

# Mechanism of single-stranded DNA annealing by RAD52–RPA complex

<https://doi.org/10.1038/s41586-024-07347-7>

Received: 6 June 2023

Accepted: 22 March 2024

Published online: 24 April 2024

Open access

 Check for updates

Chih-Chao Liang<sup>1</sup>✉, Luke A. Greenhough<sup>1</sup>, Laura Masino<sup>1</sup>, Sarah Maslen<sup>1</sup>, Ilirjana Bajrami<sup>1</sup>, Marcel Tuppi<sup>1,2</sup>, Mark Skehel<sup>1</sup>, Ian A. Taylor<sup>1</sup> & Stephen C. West<sup>1</sup>✉

RAD52 is important for the repair of DNA double-stranded breaks<sup>1,2</sup>, mitotic DNA synthesis<sup>3–5</sup> and alternative telomere length maintenance<sup>6,7</sup>. Central to these functions, RAD52 promotes the annealing of complementary single-stranded DNA (ssDNA)<sup>8,9</sup> and provides an alternative to BRCA2/RAD51-dependent homologous recombination repair<sup>10</sup>. Inactivation of RAD52 in homologous-recombination-deficient *BRCA1*- or *BRCA2*-defective cells is synthetically lethal<sup>11,12</sup>, and aberrant expression of RAD52 is associated with poor cancer prognosis<sup>13,14</sup>. As a consequence, RAD52 is an attractive therapeutic target against homologous-recombination-deficient breast, ovarian and prostate cancers<sup>15–17</sup>. Here we describe the structure of RAD52 and define the mechanism of annealing. As reported previously<sup>18–20</sup>, RAD52 forms undecameric (11-subunit) ring structures, but these rings do not represent the active form of the enzyme. Instead, cryo-electron microscopy and biochemical analyses revealed that ssDNA annealing is driven by RAD52 open rings in association with replication protein-A (RPA). Atomic models of the RAD52–ssDNA complex show that ssDNA sits in a positively charged channel around the ring. Annealing is driven by the RAD52 N-terminal domains, whereas the C-terminal regions modulate the open-ring conformation and RPA interaction. RPA associates with RAD52 at the site of ring opening with critical interactions occurring between the RPA-interacting domain of RAD52 and the winged helix domain of RPA2. Our studies provide structural snapshots throughout the annealing process and define the molecular mechanism of ssDNA annealing by the RAD52–RPA complex.

RAD52 has important roles in two pathways of recombinational repair. The highly conserved N-terminal domain (NTD; amino acids 1–209) binds to ssDNA and promotes ssDNA annealing (SSA)<sup>8,18,19</sup>, while the divergent C-terminal domain interacts with RPA (comprising RPA1, RPA2 and RPA3)<sup>9,21</sup> and, at least in *Saccharomyces cerevisiae*, promotes the loading of RAD51 recombinase<sup>22–24</sup>. In vertebrates, this latter role is the responsibility of BRCA2<sup>25–27</sup> and RAD51 paralogue complexes<sup>28,29</sup>, such that the primary role of RAD52 relates to its ability to promote the annealing of complementary ssDNAs. To determine the mechanism of annealing, we used cryo-electron microscopy (cryo-EM) to define the structure of RAD52 and the RAD52–ssDNA complex, and also obtained snapshots of the annealing process through visualization of a RAD52–RPA–ssDNA complex.

## RAD52 open rings promote SSA

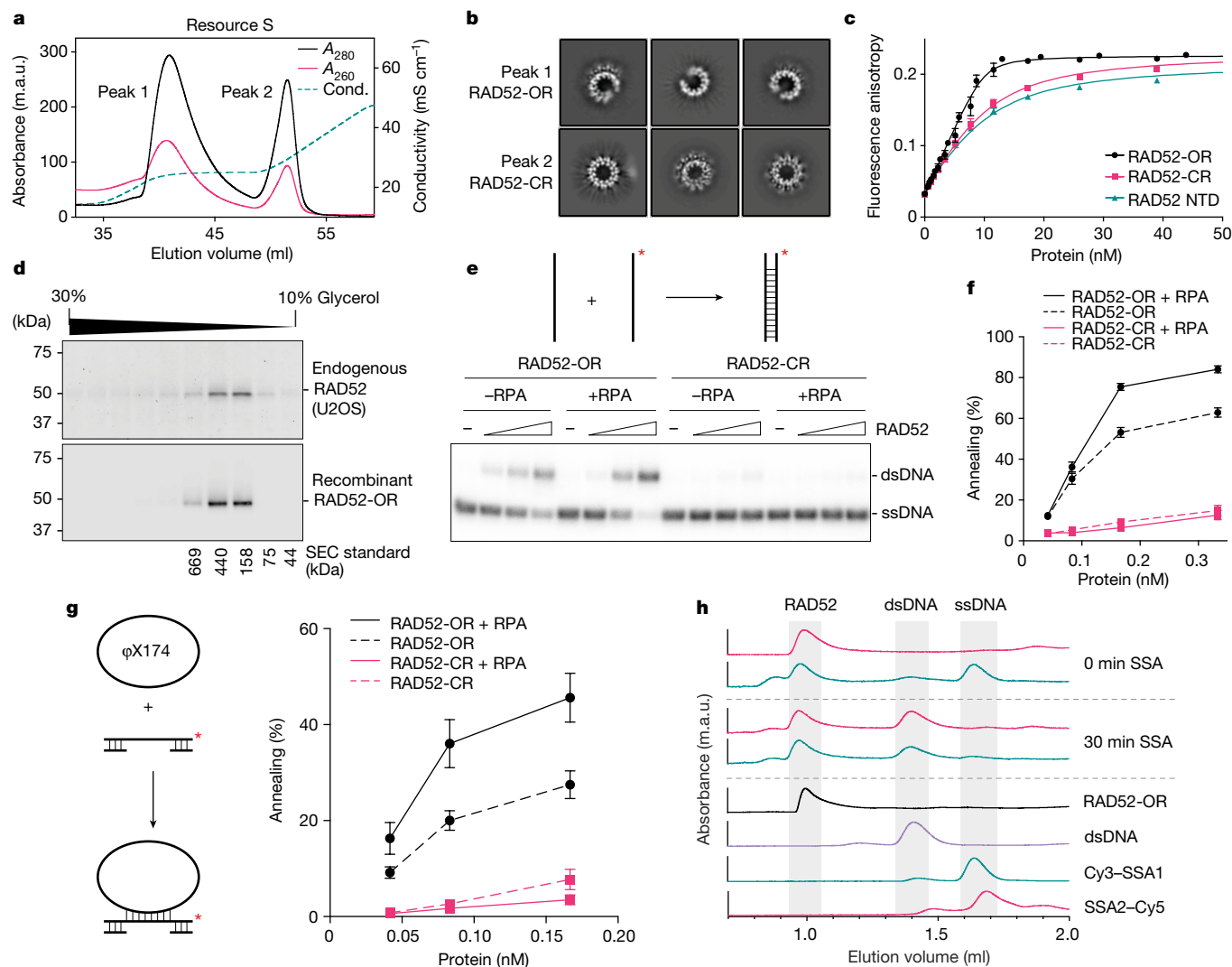
Human RAD52 was expressed in *Escherichia coli* and purified to homogeneity. During cation-exchange chromatography, the protein separated into two distinct species that appeared identical when analysed by SDS–PAGE (Fig. 1a and Extended Data Fig. 1a). Similar results were obtained with protein expressed in baculovirus-infected Sf9 insect cells (Extended Data Fig. 1b). By contrast, the N-terminal domain (NTD)

of RAD52<sup>18,19</sup> eluted as a single species from this column (Extended Data Fig. 1c,d). When visualized using cryo-EM, we found that the two species of full-length RAD52 represent open (RAD52-OR, peak 1) and closed (RAD52-CR, peak 2) ring forms (Fig. 1b). The 11-subunit closed rings exhibit features similar to the RAD52 NTD, whereas the RAD52 open rings have one or more subunits missing. Circular dichroism confirmed that both forms adopted the same secondary structure (Extended Data Fig. 1e), and intact protein mass spectrometry (MS) analysis showed that their mass approximated the calculated mass of RAD52 less the N-terminal methionine residue (Extended Data Fig. 1f).

The RAD52 open rings exhibited a high affinity ( $K_D = 0.3 \pm 0.1$  nM) for 40-nucleotide-long ssDNA but not double-stranded DNA (dsDNA), as measured by fluorescence anisotropy (Fig. 1c and Extended Data Fig. 2a). Binding was observed with linear ssDNA, ssDNAs in which the 5' or 3' ends were blocked by biotin–streptavidin, and a ssDNA/dsDNA substrate, as measured using biolayer interferometry (Extended Data Fig. 2b). By contrast, the RAD52 closed rings bound to ssDNA with reduced affinity ( $3.3 \pm 0.5$  nM) comparable to that of the RAD52 NTD ( $3.2 \pm 0.6$  nM) (Fig. 1c).

To determine whether the RAD52 oligomers, purified after overexpression in *E. coli* or insect cells, are representative of RAD52 within

<sup>1</sup>The Francis Crick Institute, London, UK. <sup>2</sup>Present address: Abcam, Cambridge Biomedical Campus, Cambridge, UK. ✉e-mail: eric.liang@crick.ac.uk; stephen.west@crick.ac.uk



**Fig. 1 | Open rings represent the active form of RAD52.** **a**, Resource S cation-exchange chromatography analysis of recombinant human RAD52. Cond., conductivity. **b**, Representative cryo-EM 2D class averages of RAD52 open (RAD52-OR) and closed rings (RAD52-CR). **c**, Single-stranded DNA (40 nucleotides: FAM-SSA4) binding by RAD52-OR, RAD52-CR or RAD52 NTD measured using fluorescence anisotropy. The lines are the best quadratic curve fits. Data are mean + s.e.m.  $n = 6$  (RAD52-OR),  $n = 3$  (RAD52-CR) and  $n = 3$  (RAD52 NTD) independent experiments. **d**, Glycerol gradient sedimentation analysis of a nuclear extract from U2OS cells compared with recombinant RAD52-OR. RAD52 was detected by western blotting. Gel-filtration protein standards are shown. **e**, Representative PAGE assay of SSA by the open or closed rings of

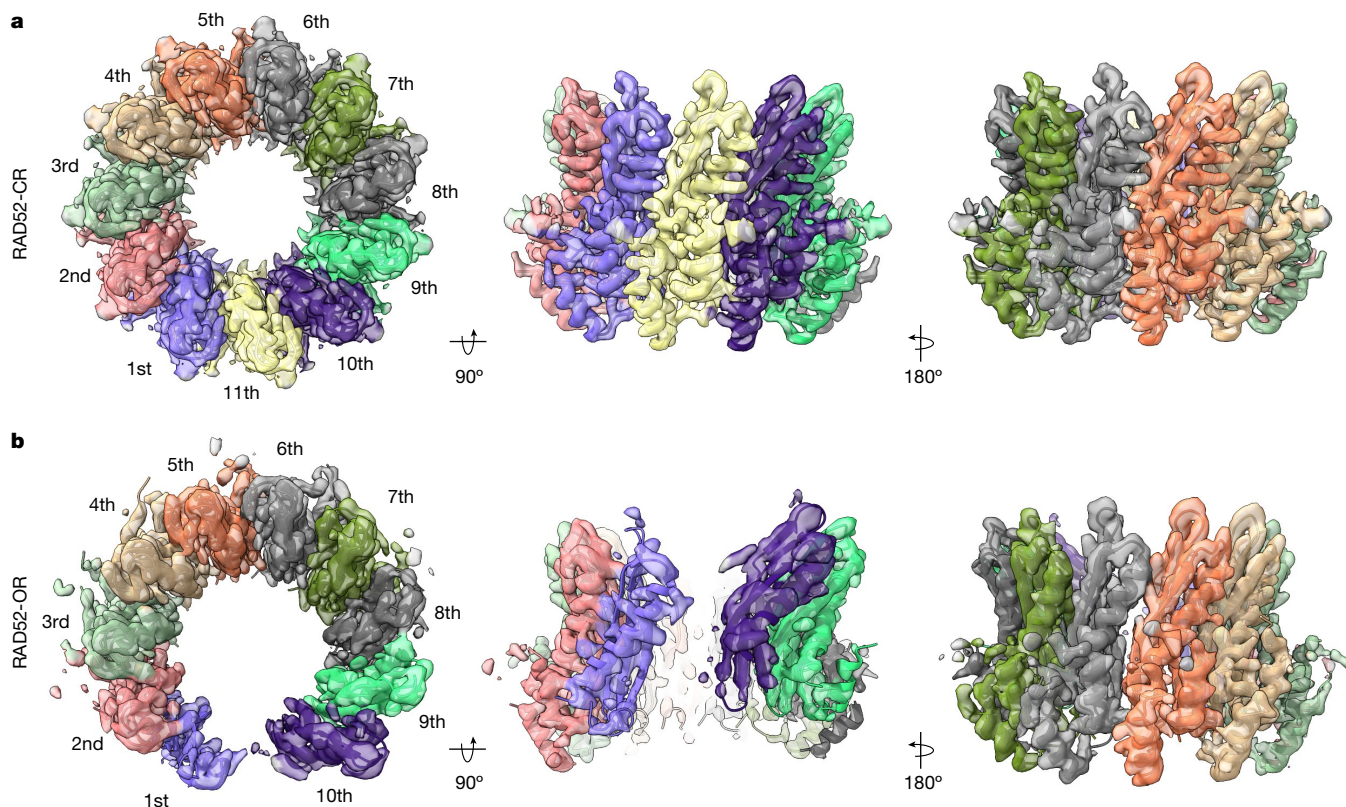
RAD52 (0, 0.08, 0.17 and 0.33 nM) using 68-nucleotide-long ssDNA (0.33 nM) with or without RPA (0.33 nM). **f**, Quantification of the SSA assays from **e**. Data are mean + s.e.m.  $n = 22$  (RAD52-OR and RAD52-OR + RPA) and  $n = 7$  (RAD52-CR and RAD52-CR + RPA) independent experiments. **g**, SSA using  $\phi$ X174 circular ssDNA and a gapped duplex by RAD52 (OR or CR) in the presence or absence of RPA. Data are mean + s.e.m.  $n = 4$  independent experiments. **h**, SEC analysis of RAD52-mediated SSA between Cy3-SSA1 (dark cyan, recorded at 647 nm) and SSA2-Cy5 (pink, 555 nm) labelled ssDNAs. RAD52 was preloaded on SSA2-Cy5 before addition of Cy3-SSA1. RAD52-OR (black) was recorded at 280 nm. In **e**,  $^{32}$ P labels are indicated with asterisks.

human cells, the recombinant RAD52 open rings were compared with endogenous RAD52 contained within a nuclear extract from U2O2 cells. Endogenous RAD52 exhibited a similar oligomeric state to the recombinant RAD52-ORs, as determined using glycerol gradient sedimentation (Fig. 1d) and size-exclusion chromatography (Extended Data Fig. 2c).

We next analysed single-strand annealing by RAD52 using in vitro assays in which  $^{32}$ P-labelled ssDNA (68 nucleotides) was incubated with its complementary strand in the presence or absence of RPA (Fig. 1e and Extended Data Fig. 2d). The RAD52-ORs annealed ssDNA in reactions stimulated by RPA, whereas the RAD52-CRs exhibited a reduced ability to promote annealing that was unaffected by the presence or absence of RPA (Fig. 1e,f). The stimulatory effect of RPA on RAD52-OR-mediated annealing was not observed with shorter ssDNAs (40 nucleotides) (Extended Data Fig. 2e). As a RAD52 ring

can bind to approximately 40 nucleotides of ssDNA<sup>18,30</sup>, these results show that efficient annealing requires the stable association of both RAD52 and RPA on ssDNA. Consistent with the DNA-binding experiments, RAD52-OR-mediated SSA did not require free ssDNA ends (Fig. 1g). Size-exclusion chromatography (SEC) revealed that dsDNA dissociated from RAD52 after completion of annealing (Fig. 1h), consistent with the low affinity of RAD52 for dsDNA (Extended Data Fig. 2a).

Denaturation of a mixed population of open and closed rings of RAD52 using guanidinium hydrochloride, followed by renaturation, resulted in the majority of the protein adopting an open-ring conformation (Extended Data Fig. 2f,g). The refolded RAD52-ORs were as active as purified RAD52 open rings (Extended Data Fig. 2h), consistent with the observation that open rings represent the active form of RAD52.



**Fig. 2 | Cryo-EM structures of RAD52 closed and open rings. a**, Top and side views of the RAD52-CR cryo-EM map (2.9 Å) and atomic model. **b**, Top and side views of the RAD52-OR cryo-EM map (3.2 Å) and atomic model. RAD52 subunits are numbered successively.

### RAD52-CR and RAD52-OR structures

The structures of RAD52-CR (2.9 Å) and RAD52-OR (3.2 Å) were determined using cryo-EM (Fig. 2, Extended Data Figs. 3a–f and 4a–f and Extended Data Table 1). The closed ring comprised 11 subunits as observed previously<sup>18,19,31</sup>. The RAD52-OR structures contained a maximum of ten subunits, although open rings with fewer subunits were also observed (Extended Data Fig. 4a,g). The structure of each subunit in the open ring was similar to that of the closed ring. As expected, the cryo-EM density around the site of ring opening was not as well resolved as other parts of the structure (Extended Data Fig. 4d).

### Interaction between RAD52 and ssDNA

In contrast to the X-ray crystal structure of the RAD52 NTD (Protein Data Bank (PDB): 1H2I)<sup>18,19</sup>, we could not visualize the ssDNA-binding domain (DBD, amino acids 46–67) in the cryo-EM model of full-length RAD52 (Extended Data Fig. 4h). One possibility is that this domain is flexible before interaction with ssDNA, enabling rapid and tight association with ssDNA (Fig. 1c). To determine the validity of this hypothesis, the structure of the RAD52-OR bound to ssDNA (2.3 Å) was solved (Fig. 3a, Extended Data Fig. 5a–e and Extended Data Table 1). An analysis of the RAD52–ssDNA complex revealed that the DBD was stabilized as ssDNA bound into the positively charged groove on the outside of the RAD52 ring. A comparison with the crystal structure of the RAD52-NTD(K102A/K133A)–ssDNA complex (PDB: 5XRZ) revealed a very similar organization of the ssDNA (Extended Data Fig. 5f).

In the RAD52-OR–ssDNA complex, the ssDNA could be visualized from the second to the seventh RAD52 subunit (Fig. 3a). There were four ssDNA nucleotides per RAD52 subunit (Fig. 3b), consistent with the observed four-nucleotide pattern of hypersensitivity to hydroxyl radicals<sup>30</sup>. Topological analysis of the structure revealed that the bound

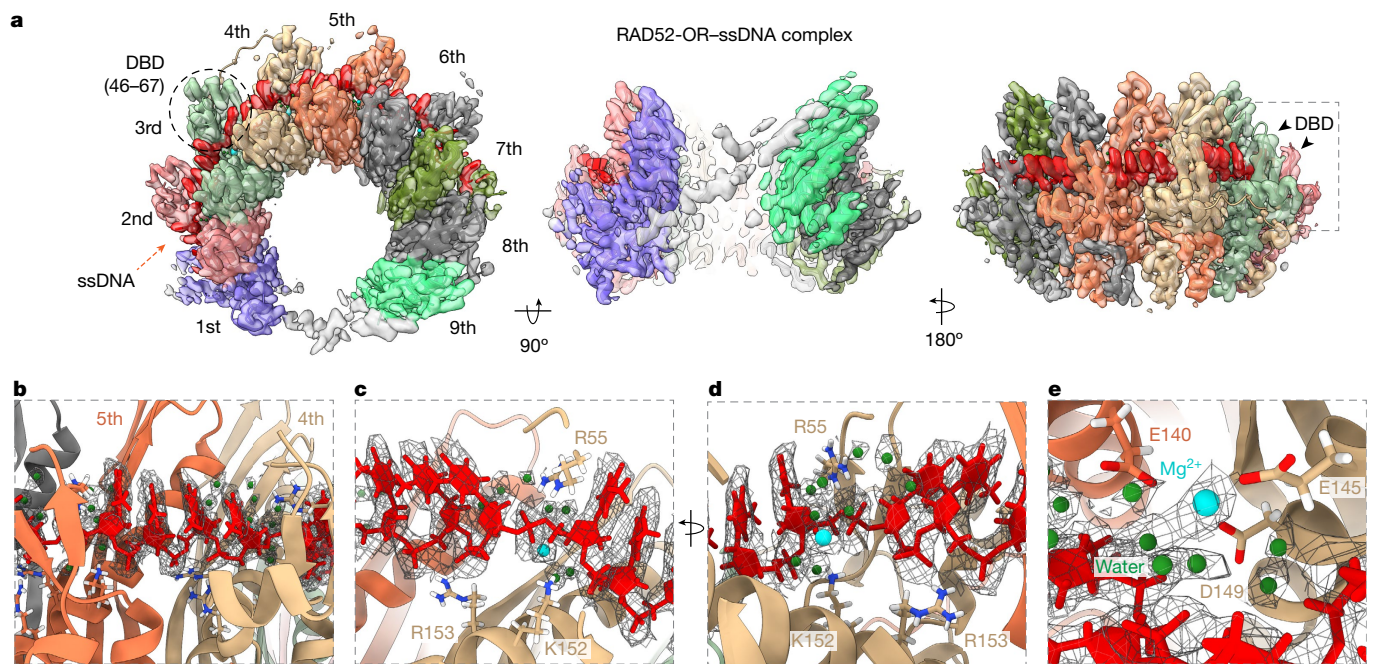
ssDNA was stretched non-uniformly, such that each quartet stack adopted the length of B-form duplex DNA. Arg153 coordinated the phosphate backbone of the two central nucleotides (Fig. 3c), whereas Arg55 and Lys152 coordinated the phosphate backbone between two quartet stacks (Fig. 3d). The Mg<sup>2+</sup>, which is required for RAD52–ssDNA interactions (Extended Data Fig. 5g), and the surrounding water molecules were coordinated by Glu145, Asp149 and Glu140 residues from the neighbouring subunit (Fig. 3e).

### Oligomeric state of human RAD52

Previously, undecameric and heptameric RAD52 closed rings were observed<sup>18,19,32</sup>, indicating a degree of structural flexibility by which subunit–subunit interactions can accommodate changes to the overall size of the protein. Consistent with this, in addition to the ten-subunit open ring, our analysis of two-dimensional (2D) averages and three-dimensional (3D) classes indicated the presence of RAD52-OR structures with fewer subunits (Extended Data Fig. 4a,g). To determine whether the oligomeric state was influenced by protein concentration, size-exclusion chromatography coupled with multi-angle laser light scattering (SEC-MALLS) was used to determine the molecular mass of RAD52-OR at a variety of concentrations and the in presence or absence of ssDNA. We found that the oligomeric state was dependent on protein concentration but was unaffected by ssDNA binding (Fig. 4a).

### Interactions modulate ring dynamics

The C-terminal region of RAD52 (amino acids 210–418) is predicted to be structurally disordered (Extended Data Fig. 6a). However, it is an important part of RAD52 as it interacts with RPA<sup>21,33</sup> and is required for nuclear localization<sup>34,35</sup>. As this region was not resolved in the cryo-EM structures, we used cross-linking MS (XL-MS) to detect dynamic



**Fig. 3 | Cryo-EM structure of the RAD52-ssDNA complex.** **a**, Top and side views of the RAD52-ssDNA complex cryo-EM map (2.3 Å) and atomic model. ssDNA is coloured red. The DBD is indicated on the third RAD52 subunit. **b**, Magnified view showing four ssDNA nucleotides (red) adopting the length of B-form DNA between two RAD52 monomers. **c, d**, Magnified view of ssDNA

interactions<sup>36</sup> between the N- and C-terminal domains of RAD52 (Fig. 4b). The N-terminal domain interacted with three hotspots in the C terminus of RAD52: one mapped to the RPA-interacting domain (RID), one is located around amino acid 35, and the other was located at the extreme C terminus.

The RID and C-terminal sequences of RAD52 are highly conserved in vertebrates (Extended Data Fig. 6b). To determine whether these regions contribute to ring conformation, two deletion mutants, RAD52( $\Delta$ RID) (deletion of amino acids 239–290) and RAD52( $\Delta$ C) (deletion of amino acids 401–418), were generated (Extended Data Fig. 6c, d). An analysis of wild-type RAD52 using cation-exchange chromatography revealed a 2:1 ratio of open to closed rings, whereas RAD52( $\Delta$ RID) and RAD52( $\Delta$ C) exhibited an increased percentage of closed rings (Fig. 4c). Thermal melting analyses confirmed that the open and closed forms of the RAD52( $\Delta$ RID) have a thermal stability similar to the wild-type RAD52, indicating that the mutant proteins have comparable open and closed architectures (Extended Data Fig. 6e). Given that the RID and the extreme C terminus are both positively charged (Extended Data Fig. 6f–h), whereas the N terminus of the exposed RAD52 subunit at site of ring opening is negatively charged (Extended Data Fig. 7a), these results indicate that these conserved regions interact through electrostatic interactions to prevent ring closure. In agreement, using focused 3D classification analysis, we found a subclass of RAD52-ssDNA particles that exhibited a low-resolution density across the site of ring opening corresponding to the C-terminal domain of RAD52 (Extended Data Fig. 7b).

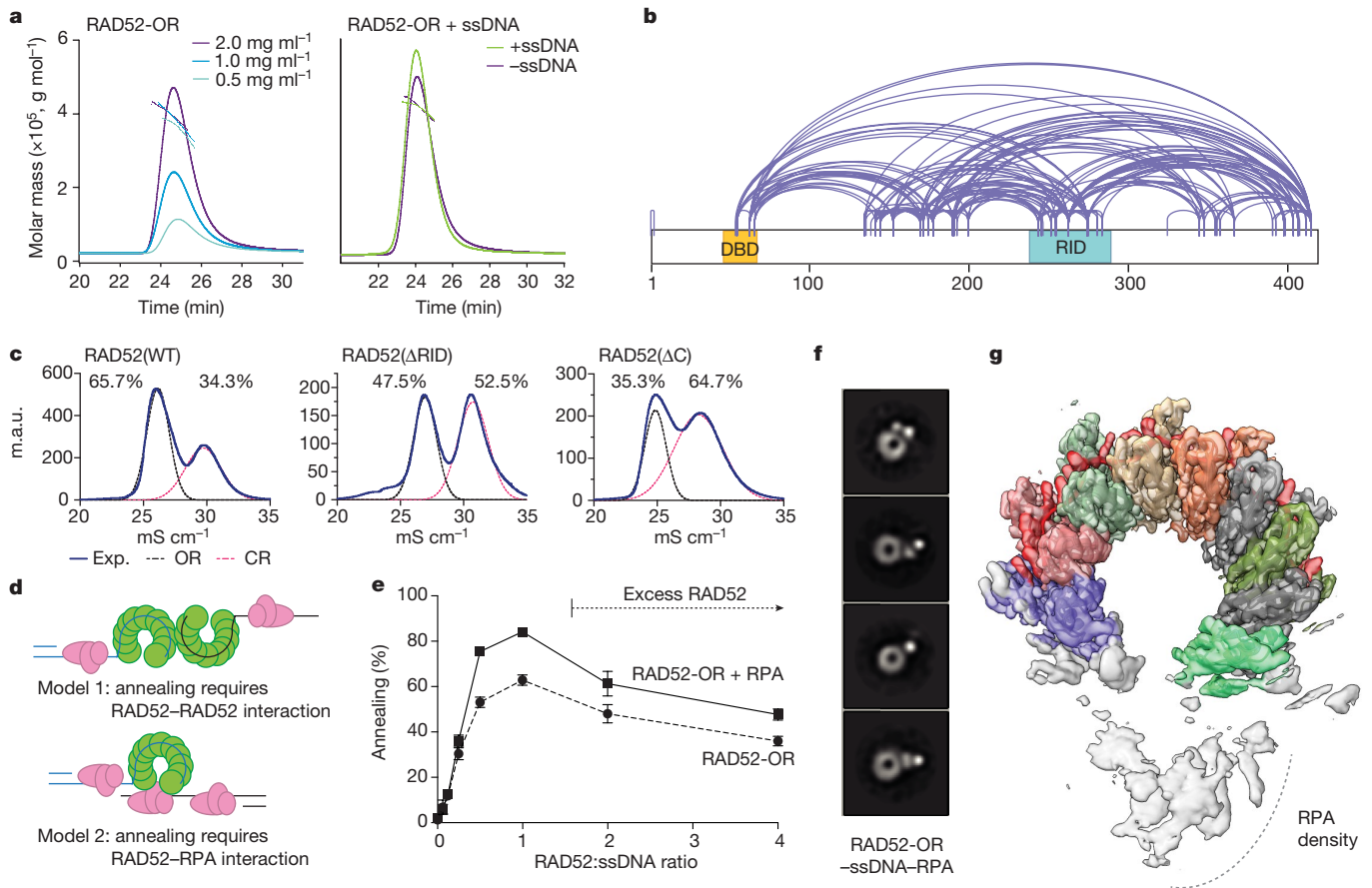
### RAD52-RPA interactions are required for SSA

Two models of SSA can be considered (Fig. 4d): (1) annealing occurs by interactions between two RAD52-ssDNA complexes<sup>20,37,38</sup>; or (2) that a RAD52-ssDNA complex anneals with naked or RPA-coated ssDNA. To help to distinguish between these models, we analysed the efficiency of SSA at different concentrations of RAD52 and found that excess RAD52, sufficient to saturate both ssDNAs, inhibited single-strand annealing

(Fig. 4e). Inhibition of SSA when the ssDNA is fully saturated by RAD52, and the requirement for RPA, leads us to suggest that model 2 should also be considered as a possible mechanism for SSA.

RAD52 open rings form more stable complexes with ssDNA (40 nucleotides) and RPA compared with RAD52 closed rings, as measured in pull-down experiments (Extended Data Fig. 7c), which may explain why the RAD52-CRs do not cooperate with RPA in SSA. To provide insights into the mechanism of annealing and to determine the role of RPA in the process, we next determined the cryo-EM structure of the RPA-ssDNA complex (3.2 Å) (Extended Data Figs. 7d and 8a–f). The density of the RPA trimeric core (47.9 kDa; RPA1 DBD-C, RPA2 DBD-D and RPA3 DBD-E) was well resolved, and we observed some additional density in RPA1 and RPA2 that corresponded to ssDNA. The molecular architecture of the human trimeric core was similar to its yeast and fungal orthologues<sup>39,40</sup>, with ssDNA bound to the positively charged groove across RPA1 DBD-C and RPA2 DBD-D (Extended Data Fig. 8g). Additional flexible densities, corresponding to DBD-A and DBD-B of RPA1 and the winged helix domain (WHD) (Extended Data Fig. 8h), were also observed, indicating that human RPA is structurally dynamic.

Analysis of the RAD52-ssDNA-RPA complex using negative-stain EM revealed no evidence of RAD52-RAD52 interactions, nor did we observe RPA bound to several RAD52 subunits around the open ring. Instead, we observed a single RPA protein sitting at the site of ring opening (Fig. 4f). These RAD52-ssDNA-RPA complexes were analysed using cryo-EM (3 Å), revealing that the RAD52 and ssDNA exhibited features similar to those present in the RAD52-OR-ssDNA cryo-EM structure (Fig. 4g and Extended Data Fig. 9a–d). Given that ssDNA was initially bound to RPA during sample preparation, these results would be consistent with the transfer of ssDNA from RPA to the RAD52 ring. The additional density corresponding to RPA at the opening of the RAD52 ring was not well resolved due to the inherent flexibility of the complex, and the local resolution around the RPA was low compared with RAD52-ssDNA (Extended Data Fig. 9d). CryoSPARC 3D classification analyses, focusing on the RPA, revealed various conformational states of the RPA density



**Fig. 4 | Mechanisms of annealing by RAD52.** **a**, SEC coupled to multi-angle laser light scattering (SEC-MALLS) analysis measuring the molecular mass of RAD52-OR at different protein concentrations and in the presence or absence of ssDNA. The solid lines are the chromatograms from the output of the differential refractometer and the scatter points are the weight-averaged molar masses determined at 1 s intervals throughout elution of the chromatographic peaks. **b**, XL-MS analysis of RAD52-ORs. **c**, Resource S cation chromatography of WT RAD52, RAD52( $\Delta$ RID) and RAD52( $\Delta$ C). Experimental ultraviolet 280 nm ( $UV_{280}$ ) absorbance is depicted as a solid blue line. Deconvoluted peaks are indicated (black lines, open ring; pink lines, closed ring). **d**, Schematic of the two

possible mechanisms for single-stranded annealing: (1) SSA by interactions between two complementary RAD52-bound ssDNAs; and (2) SSA involving interactions between the RAD52-ssDNA complex and complementary-strand ssDNA bound by RPA. **e**, The effect of RAD52 concentration on ssDNA (0.33 nM; 68 nucleotides) annealing in the presence and absence of RPA (0.33 nM). Data are mean + s.e.m.  $n = 5$  (RAD52-OR and RAD52-OR + RPA) independent experiments. **f**, Representative negative-stain EM 2D averages of reconstituted RAD52-OR-ssDNA-RPA complex (ssDNA: 68 nucleotides). **g**, Top view of the RAD52-OR-ssDNA-RPA complex cryo-EM map (3 Å) with the RAD52-ssDNA (68 nucleotides) atomic model.

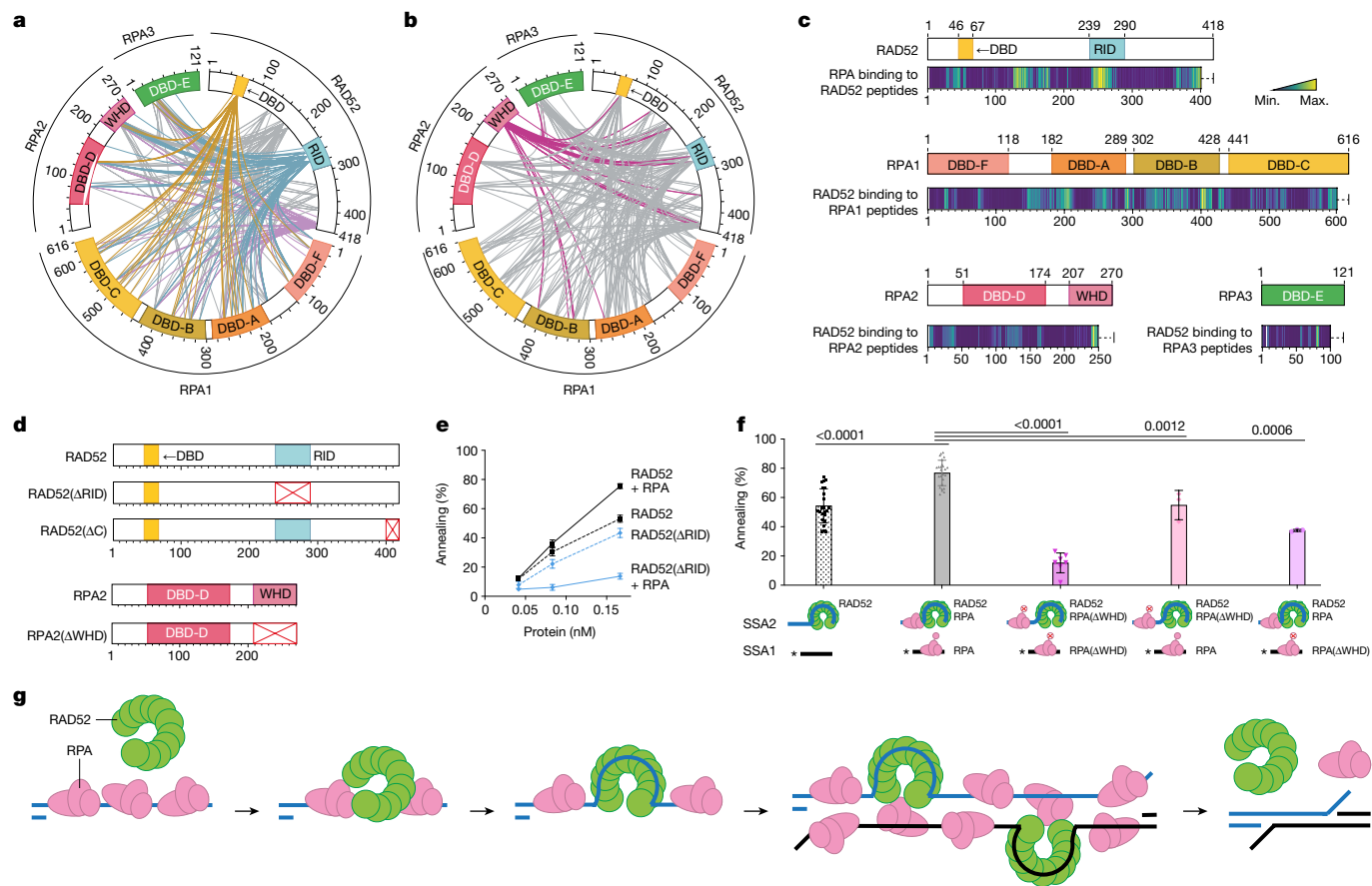
(Extended Data Fig. 9e) reflecting the heterogeneity and dynamics of the RAD52-RPA complex as observed by negative-stain EM (Fig. 4f).

### Interactions between the RID of RAD52 and RPA2

XL-MS was then used to analyse critical interactions between RAD52 and RPA. In addition to the RID of RAD52 (amino acids 239–290)<sup>21,33,37</sup>, we found that the DBD (amino acids 46–67) and the extreme C terminus (amino acids 401–418) of RAD52 interacted with RPA (Fig. 5a,b and Extended Data Fig. 10a,b). These two domains are dynamically positioned at the opening of the RAD52 ring, which allows RPA to interact with RAD52 at the site of ring opening. Notably, the DBD of RAD52 interacted specifically with DBD-A, DBD-B and DBD-C of RPA1, DBD-D of RPA2 and DBD-E of RPA3. One possibility is that the interactions between DNA-binding sites in both proteins facilitate the hand-over of ssDNA from RPA to RAD52 at, or close to, the site of ring opening. As expected from previous studies<sup>21,33,41</sup>, we observed interactions between RPA1 and RAD52, as well as interactions between the WHD of RPA2 and RAD52. The XL-MS data were well supported by quantitative information of RAD52-RPA interactions derived from peptide array analyses (Fig. 5c and Extended Data Fig. 10c–f).

To investigate the importance of these interaction hotspots, RAD52( $\Delta$ RID), RAD52(RQK/AAA) (RAD52 mutated at the key residues R260A, Q261A and K262A in the RPA-interaction domain)<sup>33,37</sup>, RAD52( $\Delta$ C), RPA( $\Delta$ FAB) (deletion of DBD-F, DBD-A and DBD-B) and RPA( $\Delta$ WHD) were purified (Fig. 5d and Extended Data Fig. 11a–d). We found that the RAD52( $\Delta$ C) mutant exhibited a similar annealing activity to that of wild-type RAD52 in the presence or absence of RPA (Extended Data Fig. 11e). However, the presence of RPA inhibited SSA by RAD52( $\Delta$ RID) (Fig. 5e). In contrast to the RID deletion, the RAD52(RQK/AAA) mutant was only inhibited when excess RPA was present (Extended Data Fig. 11f), indicating that other residues in the RID may also contribute to the interaction with RPA.

Analysis of the mutant RPAs showed that RPA( $\Delta$ FAB) stimulated RAD52-mediated SSA (Extended Data Fig. 11g), whereas RPA( $\Delta$ WHD) inhibited the reaction (Fig. 5f and Extended Data Fig. 11h). The inhibitory effect of RPA( $\Delta$ WHD) was particularly evident when bound to the complementary strand (strand 2) that interacts with the initiating RAD52-ssDNA complex (strand 1) (Fig. 5f). These data show that the RPA-interacting domain of RAD52, together with the WHD of RPA2, have critical roles in mediating interactions between a RAD52-ssDNA complex and complementary RPA-coated ssDNA leading to ssDNA annealing.



**Fig. 5 | Interactions between the RID of RAD52 and the WHD of RPA2 are important for SSA.** **a, b**, XL-MS analyses of the RAD52–ssDNA–RPA complex. The circos plots depict cross-links between RAD52–OR and RPA (RPA1, RPA2 and RPA3). Cross-links between the RAD52 DBD and RPA (yellow), the RAD52 RID and RPA (cyan), the extreme C terminus of RAD52 and RPA (light purple) and the WHD (dark pink) are highlighted. **c**, Peptide arrays of RAD52, RPA1, RPA2 and RPA3 showing interactions between RAD52 and RPA. The detected interaction intensities are shown as heat maps (yellow, maximum (max.) signal; purple, minimum (min.) signal). **d**, Schematic of RAD52, RAD52(ΔRID), RAD52(ΔC), RPA2 and RPA2(ΔWHD). **e**, SSA mediated by the indicated proteins. Data are mean + s.e.m.  $n = 22$  (RAD52 and RAD52 + RPA) and  $n = 3$  (RAD52(ΔRID) and RAD52(ΔRID) + RPA) independent experiments. The ssDNA was 68 nucleotides.

**f**, SSA catalysed by RAD52–OR and RPA or RPA(ΔWHD), as indicated in the schematics (WHD deletion is indicated by a red cross). Data are mean ± s.e.m. The concentrations were as follows: <sup>32</sup>P-labelled SSA1 and SSA2 (0.33 nM; 68 nucleotides), RPA (0.33 nM) and RAD52 (0.17 nM). All reactions were  $n = 3$ , except  $n = 22$  (RAD52 and RAD52 + RPA) and  $n = 7$  (RAD52 + RPA(ΔWHD)), where  $n$  relates to independent experiments. Statistical analysis was performed using unpaired two-tailed *t*-tests. **g**, A model for SSA by interactions between RAD52 and RPA. First, RPA binds to and protects resected ssDNA. Second, RAD52 interacts with RPA-bound ssDNA, and ssDNA wraps around RAD52. The RAD52–ssDNA complex then interacts with RPA–ssDNA, leading to complementary-strand annealing. Finally, RAD52 and RPA dissociate from the annealed dsDNA.

In summary, we have obtained high-resolution structures of RAD52, RAD52–ssDNA complex, the trimeric core of RPA and, importantly, a RAD52–ssDNA–RPA complex that provide insights into the mechanism of annealing. We observed that ssDNA wraps around the RAD52 ring, along a positively charged groove and that nucleotide quartets adopt a length equivalent to B-form duplex DNA. At the conclusion of the annealing reaction, duplex DNA is released by RAD52.

Importantly, RAD52 adopts open- and closed-ring configurations, such that the open rings represent the active form of the protein for SSA. The closed rings may be an artifact of protein overexpression. Moreover, we observed that the C terminus of RAD52 contributes to the maintenance of the open-ring conformation, providing an explanation as to why the RAD52 NTD crystallized as closed rings<sup>18,19</sup>. Without separation of the open from closed rings, achieved here using cation-exchange chromatography, the heterogeneous nature of RAD52 complicates structural analysis and the interpretation of biochemical experiments<sup>37,38</sup>.

DNA annealing by the RAD52-ORs, but not RAD52-CRs, was stimulated by RPA, and involved specific interactions between the RPA-interacting domain of RAD52 and the WHD of RPA2. Notably, visualization of the

RAD52–ssDNA–RPA complex revealed the presence of one RPA at the site of ring opening. Within this complex, the ssDNA-binding domain of RAD52 interacts specifically with DBD-A, DBD-B and DBD-C of RPA1, DBD-D of RPA2 and DBD-E of RPA3, indicating that these DNA-binding domains have an integrated role in the annealing reaction, possibly by mediating the hand-over of ssDNA from RPA to RAD52.

The observation that RPA stimulates RAD52-mediated SSA, and excess RAD52 inhibits the reaction, together with the lack of obvious RAD52–RAD52 interactions leads us to consider a new model for SSA. Initial events are likely to involve the sporadic binding of RAD52 to RPA-bound ssDNA, leading to the wrapping of ssDNA around the RAD52 ring. Then, rather than SSA being driven by interactions between two RAD52 rings<sup>20,37,38</sup>, in which the ssDNA is embedded within the DNA-binding grooves with consequential steric interactions that hinder rather than favour SSA, we suggest that SSA may involve direct interactions between RAD52–ssDNA and complementary ssDNA bound by RPA (Fig. 5g).

In human cells, RAD52 is expressed at low levels, and its concentration is around 5,000-fold lower than that of RPA<sup>42</sup>. Such a low level of expression would not be expected to support the displacement of RPA

from ssDNA. Instead, we suggest that RPA has a series of direct roles in facilitating annealing: (1) RPA promotes the removal of secondary structures from ssDNA; (2) it mediates direct interactions with RAD52 and thereby targets RAD52 to the initiating ssDNA; and (3) the localization of RPA (together with the bound complementary strand) to the site of ring opening in RAD52 brings complementary sequences into close proximity such that annealing can occur on the surface of RAD52. The resulting annealed duplex DNA will then be released due to the low affinity that RAD52 exhibits for dsDNA.

The low expression levels of RAD52 may also limit SSA, which can lead to deletions between repeat sequences and loss of genetic information<sup>43</sup>. Low levels of expression will therefore favour non-mutagenic BRCA2/RAD51-mediated homologous recombinational repair. However, in human cancers, high expression levels of RAD52 have been observed and are associated with poor survival outcome<sup>13,14</sup>. Indeed, increased genome instability resulting from elevated levels of SSA may contribute to cancer cell growth and spread.

The synthetic lethal relationship between *RAD52* and *BRCA1/2* supports the notion that there may be therapeutic opportunities to specifically inhibit RAD52 in homologous-recombination-deficient cancer cells. With this in mind, several RAD52 inhibitors have been developed and show promising results in killing *BRCA2*-deficient cancer cells<sup>15,16,44–48</sup>. However, most inhibitors have a single mechanism of action in which they target the ssDNA binding activity of RAD52. The results presented here reveal the importance of the RAD52–RPA interaction in SSA and provide a potential future strategy for the specific inhibition of the annealing reaction by developing small-molecule inhibitors that interfere with RAD52–RPA interactions.

## 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/s41586-024-07347-7>.

- Symington, L. S. Role of *RAD52* epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.* **66**, 630–670 (2002).
- Lisby, M., Rothstein, R. & Mortensen, U. H. Rad52 forms DNA repair and recombination centers during S phase. *Proc. Natl Acad. Sci. USA* **98**, 8276–8282 (2001).
- Bhowmick, R., Minocherhomji, S. & Hickson, I. D. *RAD52* facilitates mitotic DNA synthesis following replication stress. *Mol. Cell* **64**, 1171–1126 (2016).
- Sotiriou, S. K. et al. Mammalian *RAD52* functions in break-induced replication repair of collapsed DNA replication forks. *Mol. Cell* **64**, 1127–1134 (2016).
- Malacaria, E. et al. *RAD52* prevents excessive replication fork reversal and protects from nascent strand degradation. *Nat. Commun.* **10**, 1412 (2019).
- Zhang, J. M., Yadav, T., Ouyang, J., Lan, L. & Zou, L. Alternative lengthening of telomeres through two distinct break-induced replication pathways. *Cell Rep.* **26**, 955–968 (2019).
- Verma, P. et al. *RAD52* and *SLX4* act nonepistatically to ensure telomere stability during alternative telomere lengthening. *Genes Dev.* **33**, 221–235 (2019).
- Mortensen, U. H., Bendixen, C., Sunjevaric, I. & Rothstein, R. DNA strand annealing is promoted by the yeast *Rad52* protein. *Proc. Natl Acad. Sci. USA* **93**, 10729–10734 (1996).
- Sugiyama, T., New, J. H. & Kowalczykowski, S. C. DNA annealing by *Rad52* protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc. Natl Acad. Sci. USA* **95**, 6049–6054 (1998).
- Hanamshet, K., Mazina, O. M. & Mazin, A. V. Reappearance from obscurity: mammalian *Rad52* in homologous recombination. *Genes* **7**, 63 (2016).
- Feng, Z. H. et al. *RAD52* inactivation is synthetically lethal with *BRCA2* deficiency. *Proc. Natl Acad. Sci. USA* **108**, 686–691 (2011).
- Lok, B. H., Carley, A. C., Tchong, B. & Powell, S. N. *RAD52* inactivation is synthetically lethal with deficiencies in *BRCA1* and *PALB2* in addition to *BRCA2* through *RAD51*-mediated homologous recombination. *Oncogene* **32**, 3552–3558 (2013).
- Ho, V. et al. Aberrant expression of *RAD52*, its prognostic impact in rectal cancer and association with poor survival of patients. *Int. J. Mol. Sci.* **21**, 1768 (2020).
- Uhlen, M. et al. A pathology atlas of the human cancer transcriptome. *Science* **357**, 660 (2017).
- Huang, F. et al. Targeting *BRCA1*- and *BRCA2*-deficient cells with *RAD52* small molecule inhibitors. *Nucleic Acids Res.* **44**, 4189–4199 (2016).
- Hengel, S. R. et al. Small-molecule inhibitors identify the *RAD52*-ssDNA interaction as critical for recovery from replication stress and for survival of *BRCA2* deficient cells. *eLife* **5**, e14740 (2016).

- Bhat, D. S. et al. Therapeutic disruption of *RAD52*-ssDNA complexation by novel drug-like inhibitors. *NAR Cancer* **2**, zcad018 (2023).
- Singleton, M. R., Wentzell, L. M., Liu, Y., West, S. C. & Wigley, D. B. Structure of the single-strand annealing domain of human *RAD52* protein. *Proc. Natl Acad. Sci. USA* **99**, 13492–13497 (2002).
- Kagawa, W. et al. Crystal structure of the homologous-pairing domain from the human *RAD52* recombinase in the undecameric form. *Mol. Cell* **10**, 359–371 (2002).
- Saotome, M. et al. Structural basis of homology-directed DNA repair mediated by *RAD52*. *iScience* **3**, 50–62 (2018).
- Park, M. S., Ludwig, D. L., Stigger, E. & Lee, S. H. Physical interaction between human *RAD52* and RPA is required for homologous recombination in mammalian cells. *J. Biol. Chem.* **271**, 18996–19000 (1996).
- New, J. H., Sugiyama, T., Zaitseva, E. & Kowalczykowski, S. C. *Rad52* protein stimulates DNA strand exchange by *Rad51* and replication protein-A. *Nature* **391**, 407–410 (1998).
- Benson, F. E., Baumann, P. & West, S. C. Synergistic actions of *RAD51* and *RAD52* in genetic recombination and DNA repair. *Nature* **391**, 401–404 (1998).
- Deveryshty, J. et al. Yeast *Rad52* is a homodecamer and possesses *BRCA2*-like bipartite *Rad51* binding modes. *Nat. Commun.* **14**, 6215 (2023).
- Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human *BRCA2* stimulates *RAD51*-mediated recombination. *Nature* **467**, 678–683 (2010).
- Thorslund, T. et al. The breast cancer tumour suppressor *BRCA2* promotes the specific targeting of *RAD51* to single-stranded DNA. *Nat. Struct. Mol. Biol.* **17**, 1263–1265 (2010).
- Davies, A. A. et al. Role of *BRCA2* in control of the *RAD51* recombination and DNA repair protein. *Mol. Cell* **7**, 273–282 (2001).
- Greenhough, L. A. et al. Structure and function of the *RAD51B-RAD51C-RAD51D-XRCC2* tumour suppressor. *Nature* **619**, 650–657 (2023).
- Longo, M. A. et al. *RAD51C-XRCC3* structure and cancer patient mutations define DNA replication roles. *Nat. Commun.* **14**, 4445 (2023).
- Parsons, C. A., Baumann, P., Van Dyck, E. & West, S. C. Precise binding to single-stranded DNA termini by *RAD52* protein. *EMBO J.* **19**, 4175–4181 (2000).
- Kinoshita, C. et al. The cryo-EM structure of full-length *RAD52* protein contains an undecameric ring. *FEBS Open Bio* **13**, 408–418 (2023).
- Stasiak, A. Z. et al. The human *RAD52* protein exists as a heptameric ring. *Curr. Biol.* **10**, 337–340 (2000).
- Mer, G. et al. Structural basis for the recognition of DNA repair proteins UNG2, XPA, and *RAD52* by replication factor RPA. *Cell* **103**, 449–456 (2000).
- Koike, M., Yutoku, Y. & Koike, A. The C-terminal region of *RAD52* is essential for *RAD52* nuclear and nucleolar localization, and accumulation at DNA damage sites immediately after irradiation. *Biochem. Biophys. Res. Commun.* **435**, 260–266 (2013).
- Adamson, A. W. et al. The *RAD52* S346X variant reduces risk of developing breast cancer in carriers of pathogenic germline *BRCA2* mutations. *Mol. Oncol.* **14**, 1124–1133 (2020).
- O'Reilly, F. J. & Rappsilber, J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat. Struct. Mol. Biol.* **25**, 1000–1008 (2018).
- Grimme, J. M. et al. Human *RAD52* binds and wraps single-stranded DNA and mediates annealing via two h*RAD52*-ssDNA complexes. *Nucleic Acids Res.* **38**, 2917–2930 (2010).
- Rothenberg, E., Grimme, J. M., Spies, M. & Ha, T. Human *RAD52*-mediated homology search and annealing occurs by continuous interactions between overlapping nucleoprotein complexes. *Proc. Natl Acad. Sci. USA* **105**, 20274–20279 (2008).
- Fan, J. & Pavletich, N. P. Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev.* **26**, 2337–2347 (2012).
- Yates, L. A. et al. A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nat. Commun.* **9**, 5447 (2018).
- Jackson, D., Dhar, K., Wahl, J. K., Wold, M. S. & Borgstahl, G. E. O. Analysis of the human replication protein A:RAD52 complex: evidence for crosstalk between RPA32, RPA70, *RAD52* and DNA. *J. Mol. Biol.* **321**, 133–148 (2002).
- Nagaraj, N. et al. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* **7**, 548 (2011).
- Bhargava, R., Onyango, D. O. & Stark, J. M. Regulation of single-strand annealing and its role in genome maintenance. *Trends Genet.* **32**, 566–575 (2016).
- Cramer-Morales, K. et al. Personalized synthetic lethality induced by targeting *RAD52* in leukemias identified by gene mutation and expression profile. *Blood* **122**, 1293–1304 (2013).
- Chandramouly, G. et al. Small-molecule disruption of *RAD52* rings as a mechanism for precision medicine in *BRCA*-deficient cancers. *Chem. Biol.* **22**, 1491–1504 (2015).
- Sullivan, K. et al. Identification of a small molecule inhibitor of *RAD52* by structure-based selection. *PLoS ONE* **11**, e0147230 (2016).
- Yang, Q., Li, Y., Sun, R. & Li, J. Identification of a *RAD52* inhibitor inducing synthetic lethality in *BRCA2*-deficient cancer cells. *Front. Pharmacol.* **12**, 637825 (2021).
- Al-Mugotir, M. et al. Selective killing of homologous recombination-deficient cancer cell lines by inhibitors of the RPA:RAD52 protein-protein interaction. *PLoS ONE* **16**, e0248941 (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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

**Purification of recombinant human RAD52**

Human *RAD52* cDNA was codon optimized for expression in *E. coli* and cloned into pET100 (GeneArt, Thermo Fisher Scientific). Inverse PCR (primers: RAD52\_tag\_remove\_F and RAD52\_tag\_remove\_R) was performed to remove the 6×His, T7 and Xpress tags. The plasmid was transformed into BL21 Star (DE3) (Thermo Fisher Scientific) cells and a single colony was inoculated into an overnight culture using Luria broth (LB) supplemented with 0.8% glucose and 100 µg ml<sup>-1</sup> ampicillin. An aliquot was diluted into 2 l of LB containing glucose and ampicillin, to an optical density at 600 nm (OD<sub>600</sub>) of 0.1, and incubated in an orbital shaker at 37 °C and 180 rpm. When the culture reached an OD<sub>600</sub> of 0.8, IPTG (0.5 mM; Thermo Fisher Scientific) was added to induce *RAD52* expression and incubation continued for a further 3 h. The culture was collected by centrifugation at 3,300g for 15 min, and the cell pellet was resuspended in 1 vol of PBS and centrifuged again. The pellet was then resuspended in lysis buffer (25 mM MES pH 6.5, 0.5 M NaCl, 10% glycerol and 1 mM EDTA) supplemented with Halt protease inhibitor (Thermo Fisher Scientific) and 0.25 mM TCEP, and lysed with Emulsiflex C5 (Avestin) at 4 °C. The lysate was clarified by centrifugation at 60,000g and 4 °C for 10 min. The supernatant was collected and diluted dropwise with the same lysis buffer without NaCl to reach 300 mM NaCl. The lysate was then clarified again by centrifugation at 60,000g and 4 °C for 20 min and loaded onto a HiTrap SP column (Cytiva) connected to an ÄKTA pure system at 4 °C. The column was washed with 3 column volumes (CV) of buffer containing 25 mM MES pH 6.5, 0.3 M NaCl, 1 mM EDTA, 10% glycerol and 0.25 mM TCEP, and eluted with 10 CV of a linear gradient of the same buffer containing 0.3–1 M NaCl. Peak fractions were diluted 3× with buffer containing 100 mM HEPES pH 7.0, 0.25 mM TCEP and Halt protease inhibitor and loaded onto a HiTrap Q column (Cytiva), that was eluted with 10 CV of a linear gradient (0.1–1 M NaCl) of HEPES buffer containing 0.5 mM EDTA and 0.25 mM TCEP. The HiTrap Q flow-through fraction was collected as crude purified *RAD52*.

To separate the two *RAD52* conformations, *RAD52* was loaded onto a Resource S column (Cytiva). Chromatography was performed using buffer containing 25 mM HEPES pH 7.0, 0.25 mM TCEP and various concentrations of NaCl. The Resource S column was (1) washed with 3 CV of 150 mM NaCl buffer; (2) eluted with 5 CV of linear gradient of 0.2–0.278 M NaCl buffer (until the conductivity was equivalent to 24.4 mS cm<sup>-1</sup>); (3) washed with 5 CV of 0.278 M NaCl buffer; and (4) eluted with 10 CV of 0.278–0.6 M NaCl buffer. The peak fractions of the two *RAD52* forms were collected separately. *RAD52*-OR and *RAD52*-CR were loaded onto a Superose 6 Increase 10/300 GL column (Cytiva) using buffer containing 25 mM HEPES pH 8.0, 200 mM KOAc, 10% glycerol and 0.25 mM TCEP. The peak fractions were collected, aliquoted, snap-frozen in liquid nitrogen and stored at –80 °C. *RAD52* concentrations were measured at a wavelength of 280 nm using the Nanodrop (Thermo Fisher Scientific) system and calculated as an 11-subunit ring (*RAD52*-CR and *RAD52* NTD) or 10-subunit ring (*RAD52*-OR) with the exception that protomer concentration was used for circular dichroism (CD) analyses.

For the *RAD52* NTD, inverse PCR (primers: *RAD52*\_NTD\_F and *RAD52*\_NTD\_R) was used to remove the C terminus (amino acids 210–418). The *RAD52* NTD was purified using the same method as the full-length protein except that a linear gradient of 0.2–0.6 M NaCl was used for the Resource S column.

For *RAD52*( $\Delta$ RID), *RAD52*(RQK/AAA) and *RAD52*( $\Delta$ C), inverse PCR was used to remove the RPA-interacting domain (primers: *RAD52*\_RID\_F and *RAD52*\_RID\_R), extreme C terminus (primer: *RAD52*\_NTD\_F and *RAD52*\_C\_18D\_R) and introduce the R260A, Q261A and K262A mutations (primer: *RAD52*\_RQKAAA\_F and *RAD52*\_RQKAAA\_R). All mutants were purified using the same method as for the full-length protein.

**Purification of Flag–*RAD52* from Sf9 insect cells**

Human *RAD52* cDNA was codon optimized for expression in Sf9 insect cells and cloned into pFastBac1 baculovirus expression vector with an N-terminal Flag tag (GeneArt, Thermo Fisher Scientific). The plasmid was transformed into DH10Bac (Thermo Fisher Scientific), and the bacmids were isolated with PureLink HiPure Plasmid Miniprep kit (Thermo Fisher Scientific). Overall, the generation and handling of the baculovirus was performed according to the Invitrogen Bac-to-Bac Baculovirus Expression System user manual with some modifications. In brief, recombinant bacmids were transfected into Sf9 cells with FuGENE HD, and P1 viruses were collected 66–72 h after transfection. The baculovirus titre was determined by isolating the viral DNA with High Pure Viral Nucleic Acid Kit (Roche), and quantitative PCR using Platinum qPCR supermix UDG (Thermo Fisher Scientific) and Baculo-QUANT kit (Oxford Expression Technologies). The P2 baculovirus was amplified by infecting Sf9 cells at a multiplicity of infection (MOI) of 0.01 and 2 million cells per ml, and collected at 66–72 h after infection.

P2 baculovirus (MOI = 1) was used for recombinant Flag–*RAD52* expression. Sf9 cells were grown in Sf-900 III SFM (Gibco, Thermo Fisher Scientific) at 27 °C in an orbital shaker at 140 rpm. The Sf9 cells were infected for 66–72 h. Cells were collected by centrifugation at 500g for 5 min and washed once with PBS. The cell pellet was resuspended in lysis buffer (25 mM MES pH 6.5, 600 mM NaCl, 10% glycerol and 1 mM EDTA) supplemented with Halt protease inhibitor (Thermo Fisher Scientific) and 0.25 mM TCEP, and sonicated in ice/water slurry at 25 amplitude for 150 s (with 1 s intervals to prevent warming) using a qSonica Q700 sonicator. The lysate was clarified by centrifugation at 60,000g for 30 min at 4 °C.

Pre-equilibrated anti-Flag M2 agarose beads (Merck) were added to the lysate, and the mixture was incubated on a rotator at 4 °C for 1.5 h. The beads were pelleted by centrifugation at 500g for 5 min at 4 °C and transferred to a gravity flow chromatography column. The column was washed extensively with the lysis buffer, and subsequently with buffer containing 25 mM HEPES pH 7.0, 450 mM NaCl, 10% glycerol, 1 mM EDTA, 0.25 mM TCEP and Halt protease inhibitor. The last wash was performed with the same buffer at 300 mM NaCl. Flag–*RAD52* was then eluted with the buffer containing 450 mM NaCl and 0.5 mg ml<sup>-1</sup> Flag peptide. Elution was performed twice by incubating the beads with an equal volume of elution buffer for 1 h at 4 °C. The eluates were combined, and diluted 4× using the same elution buffer at 100 mM NaCl without Flag peptide to lower the NaCl concentration to 150 mM. Resource S chromatography was performed as described above.

**Purification of recombinant human RPA**

Human *RPA1*, *RPA2* and *10×His-RPA3* were synthesized and cloned into the pFastBac1 baculovirus expression vector (GeneArt, Thermo Fisher Scientific). *RPA1* (2 copies), *RPA2* and *10×His-RPA3*, together with their polyhedrin promoters, were then assembled into pBIG1a (biGBac multigene baculovirus expression vector)<sup>49</sup> using Gibson assembly (NEB). Bacmids and baculovirus were generated as described above. P2 baculovirus (MOI = 1) was used for recombinant RPA expression. Sf9 cells were grown in Sf-900 III SFM (Gibco, Thermo Fisher Scientific) at 27 °C in an orbital shaker at 140 rpm, and infected for 66–72 h. Cells were collected by centrifugation at 500g for 5 min and washed once with PBS. The cell pellet was resuspended in buffer containing 25 mM HEPES pH 8.0, 0.5 M NaCl, 10% glycerol, 0.01% Tween-20, 20 mM imidazole, Halt protease inhibitor and 0.25 mM TCEP, and sonicated in ice/water slurry at 25 amplitude for 150 s (with 1 s interval to prevent warming) with a qSonica Q700 sonicator. The lysate was clarified by centrifugation at 60,000g for 30 min at 4 °C.

Pre-equilibrated Ni-NTA beads (Qiagen) were added to the lysate and the mixture was incubated on a rotator at 4 °C for 1 h. The beads were pelleted by centrifugation at 500g for 5 min at 4 °C and transferred to a chromatography column. The column was washed extensively with



lysis buffer excluding Tween-20 while gradually decreasing the NaCl concentration from 0.5 to 0.2 M. Recombinant RPA was eluted with buffer containing 25 mM Tris-HCl pH 8.0, 0.2 M NaCl, 10% glycerol, 250 mM imidazole, Halt protease inhibitor and 0.25 mM TCEP. The RPA eluate was diluted 2× with the same elution buffer, without NaCl and imidazole, to lower the NaCl concentration to 100 mM. The diluted eluate was then loaded onto a Resource Q column (Cytiva) and eluted with linear gradient of buffer containing 0.1–0.4 M NaCl, 25 mM HEPES pH 8.0, 10% glycerol and 0.25 mM TCEP. Peak fractions containing RPA were loaded onto a Superdex 200 Increase 10/300 GL column (Cytiva) using buffer containing 25 mM HEPES pH 8.0, 200 mM KOAc, 0.5 mM EDTA, 10% glycerol and 0.25 mM TCEP. The protein was collected, aliquoted, snap-frozen in liquid nitrogen and stored at –80 °C.

For RPA1(ΔFAB) and RPA2(ΔWHD), inverse PCR was used to remove the DBD-F, DBD-A and DBD-B of RPA1 (amino acids 2–440) (primers: DBDC\_F and DBDC\_R) and the WHD of RPA2 (amino acids 207–270) (primers: RPA2\_WHD\_F and RPA2\_WHD\_R). Both deletion mutants were purified using the same method as described for the full-length protein.

### Oligonucleotides

All DNA oligonucleotides were HPLC purified (Merck and Integrated DNA Technologies). The names and sequences of the oligos were as follows where FAM is 6-carboxyfluorescein: RAD52\_tag\_remove\_F (5'-AGCGGCACCGAAGAAGCAATTTAGG-3'), RAD52\_tag\_remove\_R (5'-CATATGTATATCTCTTCTTAAAGTTAAACAAAATTTCTAGAGGG-3'), RAD52\_NTD\_F (5'-TAAAAGGGCGAGCTCAACGATCCGGCTG-3'), RAD52\_NTD\_R (5'-ACGACAGCTATTATAACGTGCTTCTTCAACGCTCGG-3'), RAD52\_RID\_F (5'-CCTCCGGCACCCTGTTAC-3'), RAD52\_RID\_R (5'-ATCCTGATCTGCCGGAATAACTGCATG-3'), RAD52\_RQKAAA\_F (5'-CGCACAGCTGCAACAGCAGTTTCGTGAACGTATG-3'), RAD52\_RQKAAA\_R (5'-GCAGCCAGTTACGCTGATGGGTTGCTTCGCTTTCAACTGCG-3'), RAD52\_C\_18D\_R (5'-ATTACCGGTGGTACGCTGATCTGCGCTATAGG-3'), DBDC\_F (5'-AACTGGAAAACCTTGATGAGTCAAATCCGAGAACCTGGG-3'), DBDC\_R (5'-CATGGATCCGCGCCGATGGTGG-3'), RPA2\_WHD\_F (5'-GCGGCCGCTTTCGAATCTAGAGCTG-3'), RPA2\_WHD\_R (5'-AGTGAGGCCATTTGCTGGCATGAACTGATTCC-3'), SSA1 (5'-TATCGAATCCGTCTAGTCAACGCTGCCGAAATTCTACAGAGTTGGGCTCCTCAACCTGCAGGTT-3'), SSA2 (5'-AACCTGCAGGTTGAGGAGCCAAACCTCACTGGTAAATTCGCAGCGTTGACTAGACGGATTCGATA-3'), FAM-SSA4 (40nt) (5'-FAM-TATCGAATCCGTCTAGTCAACGCTGCCGAATTTCTACAGT-3'), SSA5 (5'-ACTGGTGAATTCGGCAGCGTTGACTAGACGGATTCGATA-3'), SSA6 (5'-TGACCATCTTAAGCCGTCGCAACTGATCTGCCTAAGCTAT-3'), SSA7 (5'-CGCAGCTTGACTAGCAGGATTCGATA-3'), gap 1-1 (5'-CGTGAAGTCGCCGACTGAATCCAGCAATCTTTTTGAGTCTCATTTCGATCTCGGCAATCTCTTTCTGATTGTCAGATTGCATTTTAGTAAGCTCTTTTTGATTCTCAAATCCGGCG-3'), gap 1-2 (5'-CGCCGGATTTGAGAATCAAAGAGCTTAC-3') and gap 1-3 (5'-GATTGCTGGCATTTCAGTCGGCGACTTCAGC-3'). Cy3- and Cy5-labelled and biotinylated oligonucleotides were purchased (Merck). To generate FAM-SSA1/SSA2 dsDNA, equimolar concentrations of FAM-SSA1 and SSA2 were mixed in 10 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM EDTA, heated to 90 °C and gradually cooled to room temperature. Gapped DNA was annealed as described using gap 1-1, gap 1-2 and gap 1-3. Concentrations were measured using a spectrophotometer using absorbance values at 260 nm. All DNAs were stored at –20 °C.

### Fluorescence anisotropy

DNA-binding reactions (20 µl) were performed at 25 °C in buffer containing 25 mM HEPES pH 8.0, 0.2 M KOAc, 10% glycerol, 0.25 mM TCEP, 1 mM Mg(OAc)<sub>2</sub> and 0.01% Brij-35. Proteins were serially diluted and mixed with 10 nM (final concentration) of FAM-labelled DNA in 384-well microplates (Corning). The plates were measured using the CLARIOstar microplate reader (BMG Labtech). Blank-corrected anisotropy measurements were averaged and plotted against protein concentration.

RAD52 binding was curve-fitted using the following quadratic equation in GraphPad Prism 9 to determine  $K_D$  values:

$$Y = A_{\min} + (A_{\max} - A_{\min}) \times \frac{x + L + K_D - \sqrt{(x + L + K_D)^2 - 4 \times x \times L}}{2 \times L},$$

where  $Y$  is the fluorescence anisotropy,  $A_{\min}$  and  $A_{\max}$  are the minimum and maximum fluorescence anisotropy values,  $L$  is the ligand concentration (equal to 0.01 µM),  $x$  is the protein concentration and  $K_D$  is the dissociation constant. At least three independent triplicates of technical replicates were performed for each binding condition.

### Single-stranded DNA annealing

Reactions (15 µl) contained 5'-<sup>32</sup>P-labelled SSA1 (68 nucleotides) with its complementary strand SSA2 (68 nucleotides)<sup>30</sup> in 25 mM HEPES pH 8.0, 0.2 M KOAc, 1 mM Mg(OAc)<sub>2</sub>, 0.01% Brij-35, 0.25 mM TCEP and 5% glycerol. Two separate 7.5 µl reaction mixtures were set up. One contained 5'-<sup>32</sup>P-labelled SSA1 (0.33 nM) in buffer, and the second contained SSA2 (0.33 nM). RPA (0.33 nM) was added to both, as indicated. RAD52 (0.33 nM, or as indicated in figure legends) was added to SSA2 and incubated for 10 min at 25 °C. The two tubes were then mixed and incubated for 10 min at 25 °C, before being stopped by deproteinization using 3 µl of proteinase K (20 mg ml<sup>-1</sup> proteinase K in 10 mM Tris-HCl pH 7.5 and 1 mM CaCl<sub>2</sub>) and incubated at 30 °C for 30 min. The samples were supplemented with Ficoll loading buffer and analysed by PAGE with TBE as the running buffer. Gels were dried and exposed to phosphorimaging plates and images acquired using the Typhoon FLA 9500 biomolecular imager (GE) and quantified using ImageJ<sup>50,51</sup>.

For reactions using 40-nucleotide ssDNA (5'-<sup>32</sup>P-labelled SSA4 with complementary SSA5), the reactions were set up as described above except that the concentration of ssDNA was lowered to 0.13 nM to prevent self-annealing of ssDNA, and 0.13 nM of RPA was used. Concentrations of RAD52 are indicated in figure legends.

To determine whether DNA ends were required for RAD52-OR mediated annealing, interactions between 0.33 nM circular φX174 virion ssDNA and 0.33 nM <sup>32</sup>P-labelled gapped duplex DNA (a 60-nucleotide-long ssDNA that had 30-mers annealed to each end) were analysed. For these experiments, RPA (0.33 nM or 19.9 nM) was premixed with the gapped and circular ssDNAs, respectively (to provide similar coverage). RAD52 was then added to the gapped ssDNA and annealing was measured by electrophoresis through a 1% agarose gel using TAE buffer.

To analyse ssDNA annealing using size-exclusion chromatography, RAD52-OR (4 µM) was preloaded on SSA2–Cy5 (4 µM, 12.5 µl) before an equal volume of Cy3–SSA1 (4 µM) was added. After 30 min on ice, the reaction was loaded onto the Superdex 200 Increase 3.2/300 column connected to the ÄKTA pure Micro system. Chromatography was performed at 4 °C with buffer containing 25 mM HEPES pH 8.0, 200 mM KOAc, 0.25 mM TCEP and 1 mM Mg(OAc)<sub>2</sub>.

### Biolayer interferometry analysis

40-nucleotide (SSA4) ssDNA was biotinylated at either the 5' or 3' end (indicated as bio–ssDNA or ssDNA–bio, respectively). 68-nucleotide (SSA1) ssDNA was biotinylated at the 3' end (indicated as SSA1–bio), and 28 nucleotides of complementary ssDNA was annealed to the 5' end to protect the 5' ssDNA end (indicated as ds-ssDNA–bio). The experiments were performed using the Octet R8 system (Sartorius) at 25 °C in buffer containing 25 mM HEPES pH 8.0, 200 mM KOAc, 0.01% Tween-20, 1 mM Mg(OAc)<sub>2</sub> and 0.25 mM TCEP. The biotinylated DNA substrates (5 nM) were immobilized onto Octet SA streptavidin biosensors until a 0.05 threshold, and the sensors were then moved to wells containing a range of RAD52 concentrations (20, 10, 5, 2.5, 1.25, 0.625 and 0.312 nM). The association of RAD52 to DNA was recorded for 60 min and the dissociation for 5 min using the Octet BLI Discovery Software. Equilibrium dissociation constants ( $K_D$ ) were obtained by plotting association amplitudes

# Article

at equilibrium versus protein concentration (Octet Analysis Studio Software; Sartorius) and plotted in GraphPad Prism 9. The following 1:1 binding equation was used to determine  $K_D$  values: using the following quadratic equation in GraphPad Prism 9 to determine  $K_D$  values:

$$Y = B_{\max} \times X / (K_D + X),$$

where  $Y$  is the association amplitude,  $B_{\max}$  is the maximum amplitude at saturation,  $X$  is the protein concentration and  $K_D$  is the dissociation constant. Three independent triplicates were performed for each binding condition.

## CD analysis

Far-UV CD measurements were performed on a Jasco J-815 spectropolarimeter fitted with a cell holder temperature-regulated by a CDF-426S Peltier unit. Spectra were recorded at 20 °C at protein concentrations of 3.3  $\mu\text{M}$  (RAD52-OR) and 3.2  $\mu\text{M}$  (RAD52-CR) in 10 mM potassium phosphate buffer pH 8.0, 100 mM NaF and 0.25 mM TCEP. Fused silica cuvettes were used with a 1 mm path length (Hellma). Spectra were recorded at a resolution of 0.2 nm and were baseline corrected by subtraction of the appropriate buffer spectrum. CD intensities are presented as the molar CD extinction coefficient ( $\Delta\epsilon_M$ ) calculated as:

$$\Delta\epsilon_M = \frac{S}{32,980 \times c_M \times L} \text{ (units: } \text{M}^{-1}\text{cm}^{-1}\text{)},$$

where  $S$  is the signal in millidegrees,  $c_M$  is the molar concentration and  $L$  is the path length (in cm). Secondary structure content was estimated as described<sup>52</sup>.

## Intact protein MS

Proteins were diluted to 1  $\mu\text{M}$  with 0.1% (v/v) formic acid and injected onto a C4 BEH 1.7  $\mu\text{m}$ , 1.0  $\times$  100 mm, UPLC column using the Acquity I class LC (Waters) system. Proteins were eluted with a 15 min gradient of acetonitrile (2% (v/v) to 80% (v/v)) in 0.1% (v/v) formic acid using a flow rate of 50  $\mu\text{l min}^{-1}$ . The analytical column outlet was directly interfaced through an electrospray ionization source, with a time-of-flight (TOF) mass spectrometer (BioAccord, Waters). Data were acquired over a  $m/z$  range of 300–8,000, in positive-ion mode with a cone voltage of 40 V. Scans were summed together manually and deconvoluted using MaxEnt1 (Masslynx, Waters). The parameters used were as follows; input  $m/z$  range (Da): 600–2,000; output mass range (Da): 30000–60000; TOF resolution: 10000.00; and iterate to convergence.

## GuHCl denaturation and renaturation

RAD52 (purified to the HiTrap Q step) was dialysed into 25 mM HEPES pH 7.0, 6 M GuHCl, 0.5 mM EDTA and 2 mM  $\beta$ -mercaptoethanol overnight at 4 °C. The denatured protein was analysed using a Superose 6 Increase 10/300 GL column, which was run with 6 M GuHCl buffer. Protein was renatured by dialysis in native buffer (25 mM HEPES pH 7.0, 200 mM NaCl, 0.5 mM EDTA and 2 mM mercaptoethanol) for 24 h at 4 °C. The renatured RAD52 was then run on the same column using native buffer. To analyse the percentage of open and closed rings, the renatured RAD52 sample was loaded onto the Resource S column.

## Negative-stain EM sample preparation and data acquisition

Samples (4  $\mu\text{l}$ , 25  $\text{ng } \mu\text{l}^{-1}$ ) were applied for 1 min to glow discharged (25 mA, 30 s) 400-mesh carbon-coated copper grids (C400Cu100, EM Resolutions). The grids were sequentially stained in four separate 30  $\mu\text{l}$  droplets of 2% (v/v) uranyl acetate for 10, 15, 20 and 25 s. Excess uranyl acetate was blotted away from the grid using Whatmann paper, allowed to air dry and stored before imaging.

The grids were imaged on the Tecnai LaB<sub>6</sub> G2 Spirit TEM operating at 120 kV equipped with a 2K Gatan Ultrascan 1000 camera. Micrographs were acquired manually using DigitalMicrograph at a nominal

magnification of  $\times 30,000$  (3.5 Å per pixel) or  $\times 42,000$  (2.4 Å per pixel) with defocus values ranging from  $-0.7$  to  $-1.5$   $\mu\text{m}$ .

## Negative-stain EM data analysis

DM3 files were converted to MRC format using e2proc2d.py (EMAN2)<sup>53</sup>. Micrographs were imported into Relion 3.1 or 4.1<sup>54,55</sup>, CTF parameters were calculated using CTFIND4<sup>56</sup>, and particles were picked using crYOLO<sup>57</sup> or Topaz<sup>58</sup>. Particles were extracted and iteratively 2D classified (ignore CTF to first peak = yes, limit resolution E-step = 20 Å, additional arguments = --only-flip-phases).

## Cryo-EM sample preparation

Recombinant RAD52 and RPA were purified to the Resource S or Resource Q step, and freshly purified on the Superose 6 Increase 10/300 GL or Superdex 200 Increase 10/300 GL column before making the cryo-EM grids. For RAD52-CR, the protein was in a buffer containing 25 mM HEPES pH 7.0, 150 mM NaCl and 0.25 mM TCEP, diluted to 0.3  $\text{mg ml}^{-1}$ , and supplemented with 0.00005% Tween-20. A sample (4  $\mu\text{l}$ ) was applied to freshly glow-discharged (45 mA, 60 s; Quorum Emitech K100X) Quantifoil R2/1 300 mesh copper grids and vitrified using a Vitrobot Mark IV (Thermo Fisher Scientific) cooled to 4 °C with 95% humidity. Grids were double-side blotted for 0.5 s and plunge frozen in liquid ethane. For RAD52-OR, the grids were prepared as described above except Quantifoil R2/2 200 mesh copper grids were used, and the concentration was 0.25  $\text{mg ml}^{-1}$ , the Tween-20 concentration was 0.001%, and blot time was 1.5 s. For RAD52-OR–ssDNA, the protein (0.25 mM) was diluted to 0.5  $\mu\text{M}$  in 25 mM HEPES pH 8.0, 150 mM NaCl, 2 mM Mg(OAc)<sub>2</sub> and supplemented with 0.05% octyl- $\beta$ -glucoside (OG). SSA4 (1  $\mu\text{M}$ ) was added and incubated at 25 °C for 10 min. The concentration was determined by Bradford assay (Bio-Rad) and diluted to 0.15  $\text{mg ml}^{-1}$  with the same buffer. Grids were prepared as above except Quantifoil R1.2/1.3 300 mesh copper grids were used and the blot time was 2.5 s. For RPA–ssDNA, the protein (0.25 mM), in 25 mM HEPES pH 8.0, 150 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, was diluted to 3  $\mu\text{M}$ , and supplemented with 0.1 mM CHAPSO. SSA7 (6  $\mu\text{M}$ ) was added and incubated at 25 °C for 10 min. The concentration was determined using the Bradford assay (Bio-Rad) and diluted to 0.15  $\text{mg ml}^{-1}$  with the same buffer. UltrAuFoil R2/2 200 mesh gold grids (Quantifoil) were prepared as described above and the blot time was 2.5 s. The RAD52-OR–ssDNA–RPA ternary complex was assembled as indicated in the 'Reconstitution of the RAD52–ssDNA–RPA complex' section below. The concentration was determined using the Bradford assay (Bio-Rad) and diluted to 0.1  $\text{mg ml}^{-1}$  with buffer supplemented with 0.00075% Tween-20 and 0.075 mM CHAPSO. Quantifoil R2/2 200 mesh copper grids were prepared as described above, except the blot time was 3 s.

## Cryo-EM data collection, image processing and atomic model building

RAD52-CR and RAD52-OR datasets were collected on a Titan Krios Cryo-TEM equipped with a Falcon III direct electron detector (Thermo Fisher Scientific) at the Francis Crick Institute Structural Biology STP. The RAD52-OR–ssDNA dataset was collected on a Titan Krios G3i (FEI, Thermo Fisher Scientific) equipped with a Gatan K3 direct electron detector at the London consortium for cryo-EM (LonCEM). RPA–ssDNA and RAD52-OR–ssDNA–RPA datasets were collected on a Titan Krios Cryo-TEM (Thermo Fisher Scientific) equipped with a K2 direct electron detector (Gatan) at the Francis Crick Institute Structural Biology STP.

Single-particle analyses were performed within Relion (v.4.0)<sup>54</sup> and CryoSPARC<sup>59</sup>. The videos were corrected for drift and dose-weighted using RELION's own implementation of MOTIONCOR2<sup>60</sup> and subsequent contrast transfer (CTF) parameters were measured using CTFIND4<sup>56</sup>. Particles were picked automatically using crYOLO<sup>57</sup> or Topaz<sup>58</sup>. Details of image processing are illustrated in Extended Data Figs. 3, 4, 5, 8 and 9. In brief, several rounds of 2D classification were performed to remove particles that cannot be aligned to yield defined 2D averages. Several

rounds of 3D classifications were performed to separate different conformations or particles that cannot be aligned to yield high-resolution 3D volumes. 3D auto-refine, Bayesian polishing (minimum two rounds) and CTF refinement (minimum one round) were performed iteratively to achieve high resolution 3D reconstruction in RELION<sup>61,62</sup>. Polished particles were imported to CryoSPARC<sup>59</sup>, and refined using non-uniform refinement<sup>63</sup>. 3D variability<sup>64</sup> or 3D classifications were performed to detect heterogeneity within the cryo-EM densities. The cryo-EM maps were sharpened by post-processing in RELION, CryoSPARC or DeepEMhancer<sup>65</sup> if there was high variability in local resolution. The overall resolution is reported at FSC = 0.143 (ref. 66).

All model building was performed using Phenix<sup>67,68</sup>, COOT<sup>69</sup> and ISOLDE<sup>70</sup> in ChimeraX<sup>71</sup>. For RAD52-CR, the crystal structure of the RAD52 NTD (PDB: 1H21) was placed into a sharpened RAD52-CR cryo-EM map in ChimeraX<sup>71</sup> and initially refined using Namdinator<sup>72</sup>. One RAD52 subunit was removed from RAD52-CR and used for initial refinement in Namdinator for RAD52-OR. ssDNA was built manually in COOT into the RAD52-OR model using RAD52-OR-ssDNA as a starting model. RPA1, RPA2 and RPA3 AlphaFold2 models were used for Dock and rebuild in Phenix<sup>73,74</sup> and the ssDNA model was aligned and extracted from the fungal RPA structure (PDB: 4GOP)<sup>39</sup>. The RAD52-OR-ssDNA model was used as the initial model for RAD52-OR-ssDNA-RPA.

### SEC-MALLS analysis

SEC-MALLS was used to determine the molar mass composition of RAD52. Purified RAD52-OR (2.0, 1.0 or 0.5 mg ml<sup>-1</sup>) was loaded onto a Superose 6 Increase 10/300 GL column connected to a Jasco chromatography system. Chromatography was performed at 25 °C with buffer containing 25 mM HEPES pH 7.0, 150 mM NaCl, 0.25 mM TCEP and 3 mM Na<sub>3</sub>N at a flow rate of 1.0 ml min<sup>-1</sup>. RAD52-OR-ssDNA (2 mg ml<sup>-1</sup>) was analysed in a similar manner using 25 mM Bis-Tris propane pH 8.5, 200 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.25 mM TCEP and 3 mM Na<sub>3</sub>N as the running buffer. The scattered light intensity and protein concentrations of the column eluates were recorded using a DAWN-HELEOS laser photometer and an OPTILAB-rEX differential refractometer ( $dn/dc = 0.186$ ). The weight-averaged molecular mass of material contained in chromatographic peaks was determined using the combined data from both detectors in the ASTRA software v.7.3.2 (Wyatt Technology).

### Nuclear/chromatin extraction and analysis

U2OS cells (authenticated and microplasma free, as determined by the Francis Crick Institute) were grown in DMEM (Gibco) supplemented with 10% FBS (Gibco) in humidified incubators at 37 °C and 5% CO<sub>2</sub>. Cells were collected from four confluent 500 cm<sup>2</sup> square dishes and washed once with PBS. The pellet was supplemented with 5× pellet volume of CSK buffer (10 mM PIPES pH 6.8, 100 mM NaCl, 3 mM MgCl<sub>2</sub>, 300 mM sucrose, 1 mM EGTA, 0.5% Triton X-100 and 0.25 mM TCEP) supplemented with Halt protease and phosphatase inhibitors, incubated on ice for 10 min and centrifuged at 2,000g for 5 min at 4 °C. The supernatant was collected as the first CSK extract. A 3× pellet volume of CSK buffer (containing 0.1% Triton X-100) was added to the pellet, incubated on ice for 10 min and the sample was centrifuged at 3,000g for 5 min at 4 °C. The supernatant was collected as the second CSK extract. An equal volume of benzonase digestion buffer (20 mM HEPES pH 8.0, 2 mM MgCl<sub>2</sub>, 0.5% Triton X-100, 0.25 mM TCEP and 500 units benzonase/100 µl of buffer) supplemented with Halt protease and phosphatase inhibitors was added to the pellet and incubated on ice for 10 min. A 2× sample volume of high-salt buffer (20 mM HEPES pH 8.0, 600 mM NaCl and 0.25 mM TCEP) supplemented with Halt protease and phosphatase inhibitors was then added, incubated on ice for 10 min, and the sample was centrifuged at 21,000g for 10 min at 4 °C. The supernatant was collected as a nuclear/chromatin extract.

Glycerol gradients (5 ml, 10–30%) in 25 mM HEPES pH 8.0, 150 mM NaCl, 10–30% glycerol and 0.25 mM TCEP were cast in thin-wall polypropylene tubes (Beckman Coulter) using a Gradient Master (Biocomp)

and kept in the cold room overnight to equilibrate to 4 °C. U2OS nuclear/chromatin extracts (200 µl), 200 ng recombinant RAD52-OR or a gel-filtration calibration marker (Cytiva) was loaded gently onto the top of three gradients, which were then centrifuged at 4 °C and 55,000 rpm (368,000g) using SW 55 Ti rotor (Beckman Coulter) for 4 h. The fractions were collected by manual pipetting from the top of the gradients. The U2OS nuclear/chromatin extract (500 µl), 500 ng recombinant RAD52-OR or a gel-filtration calibration marker (Cytiva) were also loaded onto the pre-equilibrated Superose 6 Increase 10/300 GL column (Cytiva). Chromatography was performed with a buffer containing 25 mM HEPES pH 8.0, 150 mM NaCl, 10% glycerol and 0.25 mM TCEP at 4 °C. Fractions were collected and analysed by SDS-PAGE followed by western blotting using antibodies against RAD52 (rabbit monoclonal, 1:500, Abcam, ab124971). Alexa Fluor Plus 800 anti-rabbit secondary antibodies (1:2,000, Invitrogen, A32735) were used and the membranes were imaged using an Odyssey DLx instrument with ImageStudio software (Licor).

### RAD52 Resource S chromatogram peak fitting

Resource S chromatography was performed as described above except a linear gradient of 0.2–0.6 M NaCl was used. The UV<sub>280</sub> absorbance values were imported into GraphPad Prism 9 and curved fitted using a sum of two Gaussians equation to deconvolute open- and closed-ring peaks:

$$Y = \text{amplitude} \times \exp\left(-0.5 \times \left(\frac{X - \text{mean 1}}{\text{s.d.}}\right)^2\right) + \text{amplitude 2} \times \exp\left(-0.5 \times \left(\frac{X - \text{mean 2}}{\text{s.d. 2}}\right)^2\right)$$

### RAD52-ssDNA-RPA pull downs

The RAD52-ssDNA-RPA ternary complex (400 µl) was reconstituted in buffer containing 25 mM HEPES pH 8.0, 200 mM KOAc, 2 mM Mg(OAc)<sub>2</sub>, 0.01% Tween-20 and 0.25 mM TCEP. Biotin-labelled SSA4 (0.1 µM), with photo-cleavable linker (Integrated DNA Technologies), and recombinant RPA (0.15 µM) were mixed and incubated on ice for 10 min. RAD52-OR (0.15 µM) was then added and incubation continued for a further 10 min. Pre-washed Streptavidin Sepharose Mag beads (10 µl, Cytiva) were then added and incubated for 30 min on a head-to-toe rotator at 4 °C. The beads were washed once with reaction buffer and then with reaction buffer Tween-20. The beads were resuspended in 20 µl reaction buffer, and irradiated with 365 nm UVA on ice/water slurry to cleave the photo-cleavable linker.

### Reconstitution of the RAD52-ssDNA-RPA complex

RAD52-OR (purified to the Resource S step) and RPA (purified to the Resource Q step) were loaded onto the Superose 6 Increase 10/300 GL (Cytiva) and Superdex 200 Increase 10/300 GL (Cytiva) columns, respectively, and run with buffer containing 25 mM HEPES pH 8.0, 150 mM NaCl, 2 mM Mg(OAc)<sub>2</sub> and 0.25 mM TCEP. The reconstitution mixture for cryo-EM was supplemented with 0.00075% Tween-20 and 0.075 mM CHAPSO, whereas the XL-MS sample was supplemented with 0.05% OG. Reconstitution of the RAD52-OR-ssDNA-RPA ternary complex involved two steps: (1) RPA (1 µM final concentration) was added to SSA1 (0.5 µM final concentration) and incubated at 25 °C for 10 min; and (2) RAD52-OR (0.5 µM final concentration) was added and incubated at 25 °C for 30 min. The sample was centrifuged at 21,000g for 1 min at 4 °C before proceeding with cryo-EM grid preparation and XL-MS.

### Protein disorder prediction

The human RAD52 protein sequence (UniProt: P43351) was uploaded to the ODinPred<sup>75</sup> webserver (<https://st-protein.chem.au.dk/odinpred>). The predicted disorder probability of each residue was plotted in GraphPad Prism 9.

## Multiple-sequence alignment

RAD52 protein sequences from different organisms were aligned with Clustal Omega using the default settings<sup>76</sup>. The alignment was formatted with ESPript3.0<sup>77</sup>.

## XL-MS analysis

RAD52-OR and RAD52-OR-ssDNA-RPA ternary complexes (0.5  $\mu$ M, reconstituted as above) were supplemented with a 1:100 molar ratio of disuccinimidyl dibutyric urea (DSBU; 50  $\mu$ M) for 1 h at room temperature, before the mixture was quenched by the addition of  $\text{NH}_4\text{HCO}_3$  to a final concentration of 20 mM (15 min at room temperature). The cross-linked proteins were reduced with 10 mM dithiothreitol and alkylated with 50 mM iodoacetamide. They were then digested with trypsin at an enzyme-to-substrate ratio of 1:100, for 1 h at room temperature and further digested overnight at 37 °C after addition of trypsin at a ratio of 1:20. The peptide digests were then fractionated batch-wise by high pH reverse-phase chromatography on micro spin TARGA C18 columns (Nest Group) into four fractions (10 mM  $\text{NH}_4\text{HCO}_3$ /10% (v/v) acetonitrile pH 8.0; 10 mM  $\text{NH}_4\text{HCO}_3$ /20% (v/v) acetonitrile pH 8.0; 10 mM  $\text{NH}_4\text{HCO}_3$ /40% (v/v) acetonitrile pH 8.0; and 10 mM  $\text{NH}_4\text{HCO}_3$ /80% (v/v) acetonitrile pH 8.0). The fractions (150  $\mu$ l) were evaporated to dryness in a CentriVap concentrator (Labconco) before analysis by LC-MS/MS.

Lyophilized peptides were resuspended in 1% (v/v) formic acid and 2% (v/v) acetonitrile and analysed by nano-scale capillary LC-MS/MS using a Vanquish Neo UPLC (Thermo Fisher Scientific, Dionex) to deliver a flow of approximately 300 nl min<sup>-1</sup>. A PepMap Neo C18 5  $\mu$ m, 300  $\mu$ m  $\times$  5 mm nanoViper (Thermo Fisher Scientific, Dionex) trapped the peptides before separation on a 25 cm EASY-Spray column (25 cm  $\times$  75  $\mu$ m inner diameter, PepMap C18, 2  $\mu$ m particles, 100 Å pore size, Thermo Fisher Scientific). Peptides were eluted with a gradient of acetonitrile. The analytical column outlet was directly interfaced through a nano-flow electrospray ionization source, with a quadrupole Orbitrap mass spectrometer (Orbitrap Exploris 480, Thermo Fisher Scientific). MS data were acquired in data-dependent mode using a top ten method, where ions with a precursor charge state of 1+ and 2+ were excluded. High-resolution full scans ( $R = 60,000$ ,  $m/z$  380–1,800) were recorded in the Orbitrap followed by higher-energy collision dissociation (HCD) (stepped collision energy 30 and 32% normalized collision energy) of the ten most intense MS peaks. The fragment ion spectra were acquired at a resolution of 30,000 and a dynamic exclusion window of 20 s was applied.

For data analysis, Xcalibur raw files were converted into the MGF format using Proteome Discoverer v.2.3 (Thermo Fisher Scientific) and used directly as input files for MeroX<sup>78</sup>. Searches were performed against an ad hoc protein database containing the sequences of the proteins in the complex and a set of randomized decoy sequences generated by the software. The following parameters were set for the searches: maximum number of missed cleavages: 3; targeted residues K, S, Y and T; minimum peptide length 5 amino acids; variable modifications: carbamidomethylation of cysteine (mass shift 57.02146 Da), methionine oxidation (mass shift 15.99491 Da); DSBU modified fragments: 85.05276 Da and 111.03203 Da (precision: 5 ppm MS and 10 ppm MS/MS); false-discovery-rate cut-off: 5%. Finally, each fragmentation spectrum was manually inspected and validated.

To compare with the peptide array experiments, the number of cross-links detected for each amino acid residue was counted, and summed within an individual 20 amino acid peptide with a 1 amino acid shift, similar to the peptide array. The overlaid result was plotted using GraphPad Prism 9.

## Peptide array

Peptides (20 amino acids) with 1-amino-acid shift covering the full sequences of RAD52, RPA1, RPA2 and RPA3 were synthesized on

cellulose membranes in 3 mm spots by the Chemical Biology STP at the Francis Crick Institute. The membranes were washed with 50% ethanol and 10% acetic acid for 30 min and equilibrated with 1 $\times$  TBST (50 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.1% Tween-20) supplemented with 0.25 mM TCEP. The membrane was blocked with 5% non-fat milk in TBST (0.1% Tween-20) supplemented with 0.25 mM TCEP for 1 h at room temperature. To allow protein-peptide interactions, the membranes were incubated with RAD52-OR or RPA (1  $\mu$ g ml<sup>-1</sup>) in 1% non-fat milk in TBST (0.1% Tween-20) supplemented with 0.25 mM TCEP overnight at 4 °C. The membranes were washed in 1 $\times$  TBST (0.1% Tween-20) supplemented with 0.25 mM TCEP on an orbital shaker for 5 min at room temperature three times. The membranes were then incubated in primary antibodies (anti-His 1:1,000, Takara, 631212) in 1% non-fat milk in TBST (0.1% Tween-20) supplemented with 0.25 mM TCEP for 2 h at room temperature. The membranes were washed three times as before and incubated in Alexa-Fluor-Plus-conjugated secondary antibodies (goat anti-mouse 1:2,000, Thermo Fisher Scientific, A32730; goat anti-rabbit, 1:2,000, Thermo Fisher Scientific, A32735) in 1% non-fat milk in TBST (0.1% Tween-20) supplemented with 0.25 mM TCEP for 1 h at room temperature. The membranes were washed three times, imaged on a Li-Cor Odyssey DLx system and quantified using Image Studio Lite (Li-Cor).

## Nanoscale differential scanning fluorometry

A Prometheus NT-48 (Nanotemper) instrument was used to monitor changes in tryptophan fluorescence following thermal denaturation. Proteins were diluted to 10  $\mu$ M in 25 mM HEPES pH 8.0, 200 mM KOAc, 0.5 mM EDTA, 10% glycerol and 0.25 mM TCEP. The samples were loaded into high-sensitivity glass capillaries and the tryptophan fluorescence was monitored at 330 and 350 nm after excitation at 285 nm. Measurements were made from 25 to 95 °C with a temperature gradient of 1 °C min<sup>-1</sup>. The ratio of fluorescence intensity (350/330 nm) was plotted against temperature, and the first derivative of this curve was used to calculate thermal melting ( $T_m$ ) values.

## Statistics and reproducibility

Statistical analyses were performed using GraphPad Prism 9. Normally distributed data were compared using two-tailed unpaired *t*-tests whereas non-normally distributed data were compared using two-tailed Mann-Whitney *U*-tests. Differences were considered to be statistically significant when  $P < 0.05$ . Reported *n* values refer to independent experiments for fluorescence anisotropy, bilayer interferometry analysis and SSA assays. Glycerol gradient sedimentation analysis and size-exclusion chromatography of U2OS nuclear extract recombinant RAD52-OR were repeated independently seven times with similar results. RAD52-ssDNA-RPA pull-down experiments were repeated independently five times with similar results. RAD52 purifications were repeated independently more than 50 times with similar results. RPA purifications were repeated for ten times with similar results. Purifications of RAD52 and RPA mutants were repeated for twice with similar results.

## Reporting summary

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

## Data availability

Cryo-EM density maps and atomic models of RAD52-CR, RAD52-OR, RAD52-OR-ssDNA and RPA-ssDNA have been deposited at the Electron Microscopy Data Bank (EMDB) and PDB under the following accession codes: RAD52-CR (EMD-19189 and 8RIL), RAD52-OR (EMD-19193 and 8RJ3), RAD52-OR-ssDNA (EMD-19253 and 8RJW) and RPA-ssDNA (EMD-19255 and 8RK2). All other data and materials reported here are available on request.

49. Weissmann, F. et al. biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proc. Natl Acad. Sci. USA* **113**, 2564–2569 (2016).
50. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
51. Rueden, C. T. et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinform.* **18**, 529 (2017).
52. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–160 (2000).
53. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
54. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).
55. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
56. Rohou, A. & Grigorieff, N. CTFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
57. Wagner, T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).
58. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
59. 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).
60. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
61. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
62. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).
63. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
64. Punjani, A. & Fleet, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
65. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
66. 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).
67. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
68. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
69. Casanal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci.* **29**, 1069–1078 (2020).
70. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D* **74**, 519–530 (2018).
71. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
72. Kidmose, R. T. et al. Namdinator—automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526–531 (2019).
73. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
74. Terwilliger, T. C. et al. Improved AlphaFold modeling with implicit experimental information. *Nat. Methods* **19**, 1376–1382 (2022).
75. Dass, R., Mulder, F. A. A. & Nielsen, J. T. ODINPred: comprehensive prediction of protein order and disorder. *Sci. Rep.* **10**, 14780 (2020).
76. Madeira, F. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).
77. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
78. Gotze, M. et al. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass. Spectrom.* **26**, 83–97 (2015).

**Acknowledgements** We thank the members of the West laboratory for support and advice, in particular N. Rzechorzek for assistance with model building, N. Cronin and A. Nans for cryo-EM data collection, S. Mehmood for initial intact protein MS data collection, and R. Carzaniga, M. Wilkinson, Q. Chen, E. Punch, D. Benton, P. McIntosh and A. Nans for assistance with EM and data analysis; A. Costa for reading the manuscript; and D. Wigley, X. Zhang, P. Walker and the staff at the Crick Structure Biology STP, the Peptide Chemistry STP and the Scientific Computing STP for support. Work in S.C.W.'s laboratory was funded by the Francis Crick Institute, which receives core funding from Cancer Research UK (CC2098), the UK Medical Research Council (CC2098) and the Wellcome Trust (CC2098); the European Research Council (ERC-ADG-666400); the BBSRC (BB/W01355X/1); and the Louis-Jeantet Foundation. Cryo-EM data collection at the LonCEM was supported by the Wellcome Trust (206175/Z/17/Z).

**Author contributions** C.-C.L. and S.C.W. designed the study. C.-C.L. purified proteins, carried out biochemical analyses and performed cryo-EM and data analyses. C.-C.L., L.A.G. and L.M. performed biophysical assays. S.M. and M.S. performed intact protein MS and XL-MS. I.A.T. performed SEC-MALLS analysis. C.-C.L., I.B. and M.T. characterized the oligomeric state of cellular RAD52. S.C.W. supervised the experiments and data analyses. C.-C.L. and S.C.W. prepared the manuscript with contributions from all of the authors.

**Funding** Open Access funding provided by The Francis Crick Institute.

**Competing interests** The authors declare no competing interests.

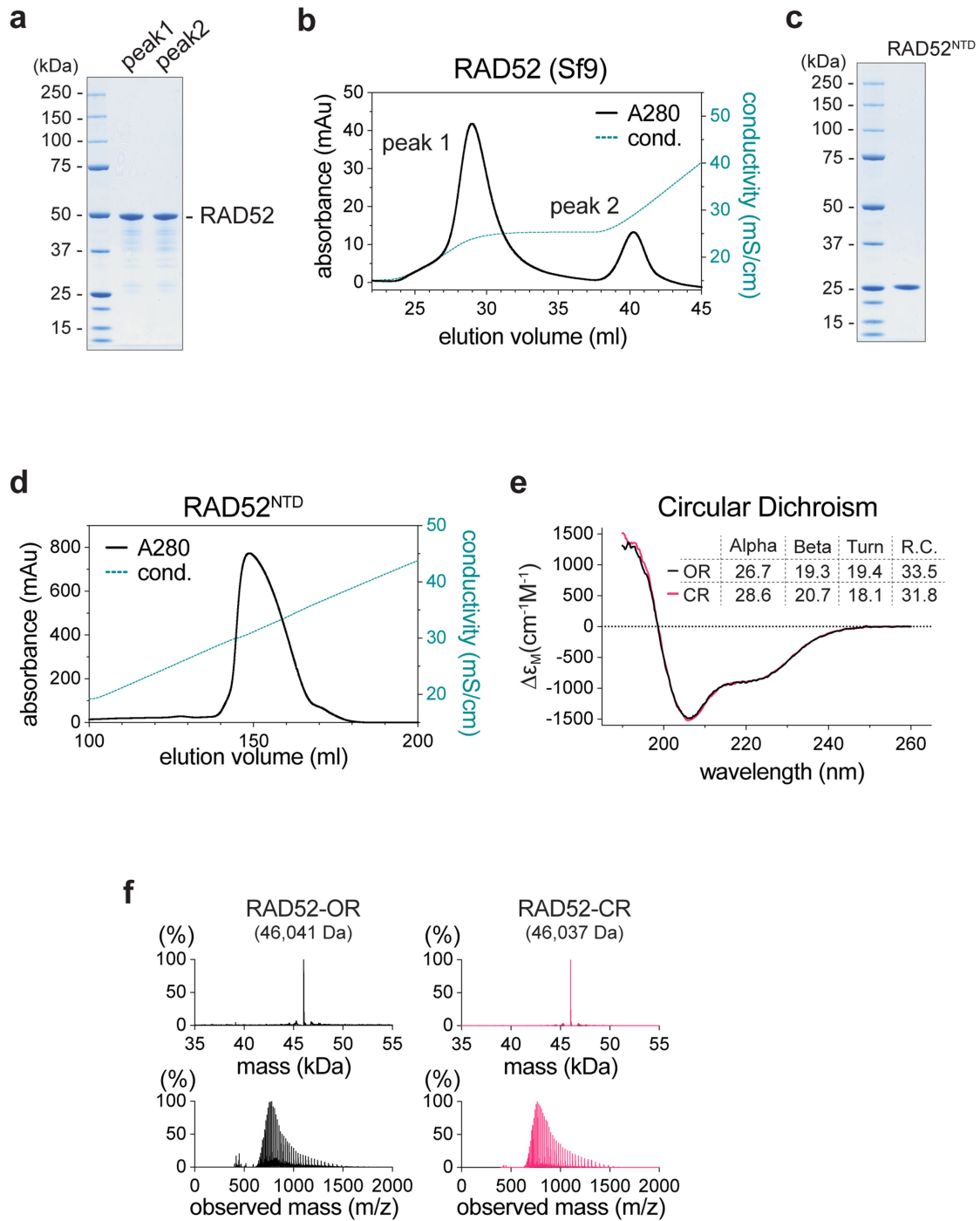
**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-024-07347-7>.

**Correspondence and requests for materials** should be addressed to Chih-Chao Liang or Stephen C. West.

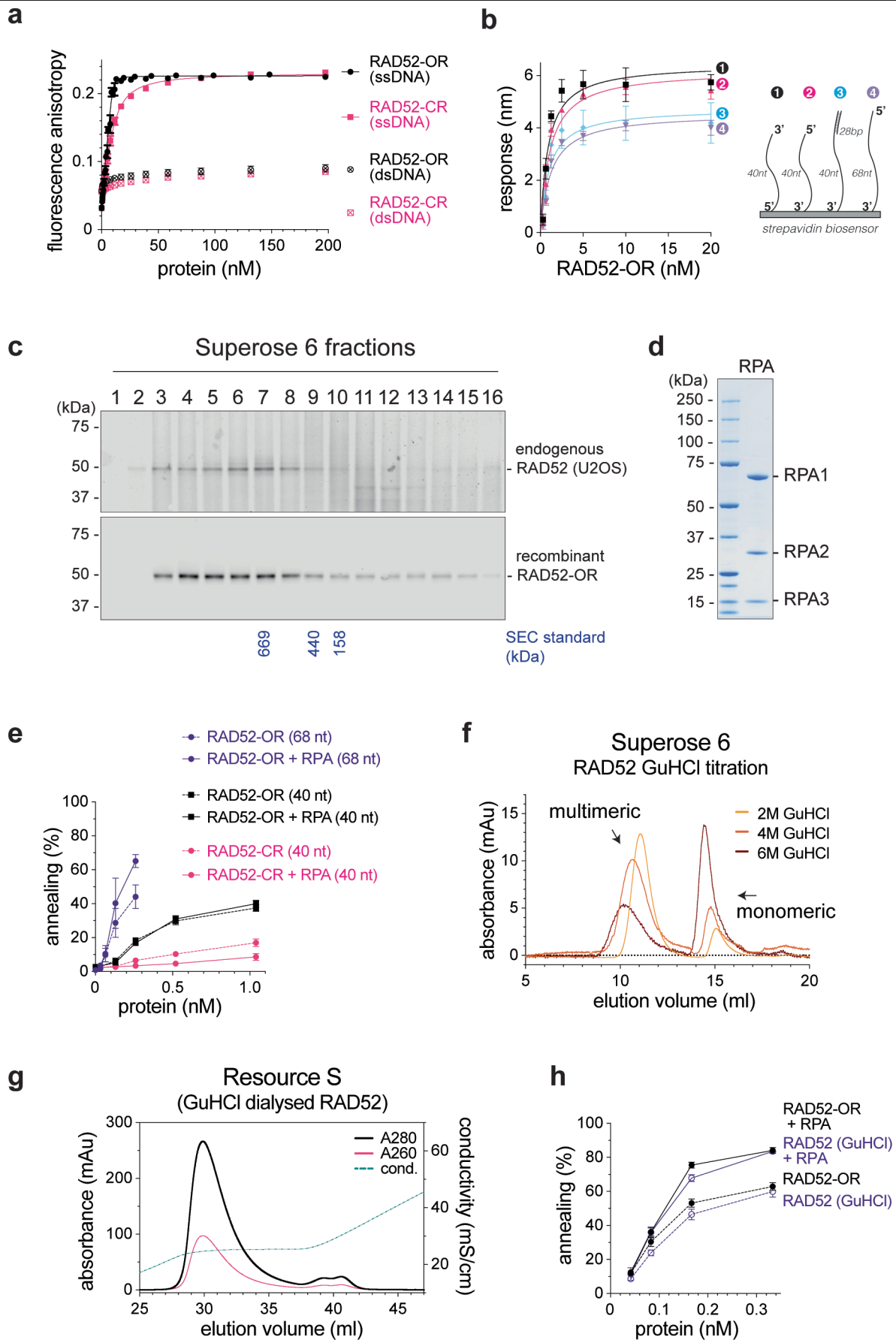
**Peer review information** Nature thanks Charles Bell, R. Scott Williams and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Characterization of human RAD52.** **a**, Coomassie blue stained SDS-PAGE of the two RAD52 peaks from Resource S cation exchange chromatography (as Fig. 1a). **b**, Resource S chromatography of human FLAG-RAD52 expressed in Sf9 insect cells. **c**, SDS-PAGE gel showing RAD52<sup>NTD</sup> (1-209aa). **d**, Resource S chromatography of RAD52<sup>NTD</sup>. **e**, Far-UV circular

dichroism (CD) spectra of RAD52 open (OR) and closed rings (CR). Secondary structure content estimates obtained by spectral deconvolution are indicated. **f**, Intact protein mass spectrometry of RAD52-OR and RAD52-CR. Upper panels: deconvoluted spectra; lower panels: raw spectra. The measured mass of each protein is shown.



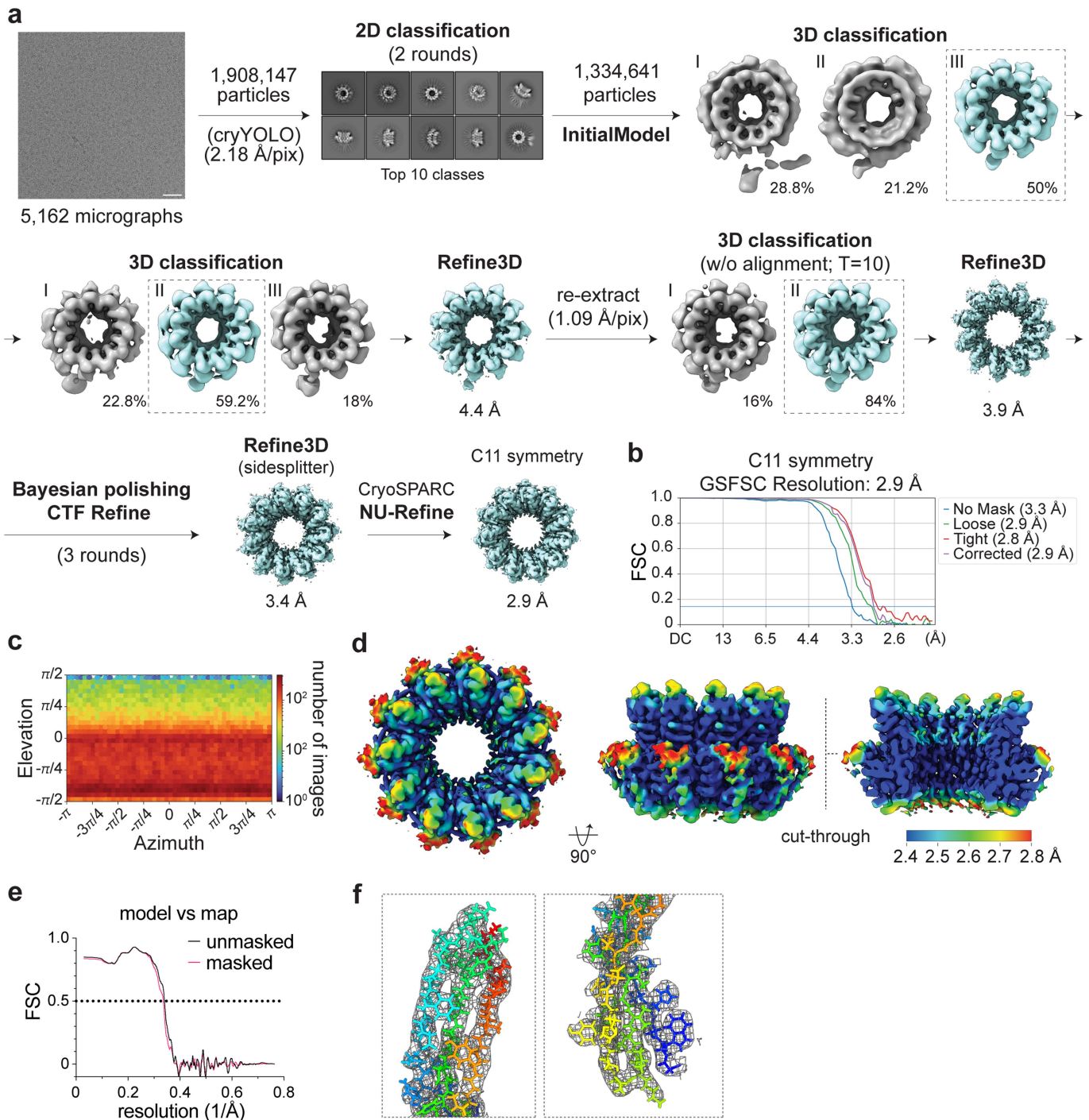
Extended Data Fig. 2 | See next page for caption.

# Article

**Extended Data Fig. 2 | Single strand DNA annealing by RAD52.** **a**, RAD52-OR and RAD52-CR binding to 40 nt long ssDNA (10 nM, FAM-SSA4) or dsDNA (10 nM, FAM-SSA4/SSA5) measured by fluorescence anisotropy. Lines denote best quadratic curve fits. Each point and error bar denotes mean + s.e.m. (RAD52-OR-ssDNA, n = 6; RAD52-CR-ssDNA, n = 3; RAD52-OR-dsDNA, n = 3; RAD52-CR-dsDNA, n = 3). n values are independent experiments. **b**, RAD52-OR binding to the indicated biotinylated DNAs, as measured by biolayer interferometry. Lines denote 1:1 binding curve fits. Each point and error bar denotes mean + s.e.m. (n = 3). Equilibrium dissociation constants ( $K_D$ ) were: (1)  $0.89 \pm 0.35$  nM, (2)  $1.11 \pm 0.4$  nM, (3)  $1.0 \pm 0.55$  nM and (4)  $1.23 \pm 0.41$  nM. **c**, Size exclusion chromatography of a nuclear extract from U2OS cells compared with

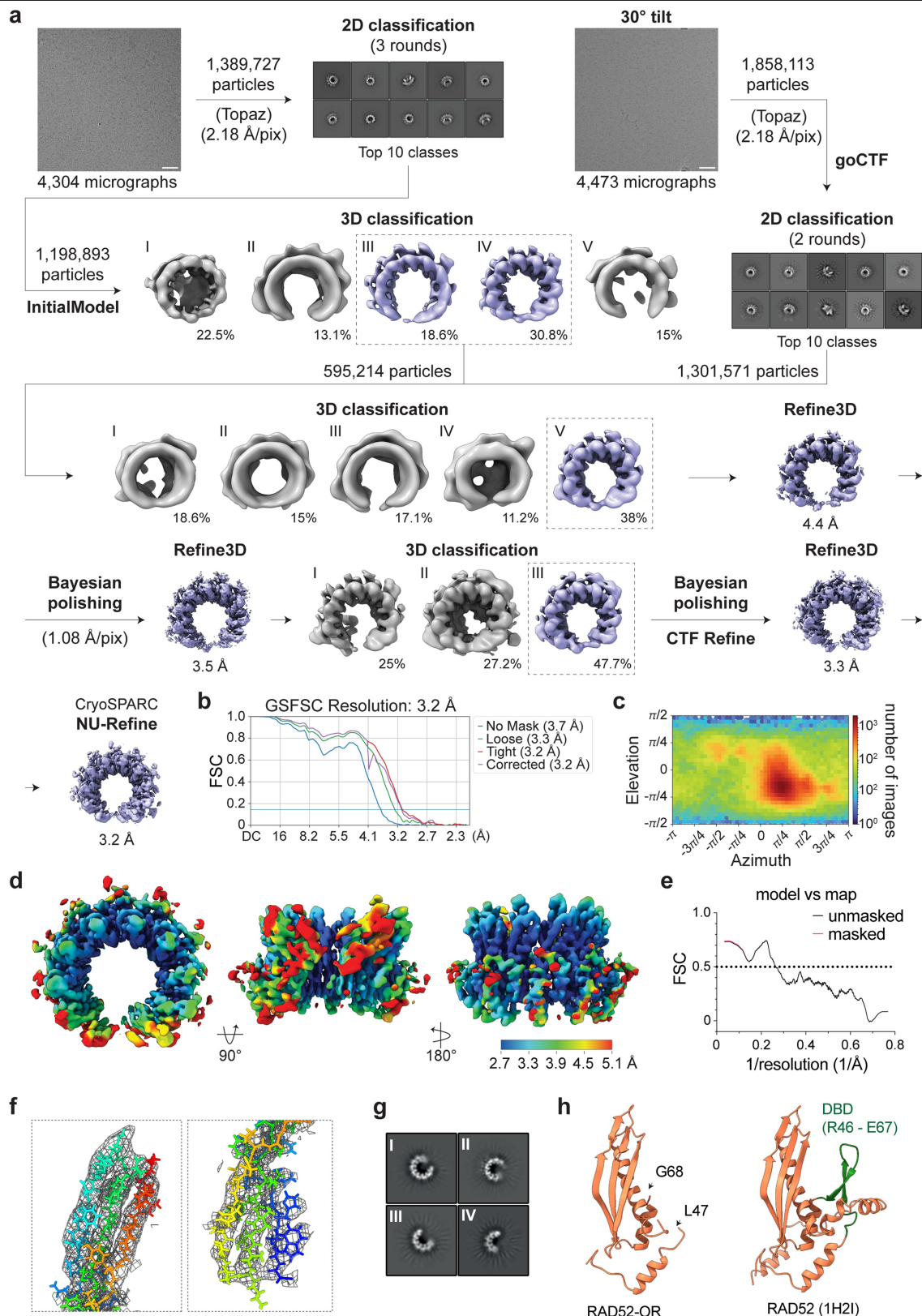
recombinant RAD52-OR. RAD52 was detected by Western blotting. **d**, SDS-PAGE of RPA. **e**, SSA using RAD52 (OR or CR) in the presence or absence of RPA with 68 or 40 nt ssDNA. Each point and error bar denotes mean + s.e.m. (RAD52-OR [68 nt] and RAD52-OR + RPA [68 nt], n = 3; RAD52-OR [40 nt] and RAD52-OR + RPA [40 nt], n = 9; RAD52-OR [40 nt] and RAD52-OR + RPA [40 nt], n = 3). n values are independent experiments. **f**, Superose 6 filtration of RAD52 following dialysis in 2, 4 and 6 M GuHCl. **g**, Resource S chromatography of RAD52 following GuHCl denaturation and refolding. **h**, SSA using renatured RAD52, as in Fig. 1f. Each point and error bar denotes mean + s.e.m. (RAD52-OR and RAD52-OR + RPA, n = 22; RAD52-OR [GuHCl] and RAD52-OR [GuHCl] + RPA, n = 3). n values are independent experiments.





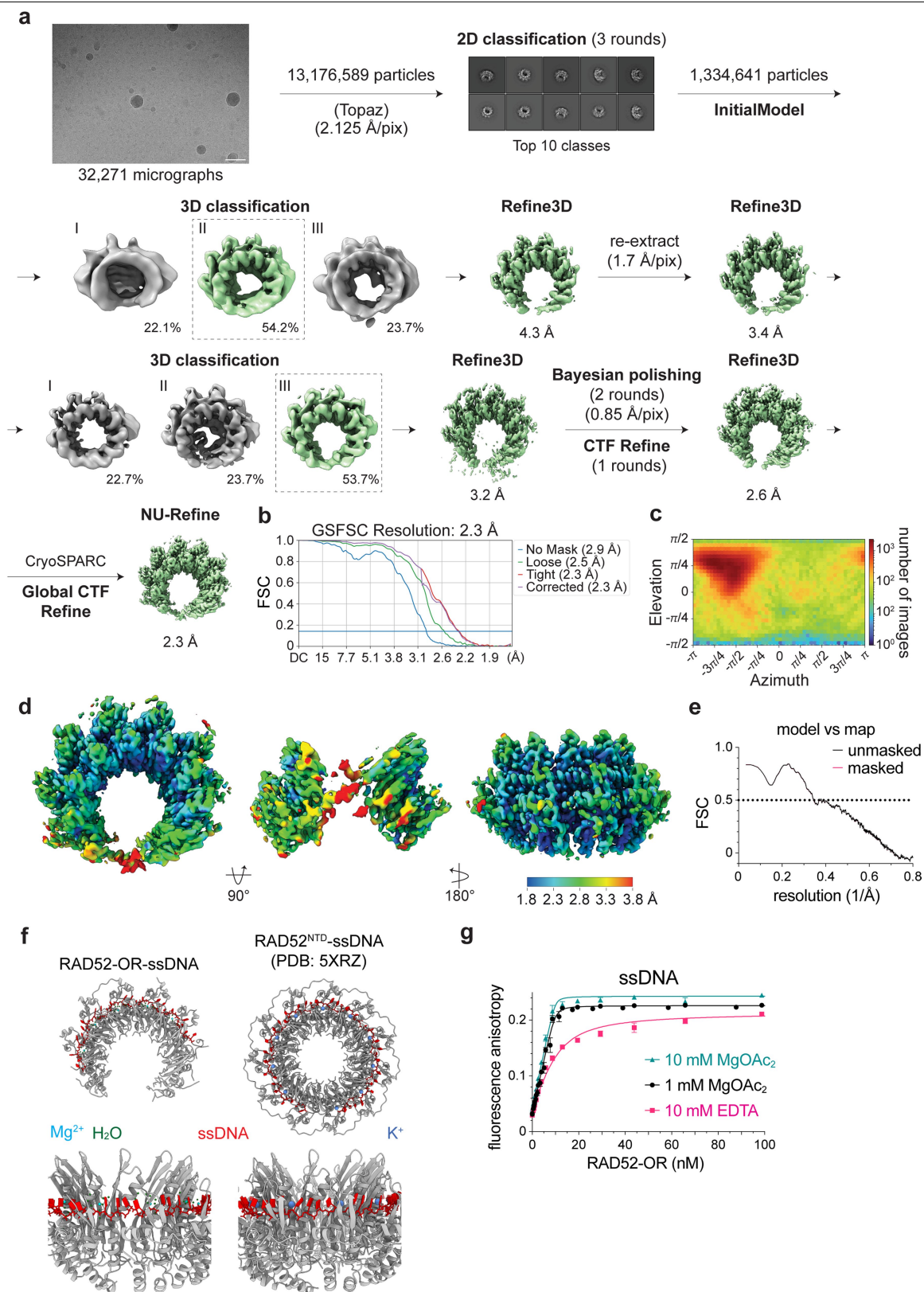
**Extended Data Fig. 3 | Cryo-EM data processing of RAD52-CR.** **a**, Schematic showing the classification and refinement steps to determine the cryo-EM structure of RAD52-CR. The software used during each processing step is indicated if outside of RELION. Scale bar for representative micrograph = 50 nm. **b**, Fourier Shell Correlation (FSC) plot for C11 symmetry refinement.

**c**, Angular distribution plot for C11 symmetry refinement. **d**, Top, side and cut-through views of RAD52-CR cryo-EM density (C11 symmetry) coloured by local resolution estimated by CryoSPARC. **e**, Model to map correlation graph. **f**, Zoom view of RAD52-CR cryo-EM map and atomic model focusing on 83-133aa of the 5<sup>th</sup> RAD52 subunit.



**Extended Data Fig. 4 | Cryo-EM data processing of RAD52-OR.** **a**, As Extended Data Fig. 3, except with the RAD52-OR. **b**, Fourier Shell Correlation (FSC) plot. **c**, Angular distribution plot. **d**, Top, front and back views of RAD52-OR cryo-EM density coloured by local resolution estimated by CryoSPARC. **e**, Model to map correlation graph. **f**, Zoom view of RAD52-OR

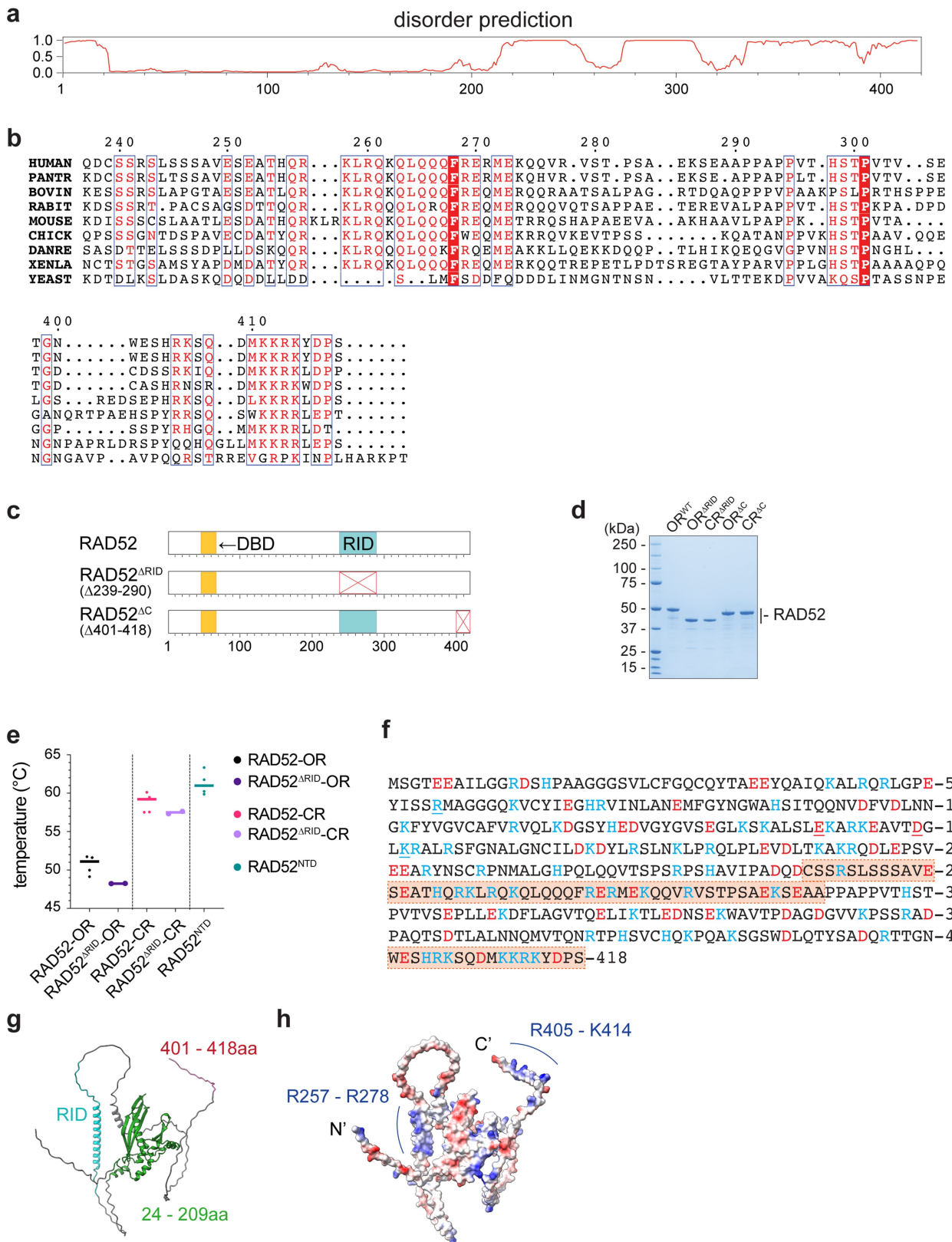
cryo-EM map and atomic model focusing on 83-133aa of the 5<sup>th</sup> RAD52 subunit. **g**, Representative cryo-EM EM 2D averages of RAD52 open ring with (I) 9, (II) 7 – 8, (III) 6 and (IV) 5 subunits. **h**, Comparison of a RAD52 protomer from 1H2I crystal structure with the open ring cryo-EM structure.



**Extended Data Fig. 5 | Cryo-EM data processing of RAD52-ssDNA complex.**

**a**, As Extended Data Fig. 3, except for the RAD52-OR-ssDNA structure. **b**, Fourier Shell Correlation (FSC) plot. **c**, Angular distribution plot. **d**, Top, front and back views of the RAD52-OR-ssDNA cryo-EM density coloured by local resolution estimated by CryoSPARC. **e**, Model to map correlation graph. **f**, Comparison

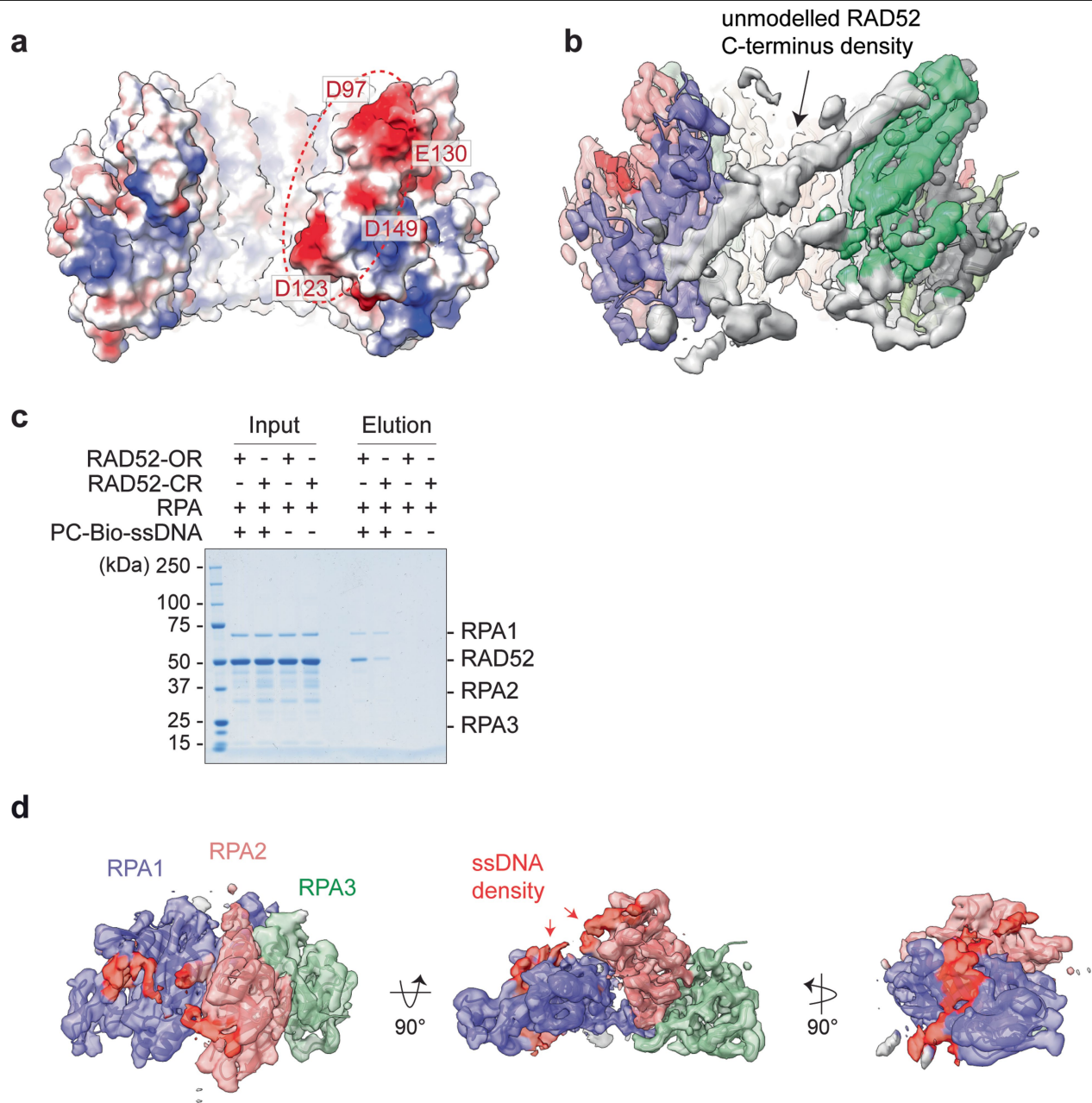
between RAD52-OR-ssDNA and RAD52<sup>NTD-K102A/K133A</sup>-ssDNA (PDB: 5XRZ) structures. **g**, Fluorescence anisotropy of RAD52-OR binding to ssDNA (FAM-SSA4; 40 nt) at different Mg<sup>2+</sup> concentrations. Lines are best quadratic curve fits. Each point and error bar denotes mean + s.e.m. (n = 3). n values are independent experiments.



Extended Data Fig. 6 | See next page for caption.

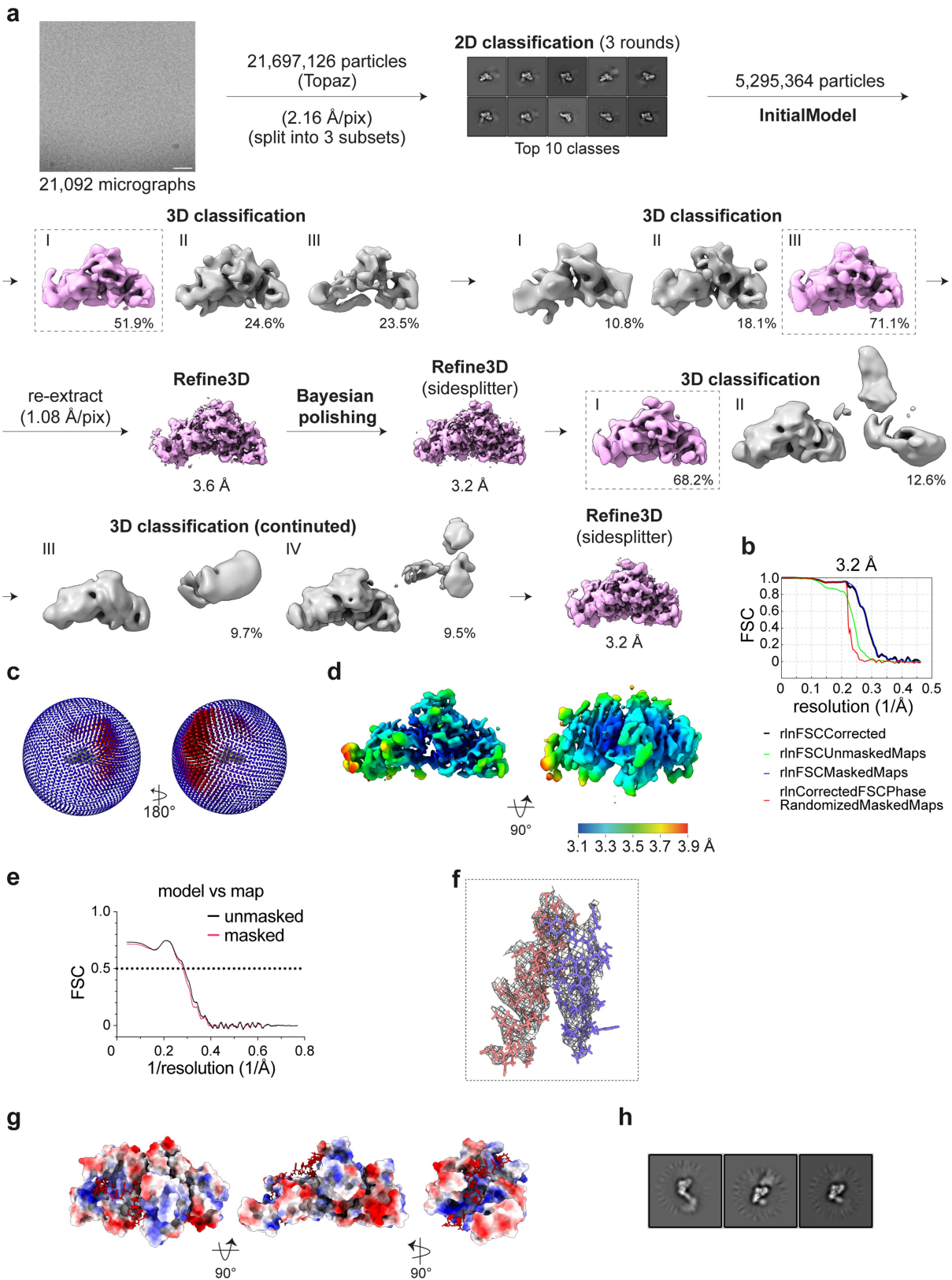
**Extended Data Fig. 6 | Analysis of RAD52 mutants.** **a**, Protein disorder prediction of human RAD52. **b**, Clustal Omega sequence alignment reveals conservation of 239-290 and 401-418aa in vertebrates. **c**, Domain architecture of RAD52, and the RAD52<sup>ARID</sup> and RAD52<sup>AC</sup> deletion mutants. **d**, SDS-PAGE of RAD52<sup>WT</sup>-OR, RAD52<sup>ARID</sup>-OR, RAD52<sup>ARID</sup>-CR, RAD52<sup>AC</sup>-OR and RAD52<sup>AC</sup>-CR, stained with Coomassie blue. **e**, Thermal melting analysis of the indicated proteins. Each point and error bar denotes mean + s.d. RAD52<sup>WT</sup>-OR (50.7 ± 1.1 °C; n = 5), RAD52<sup>ARID</sup>-OR (48.2 ± 0.01 °C; n = 2), RAD52<sup>WT</sup>-CR (58.8 ± 1.2 °C; n = 5), RAD52<sup>ARID</sup>-CR (57.5 ± 0.2 °C; n = 2) and RAD52<sup>NTD</sup>

(61.3 ± 1.6 °C; n = 4). n values are independent experiments. **f**, Human RAD52 amino acid sequence coloured by electrostatic potential (blue, positively charged; red, negatively charged). Boxed and shaded sequence highlights the RAD52 RID (239-290aa) and the extreme C-terminus (401-418aa). **g**, AlphaFold2 prediction of RAD52 structure. The N-terminal domain (24-209aa, green), RID (239-290aa, cyan) and extreme C-terminus (401-418aa, red) are indicated. **h**, AlphaFold2 predicted RAD52 structure shown in surface view coloured with electrostatic potential.



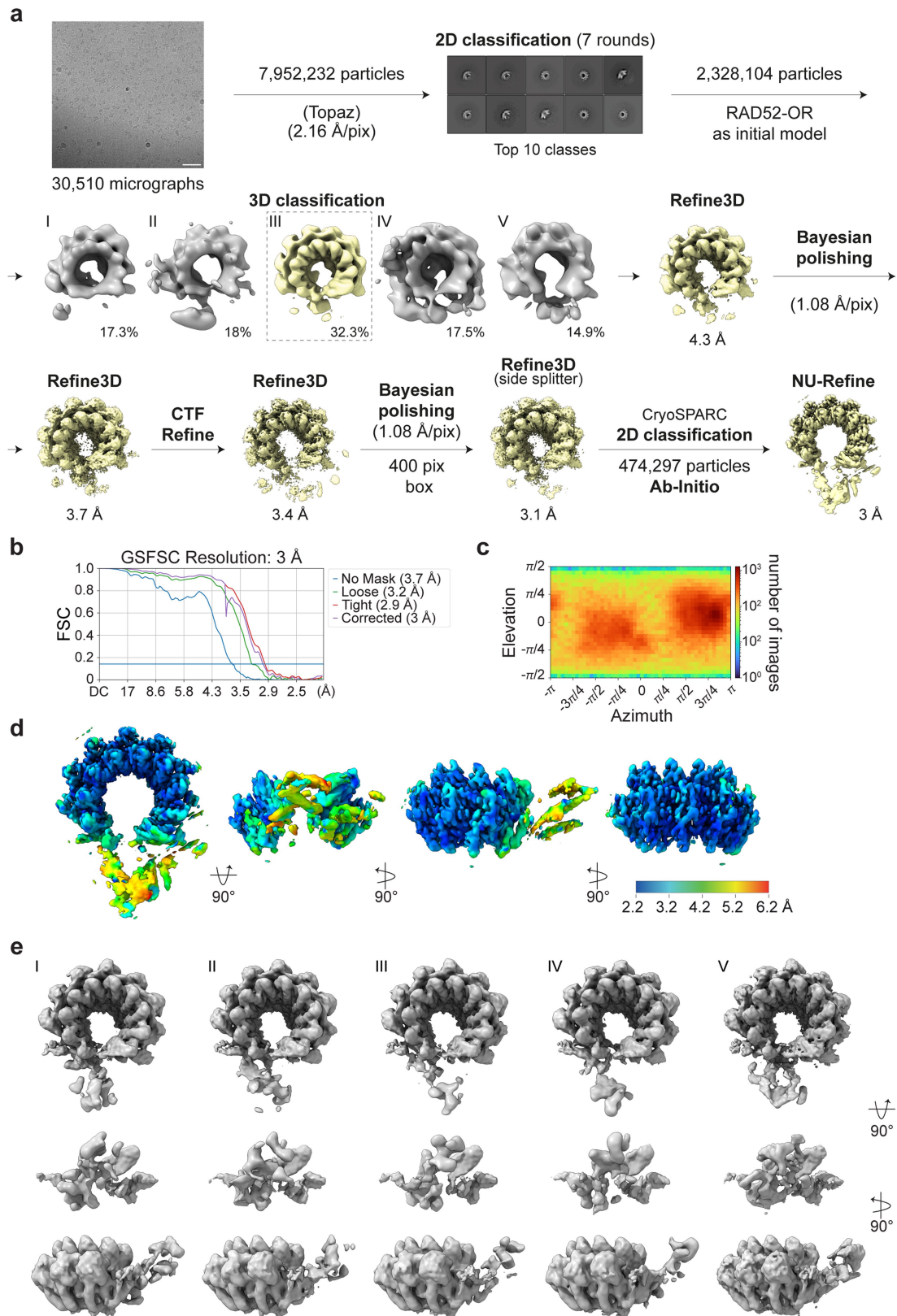
**Extended Data Fig. 7 | Structure of RPA-ssDNA and the RAD52-ssDNA-RPA complex.** **a**, Surface view of a RAD52-ssDNA atomic model with electrostatic potential indicated (blue, positively charged; red, negatively charged). D97, D123, E130 and D149 are located at the opening of the RAD52 ring. **b**, Cryo-EM density of a RAD52-ssDNA sub-class separated by CryoSPARC 3D classification analysis with additional C-terminus density occupying the opening of the

RAD52 ring. **c**, Coomassie blue stained SDS-PAGE of RAD52-ssDNA-RPA complex following streptavidin pull-down targeting biotin-labelled ssDNA with photocleavable linker (PC-Bio-ssDNA; 40 nt). **d**, Top, front and side view of the cryo-EM map of RPA-ssDNA complex (3.2 Å) with atomic model of the RPA trimeric core. ssDNA density (unmodelled) is coloured in red.



**Extended Data Fig. 8 | Cryo-EM data processing of RPA-ssDNA.** **a**, As Extended Data Fig. 3 except for the RPA-ssDNA structure. **b**, Fourier Shell Correlation (FSC) plot. **c**, Angular distribution plot. **d**, Side and top views of RPA-ssDNA cryo-EM density coloured by local resolution. **e**, Model to map correlation graph. **f**, Zoom view of RPA-ssDNA cryo-EM map and atomic model

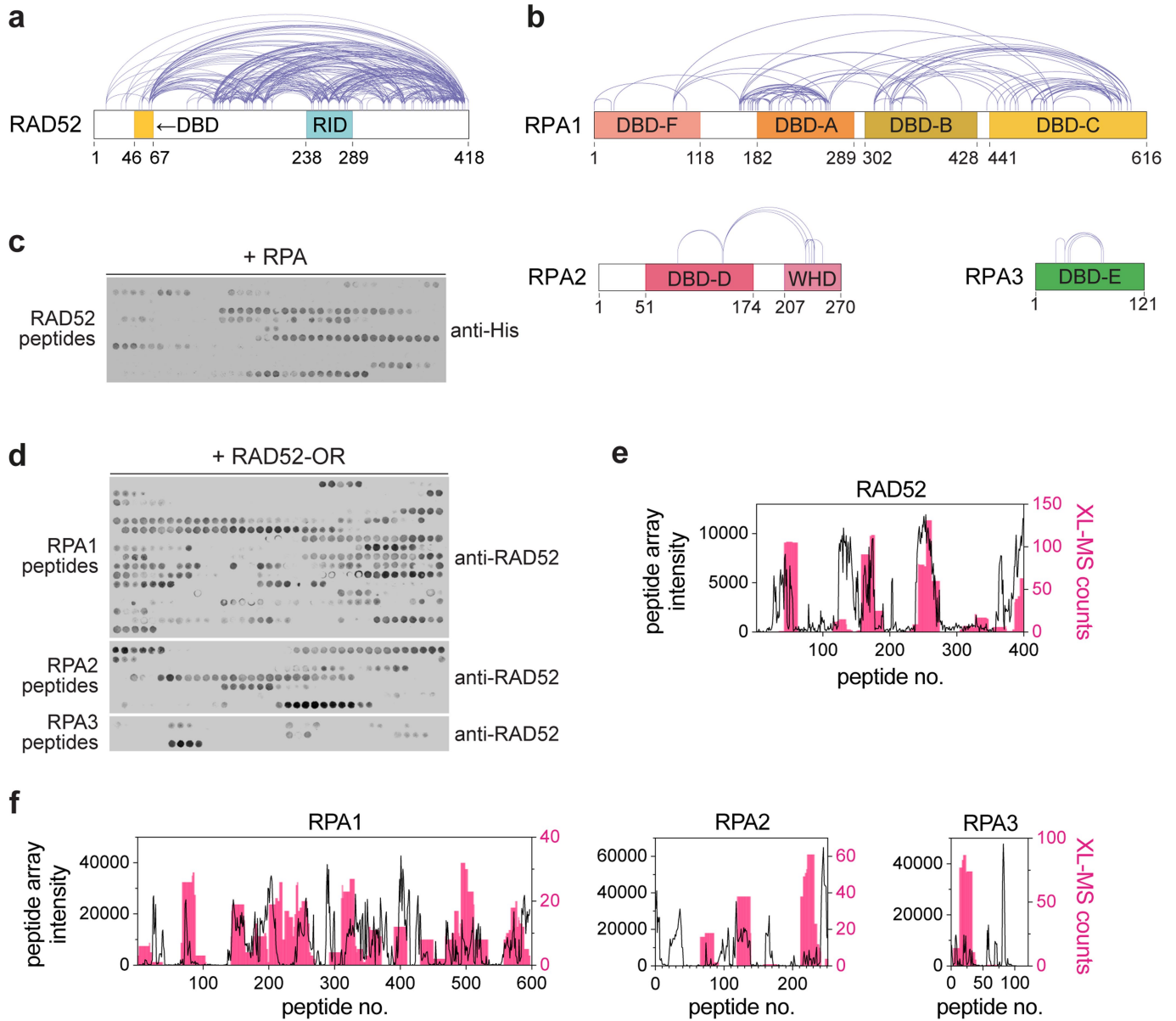
focusing on 599-616aa of RPA1 (purple) and 152-174aa of RPA2 (pink). **g**, Top, front and side views of atomic model of RPA (trimeric core) presented as an electrostatic potential coloured surface. ssDNA model is present in stick style (red). **h**, Representative cryo-EM 2D class averages of RPA trimeric core with flexible density.



**Extended Data Fig. 9 | Cryo-EM data processing of RAD52-ssDNA-RPA complex.** **a**, As Extended Data Fig. 3 except for the RAD52-OR-ssDNA-RPA structure. **b**, Fourier Shell Correlation (FSC) plot. **c**, Angular distribution plot. **d**, Top, front, side and back views of RAD52-ssDNA-RPA cryo-EM density

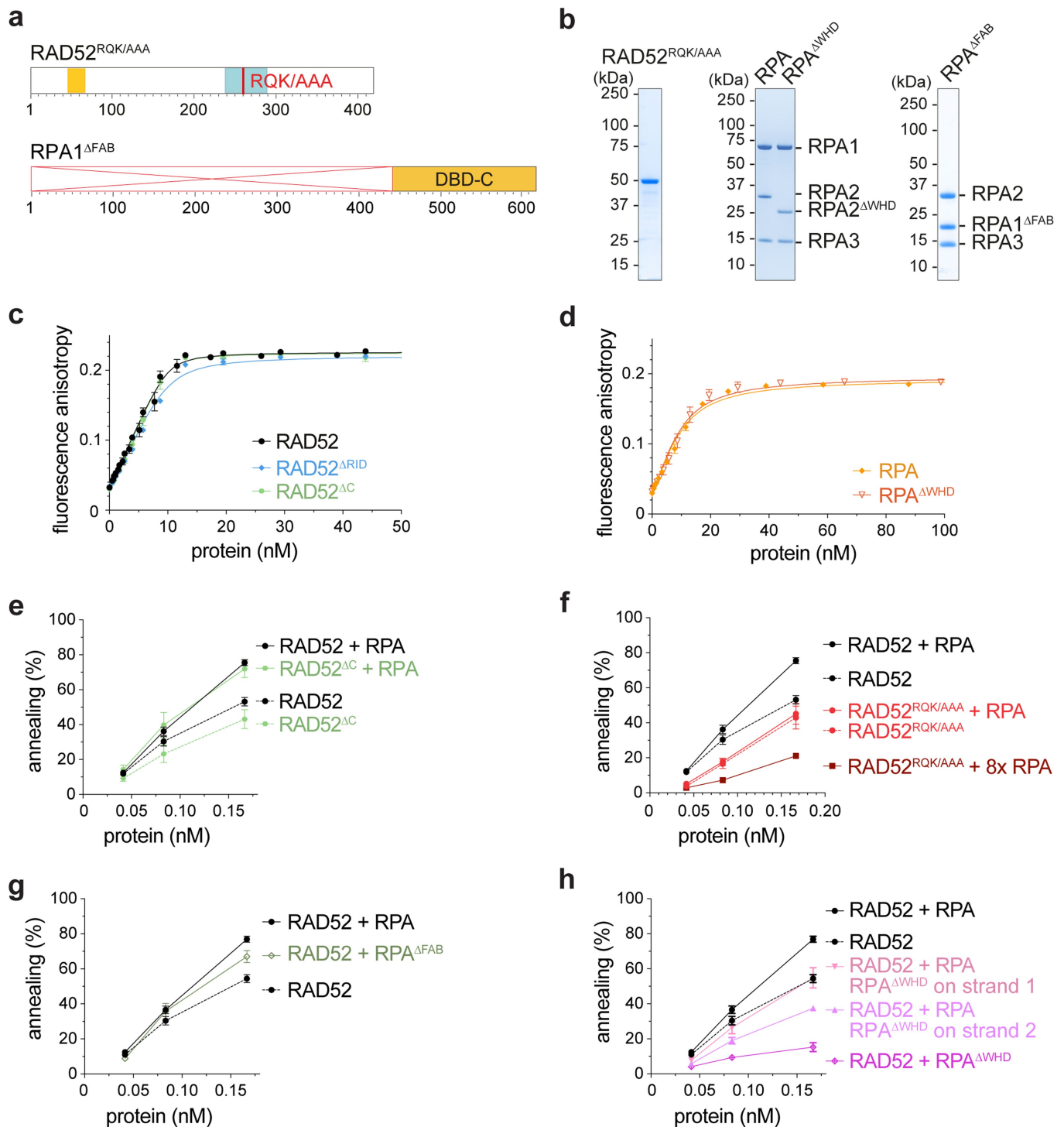
coloured by local resolution estimated by CryoSPARC. **e**, Top, front and side views of 5 classes of the RAD52-ssDNA-RPA complex separated by focused 3D classification in CryoSPARC.





**Extended Data Fig. 10 | RAD52-RPA interactions.** **a**, and **b**, Crosslinking mass spectrometry analysis of the RAD52-ssDNA-RPA complex. Diagram depicts crosslinks detected within RAD52-OR and RPA1, RPA2 or RPA3. **c**, and **d**, Peptide arrays of RAD52, RPA1, RPA2 and RPA3 showing interactions between RAD52 and RPA (related heat maps are shown in Extended Data Fig. 6c). Each spot

corresponds to a 20 aa peptide sequence, and the intensity of signal reflects the strength of interaction between target proteins and peptides. **e** and **f**, Comparison of peptide arrays and XL-MS data for RAD52-RPA interactions. Crosslinked amino acid residues were transformed into counts per peptide corresponding to peptides used in the arrays.



**Extended Data Fig. 11 | RAD52-RPA interactions involved in SSA.** **a**, Schematic of RAD52<sup>RQK/AAA</sup> (R260A, Q261A and K262A) and RPA1<sup>ΔFAB</sup> (Δ2-440aa). **b**, SDS-PAGE of RAD52<sup>RQK/AAA</sup>, RPA, RPA<sup>ΔWHD</sup> and RPA<sup>ΔFAB</sup>. **c**, Binding to ssDNA (10 nM, FAM-SSA4; 40 nt) by RAD52-OR, RAD52<sup>ΔRID</sup> and RAD52<sup>ΔC</sup> measured by fluorescence anisotropy, as in Fig. 1c. Equilibrium dissociation constants ( $K_D$ ) for ssDNA are  $0.3 \pm 0.1$  nM (RAD52-OR),  $0.7 \pm 0.2$  nM (RAD52<sup>ΔRID</sup>-OR) and  $0.2 \pm 0.2$  nM (RAD52<sup>ΔC</sup>). Lines are best quadratic curve fits. Each point and error bar denotes mean + s.e.m. (RAD52-OR,  $n = 6$ ; RAD52<sup>ΔRID</sup>,  $n = 3$ ; RAD52<sup>ΔC</sup>,  $n = 3$ ).  $n$  values are independent experiments. **d**, RPA and RPA<sup>ΔWHD</sup> binding to ssDNA (10 nM, FAM-SSA4; 40 nt) measured by fluorescence anisotropy. Equilibrium dissociation constants ( $K_D$ ) for ssDNA are  $3.6 \pm 0.5$  nM (RPA) and  $3.4 \pm 0.9$  nM (RPA<sup>ΔWHD</sup>). Lines are best quadratic curve fits. Each point and error bar denotes mean + s.d. (RPA,  $n = 3$ ; RPA<sup>ΔWHD</sup>,  $n = 2$ ).  $n$  values are independent experiments.

**e**, SSA by RAD52-OR or RAD52<sup>ΔC</sup>, with or without RPA, as in Fig. 1e. Each point and error bar denotes mean + s.e.m. (RAD52 and RAD52 + RPA,  $n = 22$ ; RAD52<sup>ΔC</sup> and RAD52<sup>ΔC</sup> + RPA,  $n = 3$ ). (ssDNA: 68 nt) **f**, SSA (68 nt) by RAD52-OR or RAD52<sup>RQK/AAA</sup>, with or without RPA. Each point and error bar denotes mean + s.e.m. (RAD52 and RAD52 + RPA,  $n = 22$ ; RAD52<sup>RQK/AAA</sup>,  $n = 6$ ; RAD52<sup>RQK/AAA</sup> + RPA and RAD52<sup>RQK/AAA</sup> + 8xRPA,  $n = 3$ ).  $n$  values are independent experiments. **g**, SSA by RAD52-OR with RPA or RPA<sup>ΔFAB</sup> (ssDNA: 68 nt). Each point and error bar denotes mean + s.e.m. ( $n = 3$ ; except RAD52 and RAD52 + RPA,  $n = 22$ ).  $n$  values are independent experiments. **h**, SSA by RAD52-OR with RPA or RPA<sup>ΔWHD</sup>, as in Fig. 5f (ssDNA: 68 nt). RAD52 was bound to strand 2. RPA was bound to both strands or replaced by RPA<sup>ΔWHD</sup> as indicated. Each point and error bar denotes mean + s.e.m. ( $n = 3$ ; except RAD52 and RAD52 + RPA,  $n = 22$ , and RAD52 + RPA<sup>ΔWHD</sup>,  $n = 7$ ).  $n$  values are independent experiments.

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

	RAD52-CR (EMD- 19189) (PDB 8RIL)	RAD52-OR (EMD- 19193) (PDB 8RJ3)	RAD52-OR- ssDNA (EMD- 19253) (PDB 8RJW)	RPA-ssDNA (EMD- 19255) (PDB 8RK2)	RAD52-OR- ssDNA-RPA
<b>Data collection and processing</b>					
Magnification	75,000	75,000	105,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	32	33	40	46	46
Defocus range (μm)	-1.5 - -2.7	-1.5 - -3	-0.75 - 2.25	-1.5 - -3.5	-1.5 - -3.5
Pixel size (Å)	1.09	1.09	0.85	1.08	1.08
Symmetry imposed	C11	C1	C1	C1	C1
Initial particle images (no.)	1,908,147	3,247,840	13,176,589	21,697,126	7,952,232
Final particle images (no.)	309,471	356,878	587,046	1,316,062	474,297
Map resolution (Å) FSC threshold = 0.143	2.9	3.2	2.3	3.2	3
Map resolution range (Å)	2.4 - 2.8	2.7 - 5.1	1.8 - 3.8	3.1 - 3.9	2.2 - 6.2
<b>Refinement</b>					
Initial model used (PDB code)	1h2i	RAD52-CR	RAD52-OR	AlphaFold	RAD52-OR- ssDNA
Model resolution (Å) FSC threshold = 0.5	2.9	3.8	2.9	3.5	
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-150	DeepEMhan cer	DeepEMhan cer or -71.5	-75	DeepEMhan cer
Model composition					
Non-hydrogen atoms	10120	8342	10304	3246	
Protein residues	1287	1061	1242	404	
Nucleotides			23		
Ligands			MG:5		
<i>B</i> factors (Å <sup>2</sup> )					
Protein	90.65	100.24	54.56	138.23	
Nucleotides			72.24		
Ligand			64.78		
R.m.s. deviations					
Bond lengths (Å)	0.005	0.003	0.005	0.003	
Bond angles (°)	0.414	0.504	0.535	0.478	
Validation					
MolProbity score	1.25	1.9	1.43	1.63	
Clashscore	3.3	9.16	5.54	6.02	
Poor rotamers (%)	0	0	0	0	
Ramachandran plot					
Favored (%)	97.35	93.87	97.32	95.71	
Allowed (%)	2.65	6.13	2.68	4.29	
Disallowed (%)	0	0	0	0	

## 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

Commercial software available with the respective instruments was used for data collection. These include Chemidoc MP Image Lab Touch Software (version 3.0.1.14, Biorad) for SDS-PAGE gel imaging, ImageStudio (v 5.2, Licor) for imaging of Western blots and peptide arrays, Clariostar BML (Labtech) for fluorescence anisotropy, (v 5.7), EPU (ThermoFisher) for cryo-EM data collection, DigitalMicrograph (Gatan) for NS-EM data, UNICORN (Cytiva) for protein purification, AcquireMP (Refeyn) for mass photometry, PR.ThermControl (V2.1.2, Nano Temper) for nanoscale differential scanning fluorometry, Octet BLI Discovery (v12.2, Sartorius) for BLI, ASTRA (v 7.3.2.19, WYATT) and ChromNAV (v 2.03.06; JASCO) for SEC-MALLS.

#### Data analysis

MARS (Labtech) for blank correction of fluorescence anisotropy data, ImageJ2 (v 2.14) for SSA quantification, MaxEnt1 (Masslynx, Waters) for intact protein mass spectrometry, Octet BLI Analysis (v 12.2, Sartorius) for BLI, PR.Stability ASAnalysis (v 1.0.2, Nano temper) for nanoscale differential scanning fluorometry, DiscoverMP (Refeyn) for mass photometry, Proteome Discoverer (v 2.3, ThermoScientific) for XL-MS and AlphaFold2 for protein structure prediction (installed locally).  
For cryo-EM and NS-EM data analysis: RELION (v 3.1 and 4.0), crYOLO (v 1.7.5), MotionCor2, CTFIND4, Topaz (v 0.2.5), ChimeraX (v 1.4), COOT (v 0.9.8.7), Phenix (v 1.2), DeepEMhancer, Matchmake  
ClustalOmega for multiple sequence alignment, ESPript (v 3.0).  
GraphPad Prism 9 (v 9.2) for statistical analysis and data plotting

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](#)

The datasets generated during and/or analysed during the current study are included alongside the Article or are available from S.C.W on reasonable request. Cryo-EM density maps and atomic models of RAD52 have been deposited in the Electron Microscopy Data Bank (EMDB) and Protein Database (PDB). Accession codes are as follows: RAD52-CR (EMDB = 19189, PDB = 8RIL), RAD52-OR (EMDB = 19193, PDB = 8RJ3), RAD52-OR-ssDNA (EMDB = 19253, PDB = 8RJW) and RPA-ssDNA (EMDB = 19255, PDB = 8RK2). All other data is archived at the Francis Crick Institute.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## 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://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	Wherever quantification is provided, a minimum 3 independent experiments were carried out to perform statistical analyses as indicated in figure legends as n.
Data exclusions	In SSA assays, some samples remained in the wells because of incomplete deproteinisation. These samples were not quantified and excluded from the datasets. In cryo-EM processing, micrographs with CTF outliers (often due to presence of contaminating ice or large areas of foil) were removed from analyses. Picked particles that did not align in 2D or 3D were removed from further analysis. For NS-EM, all datasets went through two rounds of 2D classification.
Replication	To ensure the replication, each experiment was performed at least three times in exact condition on different days. All attempts were successful. The cryo-EM structures obtained from 300 kV Krios were also observed in other lower resolution datasets obtained on the 200 kV TALOS (data not shown).
Randomization	For calculation of resolution in cryo-EM maps, FSCs were calculated using independent halves of the datasets, into which the particles were randomly allocated. In addition, Topaz particle picking models were trained on a random subset of particles. Randomization is not relevant to the in vitro analysis of recombinant proteins, and therefore was not used in this study.
Blinding	Blinding is not necessary for analysis of recombinant proteins by cryoelectron microscopy, single molecule methods or biochemistry. The researcher needs to know the protein sample used to conduct the experiments performed, and therefore blinding is not appropriate.

## 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

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" 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"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

## Methods

n/a	Involvement 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

## Antibodies

Antibodies used

For Western blotting, we used the following antibodies. RAD52 (Rabbit monoclonal, 1:500, ab124971, Abcam) and 6xHis (Mouse monoclonal, 1:1000, 631212, TakaRa) . Membranes were incubated with Alexa Fluor Plus 800 anti-rabbit secondary antibody (1:2000, Invitrogen, A32735) or Alexa Fluor Plus 800 anti-mouse secondary antibody (1:2000, Invitrogen, A32730) .

Validation

Antibodies in this study are only used for immunoblotting against recombinant proteins.

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

U2OS and Sf9 cells are available from The Francis Crick Institute, London, UK upon request.

Authentication

U2OS and Sf9 cells have been authenticated and confirmed to be the correct cell line.

Mycoplasma contamination

U2OS and Sf9 cells have tested negative for mycoplasma

Commonly misidentified lines  
(See [ICLAC](#) register)

No commonly misidentified cell lines have been used in this study.