV---KVTLISKLESE-DSVQDEAQRFLSEDEKKEEKM-----EVEEEDLDFAAE-723
LINGULA	YNKMVLFEMAKNS---GLRLINNECE-ESTYSKDAVAEFMADDVKKEKKEE---E-----EMDEEADLDDLL-725
LIMPET	KAKRDFWQIVLD---KQGLISKLECE-GSKVSKERADALLADDVKEGSQK-----IFEADADLDDMM-697
SEA_URCHIN	LLKVDFFWQDVTKD---GLFLITVTEAR-DSKVKDEADKFIQEEVPMEDP---A-----AANEEDADLDDML-729
STARFISH	IFKTFWWSQIIE---KVTLITDSSES-ESTVSKVATKFLLEETTEEEVAP-----AADADEADLDDMM-741
B_LANCELET	KGRKDFWDSMVKE---KLFLISEESS-ETSVKPEAKQFLAQKEEKVEEA-A-----MMEDEGVDLDDMM-726
HEMICHORDATE	KMRTNFWDRIVTE---SITLITSSSS-ESRVKPEAKFLKGEKKVEVPE-----IKDDDDADLDDLME-117
LAMPREY	LGLQDFWQLLVQD---RLFLITCEENS-ASSVTCBAEQFFVVEQVQEELMV-----AEPAAVDLDDLME-731
SHARK	GLYDFWDDLVQD---GVSLITQDEAG-DSAVTEAKRFLAAEERAAEPE-A-----VAEDGVDLDDMM-733
DANIO	RSLQDFWNLLAQD---ALFLISKDEVE-ASTISKANQFLSFEQKQVEVAP---AQDVTGVDLDDMM-727
GAR	RGLQDFWELLVKD---NISLISKDEVE-CSAVSKAADQFLAAEKPALLETGP---IADDEGVDLDDMM-724
XENOPUS	SALMEFWDITVQE---EISLITSESS-KGSSVTEPAKQFLAQKEEKVEEA-A-----MMEDEGVDLDDMM-726
PLATYPUS	RKLNDFWESVTE---GITLITKDEAP-GSSVTEPAQFLAPKKAATAEAA---VGEDEGVDLDDMM-725
TASMANIAN_DEVIL	KALNDFWNIVVQD---GITLITKDESP-GSSVVAEAKKFLIPKKAEM-D-----AAEEDGVDLDDMM-732
OSTRICH	KALADFWEIVMQD---RISLITKDEAG-GSLVTEPAKFLAPKQKQNTL-P-----ATDEGVDLDDMM-749
SNAKE	KNLSDFWEIVIQE---GISLITNEEAE-GSSLVAEAKKFLAPKKEPLEVS-P---VTEEGVDLDDMI-729
HUMAN	QQLNHFWIIVVQD---GILITIKKEAS-GSSVVAEAKKFLAPKDKPSGDTAA-----VFEEGGVDDLLDMI--732

such as LINP1 (Thapar et al., 2021) (Fig. 2g), though PAXX and LINP1 have both been shown to stabilize the initial synaptic complex between DNA-PKcs and Ku on DNA (Thapar et al., 2021; Wang et al., 2018).

To maintain genome integrity, it is critical that unrepaired DSB ends are not released prior to ligation, which could result in inappropriate joining and chromosomal aberrations. As seen in our structures, the Ku-XLF-XRCC4-Ligase IV flexible scaffold is a critical component of both long-range and short-range synaptic complexes. In the long-range complex, the DSB ends are protected by DNA-PKcs. We propose that after autophosphorylation-induced dissociation of DNA-PKcs from DSB ends, intramolecular “swivelling” of the Ku/XRCC4 scaffold results in the DSB ends being aligned and positioned for joining by the catalytic domain of Ligase IV (Chen et al., 2021a).

Critically, the catalytic domain of Ligase IV is attached to the XRCC4 scaffold via its tandem BRCT domains and a flexible linker (Ochi et al., 2010, 2012, 2014). Similarly, PNKP is tethered to XRCC4 via its N-terminal Fork-head associated (FHA) domain and a flexible linker (Aceytuno et al., 2017; Bernstein et al., 2005). In this way, the catalytic domains of Ligase IV and PNKP can dynamically access the DSB ends while remaining integral components of the NHEJ machine through tethering to XRCC4. A similar mechanism may be at play for DNA-PKcs, which is flexibly attached via the conserved Ku80-CTR (Fig. 2c–g), and for other processing enzymes such as Aprataxin and PNKP-like protein, APLF which is flexibly attached to both Ku and chromatin (Hammel et al., 2010, 2016) and the pol X family polymerases which interact with Ku (Zhao et al., 2020b). Critically, the intimate,

Fig. 6 Conservation of the C-terminal region of Ku80 in vertebrates and multiple invertebrates and plants that also contain DNA-PKcs, but not in Dikarya yeast, *C. elegans* and flowering plants that do not. Multiple protein sequence alignment of Ku80 was carried out using Clustal Omega and conserved amino acids were colored manually using the following criteria: acidic amino acids (D, E) are highlighted in magenta; basic amino acids (R, K, H) are highlighted in blue; small polar amino acids (T, S, G, C) are in green; aliphatic amino acids (A, L, M, I, V) are in yellow; large polar amino acids (Q, N) are in red; aromatic amino acids (F, W, Y) are in red with yellow highlight; and proline (P) is in grey. Conservation of small amino acids (A, S, T, C, V, G) is illustrated by yellow letters with dark green highlight. In bold are the amino acids present in the structure of the long-range complex (DVDDLLDM) that interact with DNA-PKcs. Organisms and accession numbers for the sequences shown are as follows: PLASMODIOPHORE (Club root pathogen), *Plasmodiophora brassicae*, CEO94558.1; HONDAEA (Protist) *Hondaia fermentalgiana*, BG27299.1; APHANOMYCESa (Crayfish plague mold) *Aphanomyces astaci*, XP_009833210.1; STYLONYCHIA (Ciliate) *Stylonychia lemnae*, CDW74323.1; NAEGLERIA (Amoeba/flagellate), *Naegleria gruberi strain*, XP_002672413.1; BARLEY, *Hordeum vulgare*, AEO86624.1; ARABIDOPSIS, *Arabidopsis thaliana*, NP_564520.1; SOYBEAN, *Glycine max*, XP_003524779.1; GREEN_ALGAE, *Klebsormidium nitens*, GAQ85406.1; LIVERWORT, *Marchantia polymorpha*, PTQ30121.1; DICTYOSTELIUM, *Dictyostelium purpureum*, XP_637846.1; S_CEREVISIAE, *Saccharomyces cerevisiae*, NP_013824.1; S_POMBE, *Schizosaccharomyces pombe*, Q9HGM8.1; NEUROSPORA, (Bread mold), *Neurospora crassa*, AFM68948.1; NEOCALLIMASTIX (Anaerobic gut fungi), *Neocallimastix californiae*, ORY73184.1; SPIZELLOMYCES (Chytridomycota fungus), *Spizellomyces punctatus*, KND04299.1; MONOSIGA, (Choanoflagellate), *Monosiga brevicollis*, XP_001750285.1; CAPSASPORA (Amoeboid protist), *Capsaspora kowczarzaki*, KJE95979.1; SPONGE, *Amphimedon queenslandica*, XP_019851059.1; HYDRA, *Hydra vulgaris*, XP_012555181.1; SOFT CORAL, *Dendronephthya gigantea*, XP_028404243.1; FLATWORM, *Macrostomum lignano*, PAA63650.1; C_ELEGANS, *Caenorhabditis elegans*, CAA83623; TRICHINELLA, Parasitic roundworm, *Trichinella pseudospiralis*, KRZ25527.1; ROTIFER, *Brachionus plicatilis*, RNA39402.1; SILKWORM, *Bombyx mori*, XP_037875756.1; THRIP, *Frankliniella occidentalis*, XP_026281764.1; DROSOPHILA, *Drosophila melanogaster*, NP_609767.2; ROBBERFLY, *Eutolmus rufibarbis*, SRR5185497; STABLEFLY, *Stomoxys calcitrans*, XP_013115123.1; MOSQUITO, *Aedes aegypti*, XP_001657128.2; TERMITE, *Cryptotermes secundus*, XP_023718673.1; DAPHNIA, *Daphnia magna*, XP_032789667.1; TICK, *Ixodes scapularis*, XP_002405506.2; PRIAPULA, Penis worm, *Priapulus caudatus*, XP_014668438.1; MILLIPEDE, *Craspedosoma sp.*, GERS01021842.1; CENTIPEDE, *Himantarium gabrielis*, GCIL01016305.1; CAPITELLA, Annelid worm, *Capitella teleta*, ELU08335.1; LINGULA, Brachiopod, *Lingula anatine*, XP_013403161.1; LIMPET, *Lottia gigantea*, XP_009064003.1; SEA URCHIN, *Strongylocentrotus purpuratus*, XP_788472.3; STARFISH, *Acanthaster planci*, XP_022103565.1; LANCELET, *Branchiostoma belcheri*, XP_019618779.1; HEMICHORDATE, Acorn worm, *Saccoglossus kowalevskii*, XP_006822840.1; LAMPREY, *Petromyzon marinus*, xp_032818808; SHARK, *Prionace glauca*, GFYY0116112.1; DANIO, *Danio rerio*, NP_001017360.2; GAR, *Lepisosteus oculatus*, XP_015214714.1; XENOPUS, *Xenopus laevis*, BAA76954.1; PLATYPUS, *Ornithorhynchus anatinus*, XP_028928834.1; TASMANIAN_DEVIL, *Sarcophilus harrisii*, XP_031813944.1; OSTRICH, *Struthio camelus australis*, XP_009666109.1; SNAKE, *Pseudonaja textilis*, XP_026563791.1. The names of organisms that do not contain DNA-PKcs are shown in red font. All others contain DNA-PKcs. See (Lees-Miller et al., 2020) for details. The human sequence is shown at the top and bottom in yellow highlight for comparison

though dynamic and flexible, interaction of the various NHEJ proteins with the core NHEJ scaffold may allow for coordinated hand-offs and efficient repair, preventing the release of toxic repair intermediates that could lead to genome instability as proposed for the base excision repair pathway (Hitomi et al., 2007).

We envision that phosphorylation by other PIKKs plays a critical role in regulating dynamic responses within NHEJ synaptic complexes, as sites on PNKP, XRCC4, Artemis and APLF, as well as DNA-PKcs itself, can be phosphorylated in an ATM-dependent manner in response to DNA damage (Chen et al., 2005b; Hammel et al., 2016; Neal et al., 2011; Schlam-Babayov et al., 2021; Segal-Raz et al., 2011; Yu et al., 2003, 2008; Zolner et al., 2011). Moreover, the DNA-PK complex can regulate the nuclease activity of MRE11 (Deshpande et al., 2020), suggesting that the NHEJ machinery regulates MRN-initiated resection and pathway choice (Shibata et al., 2014). A better understanding of not only how the NHEJ complex is regulated but how it interacts with other components of the DNA damage response such as MRN will be critical to understanding cellular mechanisms that prevent genome instability and cancer. The combination of X-ray crystal (Sibanda et al., 2017) and cryo-EM (Chaplin et al., 2021; Chen et al., 2021a, b; Sharif et al., 2017; Yin et al., 2017) structures with solution dynamics by X-ray scattering (Hammel & Tainer, 2021; Hammel et al., 2020) will provide a means to guide ongoing studies to both test and understand the functional dynamics of DNA-PK complex in orchestrating NHEJ.

Summary

In conclusion, the high-resolution structures of long-range and short-range synaptic complexes along with a new evolutionary analysis of DNA-PKcs sequence conservation provide unprecedented insight into the mechanism of NHEJ, the role of DNA-PKcs in synapsis and how DSBs are first detected, tethered then aligned for ligation. Conservation of the YRPD motif and the basic DEB-A/DEB helices suggests that the synapsis mechanism and DNA-PKcs-dependent autophosphorylation-induced regulation of NHEJ is conserved throughout evolution. DNA-PKcs is a potential cancer target, and small molecule DNA-PKcs inhibitors are currently in clinical trials (Cleary et al., 2020). The elucidation of the structure of these NHEJ synaptic complexes opens the door not only for a better understanding of NHEJ, but for design of advanced NHEJ inhibitors that could target allosteric conformations as well as its active site and thereby have improved specificity and therapeutic potential. Furthermore, the combination

of advanced allosteric DNA-PK inhibitors enabled by this mechanistic understanding with focal radiation therapy may provide new opportunities for precision oncology.

Acknowledgements We thank our colleagues and collaborators Drs. John Tainer, Michal Hammel, Katheryn Meek and David Schriemer for helpful comments and discussions.

Authors' contributions The manuscript was written by SPLM with input from all the authors. Figure 5 was generated by SC and YH. Sequence analysis for Fig. 6 was carried out by JPLM.

Funding Collaboration between the He and Lees-Miller laboratories is supported by NCI P01 CA092584 and NIH grant R01 GM135651.

Availability of data Sequences and structure codes were obtained from publicly accessible databases (NCBI and Protein Data Bank, respectively) as indicated in the text.

Code availability Not applicable.

Declarations

Conflict of interests The authors declare no conflict of interest.

Consent to participate Not applicable.

Consent for publication All the authors have read the final version of the manuscript and consent to publication.

Ethics approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aceytuno, R. D., Pieltz, C. G., Havali-Shahriari, Z., Edwards, R. A., Rey, M., Ye, R., Javed, F., Fang, S., Mani, R., Weinfeld, M., Hammel, M., Tainer, J. A., Schriemer, D. C., Lees-Miller, S. P., & Glover, J. N. M. (2017). Structural and functional characterization of the PNKP-XRCC4-LigIV DNA repair complex. *Nucleic Acids Research*, 45(10), 6238–6251. <https://doi.org/10.1093/nar/gkx275>
- Anisenko, A., Kan, M., Shadrina, O., Brattseva, A., & Gottikh, M. (2020). Phosphorylation targets of DNA-PK and their role in HIV-1 replication. *Cells*. <https://doi.org/10.3390/cells9081907>
- Arnoult, N., Correia, A., Ma, J., Merlo, A., Garcia-Gomez, S., Maric, M., Tognetti, M., Benner, C. W., Boulton, S. J., Saghatelian, A., & Karlseder, J. (2017). Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN. *Nature*, 549(7673), 548–552. <https://doi.org/10.1038/nature24023>
- Baretic, D., Maia de Oliveira, T., Niess, M., Wan, P., Pollard, H., Johnson, C. M., Truman, C., McCall, E., Fisher, D., Williams, R., & Phillips, C. (2019). Structural insights into the critical DNA damage sensors DNA-PKcs, ATM and ATR. *Progress in Biophysics and Molecular Biology*, 147, 4–16. <https://doi.org/10.1016/j.pbiomolbio.2019.06.003>
- Bernstein, N. K., Williams, R. S., Rakovszky, M. L., Cui, D., Green, R., Karimi-Busheri, F., Mani, R. S., Galicia, S., Koch, C. A., Cass, C. E., Durocher, D., Weinfeld, M., & Glover, J. N. (2005). The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. *Molecular Cell*, 17(5), 657–670.
- Blackford, A. N., & Jackson, S. P. (2017). ATM, ATR, and DNA-PK: The trinity at the heart of the DNA damage response. *Molecular Cell*, 66(6), 801–817. <https://doi.org/10.1016/j.molcel.2017.05.015>
- Block, W. D., & Lees-Miller, S. P. (2005). Putative homologues of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and other components of the non-homologous end joining machinery in *Dictyostelium discoideum*. *DNA Repair (amst)*, 4(10), 1061–1065. <https://doi.org/10.1016/j.dnarep.2005.06.008>
- Carter, T., Vancurova, I., Sun, L., Lou, W., & DeLeon, S. (1990). A DNA-activated protein kinase from HeLa cell nuclei. *Molecular and Cellular Biology*, 10(12), 6460–6471.
- Chan, D. W., Chen, B. P., Prithivirajasingh, S., Kurimasa, A., Story, M. D., Qin, J., & Chen, D. J. (2002). Autophosphorylation of the DNA-dependent protein kinase catalytic subunit is required for rejoining of DNA double-strand breaks. *Genes & Development*, 16(18), 2333–2338.
- Chan, D. W., & Lees-Miller, S. P. (1996). The DNA-dependent protein kinase is inactivated by autophosphorylation of the catalytic subunit. *Journal of Biological Chemistry*, 271(15), 8936–8941.
- Chan, D. W., Ye, R., Veillette, C. J., & Lees-Miller, S. P. (1999). DNA-dependent protein kinase phosphorylation sites in Ku 70/80 heterodimer. *Biochemistry*, 38(6), 1819–1828. <https://doi.org/10.1021/bi982584b>
- Chaplin, A. K., Hardwick, S. W., Liang, S., Kefala Stavridi, A., Hnizda, A., Cooper, L. R., De Oliveira, T. M., Chirgadze, D. Y., & Blundell, T. L. (2021). Dimers of DNA-PK create a stage for DNA double-strand break repair. *Nature Structural & Molecular Biology*, 28(1), 13–19. <https://doi.org/10.1038/s41594-020-00517-x>
- Chen, B. P., Chan, D. W., Kobayashi, J., Burma, S., Asaithamby, A., Morotomi-Yano, K., Botvinick, E., Qin, J., & Chen, D. J. (2005a). Cell cycle dependence of DNA-dependent protein kinase phosphorylation in response to DNA double strand breaks [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S.]. *Journal of Biological Chemistry*, 280(15), 14709–14715. <https://doi.org/10.1074/jbc.M408827200>
- Chen, B. P., Uematsu, N., Kobayashi, J., Lerenthal, Y., Krempler, A., Yajima, H., Lohrich, M., Shiloh, Y., & Chen, D. J. (2007). Ataxia telangiectasia mutated (ATM) is essential for DNA-PKcs phosphorylations at the Thr-2609 cluster upon DNA double strand break. *Journal of Biological Chemistry*, 282(9), 6582–6587.
- Chen, L., Morio, T., Minegishi, Y., Nakada, S., Nagasawa, M., Komatsu, K., Chessa, L., Villa, A., Lecis, D., Delia, D., & Mizutani, S. (2005b). Ataxia-telangiectasia-mutated dependent phosphorylation of Artemis in response to DNA damage. *Cancer Science*, 96(2), 134–141.
- Chen, S., Lee, L., Naila, T., Fishbain, S., Wang, A., Tomkinson, A. E., Lees-Miller, S. P., & He, Y. (2021a). Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature*, 593(7858), 294–298. <https://doi.org/10.1038/s41586-021-03458-7>

- Chen, X., Ballin, J. D., Della-Maria, J., Tsai, M. S., White, E. J., Tomkinson, A. E., & Wilson, G. M. (2009). Distinct kinetics of human DNA ligases I, IIIalpha, IIIbeta, and IV reveal direct DNA sensing ability and differential physiological functions in DNA repair. *DNA Repair (amst)*, 8(8), 961–968. <https://doi.org/10.1016/j.dnarep.2009.06.002>
- Chen, X., Xu, X., Chen, Y., Cheung, J. C., Wang, H., Jiang, J., de Val, N., Fox, T., Gellert, M., & Yang, W. (2021b). Structure of an activated DNA-PK and its implications for NHEJ. *Molecular Cell*, 81(4), 801–810.e803. <https://doi.org/10.1016/j.molcel.2020.12.015>
- Chung, J. H. (2018). The role of DNA-PK in aging and energy metabolism. *FEBS Journal*, 285(11), 1959–1972. <https://doi.org/10.1111/febs.14410>
- Cleary, J. M., Aguirre, A. J., Shapiro, G. I., & D'Andrea, A. D. (2020). Biomarker-guided development of DNA repair inhibitors. *Molecular Cell*, 78(6), 1070–1085. <https://doi.org/10.1016/j.molcel.2020.04.035>
- Connelly, M. A., Zhang, H., Kieleczawa, J., & Anderson, C. W. (1998). The promoters for human DNA-PKcs (PRKDC) and MCM4: Divergently transcribed genes located at chromosome 8 band q11. *Genomics*, 47(1), 71–83.
- Convery, E., Shin, E. K., Ding, Q., Wang, W., Douglas, P., Davis, L. S., Nickoloff, J. A., Lees-Miller, S. P., & Meek, K. (2005). Inhibition of homologous recombination by variants of the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs). *Proc Natl Acad Sci U S A*, 102(5), 1345–1350. <https://doi.org/10.1073/pnas.0406466102>
- Cui, X., Yu, Y., Gupta, S., Cho, Y. M., Lees-Miller, S. P., & Meek, K. (2005). Autophosphorylation of DNA-dependent protein kinase regulates DNA end processing and may also alter double-strand break repair pathway choice. *Molecular and Cellular Biology*, 25(24), 10842–10852. <https://doi.org/10.1128/mcb.25.24.10842-10852.2005>
- Deshpande, R. A., Myler, L. R., Soniat, M. M., Makharashvili, N., Lee, L., Lees-Miller, S. P., Finkelstein, I. J., & Paull, T. T. (2020). DNA-dependent protein kinase promotes DNA end processing by MRN and CtIP. *Science Advances*, 6(2), eaay0922. <https://doi.org/10.1126/sciadv.aay0922>
- Ding, Q., Reddy, Y. V., Wang, W., Woods, T., Douglas, P., Ramsden, D. A., Lees-Miller, S. P., & Meek, K. (2003). Autophosphorylation of the catalytic subunit of the DNA-dependent protein kinase is required for efficient end processing during DNA double-strand break repair. *Molecular and Cellular Biology*, 23(16), 5836–5848.
- Dore, A. S., Drake, A. C., Brewerton, S. C., & Blundell, T. L. (2004). Identification of DNA-PK in the arthropods Evidence for the ancient ancestry of vertebrate non-homologous end-joining [Research Support, Non-U.S. Gov't]. *DNA Repair (amst)*, 3(1), 33–41.
- Douglas, P., Cui, X., Block, W. D., Yu, Y., Gupta, S., Ding, Q., Ye, R., Morrice, N., Lees-Miller, S. P., & Meek, K. (2007). The DNA-dependent protein kinase catalytic subunit is phosphorylated in vivo on threonine 3950, a highly conserved amino acid in the protein kinase domain. *Molecular and Cellular Biology*, 27(5), 1581–1591. <https://doi.org/10.1128/mcb.01962-06>
- Douglas, P., Gupta, S., Morrice, N., Meek, K., & Lees-Miller, S. P. (2005). DNA-PK-dependent phosphorylation of Ku70/80 is not required for non-homologous end joining. *DNA Repair (amst)*, 4(9), 1006–1018. <https://doi.org/10.1016/j.dnarep.2005.05.003>
- Douglas, P., Sapkota, G. P., Morrice, N., Yu, Y., Goodarzi, A. A., Merkle, D., Meek, K., Alessi, D. R., & Lees-Miller, S. P. (2002). Identification of in vitro and in vivo phosphorylation sites in the catalytic subunit of the DNA-dependent protein kinase. *The Biochemical Journal*, 368(Pt 1), 243–251. <https://doi.org/10.1042/bj20020973>
- Douglas, P., Ye, R., Radhamani, S., Cobban, A., Jenkins, N. P., Bartlett, E., Roveredo, J., Kettenbach, A. N., & Lees-Miller, S. P. (2020). Nocodazole-induced expression and phosphorylation of anillin and other mitotic proteins are decreased in DNA-dependent protein kinase catalytic subunit-deficient cells and rescued by inhibition of the anaphase-promoting complex/cyclosome with proTAME but not apcin. *Molecular and Cellular Biology*. <https://doi.org/10.1128/MCB.00191-19>
- Douglas, P., Ye, R., Trinkle-Mulcahy, L., Neal, J. A., De Wever, V., Morrice, N. A., Meek, K., & Lees-Miller, S. P. (2014). Polo-like kinase 1 (PLK1) and protein phosphatase 6 (PP6) regulate DNA-dependent protein kinase catalytic subunit (DNA-PKcs) phosphorylation in mitosis. *Bioscience Reports*, 34(3), e00113. <https://doi.org/10.1042/bsr20140051>
- Douglas, P., Zhong, J., Ye, R., Moorhead, G. B., Xu, X., & Lees-Miller, S. P. (2010). Protein phosphatase 6 interacts with the DNA-dependent protein kinase catalytic subunit and dephosphorylates gamma-H2AX. *Molecular and Cellular Biology*, 30(6), 1368–1381. <https://doi.org/10.1128/mcb.00741-09>
- Dvir, A., Peterson, S. R., Knuth, M. W., Lu, H., & Dynan, W. S. (1992). Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II [In Vitro Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. *Proc Natl Acad Sci U S A*, 89(24), 11920–11924.
- Dvir, A., Stein, L. Y., Calore, B. L., & Dynan, W. S. (1993). Purification and characterization of a template-associated protein kinase that phosphorylates RNA polymerase II [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. *Journal of Biological Chemistry*, 268(14), 10440–10447.
- Elias-Villalobos, A., Fort, P., & Helmlinger, D. (2019). New insights into the evolutionary conservation of the sole PIKK pseudokinase Tra1/TRRAP. *Biochemical Society Transactions*. <https://doi.org/10.1042/bst20180496>
- Falck, J., Coates, J., & Jackson, S. P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature*, 434(7033), 605–611.
- Gell, D., & Jackson, S. P. (1999). Mapping of protein-protein interactions within the DNA-dependent protein kinase complex [Research Support, Non-U.S. Gov't]. *Nucleic Acids Research*, 27(17), 3494–3502.
- Ghosh, D., & Raghavan, S. C. (2021). Nonhomologous end joining: New accessory factors fine tune the machinery. *Trends in Genetics*. <https://doi.org/10.1016/j.tig.2021.03.001>
- Goodarzi, A. A., Yu, Y., Riballo, E., Douglas, P., Walker, S. A., Ye, R., Härer, C., Marchetti, C., Morrice, N., Jeggo, P. A., & Lees-Miller, S. P. (2006). DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *EMBO Journal*, 25(16), 3880–3889. <https://doi.org/10.1038/sj.emboj.7601255>
- Goodwin, J. F., Kothari, V., Drake, J. M., Zhao, S., Dylgjeri, E., Dean, J. L., Schiewer, M. J., McNair, C., Jones, J. K., Aytes, A., Magee, M. S., Snook, A. E., Zhu, Z., Den, R. B., Birbe, R. C., Gomella, L. G., Graham, N. A., Vashisht, A. A., Wohlschlegel, J. A., ... Knudsen, K. E. (2015). DNA-PKcs-mediated transcriptional regulation drives prostate cancer progression and metastasis. *Cancer Cell*, 28(1), 97–113. <https://doi.org/10.1016/j.ccell.2015.06.004>
- Gottlieb, T. M., & Jackson, S. P. (1993). The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen. *Cell*, 72(1), 131–142.
- Graham, T. G. W., Carney, S. M., Walter, J. C., & Loparo, J. J. (2018). A single XLF dimer bridges DNA ends during nonhomologous end joining. *Nature Structural & Molecular Biology*, 25(9), 877–884. <https://doi.org/10.1038/s41594-018-0120-y>

- Graham, T. G., Walter, J. C., & Loparo, J. J. (2016). Two-stage synopsis of DNA ends during non-homologous end joining. *Molecular Cell*, 61(6), 850–858. <https://doi.org/10.1016/j.molcel.2016.02.010>
- Hammel, M., Rosenberg, D. J., Bierma, J., Hura, G. L., Thapar, R., Lees-Miller, S. P., & Tainer, J. A. (2020). Visualizing functional dynamicity in the DNA-dependent protein kinase holoenzyme DNA-PK complex by integrating SAXS with cryo-EM. *Progress in Biophysics and Molecular Biology*. <https://doi.org/10.1016/j.pbiomolbio.2020.09.003>
- Hammel, M., & Tainer, J. A. (2021). X-ray scattering reveals disordered linkers and dynamic interfaces in mechanisms for DNA double-strand break repair complexes controlling cell and cancer biology. *Protein Science*. <https://doi.org/10.1002/pro.4133> in press.
- Hammel, M., Yu, Y., Mahaney, B. L., Cai, B., Ye, R., Phipps, B. M., Rambo, R. P., Hura, G. L., Pelikan, M., So, S., Abolfath, R. M., Chen, D. J., Lees-Miller, S. P., & Tainer, J. A. (2010). Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *Journal of Biological Chemistry*, 285(2), 1414–1423. <https://doi.org/10.1074/jbc.M109.065615>
- Hammel, M., Yu, Y., Radhakrishnan, S. K., Chokshi, C., Tsai, M. S., Matsumoto, Y., Kuzdovich, M., Remesh, S. G., Fang, S., Tomkinson, A. E., Lees-Miller, S. P., & Tainer, J. A. (2016). An intrinsically disordered APLF Links Ku, DNA-PKcs, and XRCC4-DNA ligase IV in an extended flexible non-homologous end joining complex. *Journal of Biological Chemistry*, 291(53), 26987–27006. <https://doi.org/10.1074/jbc.M116.751867>
- 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). *Journal of Molecular Biology*, 335(2), 573–582.
- Hartley, K. O., Gell, D., Smith, G. C., Zhang, H., Divecha, N., Connelly, M. A., Admon, A., Lees-Miller, S. P., Anderson, C. W., & Jackson, S. P. (1995). DNA-dependent protein kinase catalytic subunit: A relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. *Cell*, 82(5), 849–856.
- Hepburn, M., Saltzberg, D. J., Lee, L., Fang, S., Atkinson, C., Strynadka, N. C. J., Sali, A., Lees-Miller, S. P., & Schriemer, D. C. (2020). The active DNA-PK holoenzyme occupies a tensed state in a staggered synaptic complex. *Structure*. <https://doi.org/10.1016/j.str.2020.12.006>
- Hitomi, K., Iwai, S., & Tainer, J. A. (2007). The intricate structural chemistry of base excision repair machinery: Implications for DNA damage recognition, removal, and repair. *DNA Repair (amst)*, 6(4), 410–428. <https://doi.org/10.1016/j.dnarep.2006.10.004>
- Hung, P. J., Johnson, B., Chen, B. R., Byrum, A. K., Bredemeyer, A. L., Yewdell, W. T., Johnson, T. E., Lee, B. J., Deivasigamani, S., Hindi, I., Amatya, P., Gross, M. L., Paull, T. T., Pisapia, D. J., Chaudhuri, J., Petrini, J. J. H., Mosammaparast, N., Amarasinghe, G. K., Zha, S., ... Sleckman, B. P. (2018). MRI Is a DNA damage response adaptor during classical non-homologous end joining. *Molecular Cell*, 71(2), 332–342.e338. <https://doi.org/10.1016/j.molcel.2018.06.018>
- Jackson, S. P., MacDonald, J. J., Lees-Miller, S., & Tjian, R. (1990). GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. *Cell*, 63(1), 155–165.
- Jette, N., & Lees-Miller, S. P. (2015). The DNA-dependent protein kinase: A multifunctional protein kinase with roles in DNA double strand break repair and mitosis. *Progress in Biophysics and Molecular Biology*, 117(2–3), 194–205. <https://doi.org/10.1016/j.pbiomolbio.2014.12.003>
- Jiang, W., Crowe, J. L., Liu, X., Nakajima, S., Wang, Y., Li, C., Lee, B. J., Dubois, R. L., Liu, C., Yu, X., Lan, L., & Zha, S. (2015). Differential phosphorylation of DNA-PKcs regulates the interplay between end-processing and end-ligation during nonhomologous end-joining. *Molecular Cell*, 58(1), 172–185. <https://doi.org/10.1016/j.molcel.2015.02.024>
- Ju, B. G., Lunyak, V. V., Perissi, V., Garcia-Bassets, I., Rose, D. W., Glass, C. K., & Rosenfeld, M. G. (2006). A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. *Science*, 312(5781), 1798–1802. <https://doi.org/10.1126/science.1127196>
- Kaminski, A. M., Bebenek, K., Pedersen, L. C., & Kunkel, T. A. (2020). DNA polymerase mu: An inflexible scaffold for substrate flexibility. *DNA Repair (amst)*, 93, 102932. <https://doi.org/10.1016/j.dnarep.2020.102932>
- Lee, K. J., Lin, Y. F., Chou, H. Y., Yajima, H., Fattah, K. R., Lee, S. C., & Chen, B. P. (2011). Involvement of DNA-dependent protein kinase in normal cell cycle progression through mitosis. *Journal of Biological Chemistry*, 286(14), 12796–12802. <https://doi.org/10.1074/jbc.M110.212969>
- Lees-Miller, J. P., Cobban, A., Katsonis, P., Bacolla, A., Tsutakawa, S. E., Hammel, M., Meek, K., Anderson, D. W., Lichtarge, O., Tainer, J. A., & Lees-Miller, S. P. (2020). Uncovering DNA-PKcs ancient phylogeny, unique sequence motifs and insights for human disease. *Progress in Biophysics and Molecular Biology*. <https://doi.org/10.1016/j.pbiomolbio.2020.09.010> in Press.
- Lees-Miller, S. P., & Anderson, C. W. (1989). The human double-stranded DNA-activated protein kinase phosphorylates the 90-kDa heat-shock protein, hsp90 alpha at two NH2-terminal threonine residues. *Journal of Biological Chemistry*, 264(29), 17275–17280.
- Lees-Miller, S. P., Chen, Y. R., & Anderson, C. W. (1990). Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen. *Molecular and Cellular Biology*, 10(12), 6472–6481.
- Lees-Miller, S. P., Sakaguchi, K., Ullrich, S. J., Appella, E., & Anderson, C. W. (1992). Human DNA-activated protein kinase phosphorylates serines 15 and 37 in the amino-terminal transactivation domain of human p53. *Molecular and Cellular Biology*, 12(11), 5041–5049.
- Ma, Y., Pannicke, U., Schwarz, K., & Lieber, M. R. (2002). Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell*, 108(6), 781–794.
- Macrae, C. J., McCulloch, R. D., Ylanko, J., Durocher, D., & Koch, C. A. (2008). APLF (C2orf13) facilitates nonhomologous end-joining and undergoes ATM-dependent hyperphosphorylation following ionizing radiation. *DNA Repair (amst)*, 7(2), 292–302.
- Meek, K. (2020). Activation of DNA-PK by hairpinned DNA ends reveals a stepwise mechanism of kinase activation. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkaa614>
- Meek, K., Douglas, P., Cui, X., Ding, Q., & Lees-Miller, S. P. (2007). trans Autophosphorylation at DNA-dependent protein kinase's two major autophosphorylation site clusters facilitates end processing but not end joining. *Molecular and Cellular Biology*, 27(10), 3881–3890. <https://doi.org/10.1128/mcb.02366-06>
- Menon, V., & Povirk, L. F. (2016). End-processing nucleases and phosphodiesterases: An elite supporting cast for the non-homologous end joining pathway of DNA double-strand break repair. *DNA Repair (amst)*, 43, 57–68. <https://doi.org/10.1016/j.dnarep.2016.05.011>
- Neal, J. A., Dang, V., Douglas, P., Wold, M. S., Lees-Miller, S. P., & Meek, K. (2011). Inhibition of homologous recombination by DNA-dependent protein kinase requires kinase activity, is titratable, and is modulated by autophosphorylation. *Molecular and*

- Cellular Biology*, 31(8), 1719–1733. <https://doi.org/10.1128/mcb.01298-10>
- Nguyen, T. T., Grimm, S. A., Bushel, P. R., Li, J., Li, Y., Bennett, B. D., Lavender, C. A., Ward, J. M., Fargo, D. C., Anderson, C. W., Li, L., Resnick, M. A., & Menendez, D. (2018). Revealing a human p53 universe. *Nucleic Acids Research*, 46(16), 8153–8167. <https://doi.org/10.1093/nar/gky720>
- Nguyen, T. A., Menendez, D., Resnick, M. A., & Anderson, C. W. (2014). Mutant TP53 posttranslational modifications: Challenges and opportunities. *Human Mutation*, 35(6), 738–755. <https://doi.org/10.1002/humu.22506>
- Ochi, T., Gu, X., & Blundell, T. L. (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(4), 672–679. <https://doi.org/10.1016/j.str.2013.02.014>
- Ochi, T., Sibanda, B. L., Wu, Q., Chirgadze, D. Y., Bolanos-Garcia, V. M., & Blundell, T. L. (2010). Structural biology of DNA repair: Spatial organisation of the multicomponent complexes of nonhomologous end joining. *J Nucleic Acids*. <https://doi.org/10.4061/2010/621695>
- Ochi, T., Wu, Q., & Blundell, T. L. (2014). The spatial organization of non-homologous end joining: From bridging to end joining. *DNA Repair (amst)*, 17(100), 98–109. <https://doi.org/10.1016/j.dnarep.2014.02.010>
- Ochi, T., Wu, Q., Chirgadze, D. Y., Grossmann, J. G., Bolanos-Garcia, V. M., & Blundell, T. L. (2012). Structural insights into the role of domain flexibility in human DNA ligase IV. *Structure*, 20(7), 1212–1222. <https://doi.org/10.1016/j.str.2012.04.012>
- O'Neill, T., Dwyer, A. J., Ziv, Y., Chan, D. W., Lees-Miller, S. P., Abraham, R. H., Lai, J. H., Hill, D., Shiloh, Y., Cantley, L. C., & Rathbun, G. A. (2000). Utilization of oriented peptide libraries to identify substrate motifs selected by ATM. *Journal of Biological Chemistry*, 275(30), 22719–22727. <https://doi.org/10.1074/jbc.M001002200>
- Pannunzio, N. R., Watanabe, G., & Lieber, M. R. (2018). Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *Journal of Biological Chemistry*, 293(27), 10512–10523. <https://doi.org/10.1074/jbc.TM117.000374>
- Park, S. J., Gavrilova, O., Brown, A. L., Soto, J. E., Bremner, S., Kim, J., Xu, X., Yang, S., Um, J. H., Koch, L. G., Britton, S. L., Lieber, R. L., Philp, A., Baar, K., Kohama, S. G., Abel, E. D., Kim, M. K., & Chung, J. H. (2017). DNA-PK promotes the mitochondrial, metabolic, and physical decline that occurs during aging. *Cell Metabolism*, 25(5), 1135–1146.e1137. <https://doi.org/10.1016/j.cmet.2017.04.008>
- Pears, C. J., & Lakin, N. D. (2014). Emerging models for DNA repair: *Dictyostelium discoideum* as a model for nonhomologous end-joining. *DNA Repair (amst)*, 17, 121–131. <https://doi.org/10.1016/j.dnarep.2014.01.008>
- Petersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, 30(1), 70–82. <https://doi.org/10.1002/pro.3943>
- Ramsden, D. A., & Asagoshi, K. (2012). DNA polymerases in nonhomologous end joining: are there any benefits to standing out from the crowd? [Research Support, U.S. Gov't, P.H.S. Review]. *Environmental and Molecular Mutagenesis*, 53(9), 741–751. <https://doi.org/10.1002/em.21725>
- Schlam-Babayov, S., Bensimon, A., Harel, M., Geiger, T., Aebersold, R., Ziv, Y., & Shiloh, Y. (2021). Phosphoproteomics reveals novel modes of function and inter-relationships among PIKKs in response to genotoxic stress. *EMBO Journal*, 40(2), e104400. <https://doi.org/10.15252/embj.2020104400>
- Segal-Raz, H., Mass, G., Baranes-Bachar, K., Lerenthal, Y., Wang, S. Y., Chung, Y. M., Ziv-Lehrman, S., Strom, C. E., Helleday, T., Hu, M. C., Chen, D. J., & Shiloh, Y. (2011). ATM-mediated phosphorylation of polynucleotide kinase/phosphatase is required for effective DNA double-strand break repair. *EMBO Reports*, 12(7), 713–719. <https://doi.org/10.1038/embor.2011.96>
- Shao, Z., Flynn, R. A., Crowe, J. L., Zhu, Y., Liang, J., Jiang, W., Aryan, F., Aoude, P., Bertozzi, C. R., Estes, V. M., Lee, B. J., Bhagat, G., Zha, S., & Calo, E. (2020). DNA-PKcs has KU-dependent function in rRNA processing and haematopoiesis. *Nature*. <https://doi.org/10.1038/s41586-020-2041-2>
- Sharif, H., Li, Y., Dong, Y., Dong, L., Wang, W. L., Mao, Y., & Wu, H. (2017). Cryo-EM structure of the DNA-PK holoenzyme. *Proc Natl Acad Sci U S A*, 114(28), 7367–7372. <https://doi.org/10.1073/pnas.1707386114>
- Shibata, A., Moiani, D., Arvai, A. S., Perry, J., Harding, S. M., Genoio, M. M., Maity, R., van Rossum-Fikkert, S., Kertokallio, A., Romoli, F., Ismail, A., Ismalaj, E., Petricci, E., Neale, M. J., Bristow, R. G., Masson, J. Y., Wyman, C., Jeggo, P. A., & Tainer, J. A. (2014). DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Molecular Cell*, 53(1), 7–18. <https://doi.org/10.1016/j.molcel.2013.11.003>
- Sibanda, B. L., Chirgadze, D. Y., Ascher, D. B., & Blundell, T. L. (2017). DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. *Science*, 355(6324), 520–524. <https://doi.org/10.1126/science.aak9654>
- Sibanda, B. L., Chirgadze, D. Y., & Blundell, T. L. (2010). Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats [Research Support, Non-U.S. Gov't]. *Nature*, 463(7277), 118–121. <https://doi.org/10.1038/nature08648>
- Sipley, J. D., Menninger, J. C., Hartley, K. O., Ward, D. C., Jackson, S. P., & Anderson, C. W. (1995). Gene for the catalytic subunit of the human DNA-activated protein kinase maps to the site of the XRCC7 gene on chromosome 8. *Proc Natl Acad Sci U S A*, 92(16), 7515–7519.
- Thapar, R., Wang, J. L., Hammel, M., Ye, R., Liang, K., Sun, C., Hnizda, A., Liang, S., Maw, S. S., Lee, L., Villarreal, H., Forrester, I., Fang, S., Tsai, M. S., Blundell, T. L., Davis, A. J., Lin, C., Lees-Miller, S. P., Strick, T. R., & Tainer, J. A. (2021). Mechanism of efficient double-strand break repair by a long non-coding RNA. *Nucleic Acids Research*, 49(2), 1199–1200. <https://doi.org/10.1093/nar/gkaa1233>
- Uematsu, N., Weterings, E., Yano, K., Morotomi-Yano, K., Jakob, B., Taucher-Scholz, G., Mari, P. O., van Gent, D. C., Chen, B. P., & Chen, D. J. (2007). Autophosphorylation of DNA-PKcs regulates its dynamics at DNA double-strand breaks [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. *Journal of Cell Biology*, 177(2), 219–229. <https://doi.org/10.1083/jcb.200608077>
- Walker, A. I., Hunt, T., Jackson, R. J., & Anderson, C. W. (1985). Double-stranded DNA induces the phosphorylation of several proteins including the 90 000 mol. wt. heat-shock protein in animal cell extracts. *EMBO Journal*, 4(1), 139–145.
- Wang, C., & Lees-Miller, S. P. (2013). Detection and repair of ionizing radiation-induced DNA double strand breaks: New developments in nonhomologous end joining. *International Journal of Radiation Oncology Biology Physics*, 86(3), 440–449. <https://doi.org/10.1016/j.ijrobp.2013.01.011>
- 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. *Nature Structural & Molecular Biology*, 25(6), 482–487. <https://doi.org/10.1038/s41594-018-0065-1>
- Wang, Y. G., Nnakwe, C., Lane, W. S., Modesti, M., & Frank, K. M. (2004). Phosphorylation and regulation of DNA ligase IV stability by DNA-dependent protein kinase [Research Support,

- Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. *Journal of Biological Chemistry*, 279(36), 37282–37290. <https://doi.org/10.1074/jbc.M401217200>
- Weinfeld, M., Mani, R. S., Abdou, I., Aceytuno, R. D., & Glover, J. N. (2011). Tidying up loose ends: The role of polynucleotide kinase/phosphatase in DNA strand break repair. *Trends in Biochemical Sciences*, 36(5), 262–271. <https://doi.org/10.1016/j.tibs.2011.01.006>
- Williams, G. J., Hammel, M., Radhakrishnan, S. K., Ramsden, D., Lees-Miller, S. P., & Tainer, J. A. (2014). Structural insights into NHEJ: Building up an integrated picture of the dynamic DSB repair super complex, one component and interaction at a time. *DNA Repair (amst)*, 17, 110–120. <https://doi.org/10.1016/j.dnarep.2014.02.009>
- Xing, M., & Oksenyich, V. (2019). Genetic interaction between DNA repair factors PAXX, XLF, XRCC4 and DNA-PKcs in human cells. *FEBS Open Bio*, 9(7), 1315–1326. <https://doi.org/10.1002/2211-5463.12681>
- Yajima, H., Lee, K. J., & Chen, B. P. (2006). ATR-dependent phosphorylation of DNA-dependent protein kinase catalytic subunit in response to UV-induced replication stress. *Molecular and Cellular Biology*, 26(20), 7520–7528.
- Yang, H., Yao, F., Marti, T. M., Schmid, R. A., & Peng, R. W. (2020). Beyond DNA repair: DNA-PKcs in tumor metastasis, metabolism and immunity. *Cancers (basel)*. <https://doi.org/10.3390/cancers12113389>
- Yano, K., & Chen, D. J. (2008). Live cell imaging of XLF and XRCC4 reveals a novel view of protein assembly in the non-homologous end-joining pathway. *Cell Cycle*, 7(10), 1321–1325.
- Yin, X., Liu, M., Tian, Y., Wang, J., & Xu, Y. (2017). Cryo-EM structure of human DNA-PK holoenzyme. *Cell Research*, 27(11), 1341–1350. <https://doi.org/10.1038/cr.2017.110>
- Yoo, S., & Dynan, W. S. (1999). Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein [Research Support, Non-U.S. Gov't]. *Nucleic Acids Research*, 27(24), 4679–4686.
- Yu, Y., Mahaney, B. L., Yano, K., Ye, R., Fang, S., Douglas, P., Chen, D. J., & Lees-Miller, S. P. (2008). DNA-PK and ATM phosphorylation sites in XLF/Cernunnos are not required for repair of DNA double strand breaks. *DNA Repair (amst)*, 7(10), 1680–1692. <https://doi.org/10.1016/j.dnarep.2008.06.015>
- Yu, Y., Wang, W., Ding, Q., Ye, R., Chen, D., Merkle, D., Schriemer, D., Meek, K., & Lees-Miller, S. P. (2003). DNA-PK phosphorylation sites in XRCC4 are not required for survival after radiation or for V(D)J recombination. *DNA Repair (amst)*, 2(11), 1239–1252.
- Yue, X., Bai, C., Xie, D., Ma, T., & Zhou, P. K. (2020). DNA-PKcs: A multi-faceted player in DNA damage response. *Frontiers in Genetics*, 11, 607428. <https://doi.org/10.3389/fgene.2020.607428>
- Zhao, B., Rothenberg, E., Ramsden, D. A., & Lieber, M. R. (2020a). The molecular basis and disease relevance of non-homologous DNA end joining. *Nature Reviews Molecular Cell Biology*, 21(12), 765–781. <https://doi.org/10.1038/s41580-020-00297-8>
- Zhao, B., Watanabe, G., & Lieber, M. R. (2020b). Polymerase mu in non-homologous DNA end joining: Importance of the order of arrival at a double-strand break in a purified system. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkaa094>
- Zhou, Y., Lee, J. H., Jiang, W., Crowe, J. L., Zha, S., & Paull, T. T. (2017). Regulation of the DNA damage response by DNA-PKcs inhibitory phosphorylation of ATM. *Molecular Cell*, 65(1), 91–104. <https://doi.org/10.1016/j.molcel.2016.11.004>
- Zolner, A. E., Abdou, I., Ye, R., Mani, R. S., Fanta, M., Yu, Y., Douglas, P., Tahbaz, N., Fang, S., Dobbs, T., Wang, C., Morrice, N., Hendzel, M. J., Weinfeld, M., & Lees-Miller, S. P. (2011). Phosphorylation of polynucleotide kinase/phosphatase by DNA-dependent protein kinase and ataxia-telangiectasia mutated regulates its association with sites of DNA damage. *Nucleic Acids Research*, 39(21), 9224–9237. <https://doi.org/10.1093/nar/gkr647>