

Title: DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand-break repair

Authors: Bancinyane L. Sibanda¹, Dimitri Y. Chirgadze¹, David B. Ascher^{1,#} and Tom L. Blundell^{1,*}

Affiliations:

¹ Department of Biochemistry, University of Cambridge, Old Addenbrooke's site, 80 Tennis Court Road, Cambridge, CB2 1GA, United Kingdom.

Present address: Department of Biochemistry and Molecular Biology, University of Melbourne, 30 Flemington Road, Parkville, 3052, Australia.

*Correspondence to: tlb20@cam.ac.uk

Abstract:

DNA-dependent protein kinase catalytic subunit (DNA-PKcs) is a central component of non-homologous end joining (NHEJ), repairing DNA double-strand breaks that would otherwise lead to apoptosis or cancer. We have solved its structure in complex with the C-terminal peptide of Ku80 at 4.3 Å resolution using X-ray crystallography. We show that the 4128 amino-acid structure comprises three large structural units: the N-terminal unit, Circular Cradle, and the Head. Conformational differences between the two molecules in the asymmetric unit are correlated with changes in accessibility of the kinase active site, consistent with an allosteric mechanism to bring about kinase activation. The location of KU80ct₁₉₄ in vicinity of the BRCA1 binding site suggests competition with BRCA1, leading to pathway selection between NHEJ and homologous recombination.

One Sentence Summary: The structure shows how components of non-homologous end joining modulate kinase activation in DNA double-strand-break repair.

Main Text:

DNA double-strand breaks (DSBs), which threaten genomic stability and must be repaired promptly, may be caused either by endogenous agents, such as reactive oxygen species and failed replication forks, or by exogenous ionizing radiation and chemicals. DSBs can be repaired by using non-homologous end joining (NHEJ) (1), a process used also in V(D)J recombination (2), or by homologous recombination (HR) (3), which provides more accurate repair by retrieving lost information from the sister chromatid. DNA-dependent protein kinase catalytic subunit (DNA-PKcs), which is regulated by extensive auto-phosphorylation (4-7), and the KU70/80 heterodimer (8) play major roles in initiating NHEJ, allowing recruitment of Artemis, XRCC4, XRCC4/DNA ligase IV, XLF and PAXX. Structural information on DNA-PKcs has been limited to our low-resolution (6.6 Å) crystal structure (9) (fig. S13) and cryo-EM structures (> 13 Å) (10-13), limiting our understanding of intermolecular interactions necessary for NHEJ.

We now report crystals of seleno-methionine (Se-Met) labelled DNA-PKcs complexed with native KU80ct₁₉₄ (KU80 residues 539-732) diffracting to 4.3 Å resolution. The resulting electron density map allows chain tracing from N- to C-terminus, using 228 evenly spaced Se-Mets in two molecules of the asymmetric unit to check sequence registration (figs. S6 and S7), and definition of ~90% of the 4,128 amino acids in each chain. DNA-PKcs folds into three well-defined large structural units, within which motifs resembling HEAT repeats (Huntingtin, Elongation Factor 3, PP2 A and TOR1; helix-turn-helix motifs that recur throughout the structure, see Fig. 1A) give rise to supersecondary structures with continuous hydrophobic cores (Fig. 1A). As detailed in Fig. 1B (also fig. S14) the structural units comprise the N-terminal region (38 α -helices, residues 1 to 892, four supersecondary

structures: N1-N4), the Circular Cradle (85 α -helices, residues 893-2801; five supersecondary structures: CC1-CC5) and the C-terminal Head (64 α -helices, residues 2802-4128, FAT, FRB, kinase and FATC) (table S7). Conserved regions and phosphorylation sites are shown in Fig. 2A and fig. S11a. Also shown are acetylation sites (fig. S11b), the predicted nuclear localization sites (NLS) (fig. 11c) and the point of cleavage by Caspase 3 to deactivate DNA-PKcs (fig 11d).

The bilobal kinase (3676 to 4100) (Fig. 2B and C) is similar to that of other PI3-kinases (18, 19). It comprises a well conserved P-loop (3729-3735), catalytic loop (3919-3927) and activation loop (3940-3963) (Fig. 2C, figs. S15 and S16). There is also a conserved metal coordinating residues N3927 and D3941 in the C-lobe, the catalytic base D3922 responsible for substrate activation and a conserved histidine (H3924) that stabilizes the transition state analogue as in mTOR.

As in mTOR there is restricted access to the active site. The FAT region (FRAP, ATM, IRRAP; residues 2802-3564; with supersecondary structures FR1-5), conserved in DNA-PKcs orthologues but not in mTOR (fig. S15), folds around the kinase (fig. S17). The four α -helical bundle (residues 3582-3675) known as FRB (FKBP12-rapamycin-binding) (14) and inserted after the kinase N-terminal helix, is poorly conserved, suggesting that it does not mediate rapamycin binding in DNA-PKcs as it does in mTOR (fig. S18). An additional small region known as FATC (15) (residues 4101-4128) follows the kinase. The crystal structure of DNA-PKcs aligns well with the cryo-EM maps of DNA-PKcs (12, 16, 17), with additional density for more mobile regions observed at this higher resolution (fig. S19). The proximity of FRB to α -hairpin 1 (3836-3872) (Fig. 2B) and FATC to α -hairpin 2 (3995-4053) (Fig. 2C) results in a

deep cleft in which the substrate binds. A mutation in mice leading to severe combined immune deficiency (SCID), gives rise to a termination codon (20), and results in loss of the 83 C-terminal residues (Fig. 2C), which would be expected to destabilise the C-terminal domain leading to negligible activity. This led to the idea that mutations in DNA-PKcs would also be found in human SCID, a prediction confirmed a decade later (21).

An insertion in α -hairpin 1 in DNA-PKcs (3836-3872) (Fig. 2B), adjacent to the FATC region, allows the hairpin to reach the active center, unlike the upright and shorter loop (LBE) in mTOR that binds mammalian lethal with SEC13 protein 8 (mLST8) (19). α -hairpin 2 (residues 3995-4053), adjacent to FRB, blocks the entrance into the active site cleft and extends into the active site cleft (Fig. 2C). α -hairpin 3 (4059-4082) is unique to DNA-PKcs, and obscures the activation loop (Fig. 2C), thus burying the conserved T3950 auto-phosphorylation site which deactivates the kinase (6). In much the same way, the catalytic loop is protected by FATC.

Full access to the active site requires substantial concerted conformational changes. The FRB and the P-loop are moved further into the active site cleft in DNA-PKcs than in the intrinsically activated mTOR structure. Furthermore, comparison of the two molecules in the asymmetric unit reveals that the active site of one molecule (B) is more concealed than the other (A). These differences correlate with much larger changes in conformation and position of N1 and N2 (see movies S1 and S2). Similar but more extensive concerted changes could mediate allosteric interactions between the N-terminal and Circular Cradle units and the kinase, providing a mechanism for ensuring that DNA-PKcs selectively phosphorylates its targets.

The KU70/80 heterodimer binds and recruits DNA-PKcs to DNA ends, through the C-terminal 189 residues of KU80 (22), which contains a globular core of six α -helices (23, 24). Two regions of electron density near CC4 are candidates for the KU80ct₁₉₄ binding. Both sites are outside the cryo-EM electron density maps of apo structures, supporting their assignment to KU80ct₁₉₄ (fig. S19), and cocrystallization of Se-Met-labelled KU80ct₁₉₄ with wild-type DNA-PKcs provided further confirmation (Supplementary Material). Binding site A (Fig. 3), which shows anomalous scattering density for KU80 Se-Met, is likely the C-terminal α -helix of Ku. It lies near the 'PQR' auto-phosphorylation cluster (4, 5) and proposed BRCA1 binding site (25), raising the intriguing hypothesis that BRCA1 binding could compete with KU70/80 to change the pathway from NHEJ to HR. Binding site B (Fig. 3), could be part of KU80ct₁₉₄ between the globular region and the C-terminal helix (site A), which is predicted to have helical propensity. The lack of further density in our anomalous difference maps suggests that the globular region containing the other seleno-methionine is not bound in an ordered fashion.

Although the findings of Weterings and colleagues (26) indicate that the C-terminus of KU80 is not absolutely required for activation of DNA-PKcs, our structure indicates that it likely facilitates recruitment. Comparison to the cryo-EM maps of the DNA-PKcs-Ku70/80 complex (16) reveals that density for Ku70/80 complex is located beneath CC4 and N1 (fig. S19).

There is growing evidence that N1-N3 of DNA-PKcs mediates DNA binding (6, 7, 13, 16, 17, 27). This region, together with the Circular Cradle, forms a ring at the base of the molecule through which KU70/80 may present DNA for repair. This region has higher thermal factors and adopts different conformations between molecules A and B (movies S1 and S2),

confirming flexibility. Our results suggest that the binding of Ku or DNA activate the allosteric mechanism needed for communication between the N-terminal and Circular Cradle to communicate with the kinase in the Head.

The mechanism (fig. S21) may be mediated by two conserved residues, K881 in the extension beyond N4 and E3933 of a β -hairpin in the kinase active center, which forms a salt bridge between the N-terminal unit and the Head (Fig. 2A). Loss of this interaction could trigger conformational changes in nearby highly conserved regions (16, 28, 29) that open up the active site (fig. S21C), exposing T3950, and culminating in full kinase activity. The changes at the active center in concert with the dephosphorylation and phosphorylation of T3950 may be the ON and OFF switch to the full kinase activity.

It is possible that, once the DNA has been repaired, the N-terminus moves away from the Circular Cradle, a move analogous to a gate opening (movies S1 and S2), and thus releases the repaired and ligated DNA. Intriguingly, this movement can be inferred from comparison of the DNA-PKcs crystal structure with the cryo-EM maps of the apo structure (12, 17) (fig. S19).

Most models of the role of DNA-PKcs involve its dimerization as part of a larger complex (30) that guides the relative positions of the two broken ends. Such dimers have been observed in cryo-EM studies (28). We note that, although the DNA-PKcs is a monomer on gel filtration columns, the relative arrangement of the two molecules in the asymmetric unit of the crystals (see Fig. 4), which is consistent with the observed EM dimers (fig. S20), would align

the two broken ends ready for ligation, positioning the DNA near to where the density for the Ku70/80 complex has been observed.

In summary, the DNA-PKcs crystal structure shows that the molecule folds into three well defined units that differ in relative positions in molecules A and B, consistent with allosteric conformational changes likely mediating the regulation of activity. Its active site, which is closed relative to that in the intrinsically active mTOR, indicates that its activity is tightly regulated with the C-terminus of the N-terminal structural unit implicated as the trigger for full kinase activation. The located Se-Met site of the KU80ct₁₉₄ indicates that a portion of this domain binds close to the putative binding site for BRCA1 implying a pathway change to either NHEJ or HR. Furthermore, growing evidence points to the N-terminal region as the putative DNA binding site, a region that could also play a part in the release of the repaired DNA. Dimers, observed in EM and the crystal structure, are well positioned to present the broken DNA ends for repair.

References and Notes:

1. S. E. Critchlow, S. P. Jackson, DNA end-joining: from yeast to man. *Trends Biochem Sci* **23**, 394-398 (1998).
2. D. G. Schatz, V(D)J recombination. *Immunol Rev* **200**, 5-11 (2004).
3. S. C. West, Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol* **4**, 435-445 (2003).
4. B. P. Chen *et al.*, Cell cycle dependence of DNA-dependent protein kinase phosphorylation in response to DNA double strand breaks. *J Biol Chem* **280**, 14709-14715 (2005).
5. X. Cui *et al.*, Autophosphorylation of DNA-dependent protein kinase regulates DNA end processing and may also alter double-strand break repair pathway choice. *Mol Cell Biol* **25**, 10842-10852 (2005).
6. K. Meek, S. P. Lees-Miller, M. Modesti, N-terminal constraint activates the catalytic subunit of the DNA-dependent protein kinase in the absence of DNA or Ku. *Nucleic Acids Res* **40**, 2964-2973 (2012).
7. P. Douglas *et al.*, The DNA-dependent protein kinase catalytic subunit is phosphorylated in vivo on threonine 3950, a highly conserved amino acid in the protein kinase domain. *Mol Cell Biol* **27**, 1581-1591 (2007).
8. J. R. Walker, R. A. Corpina, J. Goldberg, Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607-614 (2001).
9. B. L. Sibanda, D. Y. Chirgadze, T. L. Blundell, Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. *Nature* **463**, 118-121 (2010).

10. C. Y. Chiu, R. B. Cary, D. J. Chen, S. R. Peterson, P. L. Stewart, Cryo-EM imaging of the catalytic subunit of the DNA-dependent protein kinase. *J Mol Biol* **284**, 1075-1081 (1998).
11. K. K. Leuther, O. Hammarsten, R. D. Kornberg, G. Chu, Structure of DNA-dependent protein kinase: implications for its regulation by DNA. *EMBO J* **18**, 1114-1123 (1999).
12. A. Rivera-Calzada, J. D. Maman, L. Spagnolo, L. H. Pearl, O. Llorca, Three-dimensional structure and regulation of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). *Structure* **13**, 243-255 (2005).
13. D. R. Williams, K. J. Lee, J. Shi, D. J. Chen, P. L. Stewart, 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 (2008).
14. J. Choi, J. Chen, S. L. Schreiber, J. Clardy, Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239-242 (1996).
15. S. A. Dames, J. M. Mulet, K. Rathgeb-Szabo, M. N. Hall, S. Grzesiek, The solution structure of the FATC domain of the protein kinase target of rapamycin suggests a role for redox-dependent structural and cellular stability. *J Biol Chem* **280**, 20558-20564 (2005).
16. L. Spagnolo, A. Rivera-Calzada, L. H. Pearl, O. Llorca, Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. *Mol Cell* **22**, 511-519 (2006).
17. S. A. Villarreal, P. L. Stewart, CryoEM and image sorting for flexible protein/DNA complexes. *J Struct Biol* **187**, 76-83 (2014).

18. E. H. Walker, O. Perisic, C. Ried, L. Stephens, R. L. Williams, Structural insights into phosphoinositide 3-kinase catalysis and signalling. *Nature* **402**, 313-320 (1999).
19. H. Yang *et al.*, mTOR kinase structure, mechanism and regulation. *Nature* **497**, 217-223 (2013).
20. R. Araki *et al.*, Nonsense mutation at Tyr-4046 in the DNA-dependent protein kinase catalytic subunit of severe combined immune deficiency mice. *Proc Natl Acad Sci U S A* **94**, 2438-2443 (1997).
21. M. van der Burg, J. J. van Dongen, D. C. van Gent, DNA-PKcs deficiency in human: long predicted, finally found. *Current Opinion in Allergy and Clinical Immunology* **9**, 503-509 (2009).
22. D. Gell, S. P. Jackson, Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic Acids Res* **27**, 3494-3502 (1999).
23. R. Harris *et al.*, The 3D solution structure of the C-terminal region of Ku86 (Ku86CTR). *J Mol Biol* **335**, 573-582 (2004).
24. Z. Zhang *et al.*, Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. *Structure* **12**, 495-502 (2004).
25. A. J. Davis *et al.*, BRCA1 modulates the autophosphorylation status of DNA-PKcs in S phase of the cell cycle. *Nucleic Acids Res* **42**, 11487-11501 (2014).
26. E. Weterings *et al.*, The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends. *Mol Cell Biol* **29**, 1134-1142 (2009).
27. M. Hammel *et al.*, Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J Biol Chem* **285**, 1414-1423 (2010).

28. J. Boskovic *et al.*, Visualization of DNA-induced conformational changes in the DNA repair kinase DNA-PKcs. *EMBO J* **22**, 5875-5882 (2003).
29. T. M. Gottlieb, S. P. Jackson, The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* **72**, 131-142 (1993).
30. N. Jette, S. P. Lees-Miller, The DNA-dependent protein kinase: A multifunctional protein kinase with roles in DNA double strand break repair and mitosis. *Prog Biophys Mol Biol* **117**, 194-205 (2015).
31. S. M. Abmayr, T. Yao, T. Parmely, J. L. Workman, in *Current Protocols in Molecular Biology*. (John Wiley & Sons, Inc., 2001).
32. J. J. Bellizzi, J. Widom, C. W. Kemp, J. Clardy, Producing selenomethionine-labeled proteins with a baculovirus expression vector system. *Structure* **7**, R263-267 (1999).
33. Q. Liu, Z. Zhang, W. A. Hendrickson, Multi-crystal anomalous diffraction for low-resolution macromolecular phasing. *Acta Crystallogr D Biol Crystallogr* **67**, 45-59 (2011).
34. W. Kabsch, Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).
35. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
36. T. C. Terwilliger, Rapid model building of alpha-helices in electron-density maps. *Acta Crystallogr D Biol Crystallogr* **66**, 268-275 (2010).
37. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
38. R. E. Smith, S. C. Lovell, D. F. Burke, R. W. Montalvao, T. L. Blundell, Andante: reducing side-chain rotamer search space during comparative modeling using

- environment-specific substitution probabilities. *Bioinformatics* **23**, 1099-1105 (2007).
39. H. C. Jubb *et al.*, Arpeggio: A web server for calculating and visualising interatomic interactions in protein structures. *Journal of Molecular Biology*. In Press.
 40. L. Willard *et al.*, VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* **31**, 3316-3319 (2003).
 41. J. U. Bowie, R. Luthy, D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, 164-170 (1991).
 42. T. A. Dobbs, J. A. Tainer, S. P. Lees-Miller, A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA Repair (Amst)* **9**, 1307-1314 (2010).
 43. A. Fujimori *et al.*, Identification of four highly conserved regions in DNA-PKcs. *Immunogenetics* **51**, 965-973 (2000).
 44. M. Le Romancer, S. C. Cosulich, S. P. Jackson, P. R. Clarke, Cleavage and inactivation of DNA-dependent protein kinase catalytic subunit during apoptosis in *Xenopus* egg extracts. *J Cell Sci* **109 (Pt 13)**, 3121-3127 (1996).
 45. Q. Song *et al.*, DNA-dependent protein kinase catalytic subunit: a target for an ICE-like protease in apoptosis. *EMBO J* **15**, 3238-3246 (1996).
 46. H. Teraoka *et al.*, CPP32/Yama/apopain cleaves the catalytic component of DNA-dependent protein kinase in the holoenzyme. *FEBS Lett* **393**, 1-6 (1996).

Acknowledgments:

Model coordinates have been deposited in the Protein Data Bank (PDB) under accession number 5LUQ.

We wish to thank Dr Joop Van den Heuvel at Helmholtz-Zentrum particularly Dr Volker Jager who prepared the HeLa cells. We are also grateful to the ESRF (beamlines ID29, ID14-1 and ID14-2 in Grenoble, France) and DIAMOND (beamlines I04 and I03 in Oxford, UK) synchrotron sources where the data were collected. This work was funded by the Wellcome-Trust 093167/Z/10/Z.

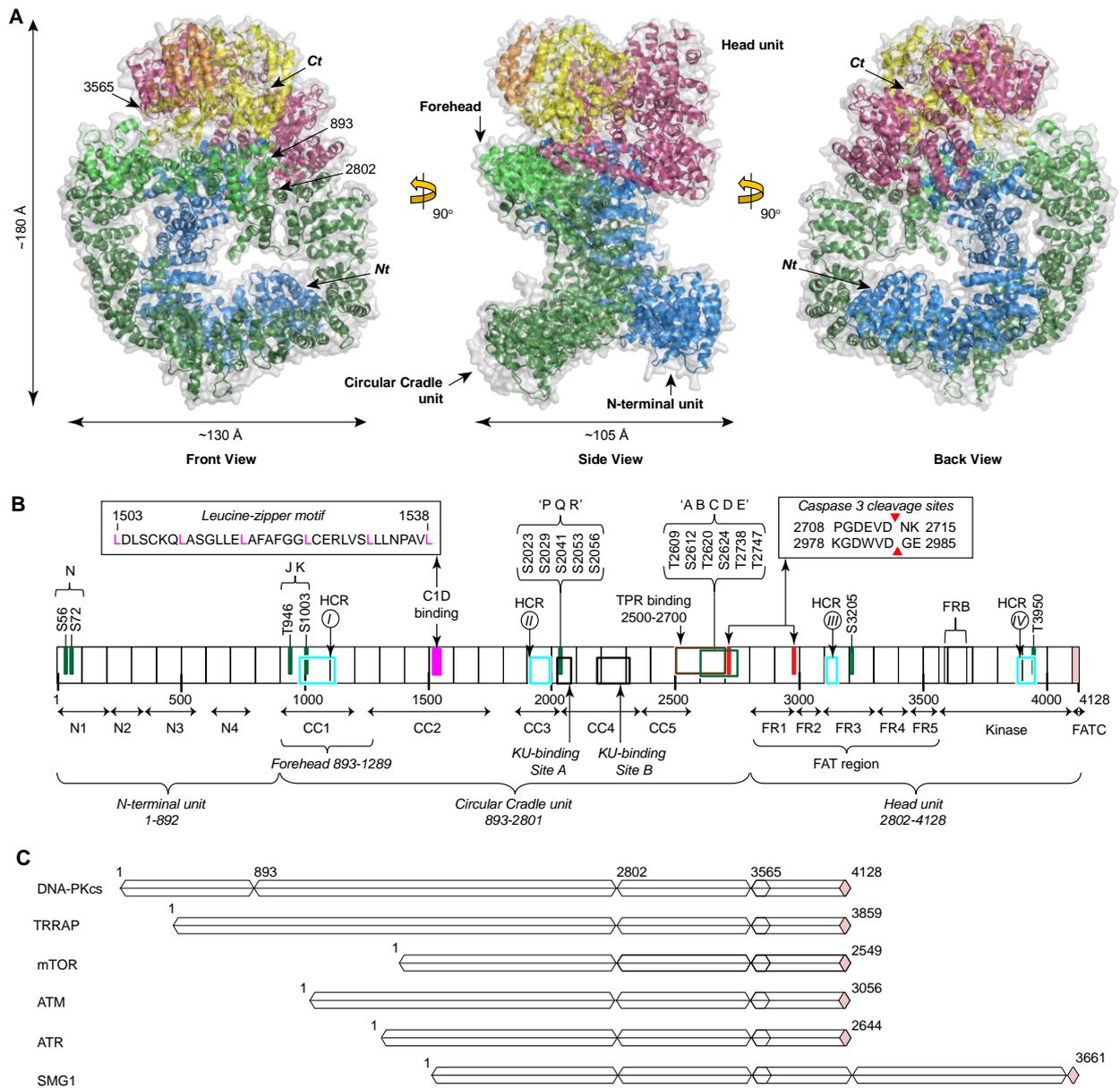


Fig. 1. The overall structure of DNA-PKcs. **(A)** Structural units of DNA-PKcs. N-terminal in blue, Circular Cradle in green, Head comprising FAT region in pink, kinase in yellow and the FATC in light pink. Also shown are the Forehead in light green and FRB (FKBP12-rapamycin-binding) in orange. The surface shown is the solvent accessible surface area calculated using Pymol. **(B)** A schematic representation highlighting the three units of DNA-PKcs and their supersecondary structural components: the N-terminal, composed of four supersecondary alpha helical structures (N1-N4); the Circular Cradle composed of five supersecondary alpha helical structures (CC1-CC5); and the Head, which contains the FAT region (FR1-FR4), kinase,

FRB and FATC regions. Highly conserved regions (HCR), KU80 binding area (Sites A and B) and interacting areas for other proteins, auto-phosphorylation sites and Caspase 3 cleavage sites are shown above the schematic. (C) Domain organization compared to other PI3-K family members, TRRAP (transformation/transcription domain associated protein), mTOR (mammalian Target of Rapamycin), ATM (ataxia-telangiectasia mutated), ATR (ATM and Rad3 related), SMG1 (human suppressor of morphogenesis in genitalia 1) colour code as in (A) with the insertion in SMG1 in lilac.

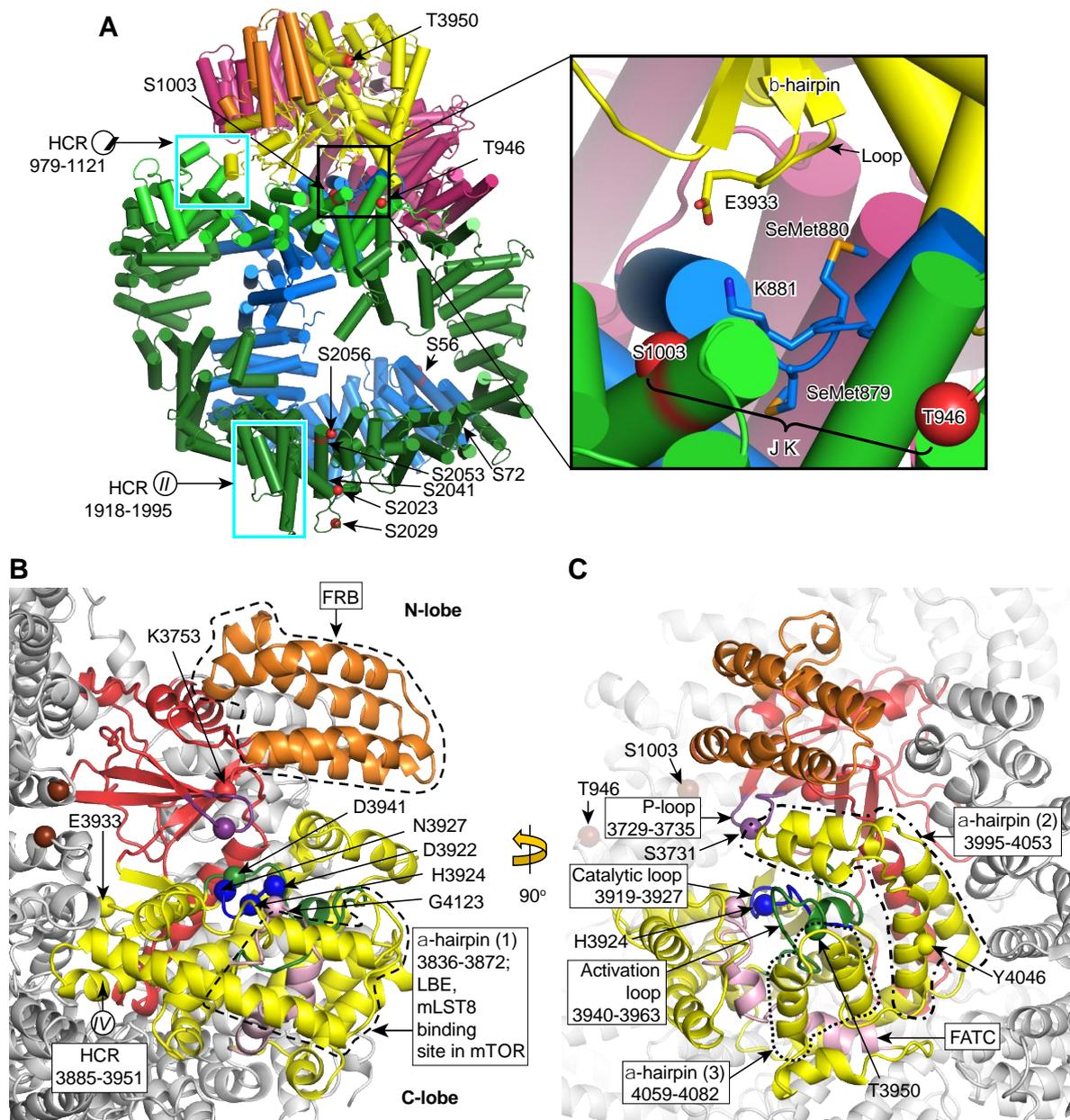


Fig. 2. The active site of DNA-PKcs/Mechanism. **(A)** The colour scheme is the same as in Fig.

1. The extension at end of N-terminal unit interacts with the active centre; the contact between E3933 and K881 is highlighted; the auto-phosphorylated residues, HCR (I) 979-1121, HCR (II) 1918-1995 and the JK phosphorylation cluster (T946 and S1003), thought to promote HR, are also highlighted. **(B)** The kinase with the N-lobe shown in red, the C-lobe yellow and HCR (IV) marked. The P-loop is shown in purple, the activation loop in green, the catalytic loop in dark blue, the FRB (orange) and the α -hairpin 1 carrying the LBE site in

mTOR shown in dashed lines. Also indicated is G4123 on the FATC region positioned on the edge of the active site cleft. (C) T946, S1003, T3950 auto-phosphorylation sites are labelled. The SCID mutation at Y4046 that leads to a termination codon resulting in the loss of the C-terminal 83 residues is also shown; this would remove a substantive part of the C-terminal lobe and very likely allosterically affect the active site cleft, so drastically reducing the kinase activity. The dashed lines in (C) depict the supersecondary structures that enclose the active site.

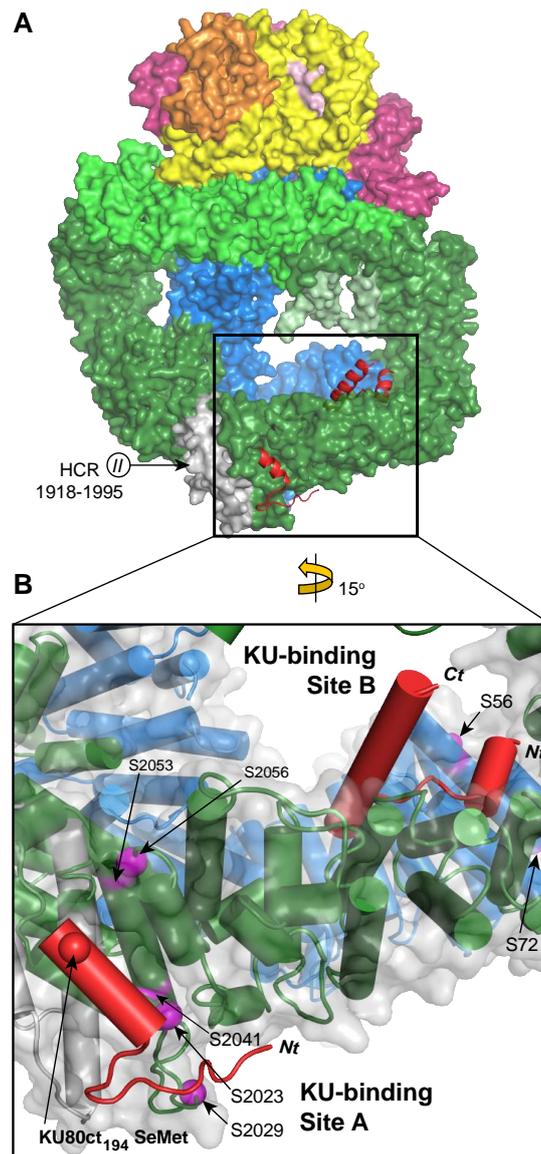


Fig. 3. KU70/80 binding area. **(A)** KU80ct194 (sites A and B in red) shown bound to DNA-PKcs, coloured according to structural units with N-terminal in blue and Circular Cradle in green as in Fig. 1, and in grey the highly conserved region HCR II that is close to the 'PQR' auto-phosphorylation site. **(B)** Shows magnified the binding sites A and B for KU80ct194, where site A is near the 'PQR' auto-phosphorylation cluster with its residues (S2023, S2029, S2041, S2053, S2056), shown as purple spheres.

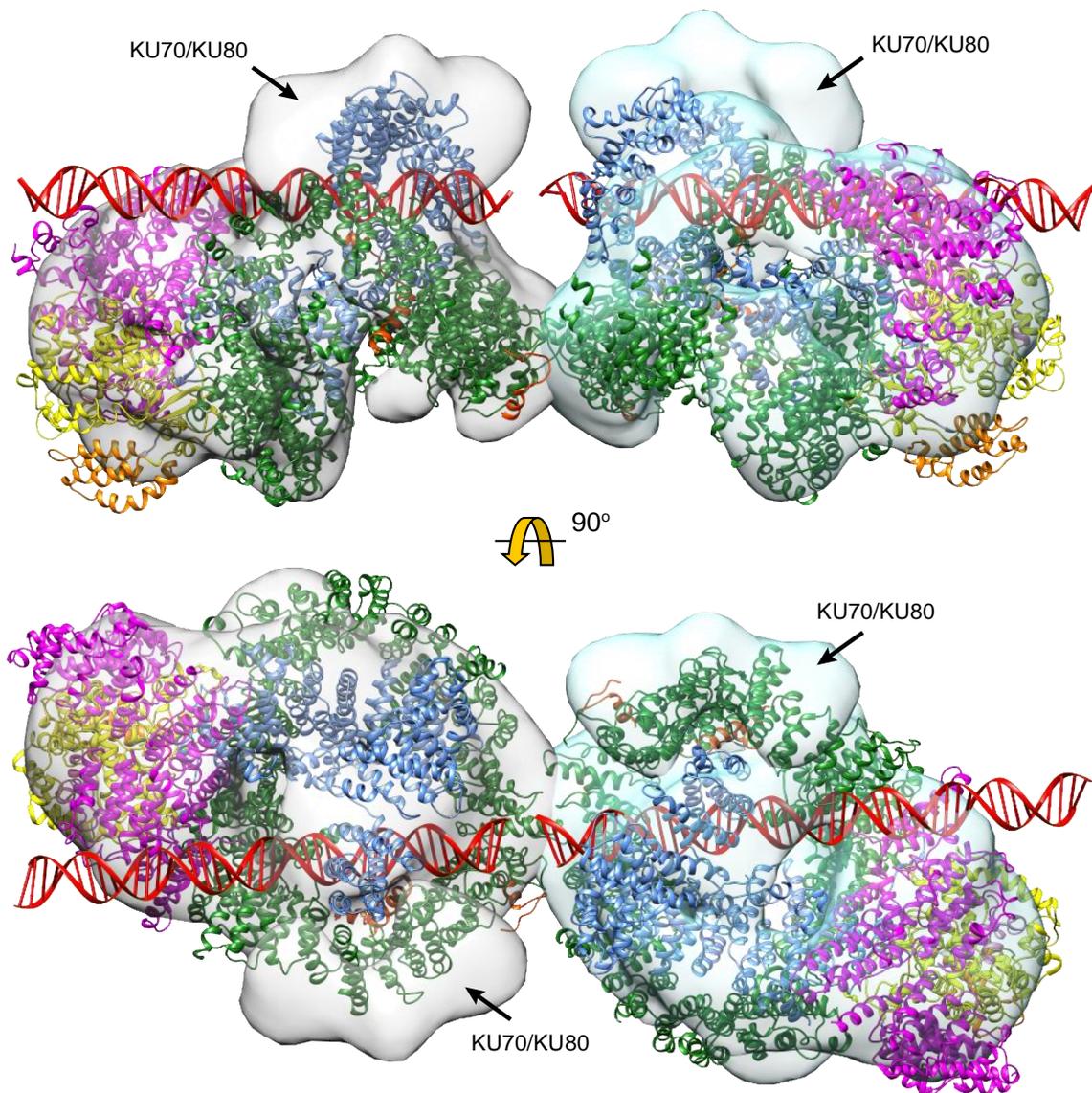


Fig. 4. The dimeric arrangement of the two DNA-PKcs molecules in the crystallographic asymmetric unit. The arrangement may allow broken ends to be brought together in a way that would facilitate double-strand break repair. The possible positions for the DNA helices bound to each DNA-PKcs molecule are shown in red, modelled in Maestro (Schrodinger Suite), making good electrostatic interactions to N1 and N2. The cryo-EM map of DNA-PKcs bound to Ku70/80 at 25 Å resolution (Spagnolo et al, 2006) is shown superimposed to each monomer, showing the position of the Ku complex relative to the overall structure.

Supplementary Materials:

Materials and Methods

Figures S1-S20

Tables S1-S7

Movies S1-S2

References (31-45)