

# Human DNA-dependent protein kinase activation mechanism

Received: 1 March 2022

Shikang Liang<sup>1,2</sup>✉ & Tom L. Blundell<sup>1,2</sup>✉

Accepted: 26 October 2022

Published online: 5 January 2023

 Check for updates

DNA-dependent protein kinase (DNA-PK), a multicomponent complex including the DNA-PK catalytic subunit and Ku70/80 heterodimer together with DNA, is central to human DNA damage response and repair. Using a DNA-PK-selective inhibitor (M3814), we identified from one dataset two cryo-EM structures of the human DNA-PK complex in different states, the intermediate state and the active state. Here we show that activation of the kinase is regulated through conformational changes caused by the binding ligand and the string region (residues 802–846) of the DNA-PK catalytic subunit, particularly the helix-hairpin-helix motif (residues 816–836) that interacts with DNA. These observations demonstrate the regulatory role of the ligand and explain why DNA-PK is DNA dependent. Cooperation and coordination among binding partners, disordered flexible regions and mechanically flexible HEAT repeats modulate the activation of the kinase. Together with previous findings, these results provide a better molecular understanding of DNA-PK catalysis.

The DNA-PK of vertebrate cells is central to the DNA damage response and the nonhomologous end-joining (NHEJ) pathway for the repair of DNA double-strand breaks, the most dangerous form of DNA damage, which may lead to genomic instability, carcinogenesis and cell death<sup>1,2</sup>. DNA-PK was first identified when addition of double-stranded DNA (dsDNA) to cell extracts was observed to stimulate the phosphorylation of certain proteins. The catalytic activity was found to be related to a very large polypeptide, now known as the DNA-PK catalytic subunit (DNA-PKcs), with 4128 amino acid residues and a molecular weight of ~470 kDa<sup>3–5</sup>. Later, it was discovered that the effective recruitment and activation of DNA-PKcs by DNA requires the heterodimer Ku70/80, which is formed by two subunits of molecular weights approximately 70 kDa and 80 kDa<sup>6,7</sup>. The kinase domain of DNA-PKcs was found to be highly similar to that of phosphoinositide 3-kinase, belonging to the family of phosphoinositide 3-kinase-related kinases (PIKKs), which includes ATM, ATR, mTOR, SMG1 and TRRAP<sup>8</sup>. PIKK family members have the greatest similarity in the kinase domain, which is surrounded by a conserved  $\alpha$ -helical FAT (FRAP, ATM and TRAPP) domain together with the PIKK regulatory domain (PRD) and the FATC domain. All PIKKs also have an  $\alpha$ -solenoid amino-terminal region, which consists mainly of HEAT repeats that have important regulatory roles.

There have been structural studies on both DNA-PKcs and the DNA-PK holoenzyme by several independent groups<sup>9–18</sup>. DNA-PKcs has been described as comprising three structural units, the N-terminal unit (1–892), the circular cradle (893–2801) and the head (2802–4128), where the kinase region is located between residues 3565 and 4100 (refs. <sup>9,10</sup>). The assembly of the DNA-PK holoenzyme and its higher-order dimerization in NHEJ complexes have also been described<sup>11–17</sup>. However, the molecular mechanism of DNA-PK activation remains unclear, with gaps between ligand-bound DNA-PKcs and ligand-bound intermediate or active DNA-PK<sup>13,18</sup>.

Our previous work suggested that inhibitors have a similar mode of structural regulation to that of ATP binding DNA-PKcs<sup>18</sup>. Using the latest-generation DNA-PK selective inhibitor M3814 to mimic ATP binding, while at the same time slowing down and inhibiting kinase catalysis, we succeeded in obtaining two different conformations of the DNA-PK–ligand complex from one cryo-EM dataset: the intermediate state, where the inhibitor binds to the ATP pocket before the closure of the kinase domain, and the active state, where the kinase domain is closed up and the ligand is brought next to the substrate-binding site. Together with the results of previous work on DNA-PKcs, DNA-PK and DNA-PKcs–ligand complexes, these findings provide structural insight

<sup>1</sup>Department of Biochemistry, University of Cambridge, Cambridge, UK. <sup>2</sup>Present address: Heart and Lung Research Institute, Department of Medicine, University of Cambridge, Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK. ✉ e-mail: [sl744@cam.ac.uk](mailto:sl744@cam.ac.uk); [tlb20@cam.ac.uk](mailto:tlb20@cam.ac.uk)

into the progression from ligand binding through the full catalytic pathway<sup>13,14,18</sup>.

## Results and discussion

### Structure overview

DNA-PKcs natively purified from HeLa cell nuclear extract was incubated on ice for 1 h with purified recombinant Ku70/80, DNA and an excess of DNA-PK selective inhibitor M3814. The dataset was collected on a Gatan K3 detector and processed (Extended Data Fig. 1 and Table 1). From the dataset, we identified an intermediate state of DNA-PK, allowing definition at 3.2 Å resolution following density modification with ResolveCryoEM, before closure of the kinase domain and catalysis (Fig. 1a). Moreover, we managed to stabilize and lock the active DNA-PK in the transient conformation adopted for catalysis with a closed kinase domain, giving a cryo-EM model at a resolution of 2.8 Å following density modification with ResolveCryoEM (Fig. 1b), in which the inhibitor occupies a position next to the substrate-binding site.

### Cryo-EM structure of DNA-PK in the intermediate state

The inhibitor docks in the center of the ATP-binding groove in the intermediate-state DNA-PK (Fig. 2a,b). The overall assembly of the kinase complex is similar to those previously defined<sup>12–14,17</sup>. The DNA–Ku70/80 core complex interacts with the N-terminal unit and circular cradle of DNA-PKcs, and the DNA passes under the N-terminal unit, while the carboxy-terminal region of Ku80 docks on the circular cradle. However, there are many differences in the conformations and structural details.

DNA-PK in the intermediate state has several structured regions that are flexible. In Ku80, these include the vWA domain (6–242), the six-helical bundle (595–706) and the 12 residues (721–732) of the C-terminal region (543–732)<sup>19,20</sup>. Within the DNA-PKcs component of DNA-PK, there are several structured regions that are flexible, including the N-terminal region, close to where the DNA end enters, and a part of the kinase region within the head unit, especially the four-helical FKBP12-rapamycin-binding (FRB) domain (3582–3675)<sup>21</sup>. Another interesting flexible region in DNA-PKcs is the helix-hairpin-helix (HhH) motif between the N terminus of DNA-PKcs and the DNA end (Fig. 2c), which is likely to come from a flexible region (802–846) hanging down in the DNA-PKcs central cavity formed by the circular cradle and N-terminal unit. It is clear that in the higher-resolution active-state DNA-PK structure, the HhH motif (816–836) is essential not only for allosteric activation and conformational transition but also for stabilizing the interaction between the DNA-PKcs N-terminal region and the head unit.

Compared with the previously published structures of the inactive-state apo-DNA-PK, the major conformational differences lie in the head unit of DNA-PKcs (Extended Data Fig. 2a,b)<sup>13,14</sup>. In the intermediate-state complex, the head unit moves towards the DNA end and the N-terminal unit of DNA-PKcs (Extended Data Fig. 2a,b). When the ligand binds, the p-loop moves close to it (Extended Data Fig. 2c). This leads to concerted movements involving conformational changes of the flanking β-sheets of the p-loop and movement in the hydrophobic core of the head unit, resulting in overall conformational changes. More importantly, the outward movement of the head unit releases the PRD (4009–4039) from its position in the apo kinase (Fig. 2b and Extended Data Fig. 2c). In the intermediate complex, there is only weak signal near the PRD position, which indicates that the PRD is destabilized. When the blockage of PRD is removed, the kinase active site becomes available for substrate peptide binding and subsequent catalytic activity.

### Cryo-EM structure of DNA-PK in the active state

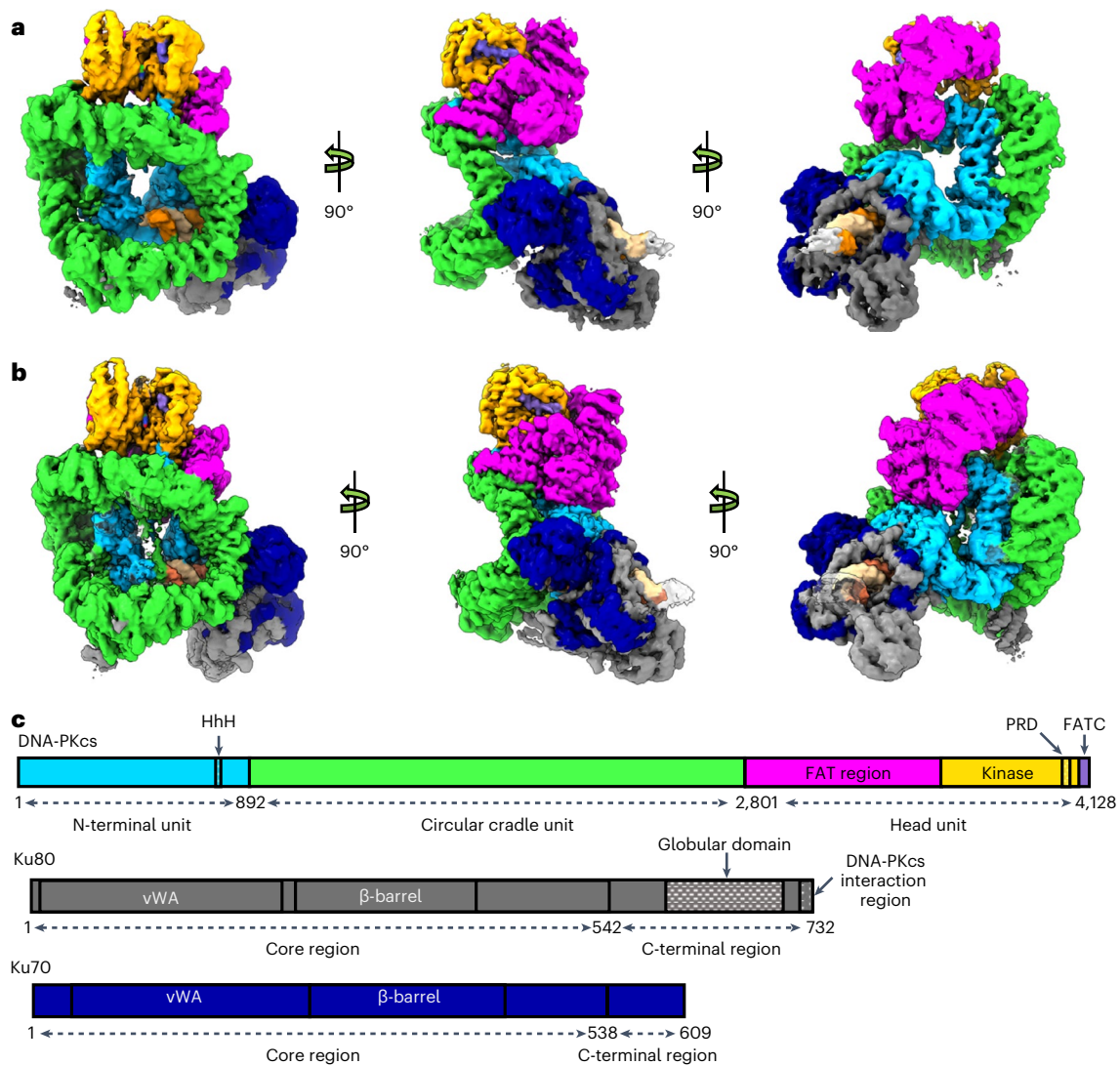
In the active state of DNA-PK, the kinase core is closed and the ligand M3814 is moved close to the substrate-binding site (Fig. 3a–c). Similar to the intermediate-state kinase complex, some structured regions are flexible in the active DNA-PK structure. Most parts of Ku80 remain

**Table 1 | Cryo-EM data collection, refinement and validation statistics**

	Intermediate DNA-PK (EMDB-14546) (PDB 7Z88)	Active DNA-PK (EMDB-14545) (PDB 7Z87)
<b>Data collection and processing</b>		
Magnification	130k	130k
Voltage (kV)	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	47.22	47.22
Defocus range (μm)	–1.1 to –2.3	–1.1 to –2.3
Pixel size (Å)	0.652	0.652
Symmetry imposed	C1	C1
Initial particle images (no.)	1094271	1094271
Final particle images (no.)	190498	275300
Map resolution (Å)	3.33	2.89
FSC threshold	0.143	0.143
Map resolution range (Å)	2.9 to >10	2.7 to >10
<b>Refinement</b>		
Initial model used (PDB code)	6ZHA, 7K1N	6ZHA, 7KOY
Model resolution (Å)	3.33	2.89
FSC threshold	0.143	0.143
Model resolution range (Å)	n/a	n/a
Map sharpening B factor (Å <sup>2</sup> )	75.2	78.3
Model composition		
Nonhydrogen atoms	38754	39708
Protein residues	4750	4840
Ligands	1	1
B factors (Å <sup>2</sup> )		
Protein	24	126
Ligand	16	101
R.m.s. deviations		
Bond lengths (Å)	0.005	0.006
Bond angles (°)	0.909	0.842
Validation		
MolProbity score	2.4	2.3
Clashscore	21	16
Poor rotamers (%)	0.17	0.13
Ramachandran plot		
Favored (%)	88.62	89.39
Allowed (%)	11.29	10.48
Disallowed (%)	0.09	0.13

flexible, including the vWA domain, the six-helical bundle (595–706) and the 12 C-terminal residues (721–732). However, in DNA-PKcs, the previously flexible N-terminal region and HhH motif (816–836) are now stabilized between the FAT region and the DNA terminus (Fig. 3d).

On one side of the HhH motif, there are electrostatic interactions between positively charged residues, including Arg820, Lys824, Lys832 and His833, and the negatively charged DNA phosphate backbone. On the other side of the HhH motif, hydrophobic interactions define the binding with the DNA-PKcs N-terminal region. Phe826 of the HhH motif docks into the groove between the two helices that pack onto the



**Fig. 1 | Overview of DNA-PK from structures defined by cryo-EM. a**, Front (left), side (middle) and back (right) views of the Coulomb potential map of DNA-PK in the intermediate state. **b**, Front (left), side (middle) and back (right) views of the Coulomb potential map of DNA-PK in the active state. **c**, Schematic representation of sequences of DNA-PKcs, Ku80 and Ku70 with structural regions

of DNA-PKcs colored differently: the N-terminal unit is in blue, the circular cradle unit in green, the FAT region of the head unit in magenta, the kinase of the head unit in yellow and the FATC of the head unit in purple; Ku80 is shown in gray and Ku70 in navy.

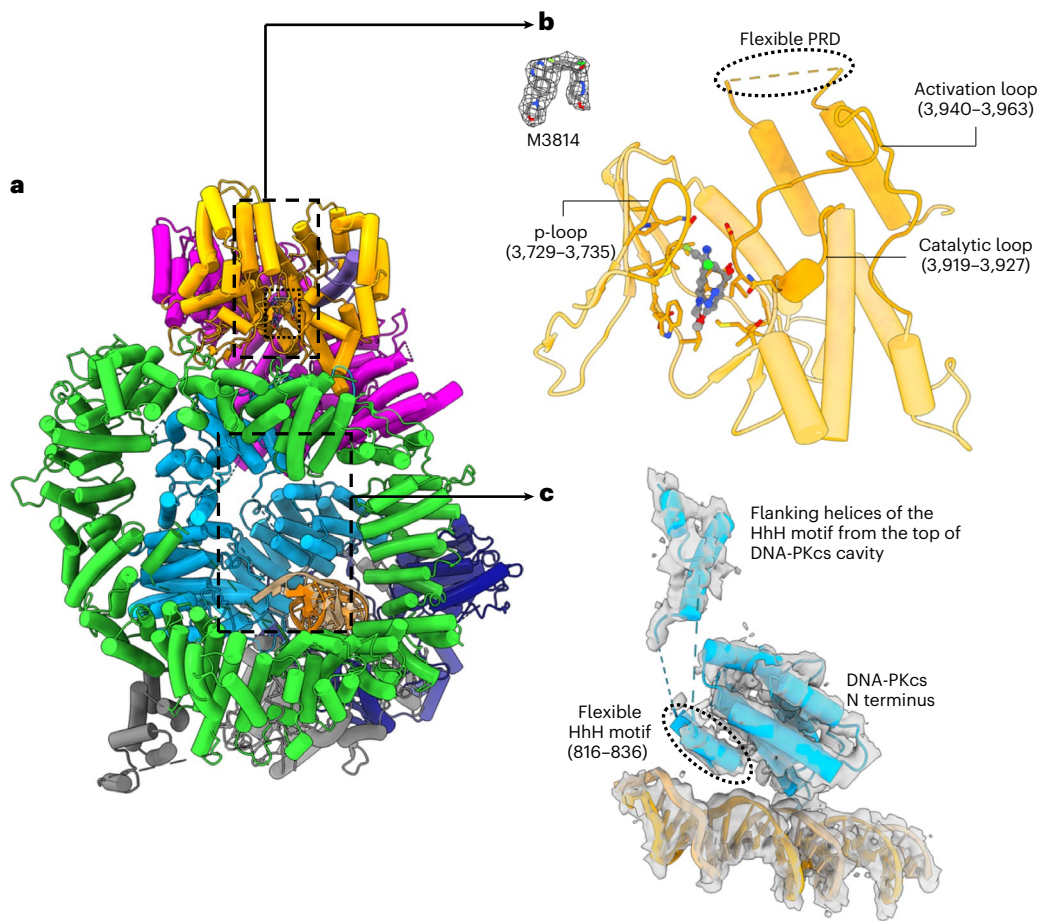
HhH motif, interacting with the side chains of Phe88 and Ile91 and the main chain atoms of Arg36, Gln40, Glu84 and Lys87. The side chains of Leu818, Val830, Leu831 and Leu834 also interact with the hydrophobic surface composed of the side chains of Val43, Leu44 and Ile91 and the main chain atoms of Glu94 and Lys95. In addition to the hydrophobic interaction, there is an electrostatic interaction between Lys835 and Glu94 and hydrogen bonds among Gln823, Arg36 and Gln40 (Fig. 3d).

The position of the N-terminal region is also defined by interactions with the FAT region of the head unit. Tyr3101 lies in the hydrophobic pocket formed by Ala49, Val50 and Leu53. There are water-mediated side chain interactions among the polar and charged residues of the interface, including Asp3066, His3070, Lys3075, Tyr3090, Arg3098 and Tyr3102 from the FAT region and Arg9, Ser60 and Asp62 from the N-terminal unit (Fig. 3f). Furthermore, in the active form of DNA-PK, a DNA-end-binding helix (2735–2768) from the flexible region of DNA-PKcs lies near the ABCDE cluster in the middle of the DNA-PKcs central cavity, restricting inward movement of the DNA (Fig. 3e). The stabilization of the DNA-end-binding helix is likely to be caused by the fixation of the activation HhH motif together with its flanking peptide string and the DNA end.

Compared with the previously proposed model of an activated DNA-PK, the structure of the active holoenzyme, described here, is more compact, showing different conformational changes (Extended Data Fig. 3a)<sup>13</sup>. The DNA end lies in the center of the holoenzyme and to different extents all regions of the holoenzyme move towards it (Extended Data Fig. 3a). Moreover, the lengths of the DNA strands coming into the DNA-PKcs central cavity are different (Extended Data Fig. 3b). A further difference is that the helix from the region around the PQR cluster continues to dock in the cleft between the N-terminal unit and the circular cradle in our active form, similar to the intermediate DNA-PK (Extended Data Fig. 3c). In addition, the PRD, instead of having a 115° outward rotation, is destabilized and removed from the substrate-binding site (Extended Data Fig. 3d), although different, destabilization and rotation have the same effect in making the substrate-binding site available for substrates to bind.

#### Intermediate-to-active-state transition of DNA-PK

The holoenzyme undergoes a series of conformational changes during the transition from the intermediate to the active state (Fig. 4a). Most changes lie in DNA-PKcs, especially in the N-terminal



**Fig. 2 | The structure of DNA-PK in the intermediate state defined by cryo-EM.** **a**, DNA-PK in intermediate state. **b**, Close-up view of the kinase catalytic core around the DNA-PK selective inhibitor (M3814). The corresponding Coulomb potential map of the inhibitor and the two-dimensional structure are shown in the upper left panel. The inhibitor sits in the ATP-binding groove, coordinated by the p-loop (3729–3735), catalytic loop (3919–3927) and activation loop (3940–3963). The PRD becomes flexible in the intermediate state of DNA-PK, opening

the substrate-binding site for catalysis. **c**, Close-up view of the N-terminal region around the HhH motif (816–836) and the corresponding Coulomb potential map. In the intermediate structure of DNA-PK, the HhH motif position is not fixed between the DNA end and N terminus but is connected through a highly flexible linking peptide to the top of the DNA-PKcs central cavity, as indicated in **a**. The color scheme is consistent with that of Fig. 1.

and head units and the part of the circular cradle unit that is close to them (893–1469) (Fig. 4a–d).

According to our structures, the main trigger for kinase activation and transition is the fixation of the HhH motif (816–836) (Fig. 4b). Stabilization of the activation HhH motif between the DNA terminus and the N-terminal region induces a conformational change of the latter. Furthermore, it restricts the conformation of the previously disordered flexible peptide (837–846) linking to the following  $\alpha$ -helix (847–862) from the top of the DNA-PKcs central cavity (Fig. 4b). The peptide string pulls the connected helix closer to the DNA end and leads to further serial downward movements of the nearby FAT region, which comes in contact with the N-terminal region. The conformational changes are also passed on to the whole molecule, especially to the kinase region (Fig. 4a–d). The peptide string acts as a ‘pull switch’ for activating the kinase, and the inhibitor in the active state is moved more than 10 Å away from its position in the intermediate state (Fig. 4d).

To understand how the kinase region is closed during activation, the kinase regions of DNA-PK in the two states were aligned by superimposing the C-lobes of the two kinase structures (Fig. 4e). This demonstrates a concerted movement of the N-lobe and neighboring FAT region (3490–3565), pushing the p-loop towards the C-lobe. The inhibitor molecule moves about 3.5 Å from its previous position relative to the C-lobe (Fig. 4e) towards the catalytic loop that contains the

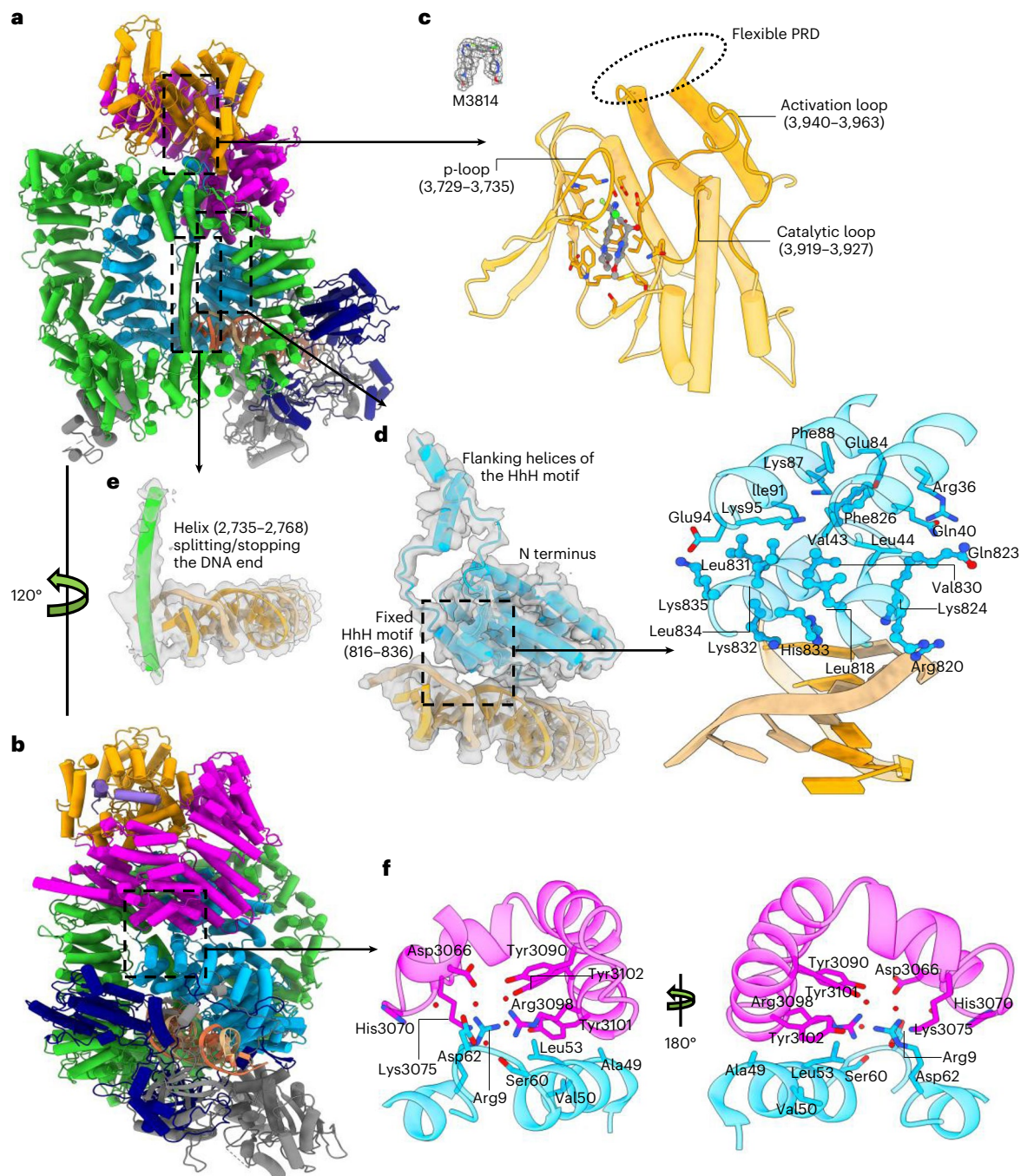
catalytic aspartate (Asp 3922) and the neighboring positively charged His 3924, which helps to localize the  $\gamma$ -phosphate of ATP together with Ser 3731 from the p-loop (Fig. 4e).

Previously, it was proposed that YRPD (2775–2795) and its interacting region (2569–2585) are important in regulation of the ABCDE cluster and autophosphorylation<sup>22,23</sup>. During the transition from the intermediate state to the active state, YRPD domain and its interacting region move closer to the catalytic site, indicating the *cis*-autophosphorylation of the ABCDE domain (Fig. 4e).

### Structural mechanism of DNA-PK activation

The structural studies of DNA-PKcs and DNA-PK, together with the intermediate and active DNA-PK, now provide a better understanding of the concerted movements of different domains during allosteric activation.

Here, we focus on comparison of the kinase domains of different DNA-PKcs-related models, aligned by the C-lobe. Comparison of apo-DNA-PKcs and apo-DNA-PK indicates that the assembly of DNA-PK holoenzyme itself leads to an opening of the N-lobe, involving outward moving of the FRB domain and partial relaxation of the PRD (Fig. 5a)<sup>13,14</sup>. When a ligand binds to the ATP-binding site in DNA-PKcs, the PRD tilts but remains docked on the substrate-binding site<sup>18</sup>. With binding of the inhibitor to the ATP pocket in DNA-PK, the FRB domain further opens



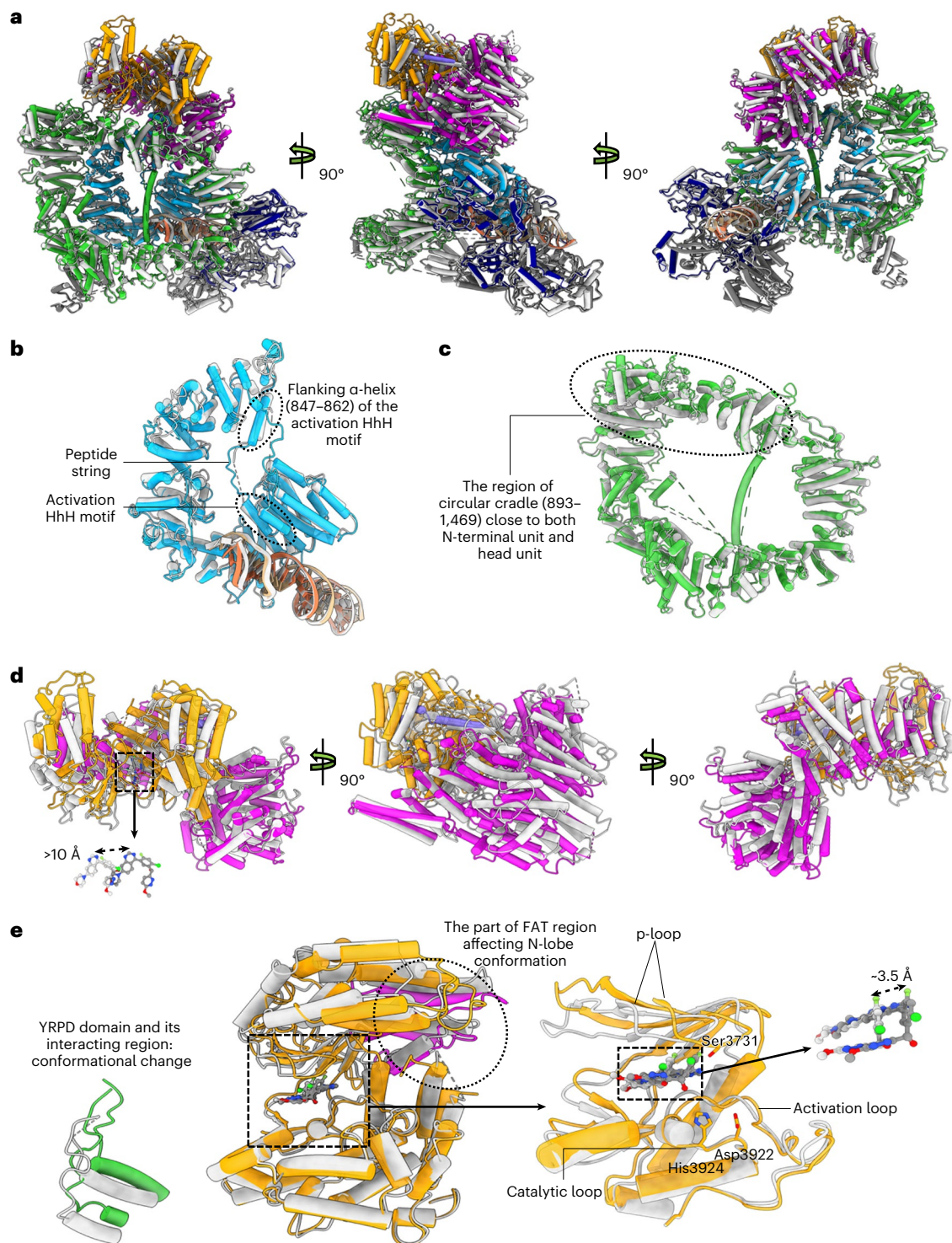
**Fig. 3 | The structure of DNA-PK in the active state defined by cryo-EM.** **a,b**, Two views, related by a rotation of 120°, of DNA-PK in the active state. **c**, Close-up view of the kinase catalytic core around the inhibitor. The corresponding Coulomb potential map of the inhibitor is shown in the upper left of this panel. **d**, The DNA end and N-terminal region around the HhH motif (816–836), the corresponding Coulomb potential map and a closer view of the

interaction interfaces. In active DNA-PK, the HhH motif is fixed between the DNA end and N terminus. The peptide string connecting HhH motif to the top of the DNA-PKcs central cavity is restricted and can be clearly traced in the map. **e**, DNA end with the DNA-end-binding helix (2735–2768) and the corresponding Coulomb potential map. **f**, Interaction interface between DNA-PKcs N terminus and the FAT region. Color scheme is consistent with that of Fig. 1.

compared with that in apo-DNA-PKcs and apo-DNA-PK, leading to the release of the PRD (Fig. 5a).

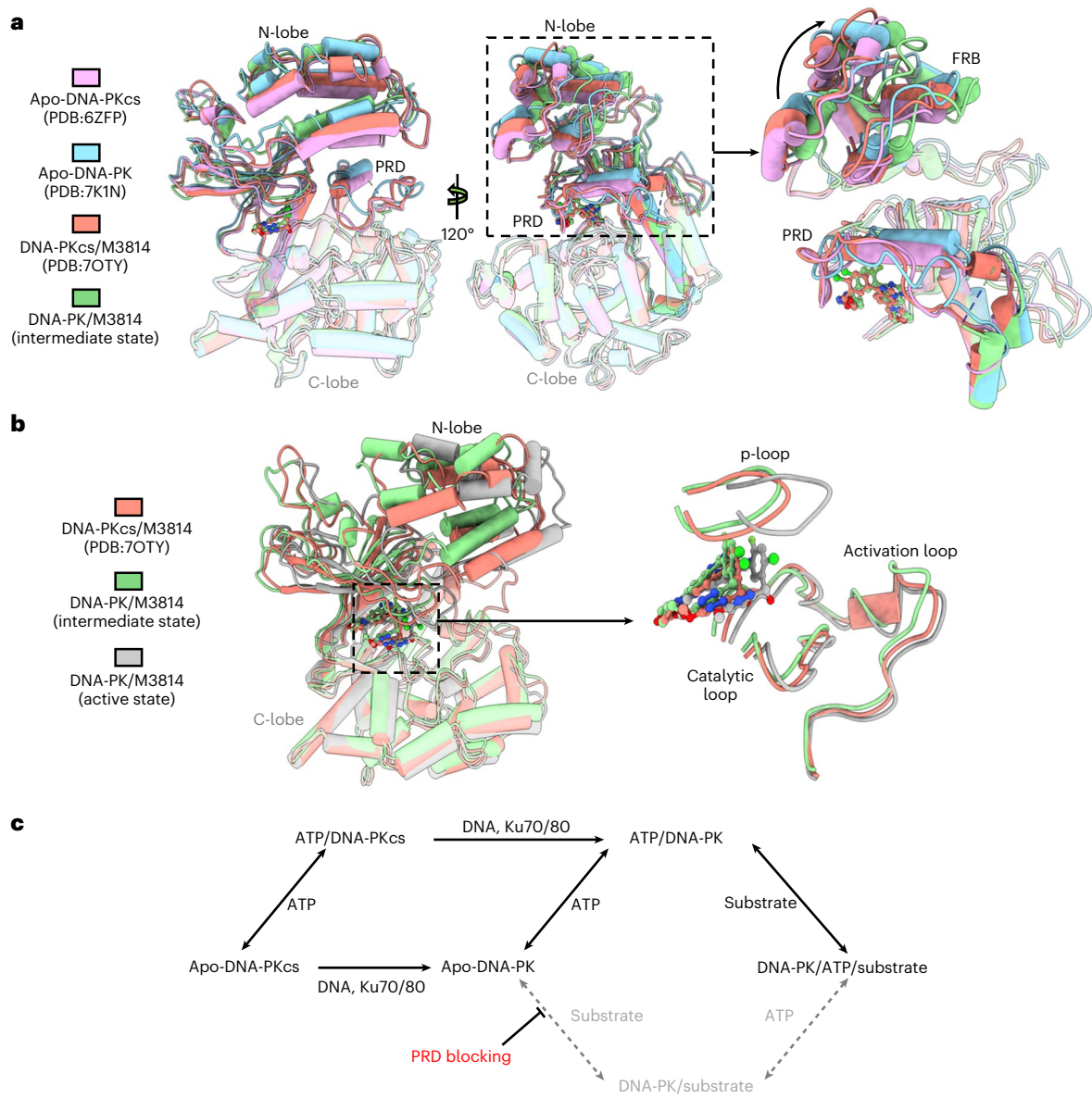
Moreover, we can now better understand the progression from DNA-PKcs to DNA-PK in the catalytic cycle. The assembly of DNA-PK from ligand-bound DNA-PKcs or ligand binding to apo-DNA-PK affect the conformation of the head unit, open the FRB domain and release PRD from the substrate-binding site. Stabilization of the activation HhH motif between the DNA end and the N-terminal bridge induces global conformational changes within the holoenzyme complex, including

N-lobe movement and closure of the kinase domain to move the ligand towards the substrate-binding site for catalysis (Fig. 5b). Of note, considering the high concentration of ATP in the cell, the *in vivo* structural progression of DNA-PK activation should start from ligand-bound DNA-PKcs and proceed to intermediate and active DNA-PK. In general, DNA, Ku70/80 and ligand binding, together with the activation HhH motif far from the DNA-PKcs head region, allosterically activate the kinase. Similar concerted movements among HEAT repeats and the kinase domain have previously been visualized in other PIKK members,



**Fig. 4 | Transition of DNA-PK from the intermediate state to the active state.** **a**, Front (left), side (middle) and back (right) views of superimposed DNA-PK molecules in the intermediate state and the active state when aligned by DNA ends. The color scheme previously used for DNA-PK in the active state is retained, whereas DNA-PK in the intermediate state is shown in silver. **b**, Close-up view of the superimposed N-terminal units and DNA ends in the intermediate state and active state when aligned by DNA ends. **c**, Magnified view of the superimposed circular cradle units in the intermediate state and the active state when aligned by DNA ends. **d**, Front (left), side (middle) and back (right) views of superimposed head units in the intermediate state and the active state when aligned by DNA ends. Close-up

views of the superimposed DNA-PK selective inhibitors in the intermediate state and the active state when aligned by DNA ends, showing that the ligand is moved more than  $10 \text{ \AA}$  during the transition. **e**, Close-up views of the superimposed kinase region, YRPD and its interacting region in the intermediate state and active state when aligned by the C-lobe of the kinase. YRPD and its interacting region become closer to the kinase catalytic site. The N-lobe moves towards the C-lobe during the activation, facilitated by the movement of the neighboring FAT region (3490–3565). Relative to the kinase C-lobe, the ligand is moved around  $3.5 \text{ \AA}$  closer to the catalytic site during activation.



and it is tempting to hypothesize that the cooperation among binding partners, the disordered flexible region and the structurally plastic HEAT repeats regulate functions of PIKK family<sup>24</sup>.

Among the phosphorylation targets of DNA-PK, self-phosphorylation is probably the most studied<sup>25–28</sup>. According to our active kinase model, the ABCDE cluster (2609–2647) is on a flexible region of polypeptide, which could have access to the kinase catalytic site and therefore could be phosphorylated in *cis*. The movement of the YRPD domain and its interacting region also suggests autophosphorylation of the ABCDE cluster in a *cis* manner. However, the PQR cluster (2023–2056) of DNA-PKcs can only be autophosphorylated in *trans* owing to its distance from the kinase catalytic region. Studies of NHEJ complexes containing DNA-PK have revealed the assembly of different dimers that probably mediate autophosphorylation in *trans*<sup>14–16</sup>. In these models, the Ku80 C-terminal region and other NHEJ components including XLF, XRCC4 and DNA ligase IV appear

to assemble as a scaffold. A recent publication has revealed an interesting new conformation of DNA-PK after self-phosphorylation on the ABCDE cluster that is in line with the previous SAX model on the displacement of the N-terminal HEAT domain, as the N-terminal HEAT domain opens greatly<sup>17,29</sup>. Of note, in the autophosphorylated DNA-PK in complex with Artemis, the activation HhH motif cannot be stabilized by interactions between the DNA end and N-terminal arm owing to the Artemis blockage and remains in a destabilization state<sup>17</sup>. Therefore, the kinase domain is not closed, and DNA-PK is in an inactive state halting further phosphorylation before end processing and removal of the Artemis catalytic domain. However, how the phosphorylation on the flexible ABCDE domain affects the overall conformation of the HEAT repeats, especially the N-terminal region, remains unclear and requires further investigation.

Nevertheless, with high-resolution DNA-PK structures in different states, our findings reveal the structural mechanism of allosteric

activation of the atypical kinase DNA-PK, the key step of its catalytic cycle, and offer improved models for future structure-guided drug discovery of new cancer therapeutics targeting the DNA-PK holoenzyme.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41594-022-00881-w>.

## References

1. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).
2. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
3. Carter, T., Vancurová, I., Sun, I., Lou, W. & DeLeon, S. A DNA-activated protein kinase from HeLa cell nuclei. *Mol. Cell. Biol.* **10**, 6460–6471 (1990).
4. Jackson, S. P., MacDonald, J. J., Lees-Miller, S. & Tjian, R. GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. *Cell* **63**, 155–165 (1990).
5. Lees-Miller, S. P., Chen, Y. R. & Anderson, C. W. Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen. *Mol. Cell. Biol.* **10**, 6472–6481 (1990).
6. Dvir, A., Peterson, S. R., Knuth, M. W., Lu, H. & Dynan, W. S. Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II. *Proc. Natl Acad. Sci. USA* **89**, 11920–11924 (1992).
7. Gottlieb, T. M. & Jackson, S. P. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* **72**, 131–142 (1993).
8. Abraham, R. T. Mammalian target of rapamycin: immunosuppressive drugs uncover a novel pathway of cytokine receptor signaling. *Curr. Opin. Immunol.* **10**, 330–336 (1998).
9. Sibanda, B. L., Chirgadze, D. Y. & Blundell, T. L. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. *Nature* **463**, 118–121 (2010).
10. Sibanda, B. L., Chirgadze, D. Y., Ascher, D. B. & Blundell, T. L. DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. *Science* **355**, 520–524 (2017).
11. Sharif, H. et al. Cryo-EM structure of the DNA-PK holoenzyme. *Proc. Natl Acad. Sci. USA* **114**, 7367–7372 (2017).
12. Yin, X., Liu, M., Tian, Y., Wang, J. & Xu, Y. Cryo-EM structure of human DNA-PK holoenzyme. *Cell Res.* **27**, 1341–1350 (2017).
13. Chen, X. et al. Structure of an activated DNA-PK and its implications for NHEJ. *Mol. Cell* **81**, 801–810.e3 (2021).
14. Chaplin, A. K. et al. Dimers of DNA-PK create a stage for DNA double-strand break repair. *Nat. Struct. Mol. Biol.* **28**, 13–19 (2021).
15. Chaplin, A. K. et al. Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Mol. Cell* **81**, 3400–3409 (2021).
16. Chen, S. et al. Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature* **593**, 294–298 (2021).
17. Liu, L. et al. Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Mol. Cell* **82**, 177–189.e4 (2022).
18. Liang, S. et al. Structural insights into inhibitor regulation of the DNA repair protein DNA-PKcs. *Nature* **601**, 643–648 (2022).
19. Zhang, Z. et al. Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. *Structure* **12**, 495–502 (2004).
20. Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607–614 (2001).
21. Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239–242 (1996).
22. Lees-Miller, J. P. et al. Uncovering DNA-PKcs ancient phylogeny, unique sequence motifs and insights for human disease. *Prog. Biophys. Mol. Biol.* **163**, 87–108 (2021).
23. Hammel, M. & Tainer, J. A. X-ray scattering reveals disordered linkers and dynamic interfaces in complexes and mechanisms for DNA double-strand break repair impacting cell and cancer biology. *Protein Sci.* **30**, 1735–1756 (2021).
24. Imseng, S., Aylett, C. H. & Maier, T. Architecture and activation of phosphatidylinositol 3-kinase related kinases. *Curr. Opin. Struct. Biol.* **49**, 177–189 (2018).
25. Douglas, P. et al. Identification of in vitro and in vivo phosphorylation sites in the catalytic subunit of the DNA-dependent protein kinase. *Biochem. J.* **368**, 243–251 (2002).
26. Cui, X. 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).
27. Douglas, P. 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).
28. Dobbs, T. A., Tainer, J. A. & Lees-Miller, S. P. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA Repair* **9**, 1307–1314 (2010).
29. Hammel, M. et al. Visualizing functional dynamicity in the DNA-dependent protein kinase holoenzyme DNA-PK complex by integrating SAXS with cryo-EM. *Prog. Biophys. Mol. Biol.* **163**, 74–86 (2021).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023



## Methods

### Purification of Ku70/80 and DNA-PKcs

Purifications of Ku70/80 and DNA-PKcs were performed as described previously<sup>30</sup>.

### DNA annealing

DNA oligonucleotides (43 nucleotides) designed for blunt-end dsDNA were purchased from Sigma Aldrich. The complementary oligonucleotides were dissolved in ultrapure water to a final concentration of 100  $\mu$ M, mixed at a molar ratio of 1:1 and annealed in a thermocycler. The mixed volume was first heated to 98 °C for 5 min and then cooled to 25 °C for 2 h to form a 43-base-pair (bp) dsDNA. The DNA sequences used for annealing can be found below.

dsDNA forward:

CCCCTGCCGATTCGCTGGAACATTTAAATTCGTCGACCTCC

dsDNA reverse:

GGAGGTCGACGAATTTAATGTTCCAGCGGAATCGGCAGCGGG

### Sample preparation and cryo-EM data collection

Purified DNA-PKcs, Ku70/80 and 43-bp DNA were mixed at a molar ratio of 1.2:1:1 in 20 mM HEPES pH 7.6, 200 mM NaCl, 0.5 mM EDTA, 2 mM MgCl<sub>2</sub>, 2 mM DTT and 1 mM M3814 and incubated on ice for 1 h. After incubation, 3  $\mu$ l of mixture was directly loaded to a Quantifoil R1.2/1.3 grid 300 Mesh Cu with graphene oxide (GO). Protein samples were left on a grid for 20 s, then blotted and plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV system (Thermo Fisher Scientific) at 4 °C and 100% humidity. The grids were then clipped and transferred to a Titan Krios electron microscope operating at a voltage of 300 kV with a K3 direct electron detector (Gatan) at the cryo-EM facility of the Department of Biochemistry, University of Cambridge. Automated data collection was performed using the Thermo-Scientific EPU software package (Table 1).

### Image processing and model refinement

Data processing was done using Warp and cryoSPARC<sup>31–33</sup>. CTF correction, motion correction and particle auto-picking were carried out using Warp. All particles were subjected to ab initio reconstruction in cryoSPARC to generate four initial three-dimensional models. Classes of DNA-PKcs only and contamination were discarded and classes of DNA-PK were selected for further optimization through iterative rounds of heterogeneous refinement to filter out unwanted particles. The two models of DNA-PK in different states were then further refined using homogeneous refinement, global CTF refinement and nonuniform refinement in cryoSPARC. The reported resolutions were determined based on the ‘gold standard’ of the 0.143 Fourier shell correlation (FSC) criterion<sup>34</sup>. Maps were further improved using ResolveCryoEM in PHENIX<sup>35</sup>. Regarding the modeling of the complexes, DNA bases were manually built in COOT, and previously published cryo-EM structures of DNA-PK (PDB: 6ZHA, 7KIN, 7KOY) were used as an initial reference for the protein modeling<sup>13,14,36</sup>. The templates were first rigid-body-fitted into the maps in CHIMERA and CHIMERAX, followed by real-space refinement in PHENIX<sup>37–39</sup>. The ligand was then modeled into the corresponding density in the cryo-EM maps in COOT, followed by multiple refinement rounds in PHENIX and model building in COOT<sup>36,38</sup>.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Cryo-EM maps have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-14546 (DNA-PK in the intermediate state) and EMD-14545 (DNA-PK in the active state). Model coordinates have been deposited in the Protein Data Bank under accession

numbers 7Z88 (DNA-PK in the intermediate state) and 7Z87 (DNA-PK in the active state). Other structures used for modeling and comparison in this study were retrieved from the Protein Data Bank with accession codes 6ZFP for DNA-PKcs; 7OTY for the DNA-PKcs–inhibitor complex; and 6ZHA, 7KIN and 7KOY for DNA-PK.

## References

- Wang, J. L. et al. Dissection of DNA double-strand-break repair using novel single-molecule forceps. *Nat. Struct. Mol. Biol.* **25**, 482–487 (2018).
- Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).
- Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
- Punjani, A., Brubaker, M. A. & Fleet, D. J. Building proteins in a day: efficient 3D molecular structure estimation with electron cryomicroscopy. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 706–718 (2017).
- Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
- Terwilliger, T. C., Ludtke, S. J., Read, R. J., Adams, P. D. & Afonine, P. V. Improvement of cryo-EM maps by density modification. *Nat. Methods* **17**, 923–927 (2020).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
- Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
- Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

## Acknowledgements

We thank L. Cooper, S. W. Hardwick and D. Y. Chirgadze from the cryo-EM facility of the Department of Biochemistry, University of Cambridge, for their help and support, and S. E. Thomas and A. K. Chaplin for help and discussion on modeling. We are also grateful for an Investigator Award from the Wellcome Trust (200814/Z/16/Z; 2016–2022) to T.L.B. for support of S.L. in this research.

## Author contributions

S.L. purified the protein, prepared the cryo-EM samples, collected and processed the cryo-EM data, and modeled the protein structures. S.L. wrote the first draft of the paper. T.L.B. directed the study, provided advice and edited the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41594-022-00881-w>.

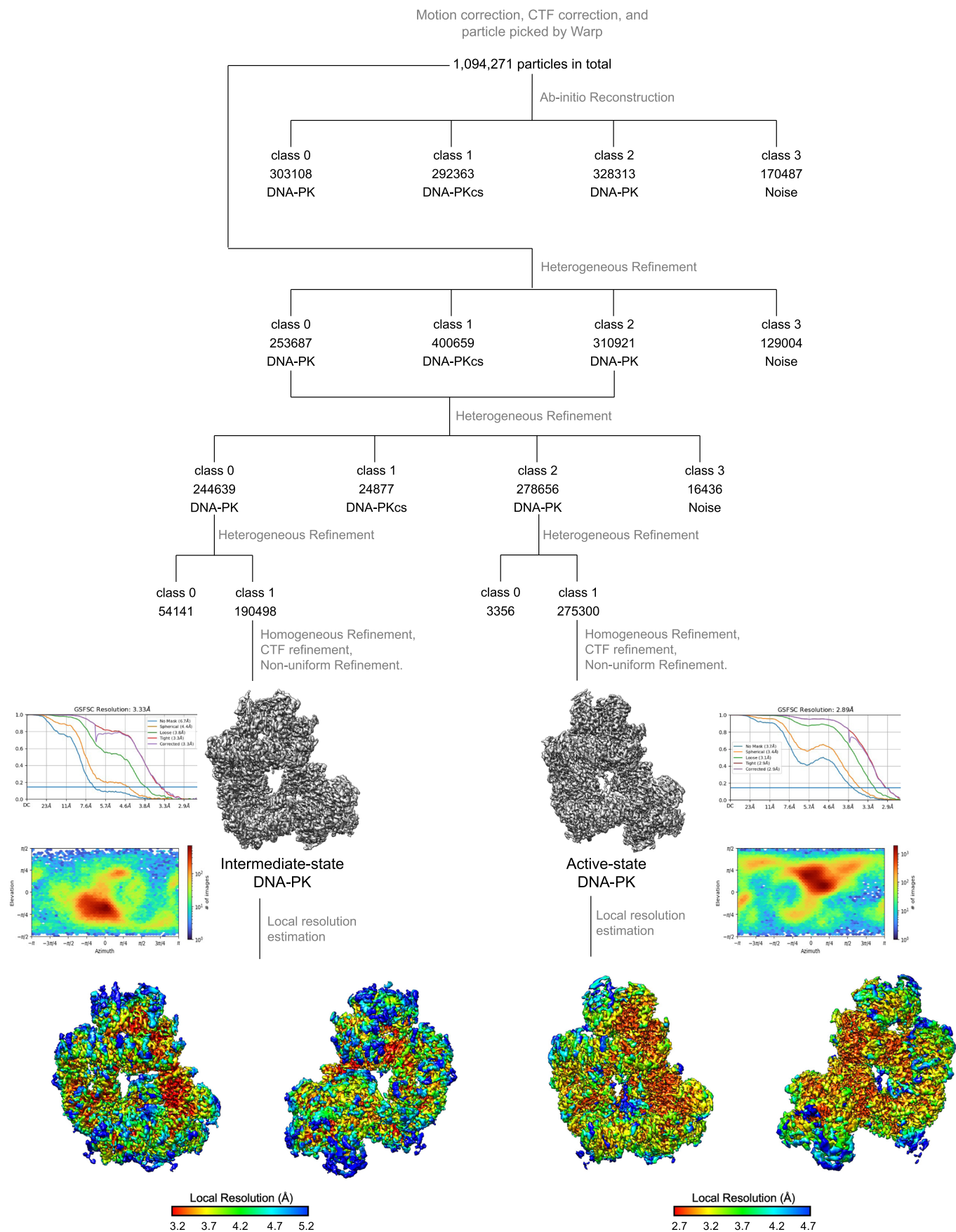
**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41594-022-00881-w>.

**Correspondence and requests for materials** should be addressed to Shikang Liang or Tom L. Blundell.

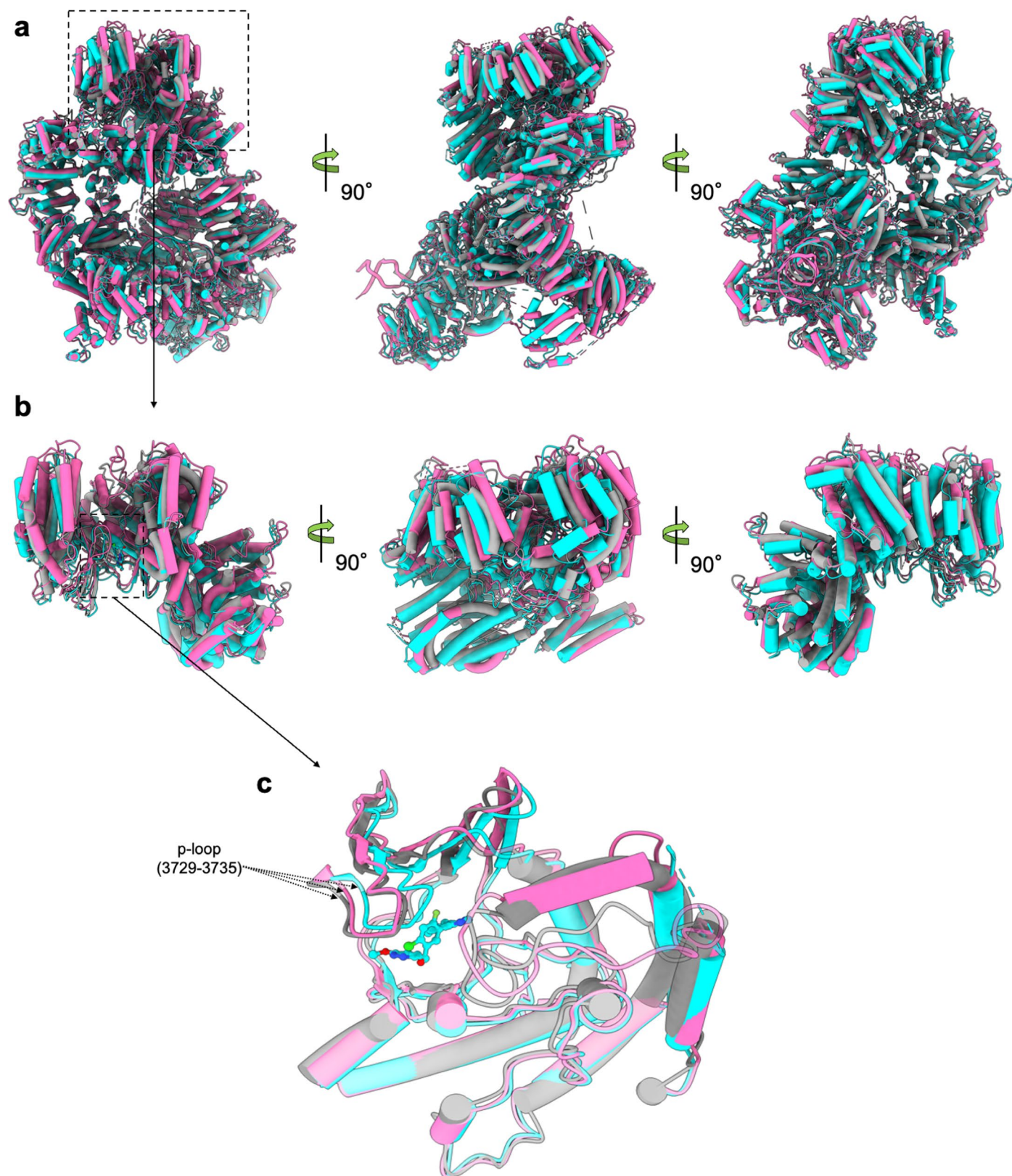
**Peer review information** *Nature Structural & Molecular Biology* thanks the anonymous reviewers for their contribution to the peer review of this

work. Primary Handling editors: Beth Moorefield and Carolina Perdigoto, in collaboration with the *Nature Structural & Molecular Biology* team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

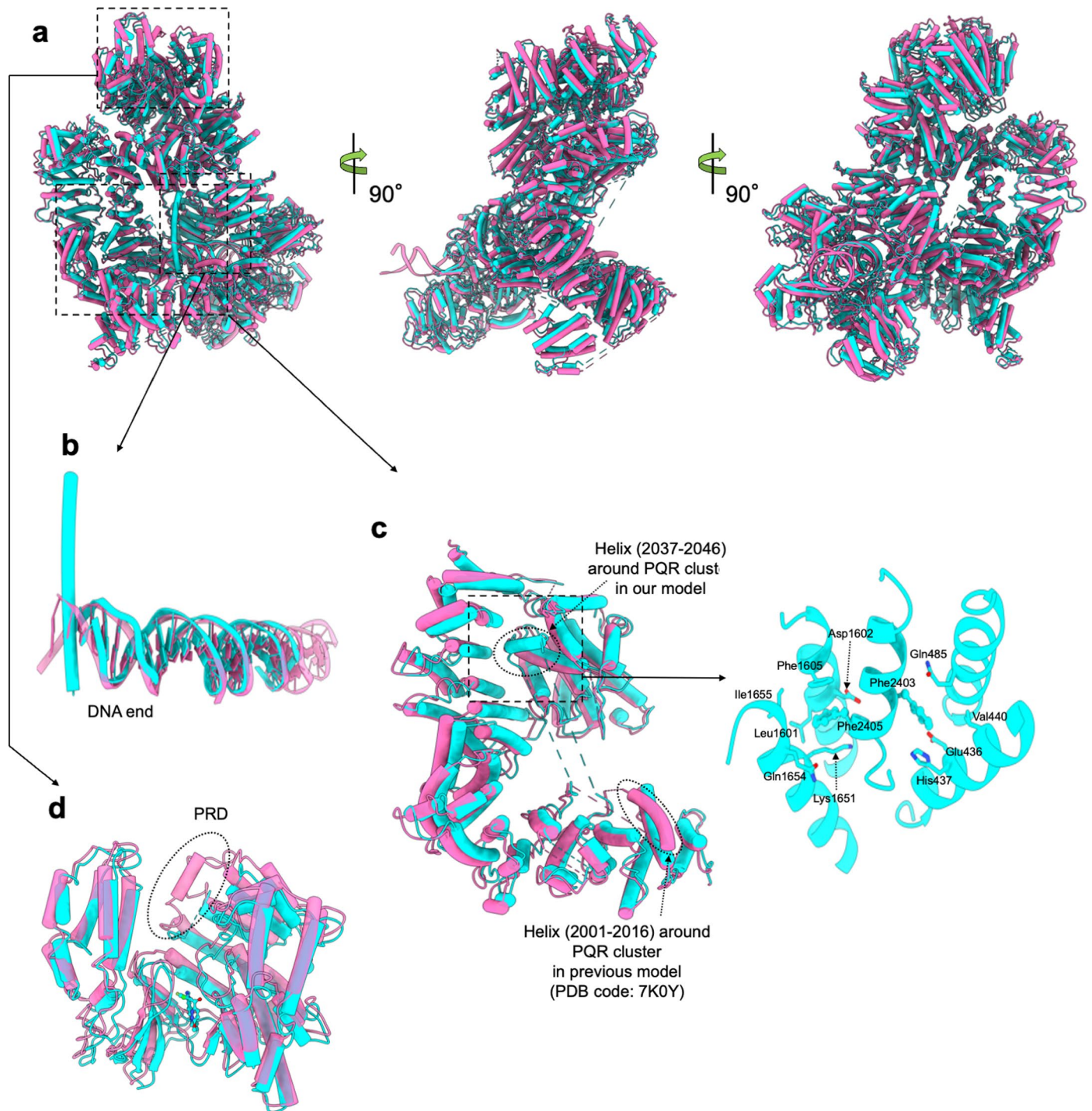


Extended Data Fig. 1 | Cryo-EM data processing of the DNA-PK/M3814 complex.



**Extended Data Fig. 2 | Comparison of DNA-PK in the intermediate state with inactive DNA-PKs. a**, Superimposed DNA-PKs in the intermediate state (in cyan) and apo DNA-PKs (PDB code: **6ZHA** in grey; PDB code: **7K1N** in pink) when aligned by the DNA ends, including front (left), side (middle) and back (right) views. The main conformational differences lie in the Head unit of DNA-PKs. **b**, Front (left), side (middle) and back (right) views of the superimposed Head units in the intermediate state and apo state when align by the DNA ends. Helices of the

Head unit of the inactive apo DNA-PKs mostly cluster together while helices of the intermediate DNA-PK tend to move toward the N terminus and DNA end. **c**, Magnified view of the superimposed kinase region in intermediate state and apo state when align by the C-lobe of the kinase. The addition of the inhibitor changes the conformation of N-lobe via interacting with the p-loop, leading to parallel movement of the p-loop flanking sheets.



**Extended Data Fig. 3 | Comparison of DNA-PK in the active state with another activated DNA-PK model. a**, Superimposed DNA-PK of our model in the active state (in cyan) and the previously published DNA-PK (PDB code: 7K0Y in pink), including front (left), side (middle) and back (right) views. There are global conformational differences. **b**, Magnified view of superimposed DNA ends of the two models. In our model, the DNA end does not go as far due to the DNA stopping helix. **c**, Magnified view of the superimposed region of Circular Cradle and N-terminal unit near the helix around PQR cluster (2037–2046) in the two

models. In our model, the helix sits in between the Circular Cradle and N-terminal unit. In the previous model, the helix (2001–2016) around PQR cluster sits on the Circular Cradle. **d**, Magnified view of the superimposed Head units in our model and the previously described apo DNA-PK close to the active state. In our model, the addition of ligand on ATP binding groove destabilizes or releases the PRD from its position in inactive state and the Head unit is further pulled down to the DNA end. The whole Head unit is also further pulled down towards the DNA end.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection Thermo-Scientific EPU 2.10.0

Data analysis WARP v1.0.8, cryoSPARC v2.5.0, Phenix v1.19, COOT 0.8.9.2, Chimera 1.15.0, ChimeraX-1.1.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Cryo-EM maps have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-14546 (DNA-PK in the intermediate state), EMD-14545 (DNA-PK in the active state). Model coordinates have been deposited in the Protein Data Bank under accession numbers 7Z88 (DNA-PK in the intermediate state), 7Z87 (DNA-PK in the active state).

Other structures used for modelling and comparison in this study were retrieved from the PDB with accession codes: 6ZFP for DNA-PKcs; 7OTY for DNA-PKcs/inhibitor complex; 6ZHA, 7K1N, 7K0Y for DNA-PK.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	There were no statistical methods used to predetermine sample size. A total of 1094271 particles was collected for this study. The final particle numbers for reconstructions are comparable to recent cryo-EM studies in the field for the final reconstruction.
Data exclusions	During the cryo-EM data analysis, bad particles and contamination, including ice and protein aggregation, were discarded through iterative 3D classifications. This procedure is standard in the single particle analysis cryo-EM field.
Replication	Refinements of the DNA-PK structures were repeated at least twice with similar 3D maps.
Randomization	For cryo-EM, particles used in refinements were randomly split into two halves automatically and cross correlation coefficients between the two half sets are calculated.
Blinding	Particles were randomly split into two halves in cryo-EM data processing and resolution estimation. Blinding is not feasible for cryo-EM processing, because visual inspection is necessary to check the data quality.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging