REPORT

DNA REPAIR

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

Bancinyane L. Sibanda, Dimitri Y. Chirgadze, David B. Ascher,* Tom L. Blundell†

DNA-dependent protein kinase catalytic subunit (DNA-PKcs) is a central component of nonhomologous 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 angstrom resolution using x-ray crystallography. We show that the 4128–amino acid structure comprises three large structural units: the N-terminal unit, the 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, which are consistent with an allosteric mechanism to bring about kinase activation. The location of KU80ct$_{194}$ in the vicinity of the breast cancer 1 (BRCA1) binding site suggests competition with BRCA1, leading to pathway selection between NHEJ and homologous recombination.

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 nonhomologous end joining (NHEJ) (*1*), a process used also in V(D)J recombination (*2*), or through 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 autophosphorylation (*4*–*7*), and the KU70/80 heterodimer (*8*) play major roles in initiating NHEJ, allowing recruitment of Artemis, x-ray repair cross-complementing protein 4 (XRCC4), XRCC4/DNA ligase IV, XRCC4-like factor (XLF), and PAXX (paralog of XRCC4 and XLF). Structural information on DNA-PKcs has been limited to our low-resolution (6.6 Å) crystal structure (fig. S13) (*9*) and cryo–electron microscopy (cryo-EM) structures (>13 Å) (*10*–*13*), limiting our understanding of intermolecular interactions necessary for NHEJ.

We now report crystals of selenomethionine (Se-Met)–labeled DNA-PKcs complexed with native KU80ct$_{194}$ (KU80 residues 539 to 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 4128 amino acids in each chain. DNA-PKcs folds into three well-defined large structural units, within which motifs resembling HEAT (Huntingtin, Elongation Factor 3, PP2 A, and TOR1) repeats recur throughout the structure and 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 to N4), the Circular Cradle (85 α-helices, residues 893 to 2801; five supersecondary structures: CC1 to CC5), and the C-terminal Head (64 α-helices, residues 2802 to 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. S11C), and the point of cleavage by caspase 3 to deactivate DNA-PKcs (fig. S11D).

The bilobal kinase (3676 to 4100) (Fig. 2, B and C) is similar to that of other PI3-kinases (*14*, *15*). It comprises a well-conserved P-loop (3729 to 3735), catalytic loop (3919 to 3927), and activation loop (3940 to 3963) (Fig. 2C and figs. S15 and S16). There is also a conserved metal, which coordinates 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 analog as in mammalian target of rapamycin (mTOR). (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.)

As in mTOR, there is restricted access to the active site. The FAT region [FRAP (FKBP12-rapamycin–associated protein), ATM (ataxia-telangiectasia mutated), TRRAP (transformation/transcription domain associated protein); residues 2802 to 3564; with supersecondary structures FR1 to FR5], conserved in DNA-PKcs orthologs but not in mTOR (fig. S15), folds around the kinase (fig. S17). The four–α-helical bundle (residues 3582 to 3675) known as FRB (FKBP12-rapamycin-binding) (*16*), 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 (*17*) (residues 4101 to 4128) follows the kinase. The crystal structure of DNA-PKcs aligns well with the cryo-EM maps of DNA-PKcs (*12*, *18*, *19*), with additional density for more mobile regions observed at this higher resolution (fig. S19). The proximity of FRB to α-hairpin 1 (residues 3836 to 3872) (Fig. 2B) and FATC to α-hairpin 2 (residues 3995 to 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 destabilize 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 (residues 3836 to 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 to 4053), adjacent to FRB, blocks the entrance into the active site cleft and extends into the active site cleft (Fig. 2C). α-hairpin 3 (residues 4059 to 4082) is specific to DNA-PKcs and obscures the activation loop (Fig. 2C), thus burying the conserved T3950 autophosphorylation site that 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 (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 KU80ct$_{194}$ binding. Both sites are outside the cryo-EM electron density maps of apolipoprotein structures, supporting their assignment to KU80ct$_{194}$ (fig. S19),

Department of Biochemistry, University of Cambridge, Old Addenbrooke's Site, 80 Tennis Court Road, Cambridge CB2 1GA, UK.
*Present address: Department of Biochemistry and Molecular Biology, University of Melbourne, 30 Flemington Road, Parkville, Victoria 3052, Australia.
†Corresponding author. Email: tlb20@cam.ac.uk

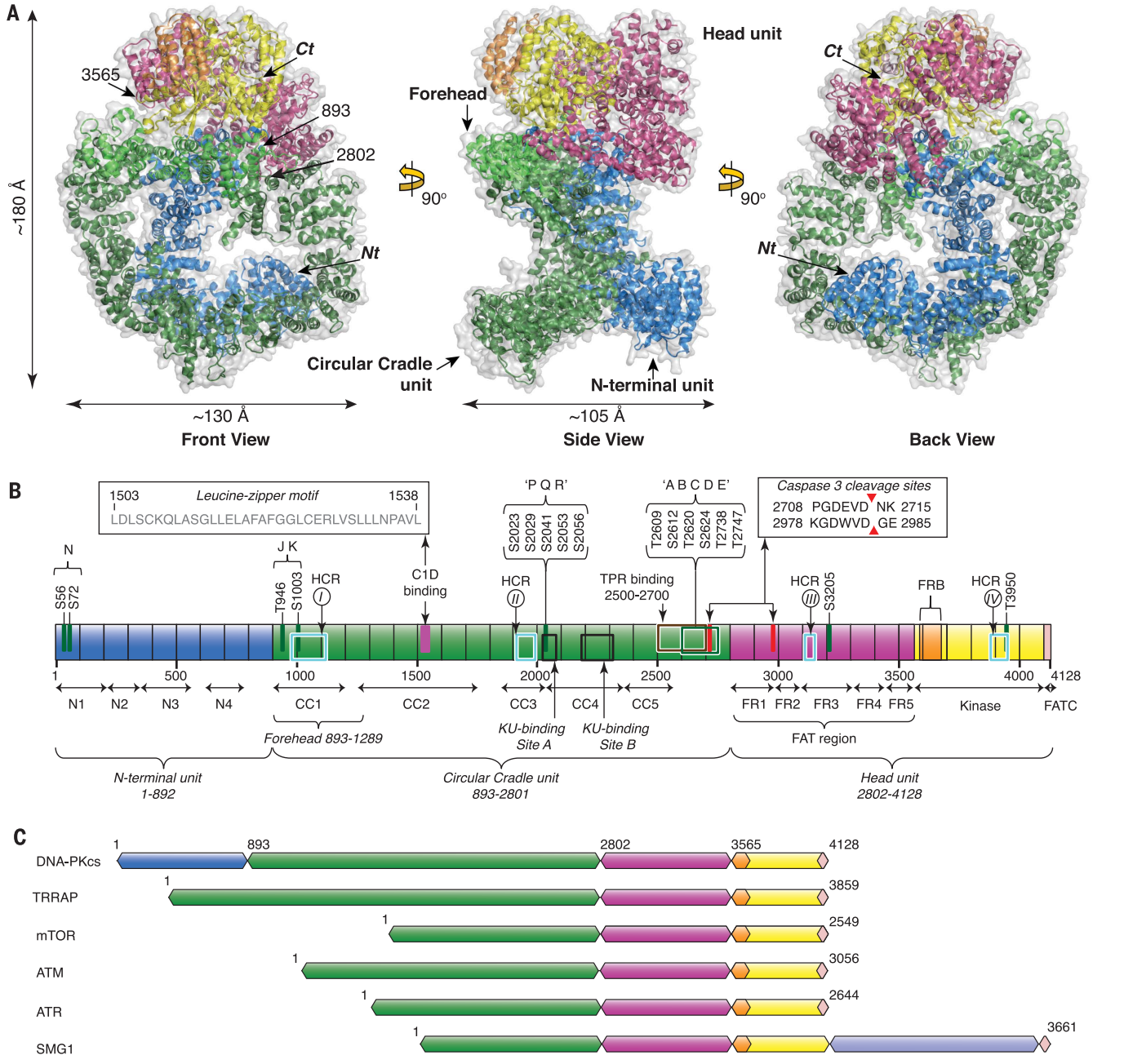

Fig. 1. The overall structure of DNA-PKcs. (**A**) Structural units of DNA-PKcs. The N terminus is in blue, the Circular Cradle is in green, Head comprising the FAT region is in pink, kinase is in yellow, and the FATC is in light pink. Also shown are the Forehead in light green and FRB (FKBP12-rapamycin-binding) in orange. The surface shown is the solvent-accessible surface area calculated by use of Pymol. (**B**) A schematic representation highlighting the three units of DNA-PKcs and their supersecondary structural components: the N terminus, composed of four supersecondary α-helical structures (N1 to N4); the Circular Cradle, composed of five supersecondary α-helical structures (CC1 to CC5); and the Head, which contains the FAT region (FR1 to FR4), kinase, FRB, and FATC regions. Highly conserved regions (HCRs), KU80 binding area (Sites A and B) and interacting areas for other proteins, autophosphorylation sites, and caspase 3 cleavage sites are shown above the schematic. (**C**) Domain organization compared with other PI3-K family members, TRRAP, mTOR, ATM, ATR (ATM- and Rad3-related), and SMG1 (human suppressor of morphogenesis in genitalia 1), color coded as in (A), with the insertion in SMG1 in light blue.

and cocrystallization of Se-Met-labeled KU80ct$_{194}$ with wild-type DNA-PKcs provided further confirmation (supplementary materials). 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" autophosphorylation cluster (*4*, *5*) and proposed BRCA1-binding site (*25*), suggesting the 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 KU80ct$_{194}$ 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 Se-Met 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 with the cryo-EM maps of the DNA-PKcs–KU70/80 complex (*18*) reveals that density for the KU70/80 complex is located beneath CC4 and N1 (fig. S19). There is

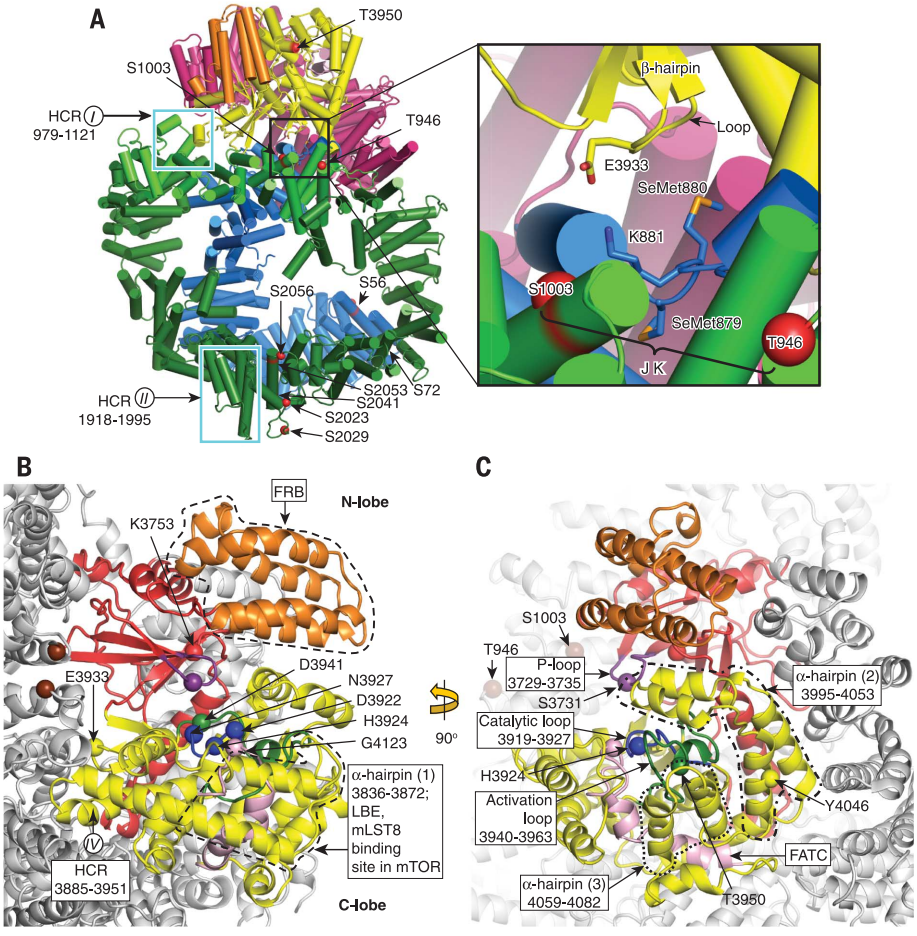

Fig. 2. The active site of DNA-PKcs/mechanism. (**A**) The color scheme is the same as in Fig. 1. The extension at the end of the N-terminal unit interacts with the active center; the contact between E3933 and K881 is highlighted; the autophosphorylated residues, HCR (I) 979 to 1121, HCR (II) 1918 to 1995, and the JK phosphorylation cluster (T946 and S1003), thought to promote HR, are also highlighted. (**B**) The kinase with the N-lobe is shown in red, the C-lobe is in yellow, and HCR (IV) is marked. The P-loop is shown in purple, the activation loop is in green, the catalytic loop is in dark blue, and the FRB (orange) and the α-hairpin 1 carrying the LBE site in mTOR are shown in dashed lines. Also indicated is G4123 on the FATC region positioned on the edge of the active site cleft. (**C**) T946, S1003, and T3950 autophosphorylation sites are labeled. 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, drastically reducing the kinase activity. The dotted and dot-dashed lines in (C) depict the supersecondary structures that enclose the active site.

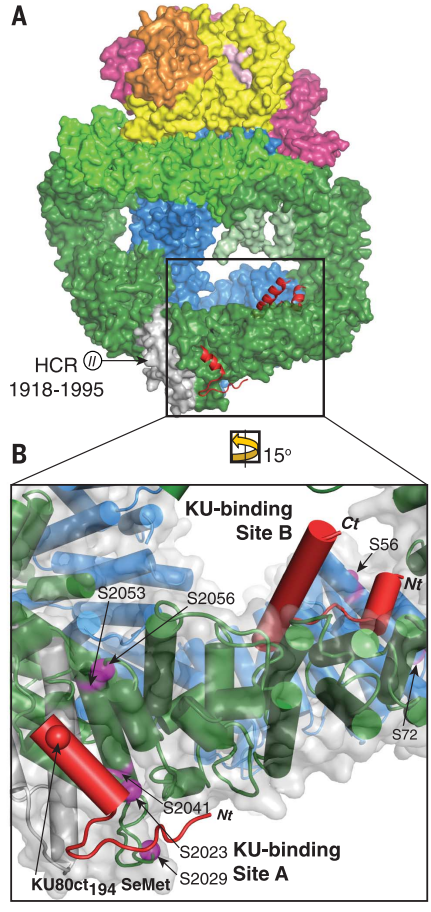

Fig. 3. KU70/80 binding area. (**A**) KU80ct194 (sites A and B in red) shown bound to DNA-PKcs, colored according to structural units, with the N terminus in blue and the Circular Cradle in green as in Fig. 1; in gray is the highly conserved region HCR II that is close to the PQR autophosphorylation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

growing evidence that N1 to N3 of DNA-PKcs mediates DNA binding (*6*, *7*, *13*, *18*, *19*, *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 activates the allosteric mechanism needed for communication between the N terminus and the 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 (*18*, *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. This movement can be inferred from comparison of the DNA-PKcs crystal structure with the cryo-EM maps of the apo structure (fig. S19) (*12*, *19*).

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*). 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 (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.

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, which is 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

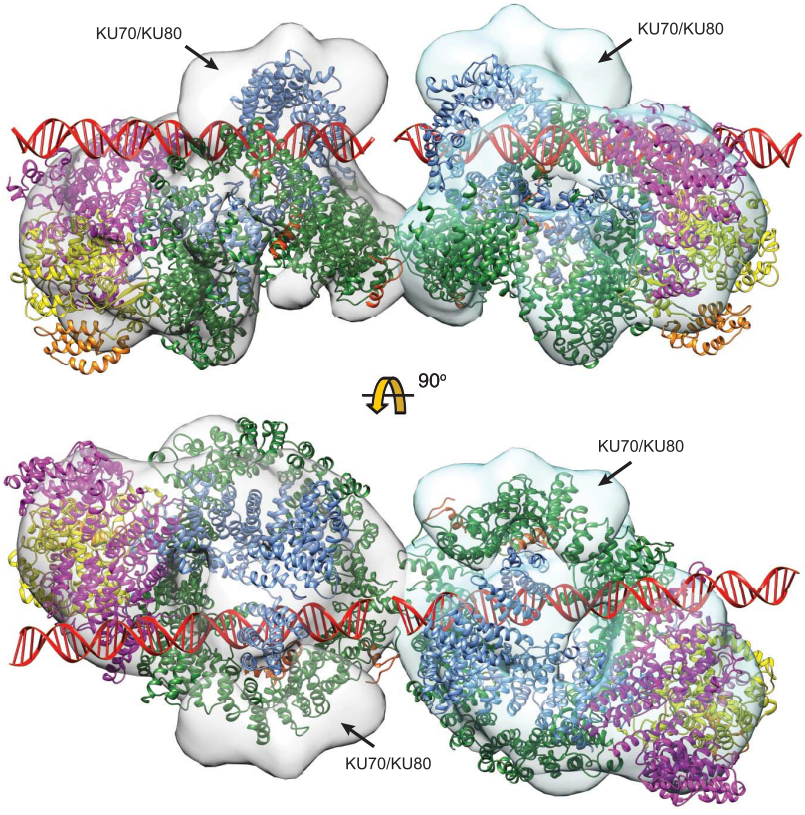

Fig. 4. The dimeric arrangement of the two DNA-PKcs molecules in the crystallographic asymmetric unit. The arrangement may allow broken ends to be brought together in a way that would facilitate DSB repair. The possible positions for the DNA helices bound to each DNA-PKcs molecule are shown in red, modeled 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 (*18*) is shown superimposed on each monomer, showing the position of the KU complex relative to the overall structure.

terminus of the N-terminal structural unit implicated as the trigger for full kinase activation. The located Se-Met site of the KU80ct$_{194}$ 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, *Trends Biochem. Sci.* **23**, 394–398 (1998).
2. D. G. Schatz, *Immunol. Rev.* **200**, 5–11 (2004).
3. S. C. West, *Nat. Rev. Mol. Cell Biol.* **4**, 435–445 (2003).
4. B. P. Chen *et al.*, *J. Biol. Chem.* **280**, 14709–14715 (2005).
5. X. Cui *et al.*, *Mol. Cell. Biol.* **25**, 10842–10852 (2005).
6. K. Meek, S. P. Lees-Miller, M. Modesti, *Nucleic Acids Res.* **40**, 2964–2973 (2012).
7. P. Douglas *et al.*, *Mol. Cell. Biol.* **27**, 1581–1591 (2007).
8. J. R. Walker, R. A. Corpina, J. Goldberg, *Nature* **412**, 607–614 (2001).
9. B. L. Sibanda, D. Y. Chirgadze, T. L. Blundell, *Nature* **463**, 118–121 (2010).
10. C. Y. Chiu, R. B. Cary, D. J. Chen, S. R. Peterson, P. L. Stewart, *J. Mol. Biol.* **284**, 1075–1081 (1998).
11. K. K. Leuther, O. Hammarsten, R. D. Kornberg, G. Chu, *EMBO J.* **18**, 1114–1123 (1999).
12. A. Rivera-Calzada, J. D. Maman, L. Spagnolo, L. H. Pearl, O. Llorca, *Structure* **13**, 243–255 (2005).
13. D. R. Williams, K. J. Lee, J. Shi, D. J. Chen, P. L. Stewart, *Structure* **16**, 468–477 (2008).
14. E. H. Walker, O. Perisic, C. Ried, L. Stephens, R. L. Williams, *Nature* **402**, 313–320 (1999).
15. H. Yang *et al.*, *Nature* **497**, 217–223 (2013).
16. J. Choi, J. Chen, S. L. Schreiber, J. Clardy, *Science* **273**, 239–242 (1996).
17. S. A. Dames, J. M. Mulet, K. Rathgeb-Szabo, M. N. Hall, S. Grzesiek, *J. Biol. Chem.* **280**, 20558–20564 (2005).
18. L. Spagnolo, A. Rivera-Calzada, L. H. Pearl, O. Llorca, *Mol. Cell* **22**, 511–519 (2006).
19. S. A. Villarreal, P. L. Stewart, *J. Struct. Biol.* **187**, 76–83 (2014).
20. R. Araki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2438–2443 (1997).
21. M. van der Burg, J. J. van Dongen, D. C. van Gent, *Curr. Opin. Allergy Clin. Immunol.* **9**, 503–509 (2009).
22. D. Gell, S. P. Jackson, *Nucleic Acids Res.* **27**, 3494–3502 (1999).
23. R. Harris *et al.*, *J. Mol. Biol.* **335**, 573–582 (2004).
24. Z. Zhang *et al.*, *Structure* **12**, 495–502 (2004).
25. A. J. Davis *et al.*, *Nucleic Acids Res.* **42**, 11487–11501 (2014).
26. E. Weterings *et al.*, *Mol. Cell. Biol.* **29**, 1134–1142 (2009).
27. M. Hammel *et al.*, *J. Biol. Chem.* **285**, 1414–1423 (2010).
28. J. Boskovic *et al.*, *EMBO J.* **22**, 5875–5882 (2003).
29. T. M. Gottlieb, S. P. Jackson, *Cell* **72**, 131–142 (1993).
30. N. Jette, S. P. Lees-Miller, *Prog. Biophys. Mol. Biol.* **117**, 194–205 (2015).

ACKNOWLEDGMENTS

Model coordinates have been deposited in the Protein Data Bank under accession number 5LUQ. We thank J. Van den Heuvel at Helmholtz-Zentrum and particularly V. Jager, who prepared the HeLa cells. We are also grateful to the European Synchrotron Radiation Facility (beamlines ID29, ID14-1, and 1D14-2 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 (grant 093167/Z/10/Z).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/355/6324/520/suppl/DC1
Materials and Methods
Figs. S1 to S21
Tables S1 to S7
References (*31*–*46*)
Movies S1 and S2

22 September 2016; accepted 5 January 2017
10.1126/science.aak9654

Science

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

Bancinyane L. Sibanda, Dimitri Y. Chirgadze, David B. Ascher and Tom L. Blundell

Science **355** (6324), 520-524.
DOI: 10.1126/science.aak9654

Activating DNA repair

DNA double-strand breaks must be repaired efficiently to avoid cell death or cancer. The break ends can either be directly ligated by nonhomologous end joining (NHEJ) or more accurately repaired by homologous recombination that uses information from the sister chromatid. Sibanda *et al.* present a high-resolution x-ray structure of a key component of the DNA repair machinery, the DNA-dependent kinase catalytic subunit (DNA-PKcs), bound to a C-terminal peptide of Ku80. The structure suggests that Ku80 presents the DNA ends for repair to a DNA-PKcs dimer and that activity is modulated by interactions between the monomers. Binding of either Ku80 or BRCA1, which may compete for the same binding site on DNA-PKcs, could provide a switch between NHEJ and homologous recombination.

Science, this issue p. 520

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.