

# HELQ is a dual-function DSB repair enzyme modulated by RPA and RAD51

<https://doi.org/10.1038/s41586-021-04261-0>

Received: 2 June 2021

Accepted: 17 November 2021

Published online: 22 December 2021

Open access

 Check for updates

Roopesh Anand<sup>1,6</sup>, Erika Buechelmaier<sup>2,3,6</sup>, Ondrej Belan<sup>1</sup>, Matthew Newton<sup>1</sup>, Aleksandra Vancevska<sup>1</sup>, Artur Kaczmarczyk<sup>4,5</sup>, Tohru Takaki<sup>1</sup>, David S. Rueda<sup>4,5</sup>, Simon N. Powell<sup>2</sup> & Simon J. Boulton<sup>1</sup>

DNA double-stranded breaks (DSBs) are deleterious lesions, and their incorrect repair can drive cancer development<sup>1</sup>. HELQ is a superfamily 2 helicase with 3' to 5' polarity, and its disruption in mice confers germ cells loss, infertility and increased predisposition to ovarian and pituitary tumours<sup>2–4</sup>. At the cellular level, defects in HELQ result in hypersensitivity to cisplatin and mitomycin C, and persistence of RAD51 foci after DNA damage<sup>3,5</sup>. Notably, HELQ binds to RPA and the RAD51-paralogue BCDX2 complex, but the relevance of these interactions and how HELQ functions in DSB repair remains unclear<sup>3,5,6</sup>. Here we show that HELQ helicase activity and a previously unappreciated DNA strand annealing function are differentially regulated by RPA and RAD51. Using biochemistry analyses and single-molecule imaging, we establish that RAD51 forms a complex with and strongly stimulates HELQ as it translocates during DNA unwinding. By contrast, RPA inhibits DNA unwinding by HELQ but strongly stimulates DNA strand annealing. Mechanistically, we show that HELQ possesses an intrinsic ability to capture RPA-bound DNA strands and then displace RPA to facilitate annealing of complementary sequences. Finally, we show that HELQ deficiency in cells compromises single-strand annealing and microhomology-mediated end-joining pathways and leads to bias towards long-tract gene conversion tracts during homologous recombination. Thus, our results implicate HELQ in multiple arms of DSB repair through co-factor-dependent modulation of intrinsic translocase and DNA strand annealing activities.

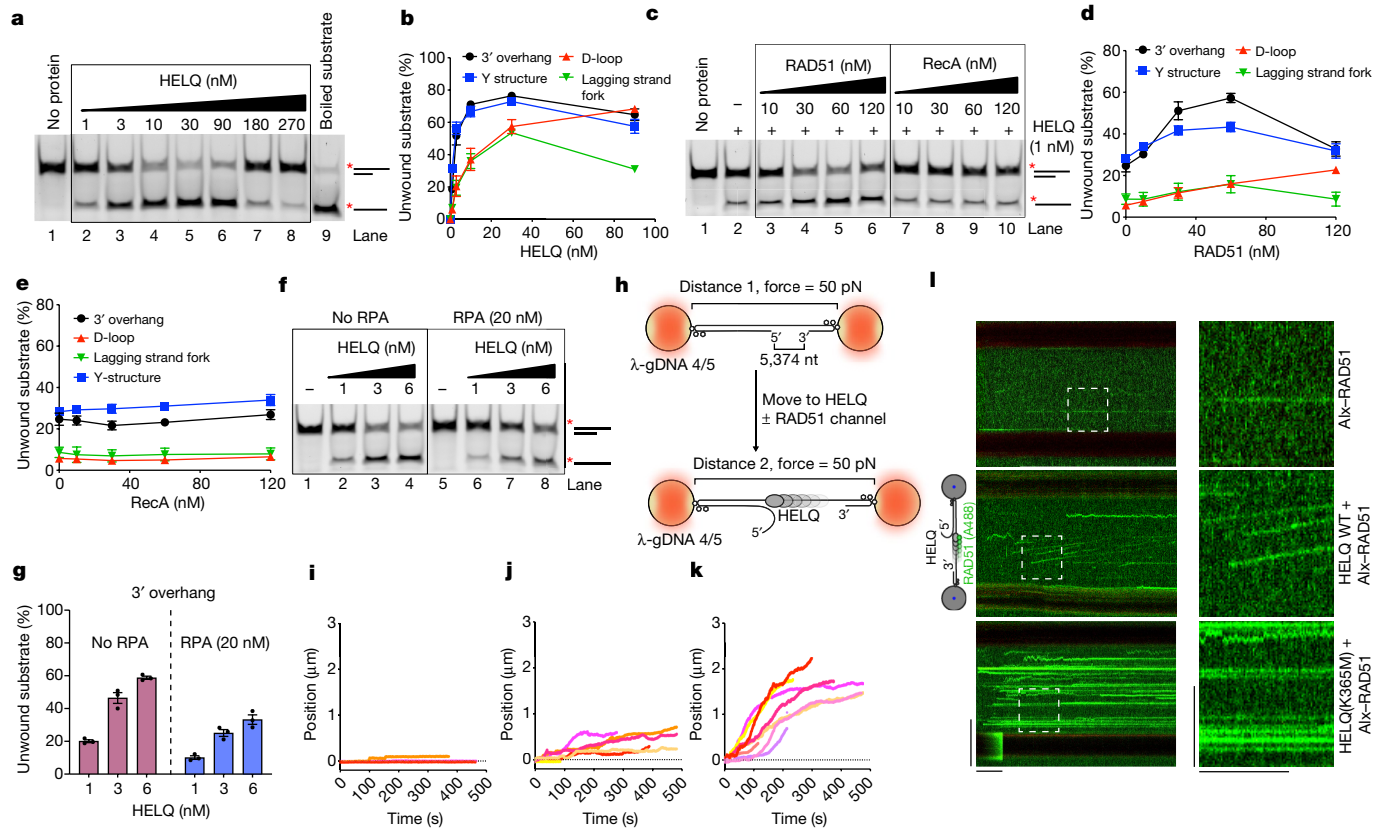
To investigate the functions of HELQ, we purified recombinant human HELQ from insect cells (Extended Data Fig. 1a), which efficiently unwound substrates containing 3' overhangs or a D-loop (Fig. 1a, b and Extended Data Fig. 1b–d). As previously reported, HELQ prefers to unwind single and double-stranded DNA junctions and therefore showed greater unwinding of 3' overhangs and Y-structures than 3' lagging strand forks and D-loops<sup>7</sup>. However, at higher concentrations of HELQ, no unwound product was observed, especially for substrates containing 3' overhangs (described below; Extended Data Fig. 1e). HELQ showed no unwinding with ATPγS, a poorly hydrolysable ATP analogue, and failed to unwind dsDNA and 5' overhang substrates (Extended Data Fig. 1f–h). The helicase-dead mutant of HELQ (HELQ(K365M)) lacked DNA-unwinding activity and showed slightly increased binding to ssDNA and dsDNA compared with the wild-type (WT) protein (Extended Data Fig. 1a, i–m).

## RAD51 stimulates HELQ unwinding activity

In vivo studies have shown that HELQ-deficient cells exhibit persistent RAD51 foci after DNA damage<sup>3,8</sup>. Furthermore, HELQ-1 from *Caenorhabditis elegans* interacts with RAD-51 (ref. <sup>8</sup>). HELQ and human RAD51,

purified from *Escherichia coli* (Extended Data Fig. 2a), also interact directly (Extended Data Fig. 2b). In unwinding assays, RAD51 strongly stimulated HELQ helicase activity with all of the tested substrates, whereas bacterial RecA—an orthologue of RAD51—did not stimulate HELQ even at higher concentrations (Fig. 1c–e and Extended Data Fig. 2c–g). We next purified and tested the BRC4 peptide (single BRCA2 BRC repeat), which prevents RAD51 from binding to DNA (Extended Data Fig. 2h–j). The BRC4 peptide did not inhibit stimulation of HELQ by RAD51 (Extended Data Fig. 2k–n), indicating that RAD51 DNA binding is not required for HELQ stimulation and excludes DNA sequestration as a possible mechanism for stimulation by RAD51. We observed that, at a higher concentration of RAD51 (that is, 120 nM), HELQ unwinding activity is inhibited (Fig. 1c, d). To investigate this, we tested the BRC4 peptide with an excess RAD51 and found that inhibiting RAD51 DNA binding also rescued DNA unwinding by HELQ (Extended Data Fig. 2o, p). Furthermore, by measuring the kinetics of DNA unwinding by HELQ, we found that addition of RAD51 resulted in a concentration-dependent increase in the HELQ DNA unwinding rate, whereas addition of RecA had no effect (Extended Data Fig. 2q–s). In cells, ssDNA generated during DNA processing is bound by RPA. To mimic these conditions, we purified fluorescently tagged human RPA–mRFP1 from *E. coli* (Extended

<sup>1</sup>DSB Repair Metabolism Laboratory, The Francis Crick Institute, London, UK. <sup>2</sup>Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>3</sup>Graduate School of Medical Sciences, Weill Cornell Medicine, New York, NY, USA. <sup>4</sup>Department of Infectious Disease, Faculty of Medicine, Imperial College London, London, UK. <sup>5</sup>Single Molecule Imaging Group, MRC-London Institute of Medical Sciences, London, UK. <sup>6</sup>These authors contributed equally: Roopesh Anand, Erika Buechelmaier. ✉e-mail: david.rueda@imperial.ac.uk; powells@mskcc.org; simon.boulton@crick.ac.uk



**Fig. 1 | RAD51 forms a co-complex with HELQ and stimulates HELQ unwinding activity.** **a**, Representative gel of the DNA unwinding assay with the indicated concentrations of HELQ with 3' overhang substrate. The asterisk indicates the position of fluorescein isothiocyanate (FITC) labelling at 5' end of the oligo. The products were resolved on a 10% native polyacrylamide gel. **b**, Quantification of the experiments such as shown in **a** and Extended Data Fig. 1b–d. HELQ concentrations of only 1–90 nM are shown.  $n = 4$  independent experiments. Data are mean  $\pm$  s.e.m. **c**, Representative gel of DNA unwinding of 3' overhang substrate with HELQ and the indicated concentrations of RAD51 or RecA. **d**, **e**, Quantification of the experiments shown in **c** and Extended Data Fig. 2c, d for RAD51 (**d**) and RecA (**e**).  $n = 3$  (3' overhang),  $n = 4$  (Y-structure),  $n = 3$  (D-loop) and  $n = 3$  (lagging strand fork) independent experiments. Data are mean  $\pm$  s.e.m. **f**, Representative gel of the DNA unwinding assay of 3' overhang

Data Fig. 3a). Addition of RPA inhibited DNA unwinding by HELQ, especially for 3' overhang substrates (Fig. 1f, g and Extended Data Fig. 3b, c). At lower concentrations that were insufficient to cover the entire ssDNA region, RPA still inhibited HELQ unwinding of 3' overhang substrates (Extended Data Fig. 3d, e). Despite the inhibitory effect of RPA, RAD51 still stimulated HELQ helicase activity in the presence of RPA (Extended Data Fig. 3f–i).

### Visualization of HELQ DNA unwinding

To better understand HELQ stimulation by RAD51, we used an optical tweezer set-up combined with microfluidics and confocal microscopy (C-TRAP) for single-molecule imaging (SMI) analysis. As shown in Fig. 1h, a single dsDNA molecule ( $\lambda$ -DNA) containing a ssDNA gap<sup>9</sup> was tethered between two optically trapped beads and held at constant force (50 pN) to prevent the reannealing of unwound DNA. After addition of HELQ, DNA unwinding was observed as an increase in the distance between the beads due to the expansion of the ssDNA region. Neither RAD51 alone nor HELQ(K365M) showed evidence of unwinding (Fig. 1i and Extended Data Fig. 4a). Combining WT HELQ and RAD51 resulted in a considerable increase in overall DNA unwinding, whereas

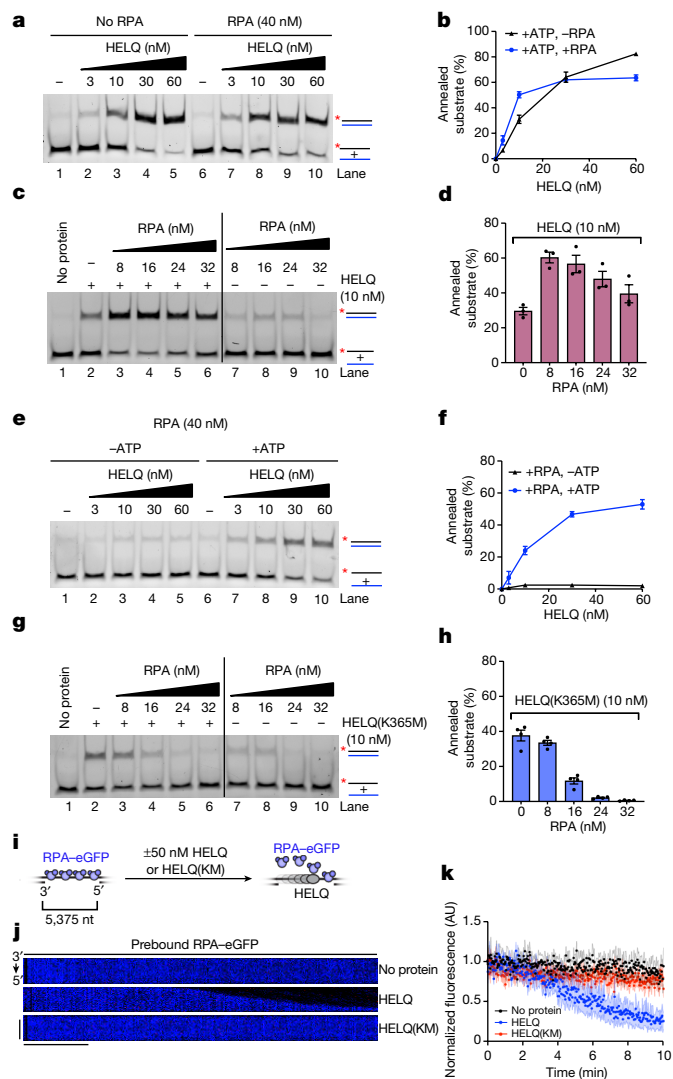
substrate with the indicated concentrations of HELQ in the absence and presence of RPA (20 nM). **g**, Quantification of the experiments shown in **f**.  $n = 3$  independent experiments. Data are mean  $\pm$  s.e.m. **h**, Schematics of the experimental set-up of the optical tweezer (C-Trap) system to observe DNA unwinding. These experiments were performed at room temperature. **i–k**, Bead centre displacement measured between the traps as a function of time with 25 nM RAD51 (**i**), 50 nM HELQ (**j**), and 25 nM RAD51 and 50 nM HELQ (**k**). The traces represent individual DNA molecules ( $n \geq 4$ ). **l**, Representative kymographs of single Alx–RAD51 binding events on gapped DNA in the presence or absence of 50 nM HELQ or HELQ(K365M). Unidirectional movement of Alx–RAD51 indicates translocation of Alx–RAD51–HELQ complex. Scale bars, 60 s (horizontal), 10  $\mu$ m (vertical, left), 5  $\mu$ m (vertical, right).

no such stimulation was observed with HELQ(K365M) (Fig. 1j, k and Extended Data Fig. 4b). Within unwinding traces for individual DNA molecules, rapid unwinding bursts interspersed by pauses can be distinguished (Extended Data Fig. 4c–e) and corresponded to a mean rate of  $3.3 \pm 0.4$  nm  $s^{-1}$  (mean  $\pm$  s.e.m. Extended Data Fig. 4f). In the presence of RAD51, an increased number of molecules showed greater unwinding rates (Extended Data Fig. 4g). To directly visualize RAD51 during DNA unwinding with HELQ, mutant RAD51(C319S) was purified and efficiently labelled with Alexa Fluor 488 C5 maleimide dye (Alx–RAD51) (Extended Data Fig. 4h). Whereas Alx–RAD51 alone displayed mostly static binding traces with occasional diffusing species, addition of HELQ showed unidirectional translocating traces indicating active movement of an Alx–RAD51–HELQ complex along the ssDNA backbone (Fig. 1l and Extended Data Fig. 4i). After analysis, we found that HELQ with RAD51 translocates fastest at the rate of  $14 \pm 5$  nm  $s^{-1}$  in gapped substrate (Extended Data Fig. 4j, k). By contrast, HELQ(K365M) retained the ability to bind to RAD51 but showed no translocation with only static or diffusing traces. Together, these results indicate that RAD51 and HELQ form a complex that unwinds DNA at a rate of approximately threefold faster compared with HELQ alone.

## DNA strand annealing by HELQ

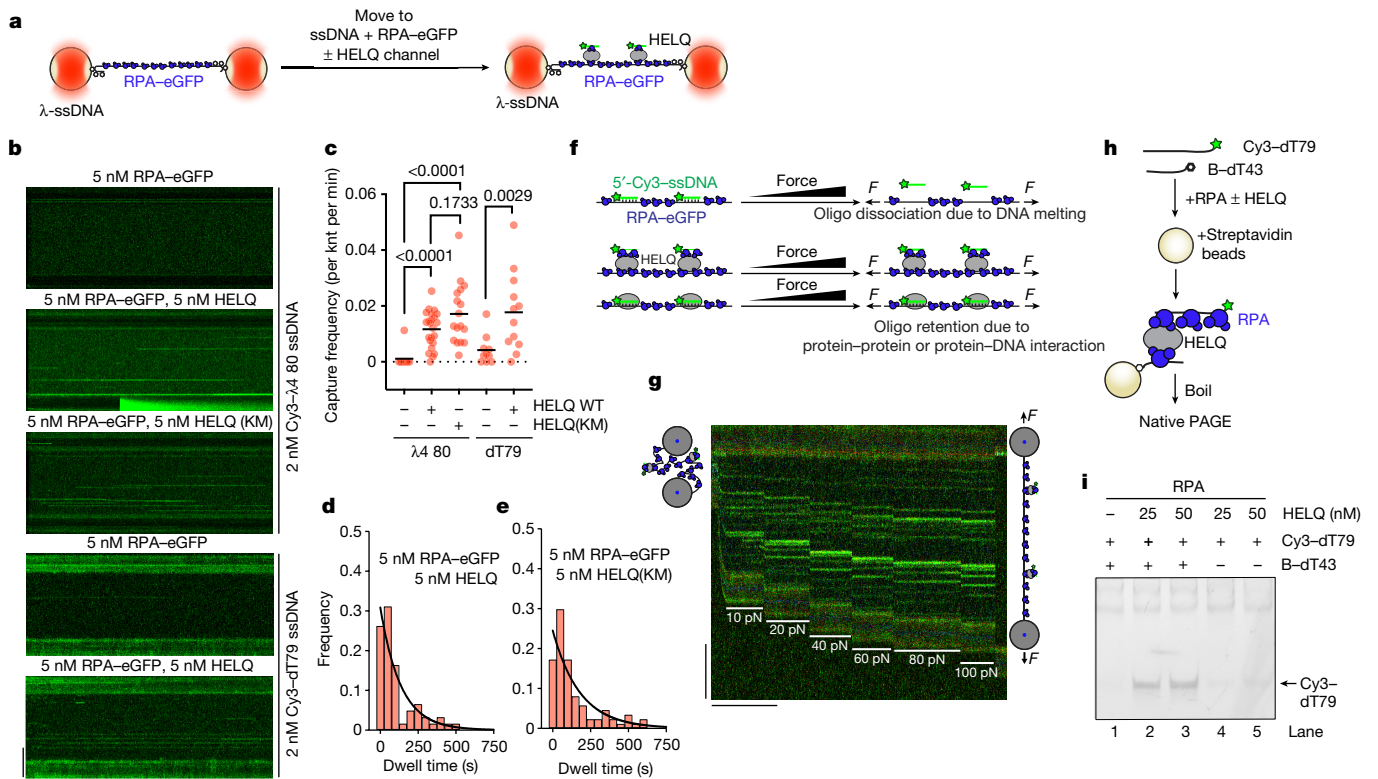
As shown above, a strong reduction in unwound product was observed at higher concentrations of HELQ (Fig. 1a, b and Extended Data Fig. 1b–e); we reasoned that this could result from the reannealing of the unwound product. Indeed, reactions containing an unlabelled ‘cold’ oligonucleotide (oligo) yielded an increase in unwound product with excess HELQ (Extended Data Fig. 5a (compare lanes 3 and 4 with 7 and 8)). Kinetic analysis also showed that HELQ initially unwinds the substrate but then reanneals it back together at later time points (Extended Data Fig. 5b). Prompted by this, we directly tested HELQ for DNA strand annealing activity without and with an excess of RPA needed for 100% ssDNA coverage (theoretically, 16 nM RPA covers 10 nM ssDNA). We found that HELQ efficiently anneals complementary DNA strands either without or with RPA (Fig. 2a, b). At lower concentrations of HELQ, RPA stimulated HELQ DNA annealing activity by around twofold. However, at higher concentrations, HELQ showed greater DNA annealing activity in the absence of RPA. This raised the possibility that RPA aids HELQ loading on ssDNA when HELQ is present in limiting amounts. Titration experiments confirmed that substoichiometric levels of RPA are sufficient to stimulate HELQ annealing activity (Fig. 2c, d). Notably, HELQ could still anneal complementary DNA strands in the presence of excess RPA (Extended Data Fig. 5c, d).

We next tested the requirement of ATP binding and hydrolysis for DNA annealing by HELQ. Surprisingly, in the presence of RPA, HELQ showed no DNA annealing without ATP, whereas ATP became dispensable when RPA was excluded from the reaction (Fig. 2e, f and Extended Data Fig. 5e, f). Even in the absence of RPA, ATP stimulated the DNA annealing activity of HELQ (Extended Data Fig. 5e, f). HELQ also failed completely to anneal DNA with ATPγS in the presence of RPA (Extended Data Fig. 5g). Collectively, these data suggest that HELQ possesses intrinsic DNA annealing activity that requires ATP binding and hydrolysis when ssDNA is coated with RPA. We next tested the helicase-inactive HELQ(K365M) mutant for DNA annealing activity and found that HELQ(K365M) is defective for DNA annealing in the presence of excess RPA but could anneal ssDNAs when RPA is excluded from reactions (Extended Data Fig. 5h, i). A titration experiment showed that, in contrast to the WT, HELQ(K365M) becomes progressively impaired by increasing concentrations of RPA (Fig. 2g, h and Extended Data Fig. 5j, k). RPA also failed to stimulate HELQ(K365M) (Fig. 2g, h). We also tested *E. coli* SSB protein and found that it only weakly stimulates HELQ annealing activity (Extended Data Fig. 5l, m). The N-terminal fragment of HELQ was previously shown to displace RPA from ssDNA<sup>6</sup>. However, full-length HELQ was not analysed for such activity. Thus, we directly tested RPA displacement from ssDNA during DNA strand annealing by omitting the deproteination step. We observed that HELQ can strip an excess of RPA from ssDNA, which occurred coincidentally with the appearance of the annealed products (Extended Data Fig. 6a, b). To directly visualize RPA stripping from ssDNA by HELQ, we measured RPA–eGFP displacement by SMI analysis and found that WT HELQ ( $k = 0.136 \pm 0.008 \text{ min}^{-1}$ ) could efficiently strip RPA from ssDNA but HELQ(K365M) ( $k = 0.017 \pm 0.004 \text{ min}^{-1}$ ) could not (Fig. 2i–k). Using a single-molecule Förster resonance energy transfer (FRET)-based assay (Extended Data Fig. 6c–j), we observed concentration-dependent RPA stripping by WT HELQ, followed by rebinding of RPA (Extended Data Fig. 6k–n). The RPA rebinding is independent of HELQ concentration, indicating a constant transition rate from the free ( $t_{\text{off}}$ ) to bound ( $t_{\text{on}}$ ) state at various HELQ concentrations (Extended Data Fig. 6o). HELQ(K365M) did not show any RPA stripping (Extended Data Fig. 6p), indicating that active RPA stripping has a critical role in HELQ-mediated DNA annealing. Finally, we found that the addition of RAD51 had no effect on HELQ-dependent DNA annealing activity (Extended Data Fig. 6q, r). To study whether, like RPA, HELQ can also strip RAD51 from ssDNA, we directly tested RAD51 removal



**Fig. 2 | HELQ strips RPA from and anneals ssDNA.** **a**, Representative gel of DNA annealing assay with the indicated concentrations of HELQ in the absence and presence of RPA (40 nM). The black and blue colours of substrate represent complementary DNA strands. The asterisk indicates the position of FITC labelling at the 5' end of the oligo. The products were resolved with 10% native polyacrylamide gel. **b**, Quantification of experiments such as shown in **a**.  $n = 6$  independent experiments. Data are mean  $\pm$  s.e.m. **c**, Representative gel of the DNA annealing assay with HELQ (10 nM) and the indicated concentrations of RPA. **d**, Quantification of the experiments shown in **c**.  $n = 3$  independent experiments. Data are mean  $\pm$  s.e.m. **e**, Representative gel of the DNA annealing assay with the indicated concentrations of HELQ and RPA (40 nM), in the absence and presence of ATP. **f**, Quantification of the experiments shown in **e**.  $n = 3$  independent experiments. Data are mean  $\pm$  s.e.m. **g**, Representative gel of DNA annealing assay with HELQ(K365M) (10 nM) and the indicated concentrations of RPA. **h**, Quantification of the experiments shown in **g**.  $n = 4$  independent experiments. Data are mean  $\pm$  s.e.m. **i**, Schematics of the experimental set-up of the optical tweezer (C-Trap) system to observe RPA–eGFP stripping from ssDNA. These experiments were performed at room temperature. **j**, Representative kymographs of stripping of RPA–eGFP prebound to gapped DNA in the presence or absence of 50 nM HELQ or HELQ(K365M). Unidirectional stripping of RPA–eGFP with HELQ WT indicates 3' to 5' translocation of HELQ. **k**, Removal of RPA–eGFP measured from gapped ssDNA as function of time in indicated conditions. The traces represent individual DNA molecules ( $n \geq 3$ ). Scale bars, 2 min (horizontal), 2  $\mu\text{m}$  (vertical).

from ssDNA and dsDNA. HELQ did not remove RAD51 from either ssDNA or dsDNA in bulk assays (Extended Data Fig. 7a–d) and only weakly displaced RAD51 from ssDNA in our SMI setup (Extended Data



**Fig. 3 | HELQ exhibits sequence-independent RPA-coated DNA capture activity.** **a**, Schematics of the experimental optical tweezer set-up for observing the capture of Cy3-labelled DNA oligos. **b**, Kymographs showing the capture of targeted Cy3-labelled 80-mer oligo ( $\lambda$ 4) in *trans* with HELQ and HELQ(K365M) in the presence of RPA-eGFP at multiple positions of RPA-eGFP-coated  $\lambda$ -ssDNA (top). KM, HELQ(K365M). Kymographs showing the capture of non-targeted Cy3-labelled dT79 homopolymer in *trans* with HELQ in the presence of RPA-eGFP at multiple positions of RPA-eGFP-coated  $\lambda$ -ssDNA (bottom). Scale bars, 60 s (horizontal), 10  $\mu$ m (vertical). **c**, Quantification of the experiments shown in **b**. Each datapoint represents a single DNA molecule. Data are mean  $\pm$  s.d. Statistical analysis was performed using two-sided Mann-Whitney *U*-tests. No adjustments were made for multiple comparisons. **d**, The dwell times of

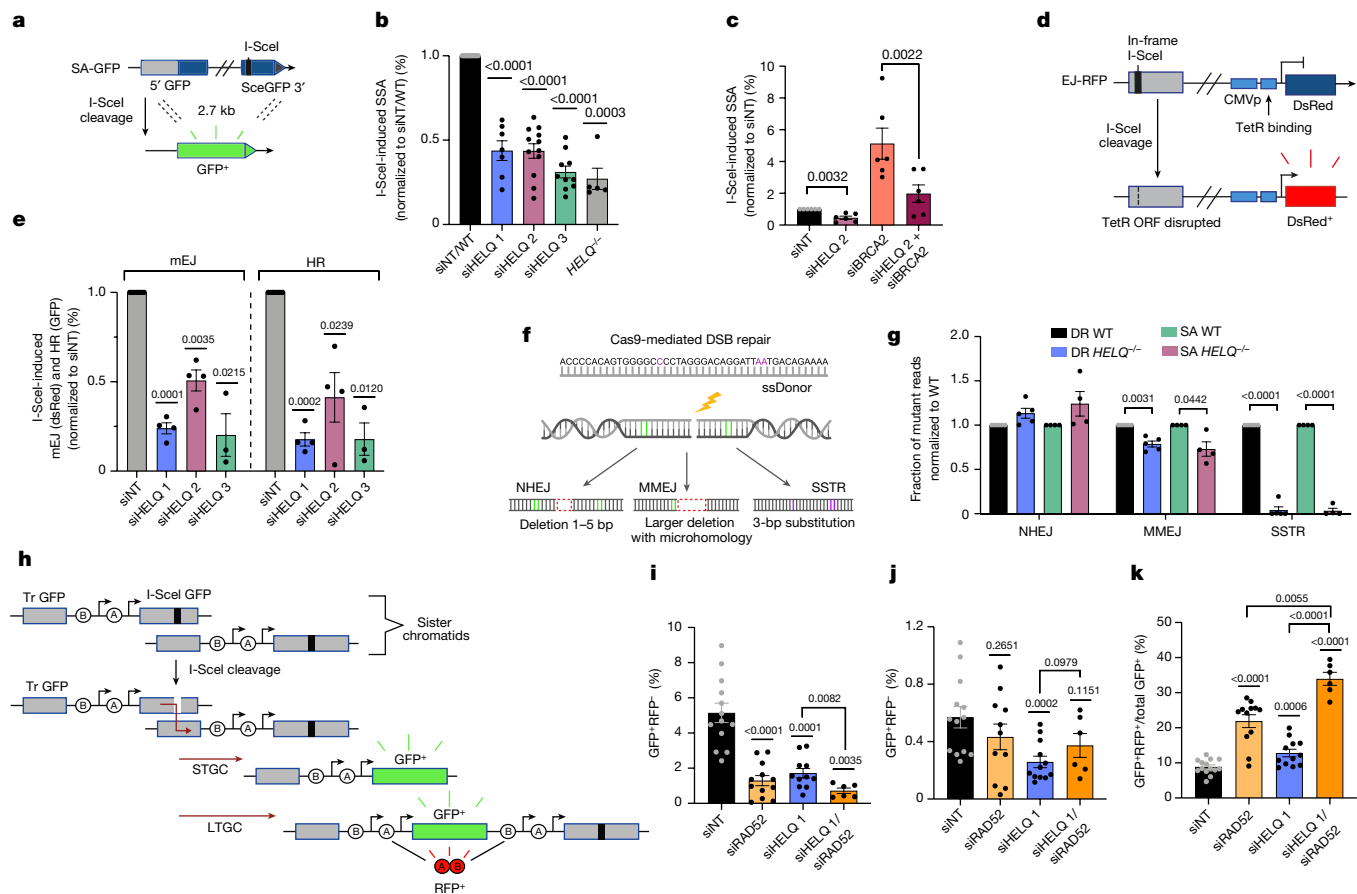
captured  $\lambda$ 4 by HELQ in the presence of RPA-eGFP.  $n = 61$ . The black line represents the exponential fit.  $\tau = 134$  s. **e**, The dwell times of captured  $\lambda$ 4 by HELQ(K365M) in the presence of RPA-eGFP.  $n = 87$ . The black line represents the exponential fit.  $\tau = 179$  s. **f**, The experimental conditions for the optical tweezer experiment in **g**. **g**, Kymograph showing the Cy3- $\lambda$ 4 oligo captured on RPA-eGFP-coated  $\lambda$ -ssDNA in the presence of HELQ and RPA-eGFP after stretching of tethered  $\lambda$ -ssDNA by a gradual increase of force. Scale bars, 60 s (horizontal), 10  $\mu$ m (vertical). **h**, Schematic of the bulk capture assay. **i**, Native gel showing the capture assay with the indicated concentrations of HELQ and RPA (82 nM). B indicates biotin at the 3' end of B-dT43 oligo. The experiment was performed twice with similar results.

Fig. 7e–g). This suggests that HELQ is unlikely to disrupt fully formed RAD51 nucleofilaments but might remove RAD51 that is bound to short-resected DNA, as is present during microhomology-mediated end joining (MMEJ).

### HELQ captures RPA-ssDNA

RAD52 has been shown to possess robust DNA strand annealing activity and has a central role in single-strand annealing (SSA) repair of DSBs<sup>10–14</sup>. To investigate the mechanistic basis of the strand annealing ability of HELQ, we modelled our experiments on RAD52 annealing activity. Using optical tweezers, it was previously shown that RAD52 can *trans*-capture labelled oligos at multiple sites along  $\lambda$ -DNA, independent of DNA sequence<sup>15</sup>. Using a similar set-up, we tested the ability of HELQ to capture a Cy3-labelled 80-mer oligo ssDNA in the presence of RPA-eGFP (around 100% coverage; Fig. 3a and Extended Data Fig. 8a). HELQ facilitated the capture of  $\lambda$ 4 oligo<sup>16,17</sup> at multiple sites (Fig. 3b, c). Notably, in contrast to its annealing or unwinding activity, HELQ(K365M) exhibited efficient DNA capture activity (Fig. 3b, c). We also analysed the dwell times of captured oligos and found that HELQ(K365M) showed moderately increased dwell times compared with the WT (Fig. 3d, e), most likely due to slightly higher DNA binding of HELQ(K365M). We also tested a 79-nucleotide T-homopolymer (dT79) and obtained similar capture frequencies as with the  $\lambda$ 4 oligo (Fig. 3c).

HELQ showed capture of both homologous ( $\lambda$ 4) and T-homopolymer oligos at various sites on  $\lambda$ -DNA. To determine whether microhomologies present in ~50 kb  $\lambda$ -DNA could explain the capture pattern of both oligos, we performed forced-stretching experiments as shown in Fig. 3f. As DNA starts to melt at forces of >60 pN, we reasoned that, if HELQ oligo capture involves base-pairing interactions, short microhomologies should dissociate faster than the ones with longer homology. However, even at very high force (90–100 pN), all oligos remained engaged with ssDNA, irrespective of position (Fig. 3g). We next performed pulling experiments<sup>18</sup> and observed characteristic force spikes when beads were pulled apart at a low force (10–15 pN) (Extended Data Fig. 8b, c). These spikes correspond to the disruption of HELQ complexes capturing RPA-coated ssDNA *in cis*. To examine this further, we developed a bulk capture assay, in which we attempted to pull out labelled non-complementary DNA (Cy3-dT79) with a biotinylated dT43 oligo with HELQ. We found that both WT HELQ and the HELQ(K365M) mutant could capture non-complementary oligos bound to RPA (Fig. 3h, i and Extended Data Fig. 8d). Thus, HELQ and HELQ(K365M) can both capture DNA strands independent of sequence, probably through DNA tethering, but only the WT can anneal RPA-coated complementary DNA strands through RPA stripping. Interestingly, yeast Rad52, when bound to RPA-coated ssDNA clusters, can capture additional free



**Fig. 4 | HELQ functions in DSB repair, SSA and MMEJ.** **a**, Schematic of the SA-GFP reporter assay. **b**, I-SceI-induced SSA frequency in U2OS-SA *HELQ*<sup>-/-</sup> cells and WT cells treated with the indicated short interfering RNA (siRNA). *n* = 17 (non-targeting siRNA (siNT)/WT), *n* = 7 (siHELQ1), *n* = 12 (siHELQ2), *n* = 10 (siHELQ3) and *n* = 5 (*HELQ*<sup>-/-</sup>) independent experiments. Data are mean ± s.e.m. **c**, SSA frequency in U2OS-SA cells treated with the indicated siRNA. *n* = 6 independent experiments. Data are mean ± s.e.m. **d**, Schematic of the EJ-RFP reporter assay. **e**, I-SceI-induced mutagenic end-joining (mEJ-RFP) and HR (DR-GFP) frequencies in U2OS-EJDR cells treated with the indicated siRNA. *n* = 7 (siNT), *n* = 4 (siHELQ1), *n* = 2 (siHELQ2) and *n* = 3 (siHELQ3) independent experiments. Data are mean ± s.e.m. **f**, Schematic of Cas9-mediated DSB repair assay. **g**, Cas9-mediated DSB repair assay showing the frequencies of NHEJ, MMEJ and SSTR repair in *HELQ*<sup>-/-</sup> and WT U2OS-DR or U2OS-SA cells. *n* = 5

(U2OS-DR WT and DR*HELQ*<sup>-/-</sup>) and *n* = 4 (U2OS-SA and SA*HELQ*<sup>-/-</sup>) independent experiments. Data are mean ± s.e.m. **h**, Schematic of the RFP-SCR reporter assay. Tr GFP, truncated GFP. **i**, I-SceI-induced STGC frequency in U2OS-RFP-SCR cells treated with the indicated siRNA. *n* = 13 (siNT) *n* = 11 (siHELQ1), *n* = 12 (siRAD52) and *n* = 6 (siHELQ1/siRAD52) independent experiments. Data are mean ± s.e.m. **j**, LTGC frequency as shown in **i**. *n* = 13 (siNT), *n* = 12 (siHELQ1), *n* = 11 (siRAD52) and *n* = 6 (siHELQ1/siRAD52) independent experiments. Data are mean ± s.e.m. **k**, The ratio of LTGC/total GC from the experiments in **i** and **j**. *n* = 14 (siNT), *n* = 12 (siHELQ1), *n* = 12 (siRAD52) and *n* = 6 (siHELQ1/siRAD52) independent experiments. Data are mean ± s.e.m. For **b**, **c**, **e**, **g**, **i**–**k**, statistical analysis was performed using two-tailed paired *t*-tests as compared with siNT, unless otherwise indicated.

RPA in pre-existing Rad52–RPA–ssDNA clusters<sup>19</sup>. This activity was postulated to be important for second-end capture during homologous recombination (HR).

### HELQ functions in SSA and MMEJ

To extend our findings with HELQ to DSB repair in vivo, we first confirmed that deletion of *HELQ* or protein depletion in cells inhibits HR (Extended Data Fig. 9a–e). As DNA annealing is required for SSA repair, we investigated a potential role for HELQ in this process. Strikingly, HELQ depletion or *HELQ* deletion impaired SSA repair of an integrated SSA reporter (SA-GFP; Fig. 4a, b). Although depletion of the HR factor, BRCA2, increases SSA repair, this was strongly reduced by HELQ depletion (Fig. 4c and Extended Data Fig. 9f). Consistent with an epistatic role in SSA, co-depletion of RAD52 and HELQ did not further decrease SSA repair compared with individual depletions (Extended Data Fig. 9g). We also assessed whether HELQ functions in alternative end joining repair, which also involves an annealing step. Using cells containing both EJ-RFP and DR-GFP reporter systems for simultaneous detection of

alternative end joining and HR, respectively, we observed a significant reduction in both DSB repair pathways after HELQ depletion (Fig. 4d, e). Alternative end joining encompasses broad non-homologous end joining (NHEJ) repair events including MMEJ. To specifically study the role of HELQ in MMEJ, we used a Cas9-mediated DSB repair assay in which the fate of DSB repair by MMEJ, NHEJ or single-strand templated repair (SSTR) can be determined<sup>20</sup> (Fig. 4f). The loss of *HELQ* resulted in a slight increase in NHEJ, whereas MMEJ was significantly impaired and SSTR was completely abolished (Fig. 4g). Interestingly, DNA annealing is important for SSTR<sup>21,22</sup>. A role for HELQ in DNA end resection could potentially explain defects in HR, SSA and MMEJ. However, we did not observe any reduction in RPA or RAD51 (refs. 3,23,24) focus formation after irradiation or camptothecin treatment in *HELQ*<sup>-/-</sup> or depleted cells, excluding a role for HELQ in resection (Extended Data Fig. 9h–i). Impaired DNA strand annealing during second-end capture in DSB repair or failure to capture the repaired strand in synthesis-dependent strand annealing can result in a shift towards long-tract gene conversion (LTGC)<sup>25–29</sup>. Using the same reporter system, we found that, similar to RAD52, HELQ deficiency results in a strong and

moderate reduction in short-tract gene conversion (STGC) and LTGC, respectively, with the LTGC/total gene conversion (GC) ratio showing a significant bias towards LTGC (Fig. 4h–k). Finally, co-depletion of HELQ with RAD52 showed a further bias towards LTGC, implying that they have roles in parallel pathways during GC (Fig. 4k).

In summary, our study implicates HELQ in several distinct DSB repair pathways, including HR, SSA and MMEJ, casting light on its role in genome stability and tumour avoidance. As these repair pathways each require DNA annealing steps, we propose that HELQ functions in these pathways through its intrinsic ability to capture RPA-bound ssDNA and then displace RPA to facilitate annealing of complementary DNA strands. HELQ alone or in complex with RAD51 might unwind D-loops before annealing complementary strands (Extended Data Fig. 10). The bias towards LTGC events after HELQ depletion is consistent with a role for DNA annealing by HELQ during second-strand capture and/or synthesis-dependent strand annealing during HR, which may explain why HR is reduced in HELQ-deficient cells. Finally, our finding that HELQ is epistatic with RAD52 for SSA but additive for GC during HR is surprising, as this implicates two DNA strand annealing enzymes in several distinct DSB repair processes in cells.

## Online content

Any methods, additional references, Nature Research 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/s41586-021-04261-0>.

1. Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.* **27**, 247–254 (2001).
2. Marini, F. & Wood, R. D. A human DNA helicase homologous to the DNA cross-link sensitivity protein Mus308. *J. Biol. Chem.* **277**, 8716–8723 (2002).
3. Adelman, C. A. et al. HELQ promotes RAD51 paralogue-dependent repair to avert germ cell loss and tumorigenesis. *Nature* **502**, 381–384 (2013).
4. Luebben, S. W. et al. *Helq* acts in parallel to *Fancc* to suppress replication-associated genome instability. *Nucleic Acids Res.* **41**, 10283–10297 (2013).
5. Takata, K., Reh, S., Tomida, J., Person, M. D. & Wood, R. D. Human DNA helicase HELQ participates in DNA interstrand crosslink tolerance with ATR and RAD51 paralogs. *Nat. Commun.* **4**, 2338 (2013).
6. Jenkins, T. et al. The HelQ human DNA repair helicase utilizes a PWI-like domain for DNA loading through interaction with RPA, triggering DNA unwinding by the HelQ helicase core. *NAR Cancer* **3**, zcaa043 (2021).
7. Tafel, A. A., Wu, L. & McHugh, P. J. Human HEL308 localizes to damaged replication forks and unwinds lagging strand structures. *J. Biol. Chem.* **286**, 15832–15840 (2011).
8. Ward, J. D. et al. Overlapping mechanisms promote postsynaptic RAD-51 filament disassembly during meiotic double-strand break repair. *Mol. Cell* **37**, 259–272 (2010).
9. Belan, O. et al. Single-molecule analysis reveals cooperative stimulation of Rad51 filament nucleation and growth by mediator proteins. *Mol. Cell* **81**, 1058–1073 (2021).
10. 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).

11. Kagawa, W., Kurumizaka, H., Ikawa, S., Yokoyama, S. & Shibata, T. Homologous pairing promoted by the human Rad52 protein. *J. Biol. Chem.* **276**, 35201–35208 (2001).
12. Ivanov, E. L., Sugawara, N., Fishman-Lobell, J. & Haber, J. E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* **142**, 693–704 (1996).
13. 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).
14. Reddy, G., Golub, E. I. & Radding, C. M. Human Rad52 protein promotes single-strand DNA annealing followed by branch migration. *Mutat. Res.* **377**, 53–59 (1997).
15. Brouwer, I. et al. Human RAD52 captures and holds DNA strands, increases DNA flexibility, and prevents melting of duplex DNA: implications for DNA recombination. *Cell Rep.* **18**, 2845–2853 (2017).
16. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
17. Newton, M. D. et al. DNA stretching induces Cas9 off-target activity. *Nat. Struct. Mol. Biol.* **26**, 185–192 (2019).
18. Barroso-González, J. et al. RAD51API is an essential mediator of alternative lengthening of telomeres. *Mol. Cell* **76**, 11–26 (2019).
19. Gibb, B. et al. Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules. *Nat. Struct. Mol. Biol.* **21**, 893–900 (2014).
20. Hussain, S. S. et al. Measuring nonhomologous end-joining, homologous recombination and alternative end-joining simultaneously at an endogenous locus in any transfectable human cell. *Nucleic Acids Res.* **49**, e74 (2021).
21. Gallagher, D. N. et al. A Rad51-independent pathway promotes single-strand template repair in gene editing. *PLoS Genet.* **16**, e1008689 (2020).
22. Gallagher, D. N. & Haber, J. E. Single-strand template repair: key insights to increase the efficiency of gene editing. *Curr. Genet.* **67**, 747–753 (2021).
23. Moldovan, G. L. et al. DNA polymerase POLN participates in cross-link repair and homologous recombination. *Mol. Cell Biol.* **30**, 1088–1096 (2010).
24. Takata, K. et al. Conserved overlapping gene arrangement, restricted expression, and biochemical activities of DNA polymerase  $\nu$  (POLN). *J. Biol. Chem.* **290**, 24278–24293 (2015).
25. Nagaraju, G., Odate, S., Xie, A. & Scully, R. Differential regulation of short- and long-tract gene conversion between sister chromatids by Rad51C. *Mol. Cell Biol.* **26**, 8075–8086 (2006).
26. Nagaraju, G., Hartlerode, A., Kwok, A., Chandramouly, G. & Scully, R. XRCC2 and XRCC3 regulate the balance between short- and long-tract gene conversions between sister chromatids. *Mol. Cell Biol.* **29**, 4283–4294 (2009).
27. Mcllwraith, M. J. & West, S. C. DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair. *Mol. Cell* **29**, 510–516 (2008).
28. Lao, J. P., Oh, S. D., Shinohara, M., Shinohara, A. & Hunter, N. Rad52 promotes postinvasion steps of meiotic double-strand-break repair. *Mol. Cell* **29**, 517–524 (2008).
29. Mehta, A., Beach, A. & Haber, J. E. Homology requirements and competition between gene conversion and break-induced replication during double-strand break repair. *Mol. Cell* **65**, 515–526 (2017).

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



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

© The Author(s) 2021

## Methods

### Preparation of expression vectors

Custom HELQ ORF was purchased from GeneArt and used as a template during PCR to prepare plasmid (pFastbac1) compatible for bacmid preparation for expression in insect cells. To prepare the MBP-HELQ-Flag construct, HELQ was amplified by PCR using the primers HELQ\_F and HELQ\_FLAG\_R. The amplified insert was digested with BamHI and XmaI and inserted into the pFastbac1 vector containing MBP (maltose binding protein tag; previously inserted using the EcoRV and HindIII restriction sites). The resulting construct was pFB-MBP-HELQ-FLAG. To prepare helicase-dead HELQ(K365M), pFB-MBP-HELQ-Flag was mutagenized with the primers HELQ\_K365M\_F and HELQ\_K365M\_R using the Q5 site-directed mutagenesis kit according to the manufacturer's instructions. The constructs for pET11c-RAD51 and, RPA-eGFP (pMM801) and RPA-mRFP1 (pMM802) were gifts from L. Krejci and M. Modesti, respectively.

### Recombinant protein purification

To express proteins in insect cells, bacmid, primary and secondary baculoviruses were prepared according to the manufacturer's instructions (Bac-to-bac system, Life technologies). To express recombinant MBP-HELQ-FLAG, *Spodoptera frugiperda* (Sf9) insect cells were seeded at 500,000 cells per ml and, after around 24 h, cells were infected with MBP-HELQ-Flag baculovirus. The infected cells were incubated at 27 °C for 56 h with continuous agitation. Cells were collected by centrifugation at 500g for 10 min and washed once with 1× PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>). The collected pellets were snap-frozen in liquid nitrogen and stored at -80 °C until further use. All subsequent steps were performed either on ice or at 4 °C. The cells pellets were resuspended in 3 volumes of lysis buffer containing 50 mM Tris-HCl pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA), protease inhibitor cocktail tablets (Roche), 30 µg ml<sup>-1</sup> leupeptin (Merck), 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM dithiothreitol (DTT), 0.1% NP-40 substitute (NP-40) and incubated for 15 min with continuous agitation. Next, 50% glycerol and 5 M NaCl were added sequentially to the final concentrations of 16.7% and 310 mM, respectively, and the suspension was further incubated for 30 min with continuous agitation. The suspension was centrifuged at -48,000g for 30 min to obtain the soluble extract. The amylose resin (NEB) was pre-equilibrated with amylose wash buffer I (50 mM Tris-HCl pH 7.5, 1 mM 2-mercaptoethanol (β-ME), 1 M NaCl, 1 mM PMSF, 10% glycerol and 0.1% NP-40) and added to 50 ml tubes containing soluble extract. These tubes were subsequently incubated for 1 h with continuous rotation. After incubation, the resin was washed batchwise four times by centrifugation at 2,000g for two min and twice on column with amylose wash buffer I. Resin was washed twice more on column with amylose wash buffer II (same as wash buffer I but with 0.5 mM β-ME and 0.8 M NaCl). Protein was eluted from the resin using amylose elution buffer (same as amylose wash buffer II supplemented with 10 mM maltose) and total protein was estimated using the Bradford assay. To remove the MBP tag, 1/8 (w/w) of PreScission protease to the total protein was added to amylose eluate and incubated for 2 h without rotation but with gentle agitation at regular intervals. The Flag resin (anti-Flag M2 resin, Sigma-Aldrich), which was pre-equilibrated with Flag wash buffer (50 mM Tris-HCl pH 7.5, 0.8 M NaCl, 1 mM PMSF, 10% glycerol), was added to amylose eluate containing PreScission protease and incubated for 2 h with continuous rotation. Flag resin was collected directly on column and washed six times with Flag wash buffer. The protein was eluted from Flag resin with Flag elution buffer (50 mM Tris-HCl pH 7.5, 0.5 mM β-ME, 150 mM NaCl, 1 mM PMSF, 10% glycerol, 150 ng µl<sup>-1</sup> 3×Flag peptide (Sigma-Aldrich)), aliquoted, frozen in liquid nitrogen and stored at -80 °C. The same purification procedure was used to purify HELQ(K365M).

Recombinant human RAD51 was purified as described previously with a few modifications<sup>30</sup>. The pET11c-RAD51 expression vector was transformed into *E. coli* BLR(DE3)pLysS cells and subsequent culture containing ampicillin (100 mg l<sup>-1</sup>) and chloramphenicol (33 mg l<sup>-1</sup>) was grown to an optical density (OD) at 600 nm of 0.7. RAD51 expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37 °C for 3–4 h. All of the subsequent steps were performed either on ice or at 4 °C. Cells were collected by centrifugation at 5,000g. Cell pellets were resuspended in cell breakage buffer (50 mM Tris-HCl pH 7.5, 10% sucrose, 0.5 mM EDTA, 1 M KCl, 1 tablet per 50 ml of protease inhibitor cocktail tablets (Roche), 1 mM PMSF, 1 mM DTT and 0.01% NP-40), sonicated and centrifuged at 100,000g for 1 h. To precipitate RAD51, 0.242 g ml<sup>-1</sup> ammonium sulphate was mixed with clarified supernatant and centrifuged for 20 min at 10,000g. The pellet was resuspended with buffer K (20 mM K<sub>2</sub>HPO<sub>4</sub> pH 7.5, 10% glycerol, 0.5 mM EDTA, 1 mM DTT, 0.01% NP-40) and loaded onto the Q Sepharose Fast flow column (Cytiva), pre-equilibrated with K buffer-low (K buffer supplemented with 175 mM KCl). The column was washed extensively with K buffer-low and protein was subsequently eluted with a KCl gradient using K buffer-high (K buffer supplemented with 0.6 M KCl). The eluted fractions containing RAD51 were pooled and diluted with 6 volumes of dilution buffer (25 mM Tris HCl pH 7.5, 0.5 mM EDTA, 1 mM DTT, 0.01% NP-40). The diluted sample was loaded onto the HiTrap Heparin HP affinity column (Cytiva), which was pre-equilibrated with buffer H without glycerol (25 mM Tris HCl pH 7.5, 0.5 mM EDTA, 1 mM DTT, 0.01% NP-40, 150 mM KCl) and washed with buffer H containing 10% glycerol. Protein was eluted in buffer H with a KCl gradient (0.1 M to 1 M KCl) and the fractions containing RAD51 were pooled and dialysed in buffer H without glycerol. The dialysed sample was loaded onto the Mono Q5/50 GL column (Cytiva), equilibrated with buffer Q (25 mM Tris HCl pH 7.5, 0.5 mM EDTA, 1 mM DTT, 0.01% NP-40, 100 mM KCl, 10% glycerol) and the column was further washed with buffer Q containing 50 mM KCl but lacking glycerol. RAD51 was eluted from the Mono Q column with a KCl gradient (0.05 M to 1 M) in buffer Q lacking glycerol. The eluted fractions containing RAD51 were pooled and further concentrated with the Vivaspin Centrifugal Concentrator (30 kDa molecular weight cut-off (MWCO)). Glycerol was added to the concentrated sample to a final concentration of 10%. Finally, the samples were aliquoted, frozen in liquid nitrogen and stored at -80 °C. The RAD51(C319S) mutant was purified using same procedure. RPA-mRFP1 and RPA-eGFP were purified as described previously<sup>31</sup>. Recombinant RecA (M0249) and ET SSB (M2401) were commercially purchased from NEB, England.

To purify GST and GST-BRC4, the BRC4 construct was cloned into the pGEX6P-1 vector containing a GST tag using BamHI and EcoRI restriction sites. The pGEX6P-1 and pGEX6P-1-BRC4 constructs were transformed into BL21 DE3 cells and plated onto an agar plate containing ampicillin (100 µg ml<sup>-1</sup>). A single colony was isolated and inoculated into 6 ml preculture overnight. The next day, 4.5 ml preculture medium was added to 450 ml LB medium containing ampicillin (100 µg ml<sup>-1</sup>), and the OD at 600nm was monitored at regular intervals. Proteins were induced with 1 mM IPTG at 0.6 OD and cultures were incubated for 4 h. The cell pellets were collected, washed with cold PBS and stored at -80 °C. For purification, the cell pellets were lysed by sonication in PBS and the samples were centrifuged at 75,600g for 30 min at 4 °C. Next, the supernatant was collected and incubated with 1.2 ml glutathione resin for 1 h at 4 °C. The resin was washed three times with PBS and the proteins were eluted with elution buffer (20 mM Tris-HCl pH 7.5, 20 mM glutathione). The eluted proteins were aliquoted, snap-frozen in liquid nitrogen and stored at -80 °C.

### Preparation of labelled proteins

The RAD51(C319S) variant was expressed and purified as described earlier for WT RAD51 (refs. <sup>30,32</sup>). After purification, the protein was fluorescently labelled with Alexa Fluor 488 C<sub>5</sub> maleimide (Thermo Fisher Scientific, A10254) according to previously described protocol<sup>31</sup>.







To estimate the average velocity of the translocating molecule, the total route of the molecule (a sum of frame-to-frame displacements) was divided by the total trajectory time. Here, every trajectory was smoothed using the Savitzky–Golay filter (smoothing factor = 51) to eliminate tracking inaccuracies and the molecule's thermal fluctuations.

To estimate the loop size formed by the HELQ–RPA–DNA complex, the contour length after each unfolding event in the force–distance curve was fitted by the worm-like chain model. The difference between contour lengths of the neighbouring events corresponds with the loop size.

### RPA stripping using single-molecule FRET

Flow chambers were prepared as described previously<sup>1,2</sup>. Quartz slides and coverslips were passivated with polyethylene glycol (5% biotinylated) and flow chambers were constructed using double-sided sticky tape and sealed with epoxy. 5'-biotin- and internal amino linker-modified DNA oligonucleotides were labelled with Cy3-NHS ester and HPLC purified as previously described<sup>3</sup>. DNA (6 pM) was immobilized through biotin–streptavidin interactions. All of the experiments were performed in the standard HELQ buffer with addition of the PCA/PCD oxygen scavenger system with 5 mM PCA, 100 nM PCD and saturating Trolox. The flow chambers were imaged on a home-built, prism-based total internal reflection microscope with a 532 nm excitation laser (~3.8 mW), and images were acquired on an EM-CCD camera (Andor) with a 30 ms exposure time. FRET efficiencies were calculated from integrated donor ( $I_D$ ) and acceptor ( $I_A$ ) intensities as  $FRET = I_A / (I_D + I_A)$  (refs.<sup>1,3</sup>). The images and data were analysed using custom IDL, MATLAB and R scripts, which are available on request. FRET efficiency histograms were constructed by averaging the first 10 frames of each trajectory, with bins of 0.1. The dwell times of the free (high FRET) and bound (low FRET) states were measured, and dwell-time histograms were plotted. These were fit with single exponential fits to obtain average dwell times.

### Cell culture

The U2OS human osteosarcoma cell line was grown in DMEM supplemented with 10% bovine growth serum, 2 mM L-glutamine, 100  $\mu$ g ml<sup>-1</sup> streptomycin and 100 U ml<sup>-1</sup> penicillin. U2OS-EJDR cells were cultured in DMEM supplemented with 10% tetracycline-free fetal bovine serum, 2 mM L-glutamine, 100  $\mu$ g ml<sup>-1</sup> streptomycin and 100 U ml<sup>-1</sup> penicillin. U2OS-DR cells contain a stably integrated DR-GFP reporter to measure DSB repair by HR as previously described<sup>35</sup>. U2OS-SA cells contain a stably integrated SA-GFP reporter to measure DSB repair by SSA as previously described<sup>36</sup>. U2OS-DR cells containing a stably integrated EJ-RFP reporter to measure DSB repair by mutagenic end-joining constitute the U2OS-EJDR cell line as previously described<sup>37</sup>. U2OS-RFP-SCR cells contain a stably integrated RFP-SCR reporter for quantifying STGC and LTGC in HR as previously described<sup>38</sup>.

U2OS-DR *HELQ*<sup>-/-</sup> and U2OS-SA *HELQ*<sup>-/-</sup> cells were generated using the CRISPR–Cas9 system. Knockouts were verified using Sanger sequencing and immunoprecipitation/western blot.

### siRNA

The following siRNA oligonucleotides were used to transiently deplete *HELQ*: HELQ 1 (Qiagen FlexiTube siRNA, S100435449); HELQ 2 (Horizon siGENOME SMARTpool, M-015379-01-0005); HELQ 3 (HelQ\_M<sup>23</sup>; CAAAGGAAGATTTCCTCCAATAA).

*RAD52* was depleted using the On-Target plus SMART pool siRNA L-011760-00-0005 (Horizon). *BRCA2* was depleted using the On-Target plus SMART pool siRNA L-003462-00-0005 (Horizon). The On-Target plus non-targeting siRNA pool was used for non-targeting controls (D-001810-10-05, Horizon).

### DSB repair assays

Cells ( $0.25 \times 10^6$ ) were reverse-transfected with 30 pmol siRNA using Lipofectamine RNAiMAX Reagent (Invitrogen) according to the manufacturer's instructions. After 48 h, cells were transfected with 2  $\mu$ g of

pCMV-I-SceI-3 $\times$ NLS or pCMV 3 $\times$ NLS empty vector and 30 pmol siRNA using Lipofectamine 2000 (Invitrogen). Cells were collected for analysis by flow cytometry at 72 h using the LSR Fortessa instrument (BD Biosciences). For each experiment, the percentage of GFP- or RFP-positive cells in the empty vector control was subtracted from the I-SceI-transfected cells. Data from each reporter assay represent the mean  $\pm$  s.e.m. of at least three independent experiments, and statistical analysis was performed using two-tailed paired *t*-tests. The Cas9-mediated DSB repair assay was used to introduce a Cas9-mediated site-specific DSB and break repair outcomes were detected using next-generation sequencing as previously described<sup>39</sup>. In brief, cells were transfected with AAVS1 T2 CRISPR in pX330 (Addgene plasmid, 72833) using Lipofectamine 3000 (Invitrogen) and collected for genomic DNA extraction after 48 h. To measure SSTR, 1  $\mu$ l of a 10  $\mu$ M 121 bp donor oligonucleotide with three substitutions (purchased from IDT) was co-transfected with the CRISPR plasmid. A 201 bp PCR amplicon covering the expected Cas9 break site was sent for next-generation sequencing and reads were analysed for insertions, deletions and substitutions. NHEJ is defined as 1–5 bp deletion, MMEJ as >5 bp deletion with at least 2 bp microhomology, and SSTR as the introduction of three 1 bp substitutions. Data represent the mean  $\pm$  s.e.m. of at least four independent experiments, and statistical analysis was performed using two-tailed paired *t*-tests.

### Immunoprecipitation and western blot

Cells were lysed in RIPA buffer (Teknova, R3792) with HALT protease inhibitor cocktail (Thermo Scientific Scientific). As *HELQ* is expressed at low levels in human cell lines and the commercially available antibodies tested did not dependably detect endogenous *HELQ* by western blot, we validated siRNA-mediated *HELQ* knockdown using *HELQ* immunoprecipitation. Cells ( $10^6$ ) were transfected with 2  $\mu$ g siRNA by electroporation using the Amaxa nucleofector system, and plated into 150 mm dishes. After 72 h, cells were collected and whole-cell extracts were used for immunoprecipitation. Protein (2 mg) was incubated with 1  $\mu$ g anti-HEL308 antibodies (Novus Biologicals, NBPI-91842) at 4 °C overnight with rotation. After washing 0.25 mg Pierce Protein A/G Magnetic Beads (Thermo Fisher Scientific, 88802), the antigen sample–antibody mix was added to the beads and incubated for 1 h at room temperature with rotation. The beads were washed four times and eluted in SDS–PAGE reducing sample buffer (Invitrogen) for 10 min at 96 °C. The samples were loaded onto 4–12% Bis-Tris precast gels (Invitrogen) for SDS–gel electrophoresis and transferred onto Immobilon-P PVDF membrane (Millipore). Membranes were blocked for 1 h at room temperature with Pierce clear milk blocking buffer. For western blot analysis of *RAD52* and *BRCA2*, 50  $\mu$ g of protein was loaded onto 10% Bis-Tris or 3–8% Tris-acetate precast gels (Invitrogen) for SDS–gel electrophoresis. Proteins were transferred onto a PVDF (polyvinylidene difluoride) membrane (BioRad) for *BRCA2* detection. The iBlot Gel Transfer System (Invitrogen) was used to perform dry blotting of proteins onto nitrocellulose membranes for *RAD52* detection. Membranes were blocked for 1 h at room temperature with Pierce clear milk blocking buffer. Antibodies for the western blot analysis were as follows: anti-HEL308 (Santa Cruz Biotechnology, sc-81095), anti-*RAD52* (Santa Cruz Biotechnology, sc-365341), anti-*BRCA2* Ab-1 (Millipore Sigma-Aldrich, OP95), anti-SMC1 (Bethyl laboratories, A300-055A).

### Gene expression

Cells were collected 72 h after siRNA transfection. Cell lysis, reverse transcription and quantitative PCR were performed using the TaqMan Gene Expression Cells-to-CT kit from Invitrogen according to the manufacturer's instructions. TaqMan Gene Expression Assays for *HELQ* (Hs01127906\_m1) and *ACTB* endogenous control (Hs99999903\_m1) were run in triplicate on the QuantStudio 6 Pro real-time PCR instrument (Applied Biosystems). Relative fold gene expression was calculated using the  $\Delta\Delta C_t$  method. Data represent the mean  $\pm$  s.e.m. of at least three independent experiments, and statistical analysis was performed using two-tailed paired *t*-tests.

## Immunofluorescence

WT U2OS-DR and *HELQ*<sup>-/-</sup> U2OS-DR cells were seeded onto four-chamber tissue culture slides (Millipore) and treated the next day with 4 μM camptothecin or 10 Gy irradiation. For siRNA knockdown, cells were transfected with non-targeting or *HELQ* siRNA, incubated overnight and seeded onto chamber slides the next day. Cells were treated with the designated damaging agent 48 h after siRNA transfection. At the designated time points, cells were fixed, blocked and permeabilized. Cells were stained with the following antibodies: anti-phosphorylated-histone H2A.X (Ser139) (05-636, Millipore Sigma-Aldrich), anti-RAD51 (PC130 Millipore Sigma-Aldrich), anti-RPA32 (2208, Cell Signaling Technology), anti-phosphorylated-RPA32 (S4/S8) (ab87277, Abcam). Secondary antibodies were as follows: AlexaFluor 488-labelled goat anti-rabbit IgG, AlexaFluor 568-labelled donkey anti-mouse IgG, AlexaFluor 568-labelled goat anti-rat (Invitrogen). Images were obtained using a Zeiss LSM 880 confocal laser scanning microscope and analysed using ImageJ. At least 100 nuclei were counted per experiment and nuclei with >5 foci were scored as positive. Data represent the mean ± s.e.m. of at least three independent experiments, and statistical analysis was performed using two-tailed paired *t*-tests.

## Software

Chemidoc MP Image Lab Touch software (Bio-Rad, v.2.2.0.08) was used for gel imaging. Quenching-based kinetic unwinding assay data were collected using Clariostar BML Labtech (v.5.20 R5). Bluelake software from LUMICKS was used for data collection of SMI (LUMICKS). Similarly, for smFRET-based assays, NimOS software from ONI was used. BD FACSDiva software was used with the BD Biosciences LSR Fortessa Analyzer for flow cytometry data acquisition. Zen 2.3 SP1 FP3 (black) v.14.0.18.201 was used for confocal microscopy image acquisition.

The quantification of gel-based data was carried using ImageJ (NIH v.1.52k). Mars Data analysis software (BML Labtech v.3.10 R6) was used for quenching-based kinetic unwinding assays. To analyse the SMI by optical tweezer, Pylake software from Lumicks was used. Furthermore, custom scripts were used for analysing some of unwinding/translocation assays carried out by SMI (<https://github.com/singlemoleculer-group>). smFRET analysis was performed using iSMS software (open source)<sup>39</sup>. For microscopy data analysis, ImageJ (NIH, v.1.53e) was used. QuantStudio Design and Analysis Software v.2 was used with the QuantStudio 6 Pro real-time PCR instrument for relative gene expression analysis. Flow cytometry data were analysed using BD FlowJo (v.10.6.2). FlowJo was used to gate single cells and then select for GFP<sup>+</sup> and/or RFP cells depending on the assay as shown. The background signal from the untransfected control was subtracted from each experiment. Representative plots from FlowJo showing the gating strategy are provided in Supplementary Fig. 2. Cas9 DSB repair assay sequencing data were analysed as in Hussain et al.<sup>40</sup>, using PEAR software for read stitching, BLOSUM62 for alignment and code for microhomology/deletion analysis available on GitHub (<https://github.com/cjsifuen/delmh>). Statistical analysis was performed using GraphPad Prism (v.8.2.1 and v.8.4.2). All schematics except for those of the quenching-based kinetic unwinding assays were generated using Adobe Illustrator v.2.3. For the schematics of quenching-based kinetic unwinding assays, Biorender.com was used.

## Reporting summary

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

## Data availability

The datasets generated during and/or analysed during the current study are included alongside the Article. Raw datasets for experiments performed on the C-TRAP are not included and are available from S.J.B. on reasonable request. Gel source data are provided in the Supplementary Information. All data are archived at the Francis Crick Institute or Sloan Kettering Institute.

## Code availability

The custom script made for determining the DNA unwinding/translocation by SMI is available at GitHub (<https://github.com/singlemoleculer-group>).

- Spirek, M. et al. Human RAD51 rapidly forms intrinsically dynamic nucleoprotein filaments modulated by nucleotide binding state. *Nucleic Acids Res.* **46**, 3967–3980 (2018).
- Modesti, M. Fluorescent labeling of proteins. *Methods Mol. Biol.* **1665**, 115–134 (2018).
- Modesti, M. et al. Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. *Structure* **15**, 599–609 (2007).
- Candelli, A. et al. Visualization and quantification of nascent RAD51 filament formation at single-monomer resolution. *Proc. Natl Acad. Sci. USA* **111**, 15090–15095 (2014).
- King, G. A., Burla, F., Peterman, E. J. G. & Wuite, G. J. L. Supercoiling DNA optically. *Proc. Natl Acad. Sci. USA* **116**, 26534–26539 (2019).
- Pierce, A. J., Johnson, R. D., Thompson, L. H. & Jasin, M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev.* **13**, 2633–2638 (1999).
- Stark, J. M., Pierce, A. J., Oh, J., Pastink, A. & Jasin, M. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol. Cell. Biol.* **24**, 9305–9316 (2004).
- Bindra, R. S., Goglia, A. G., Jasin, M. & Powell, S. N. Development of an assay to measure mutagenic non-homologous end-joining repair activity in mammalian cells. *Nucleic Acids Res.* **41**, e115 (2013).
- Chandramouly, G. et al. BRCA1 and CtIP suppress long-tract gene conversion between sister chromatids. *Nat. Commun.* **4**, 2404 (2013).
- Preus, S., Noer, S. L., Hildebrandt, L. L., Gudnason, D. & Birkedal, V. iSMS: single-molecule FRET microscopy software. *Nat. Methods* **12**, 593–594 (2015).
- Hussain, S. S. et al. Measuring nonhomologous end-joining, homologous recombination and alternative end-joining simultaneously at an endogenous locus in any transfectable human cell. *Nucleic Acids Res.* **49**, e74 (2021).

**Acknowledgements** We thank L. Krejci (MUNI, Masaryk University, Bruno) and M. Modesti (CRCM, Marseille) for providing the expression constructs for RAD51 and RPA. R.A. was supported by an EMBO long-term fellowship (ALTF975-2018) and an SNF postdoc.mobility fellowship (P400PB-186718).

**Author contributions** R.A., E.B., S.N.P. and S.J.B. developed the concept for the paper, and D.S.R. advised on SMI experiments. R.A. purified all proteins and performed all bulk biochemical assays, including DNA unwinding, strand annealing and DNA capture experiments. T.T. and A.V. expressed and purified GST-BRC4 peptides. E.B. conducted all cellular assays including HR, SSA, MMEJ and GC tract-length assays. O.B. and M.N. performed and analysed SMI and FRET-based experiments. A.K. and D.S.R. analysed part of the SMI-based DNA capture experiments. S.J.B. and R.A. wrote the paper with editorial input from all of the other authors.

**Competing interests** S.J.B. is a co-founder and VP Science Strategy at Artios Pharma Ltd., Babraham Research Campus, UK.

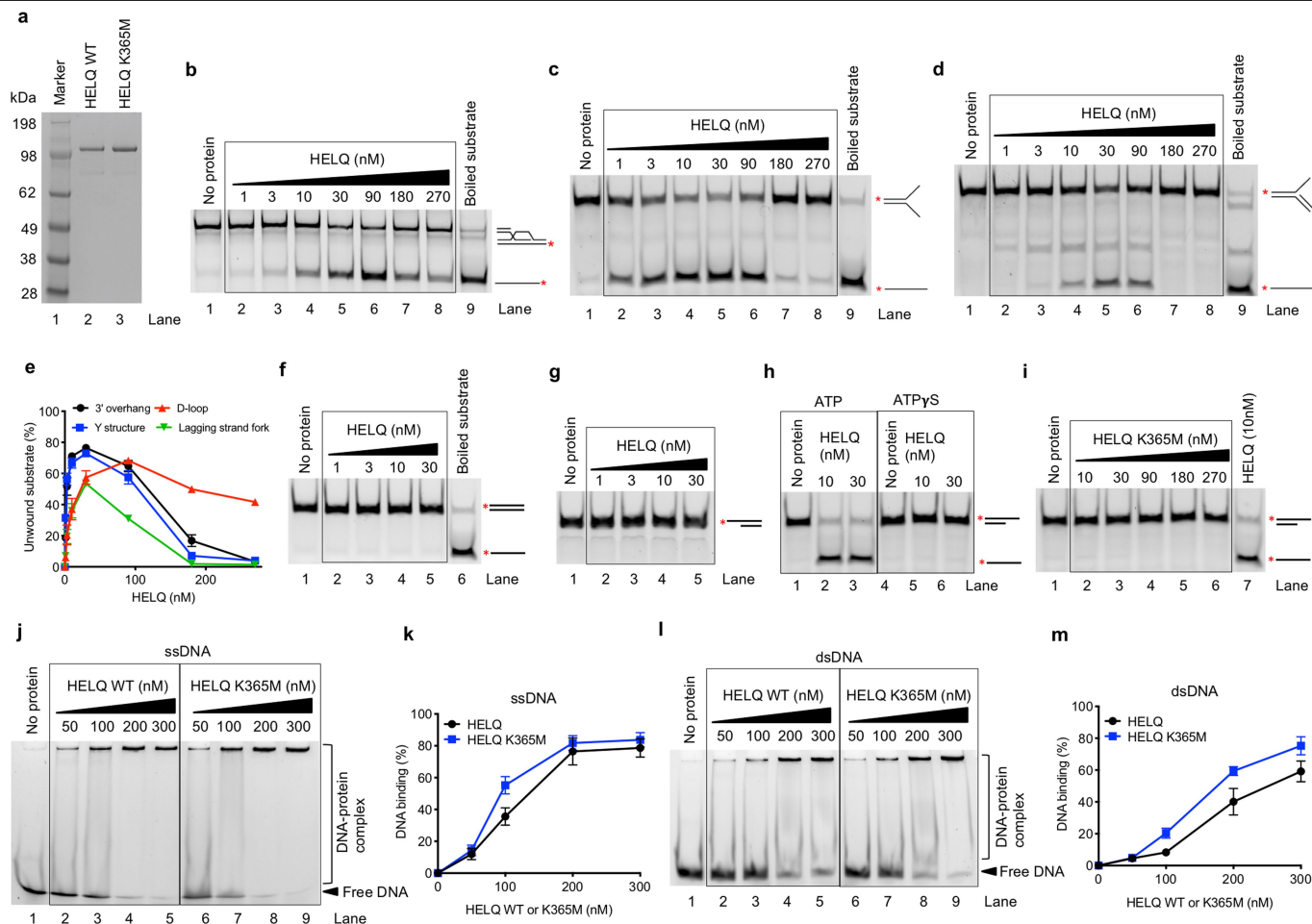
## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-04261-0>.

**Correspondence and requests for materials** should be addressed to David S. Rueda, Simon N. Powell or Simon J. Boulton.

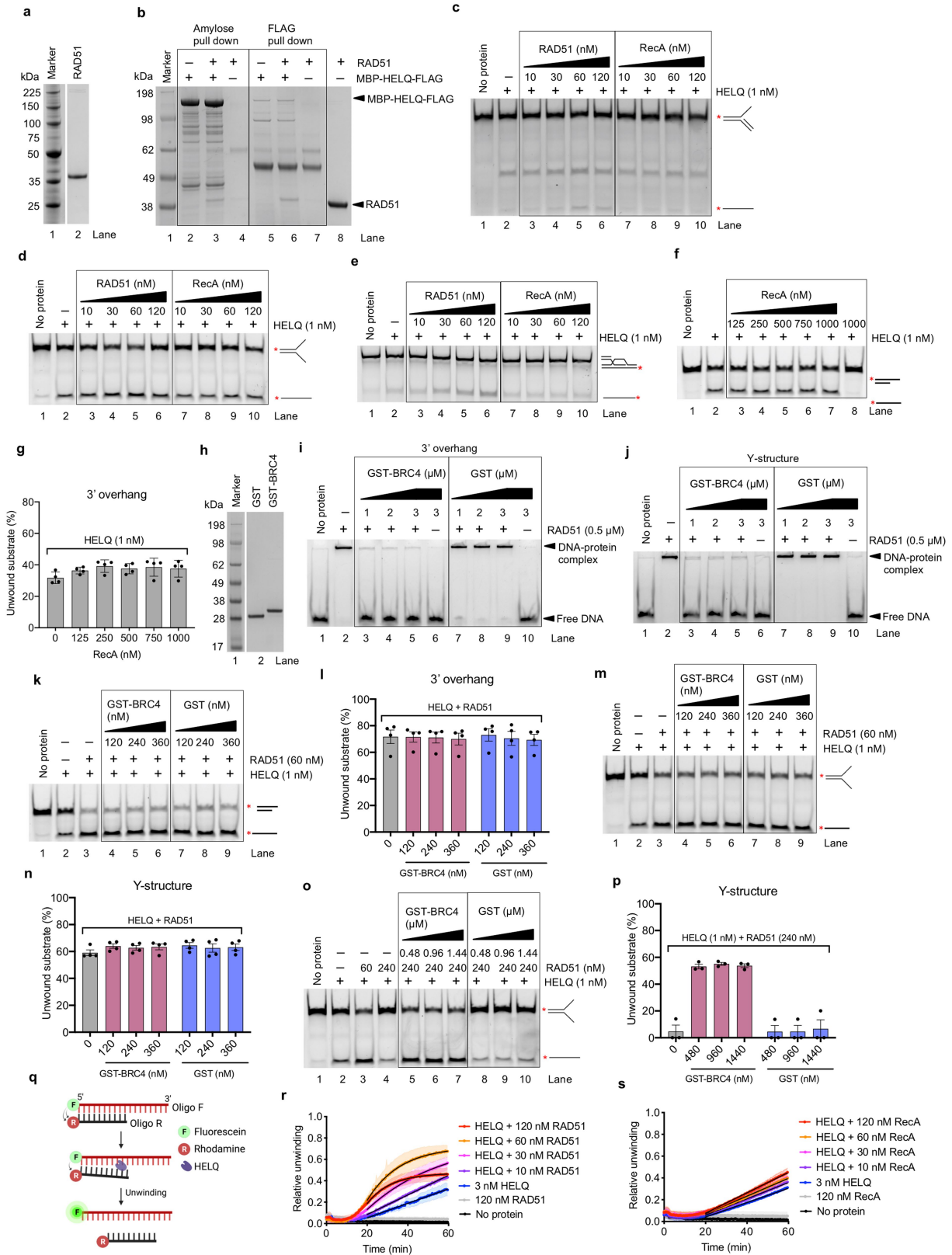
**Peer review information** Nature thanks James Haber and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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



**Extended Data Fig. 1 | HELQ specifically unwinds substrates with 3' overhang and D-loop.** **a**, SDS-PAGE gel (4–12% polyacrylamide) showing purified recombinant human HELQ WT and HELQ(K365M) from insect cells. The gel was stained with Coomassie brilliant blue (CBB). We used two and single preparations of HELQ WT and HELQ(K365M) respectively in this study. **b–d**, Representative native gels (10% polyacrylamide) of DNA unwinding assay of D-loop, Y-structure and lagging strand fork with indicated concentrations of HELQ. The asterisk indicates the position of FITC labelling at 5' end of oligo. **e**, Quantification of experiments such as shown in b–d and Fig. 1a with HELQ concentration ranging from 1–270 nM. n = 4 independent experiments; mean ± S.E.M. **f–g**, Representative native gels of DNA unwinding assay of dsDNA and

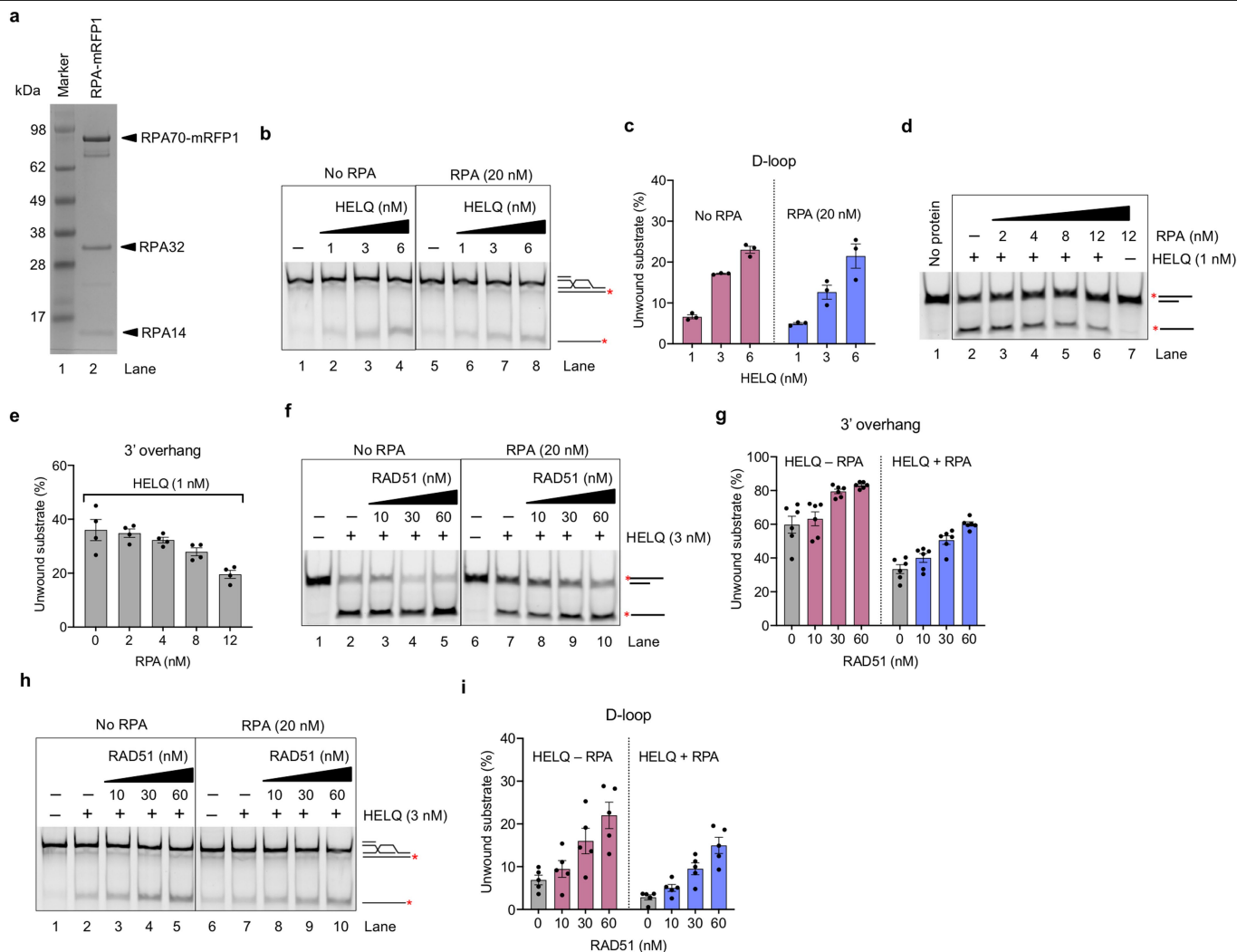
5'-overhang with indicated concentrations of HELQ. **h**, Native gel showing the DNA unwinding assay of 3'-overhang by HELQ in the presence of ATP (2 mM) and ATPγS (2 mM), a poorly hydrolysable ATP analogue (n=2). **i**, Native gel showing DNA unwinding assay of 3'-overhang with indicated concentrations of HELQ(K365M) (n=2). **j**, Electrophoretic mobility shift assay (EMSA) with ssDNA and indicated concentrations of HELQ and HELQ(K365M). The final products were resolved with native 6% polyacrylamide gels. **k**, Quantification of experiments such as shown in j. n = 3 independent experiments; mean ± S.E.M. **l**, EMSA with dsDNA and indicated concentrations of HELQ and HELQ(K365M). **m**, Quantification of experiments such as shown in l. n = 3 independent experiments; mean ± S.E.M.



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

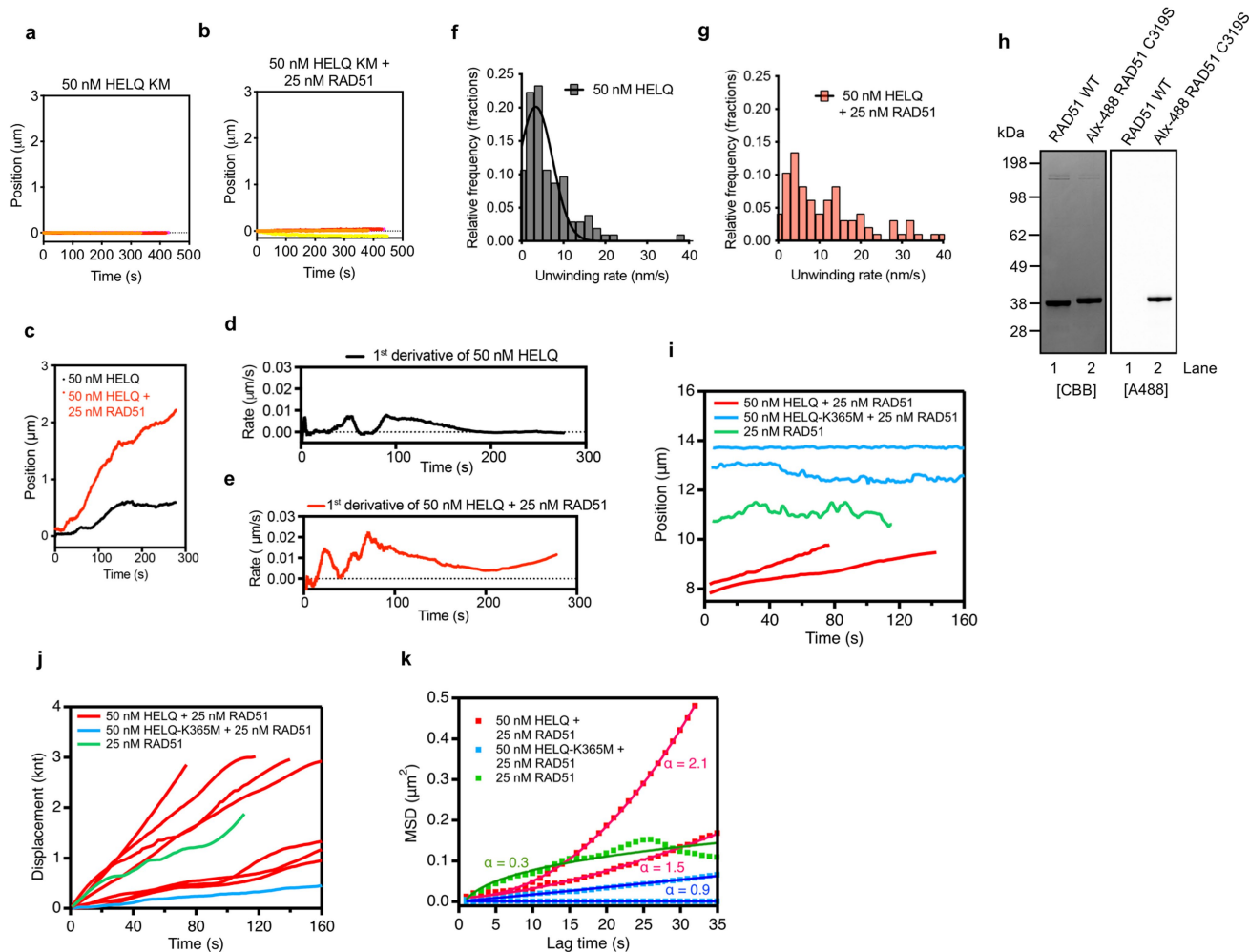
**Extended Data Fig. 2 | RAD51 interacts directly to HELQ and promotes its helicase activity.** **a**, SDS-PAGE gel (4-12%) showing purified recombinant human RAD51 from *E. coli*. The gel was stained with CBB. We used two preparations of RAD51 in this study. **b**, Protein interaction analysis of MBP-HELQ-Flag and RAD51 using both amylose and Flag pull-down assay. The final eluates were run on SDS-PAGE gel (4-12%) and stained with CBB. The interaction analysis was repeated 3 times with similar results. **c-e**, Representative native gels of DNA unwinding of D-loop, Y-structure and lagging strand fork substrates with HELQ (1nM) and indicated concentrations of either RAD51 or RecA. **f**, Representative native gel of DNA unwinding of 3'-overhang with indicated concentrations of RecA and HELQ (1nM). **g**, Quantification of experiments such as shown in f. n = 4 independent experiments; mean  $\pm$  S.E.M. **h**, SDS-PAGE gel (4-12%) showing purified recombinant GST and GST-BRC4 peptides from *E. coli*. The gel was stained with CBB. Single preparation of GST and GST-BRC4 used in this study. **i-j**, EMSA gels (6% polyacrylamide) showing RAD51 binding to 3'-overhang and Y-structure in the presence and absence of indicated concentration of GST-BRC4 and GST peptides. **k**, Representative native gel of DNA unwinding of 3'-overhang by HELQ with RAD51 and indicated concentrations of GST-BRC4 and GST peptides. **l**, Quantification of experiments such as shown in k. n = 4 independent experiments; mean  $\pm$  S.E.M. **m**, Representative native gel of DNA unwinding of

Y-structure by HELQ with RAD51 and indicated concentrations of GST-BRC4 and GST peptides. **n**, Quantification of experiments such as shown in m. n = 4 independent experiments; mean  $\pm$  S.E.M. **o**, Representative native gel of DNA unwinding of Y-structure by HELQ with inhibitory concentration of RAD51 and indicated concentrations of for substrates GST-BRC4 and GST peptides. **p**, Quantification of experiments such as shown in o. n = 3 independent experiments; mean  $\pm$  S.E.M. **q**, Schematics representation of quenching-based kinetic DNA unwinding assay of 3'-overhang. Initially, oligo F (49-mer), labelled at 5' end with fluorescein (F), is annealed with oligo R, which is labelled with rhodamine (R) at 3' end. Due to close proximity, FRET from fluorescein is quenched by rhodamine constitutively resulting in low FRET signal. Upon DNA unwinding, DNA strands are separated and thus rhodamine no longer able to quench fluorescein, results in higher FRET signal. **r**, Relative unwinding of 3'-overhang with HELQ and indicated concentrations of RAD51 as determined by quenching-based kinetic assay. n = 3 independent experiments; shaded area represents mean  $\pm$  S.E.M.; black lines represent exponential or linear fits. **s**, Relative unwinding of 3'-overhang with HELQ and indicated concentrations of RecA as determined by quenching-based kinetic assay. n = 3 independent experiments; shaded area represents mean  $\pm$  S.E.M.; black lines represent exponential or linear fits.



**Extended Data Fig. 3 | RPA inhibits HELQ DNA unwinding activity. a**, SDS-PAGE gel (4-12%) showing purified recombinant human RPA-mRFP1 from *E. coli*. The gel was stained with CBB. Single preparation of RPA-mRFP1 was used in this study. **b**, Representative gel of DNA unwinding assay of D-loop with HELQ in the absence and presence of RPA (20 nM). **c**, Quantification of experiments such as shown in b. n = 3 independent experiments; mean  $\pm$  S.E.M. **d**, Representative gel of DNA unwinding assay of 3'-overhang with HELQ (1 nM) and indicated concentrations of RPA. **e**, Quantification of experiments such as shown in d.

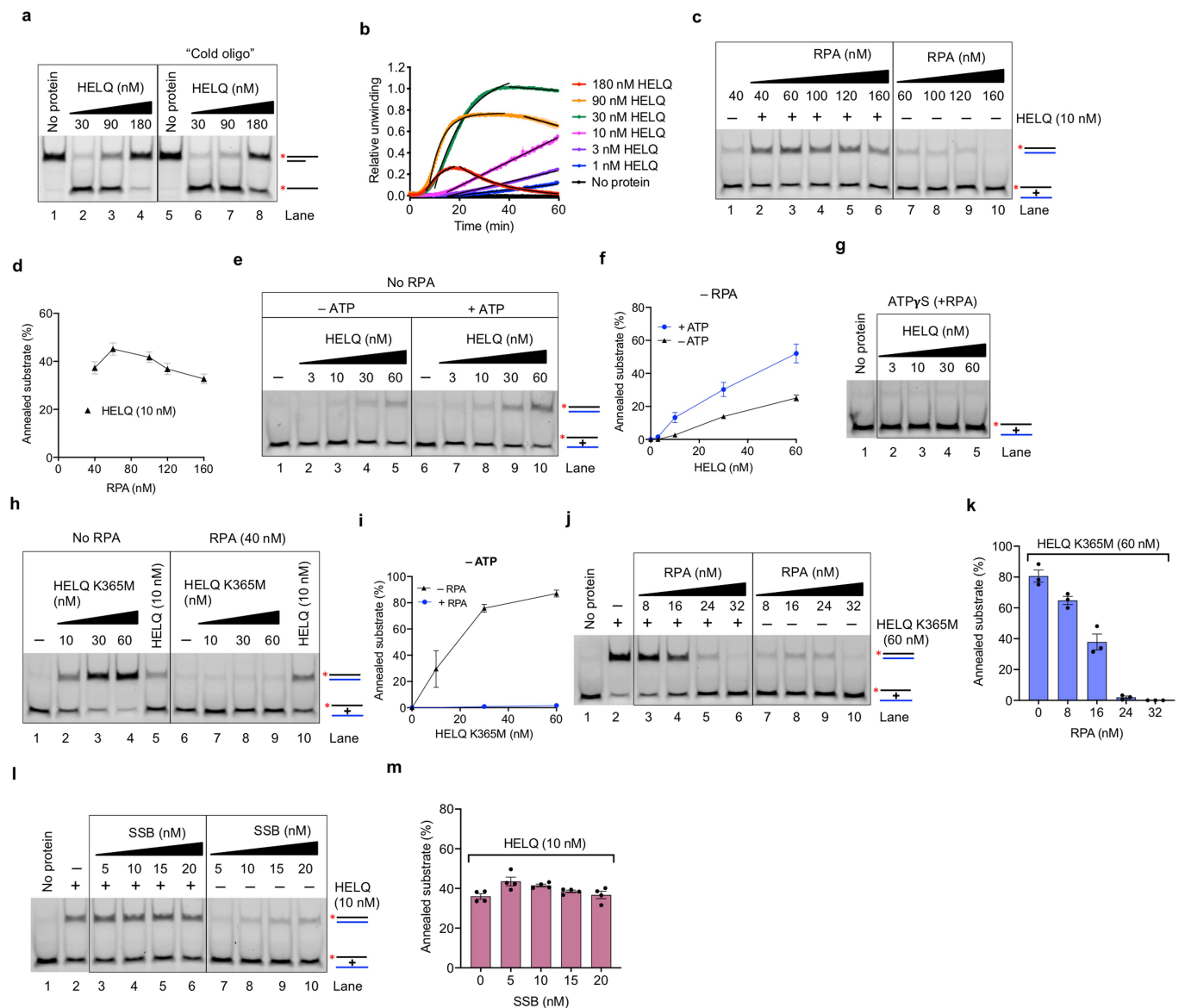
n = 3 independent experiments; mean  $\pm$  S.E.M. **f**, Representative gel of DNA unwinding assay of 3'-overhang with HELQ (3 nM) and indicated concentrations of RAD51, in the absence and presence of RPA (20 nM). **g**, Quantification of experiments such as shown in f. n = 6 independent experiments; mean  $\pm$  S.E.M. **h**, Representative gel of DNA unwinding assay of D-loop with HELQ (3 nM) and indicated concentrations of RAD51, in absence and presence of RPA (20 nM). **i**, Quantification of experiments such as shown in h. n = 5 independent experiments; mean  $\pm$  S.E.M.



**Extended Data Fig. 4 | RAD51-HELQ complex translocates along ssDNA backbone.** **a-b.** Bead centre displacement measured between the traps as a function of time in indicated conditions. Traces represent individual DNA molecules ( $n \geq 3$ ). **c.** Example of two real-time bead displacement traces for indicated conditions. Unwinding bursts of linearly increasing bead distance are interspersed by pauses where no distance change is observed. **d-e.** First derivative with applied smoothing of traces shown in **c**. Height of individual peaks corresponds to the rate of individual unwinding bursts. **f-g.** Histogram showing HELQ unwinding burst rate distribution in the absence ( $n = 103$ ) and presence ( $n = 97$ ) of RAD51. Black line in **f** represents single Gaussian fits. **h.** Chemical labelling of RAD51. RAD51 C319S was labelled in pH 7.0 using maleimide esters of Alexa Fluor 488. After reaction termination and purification of labelled species, labelling efficiency was assessed, and free dye component was evaluated using SDS-PAGE (4-12%) and subsequent fluorescent imaging. 1:1 labelling stoichiometry was achieved as measured, spectrophotometrically. Proteins were labelled typically with 80-100% labelling efficiency. Single preparation of RAD51 C319S was used in this study. **i.** Displacement of proteins bound to individual DNA tethers. The representative trajectories illustrate either

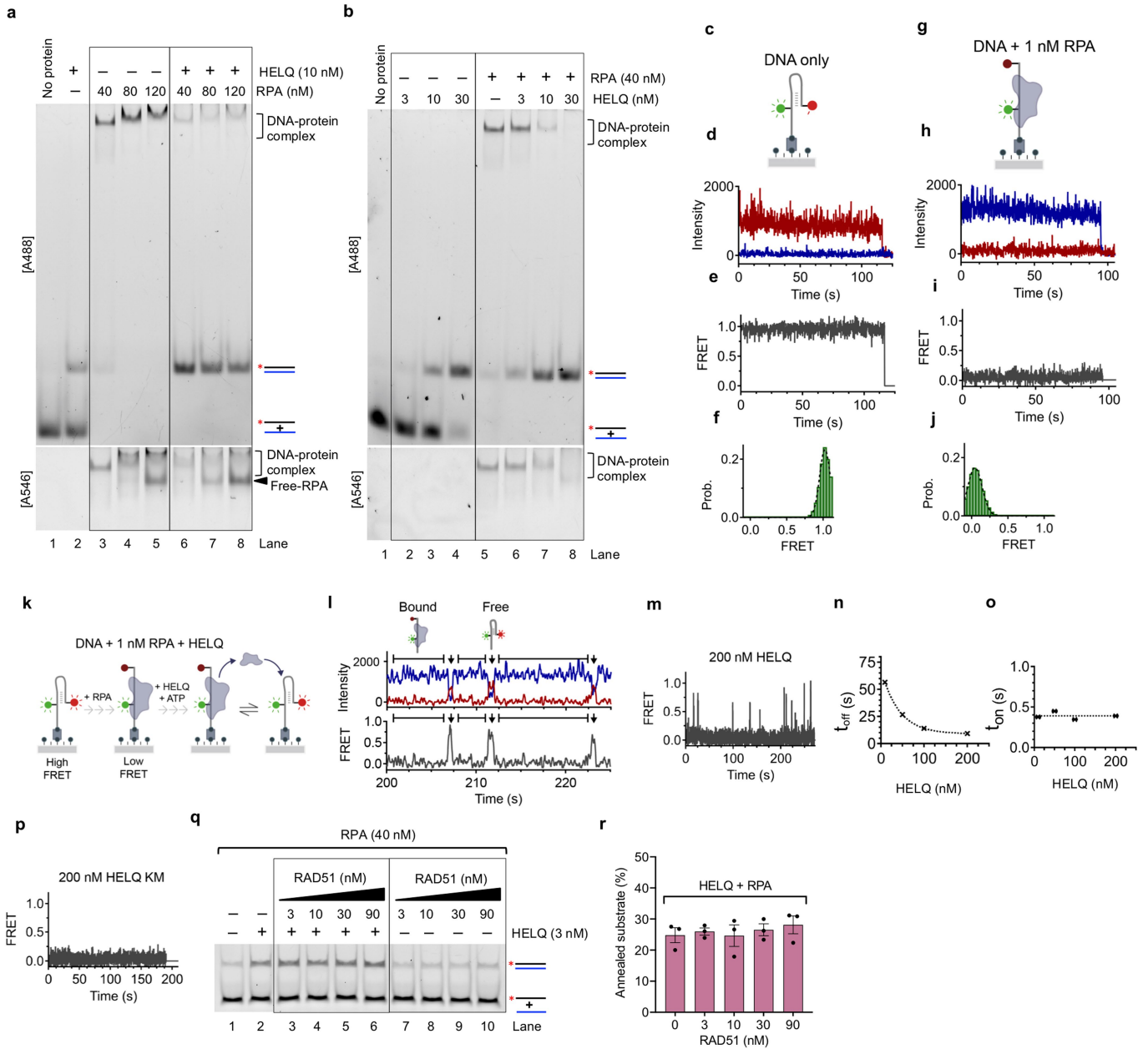
unidirectional movement (red), 1D diffusion (green) and static binding with occasional diffusion (blue). The positions of molecules in time were measured by fitting a moving window of three kymograph frames with Gaussian function. **j.** Total displacement of translocating HELQ-Alx-RAD51 complexes (red) used to calculate the translocation rate. Two populations of translocating molecules can be distinguished: faster with the mean rate of  $14 \pm 5$  nm/s and slower with the mean rate of  $4 \pm 1$  nm/s. The slower rate might result due to “pushing” of sparsely bound Alx-RAD51 on DNA by HELQ alone species. The total displacement of a representative RAD51 filament (green) includes all frame-to-frame displacements, regardless of directionality. The stationary molecule apparent displacement (blue) of  $2 \pm 1$  nm/s results from thermal fluctuations of the tethered DNA. **k.** Mean Square Displacement calculated from the trajectories shown in panel **i**, plotted as a function of time interval for a period up to 35 sec. Fitting the MSD plot with a power law  $D\Delta t^\alpha$  resulted in the factor  $\alpha > 1$  for HELQ-RAD51 complexes (red) indicating their diffusive (directed) motion, in contrast to diffusing RAD51 and HELQ(K365M)-RAD51 complexes ( $\alpha \leq 1$ ) (green, blue).





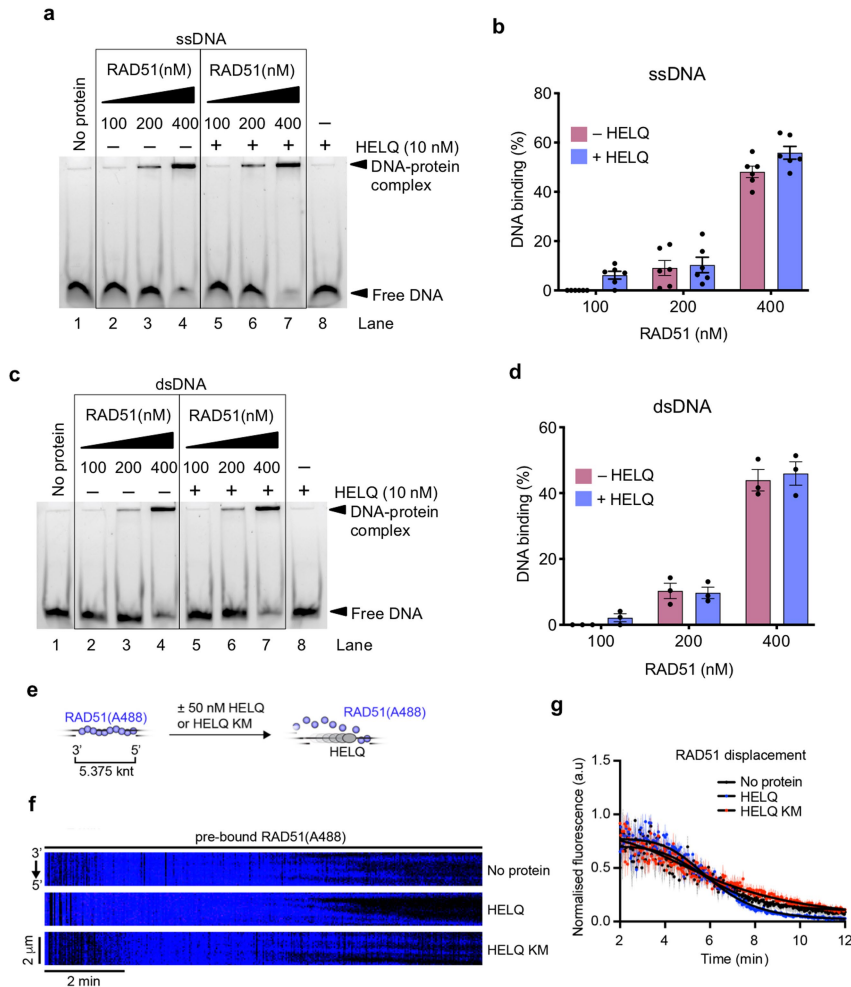
**Extended Data Fig. 5 | ATP is important for HELQ annealing activity in presence of RPA.** **a**, Representative gel of DNA unwinding assay of 3'-overhang with HELQ, in the presence and absence of "cold oligo" (25 nM) i.e., unlabelled oligo with identical DNA sequence as FITC-labelled oligo. **b**, Relative unwinding of 3'-overhang with indicated concentrations of HELQ as determined by quenching-based kinetic assay. n = 3 independent experiments; shaded area represents mean  $\pm$  S.E.M.; black lines represent exponential or linear fits. **c**, Representative gel of DNA annealing assay with HELQ (10 nM) and indicated concentrations of RPA. The black and blue colours of substrate represent complementary DNA strands. The asterisk indicates the position of FITC labelling at 5' end. The products were resolved on 10% native polyacrylamide gel. **d**, Quantification of experiments such as shown in c. n = 5 independent experiments; mean  $\pm$  S.E.M. **e**, Representative gel of DNA annealing assay with

indicated concentrations of HELQ in the absence and presence of ATP. **f**, Quantification of experiments such as shown in e. n = 3 independent experiments; mean  $\pm$  S.E.M. **g**, Native gel showing DNA annealing assay with indicated concentrations of HELQ and RPA (40 nM) in the presence of ATP $\gamma$ S (n=2). **h**, Representative gel of DNA annealing assay with indicated concentrations of HELQ(K365M) in the absence and presence of RPA (40 nM). **i**, Quantification of experiments such as shown in h. n = 4 independent experiments; mean  $\pm$  S.E.M. **j**, Representative gel of DNA annealing assay with HELQ(K365M) (60 nM) and indicated concentrations of RPA. **k**, Quantification of experiments such as shown in j. n = 3 independent experiments; mean  $\pm$  S.E.M. **l**, Representative gel of DNA annealing assay with HELQ (10 nM) and various concentrations of SSB. **m**, Quantification of experiments such as shown in l. n = 4 independent experiments; mean  $\pm$  S.E.M.



**Extended Data Fig. 6 | HELQ strips RPA from ssDNA.** **a**, Native gel (longer 6% polyacrylamide) showing RPA stripping assay with HELQ (10 nM) alone and with indicated concentrations of RPA. **b**, Native gel (longer 6% polyacrylamide) showing RPA stripping assay with indicated concentrations of HELQ in the absence and presence of RPA (40 nM). **c**, Schematic of immobilized dual labelled (Cy3 and Cy5) DNA in the absence of RPA. **d–e**, Representative intensity trajectory (top) and corresponding FRET trajectory (bottom) of dual labelled DNA in the absence of RPA. **f**, Time-binned FRET histogram of DNA only, fit with gaussian. **g**, Schematic of immobilized dual labelled DNA in the presence of RPA. **h–i**, Representative intensity trajectory (top) and corresponding FRET trajectory (bottom) of dual labelled DNA in the presence of RPA. Stable high FRET is observed. **j**, Time-binned FRET histogram of DNA in the presence of RPA. Fit with gaussian. **k**, Schematic of the experimental set up of single-molecule FRET-based RPA stripping assay. DNA dual-labelled with the FRET pair Cy3 and Cy5 is immobilized on the microscope slide. In the absence of RPA, a short 6 nt

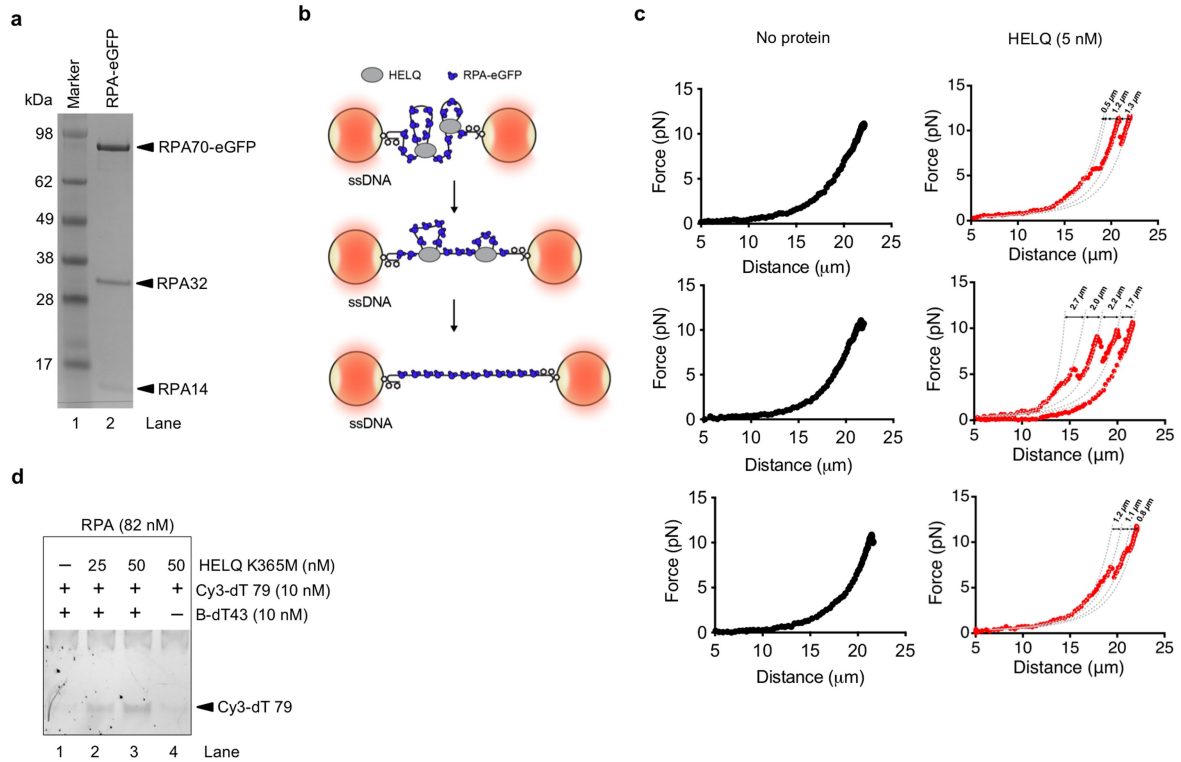
sequence of homology causes the DNA to fold into a high FRET state. Upon RPA binding, the DNA unfolds resulting in a low FRET state. Addition of HELQ results in cycling between the low (bound) and high (free) FRET states as RPA is bound and removed respectively. **l**, Example of single-molecule fluorescence trajectory (top, Cy3 in blue, Cy5 in red) and corresponding FRET (bottom) showing the transition from low FRET (bound) to high FRET (free). **m**, Representative FRET trajectory of DNA in the presence of 1 nM RPA and 200 nM HELQ, spikes of high FRET correspond to RPA removal events. **n**, Plot of dwell time of RPA bound ( $t_{off}$ ), low FRET, state with increasing HELQ concentration. **o**, Plot of dwell time of free state ( $t_{on}$ ), high FRET, on RPA removal with increasing HELQ concentration. **p**, Representative FRET trajectory of DNA in the presence of 1 nM RPA and 200 nM HELQ KM. **q**, Representative gel of DNA annealing assay with HELQ (3 nM), RPA (40 nM) and indicated concentrations of RAD51. **r**, Quantification of annealed substrate (%) of experiments such as shown in **q**.  $n = 3$  independent experiments; mean  $\pm$  S.E.M.



**Extended Data Fig. 7 | HELQ weakly displaces RAD51 from ssDNA.**

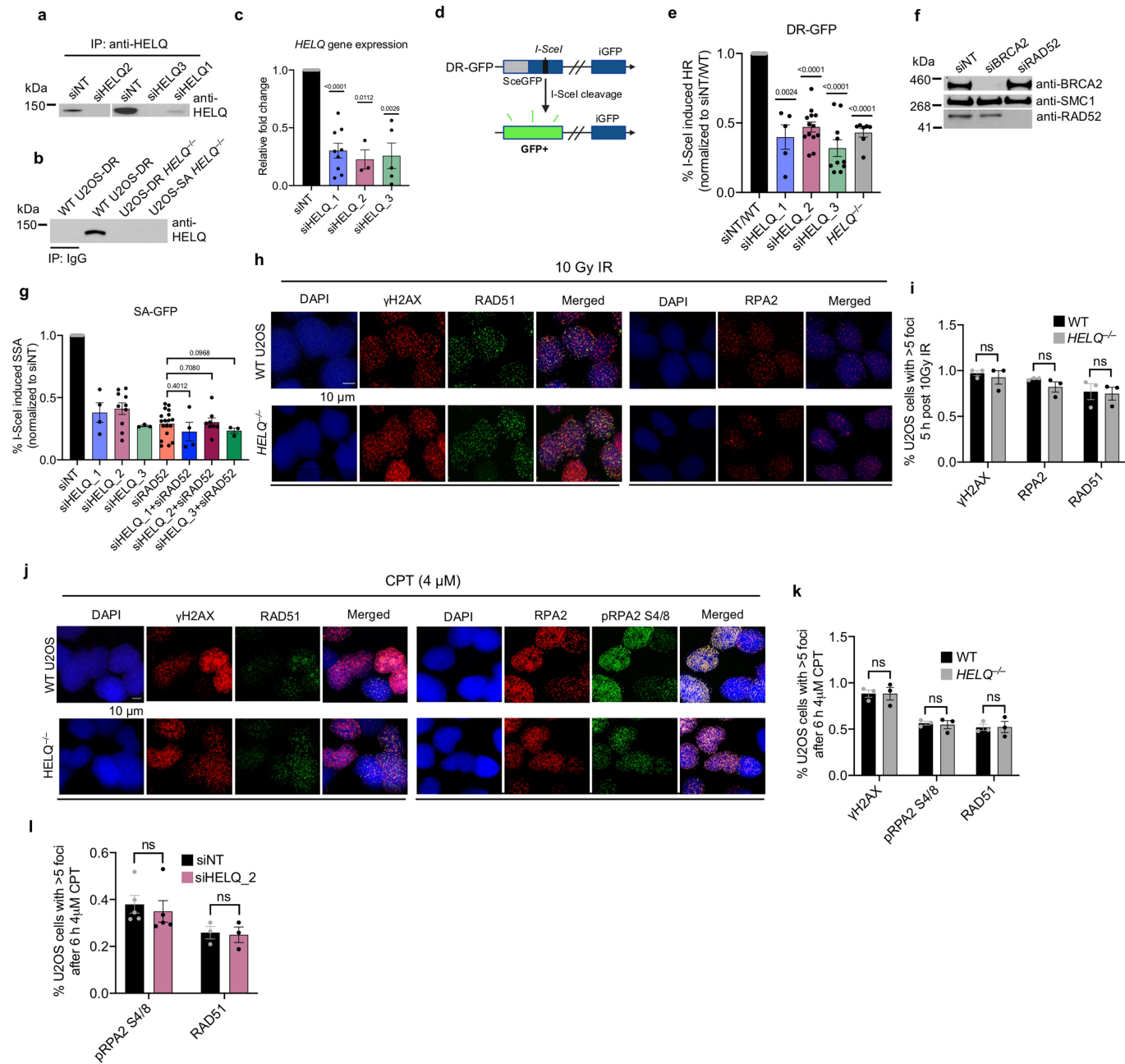
**a**, Representative EMSA gel showing RAD51 binding to ssDNA in the absence and presence of HELQ. Reactions were incubated at 37 °C for 20 min. **b**, Quantification of experiments such as shown in a. n = 6 independent experiments; mean ± S.E.M. **c**, Representative EMSA gel showing RAD51 binding to dsDNA in the absence and presence of HELQ. Reactions were incubated at 37 °C for 20 min. **d**, Quantification of experiments such as shown

in c. n = 3 independent experiments; mean ± S.E.M. **e**, Schematic of the experimental set up of single-molecule FRET-based RAD51 removal assay. **f**, Representative kymographs of removal of Alx-RAD51 pre-bound to gapped DNA in the presence and the absence of 50 nM HELQ or HELQ(K365M). **g**, Removal of Alx-RAD51 measured from gapped ssDNA as a function of time in indicated conditions. Traces represent individual DNA molecules (n ≥ 3).



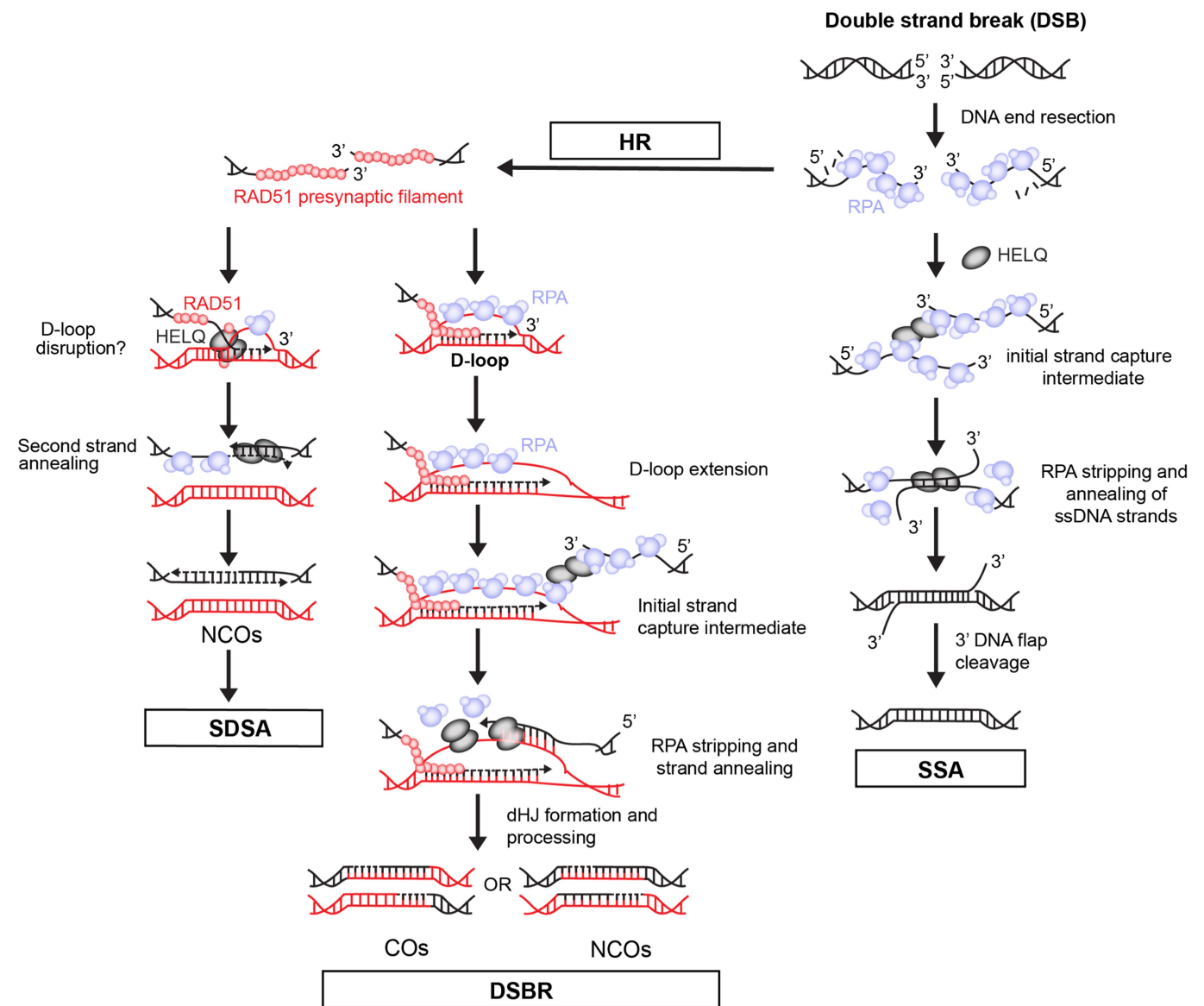
**Extended Data Fig. 8 | HELQ can capture non-complementary DNA strands in presence of RPA.** **a**, SDS-PAGE gel (4-12%) showing purified recombinant human RPA-eGFP from *E. coli*. The gel was stained with CBB. Single preparation of RPA-eGFP used in this study. **b**, A schematic of DNA pulling process. RPA-eGFP-coated ss- $\lambda$  DNA tethered between the two streptavidin beads ( $\sim 4.8 \mu\text{m}$  diameter) was collapsed by bringing beads at  $< 5 \mu\text{m}$  distance. Beads were subsequently pulled apart at constant speed (step size =  $0.2 \mu\text{m}$ , frequency = 500 Hz). Force-extension curves were then recorded. **c**, Force-distance curves of individual

eGFP-RPA-coated ss- $\lambda$  DNA molecules recorded in the absence or presence of HELQ (left and right panels, respectively). Sawtooth-like patterns in the FD curves (red) indicate disruption of ssDNA loops held together by HELQ. The average loop sizes ( $1.5 \pm 0.5 \mu\text{m}$ ,  $N = 10$ ) were estimated from the differences in the fitted contour lengths between the disruption events (dashed lines). **d**, Native gel showing the capture assay with the indicated concentrations of HELQ(K365M) in the presence of RPA (82 nM). The experiment performed two times with similar results.



**Extended Data Fig. 9 | HELQ is epistatic with RAD52 for SSA and additive for second-end capture.** **a**, HELQ protein levels in U2OS WT cells, post 72 h post siRNA transfection, confirmed by HELQ-immunoprecipitation and western blot. **b**, HELQ protein levels in U2OS WT and *HELQ*<sup>-/-</sup> cells, confirmed by HELQ-immunoprecipitation and western blot. **c**, *HELQ* gene expression after treatment with siRNA post 72 h in U2OS cells. siNT n = 9, siHELQ\_1 n = 9, siHELQ\_2 n = 3, siHELQ\_3 n = 3 independent experiments ± S.E.M, compared to siNT. **d**, Schematic representation of DR-GFP reporter assay for measuring DSB repair by HR. **e**, I-SceI-induced HR frequency in U2OS DR *HELQ*<sup>-/-</sup> cells and in U2OS-DR cells treated with the indicated siRNA. siNT/WT, n = 20; siHELQ\_1, n = 5; siHELQ\_2, n = 13; siHELQ\_3, n = 10; *HELQ*<sup>-/-</sup>, n = 7 independent experiments ± S.E.M, compared to siNT or WT cells. **f**, BRCA2 and RAD52 proteins levels 72 h post siRNA transfection confirmed by western blot. siNT, n = 21;

siRAD52, n = 4; siHELQ\_1, n = 4; siHELQ\_2, n = 10; siHELQ\_3, n = 3; siHELQ\_1/RAD52, n = 4, siHELQ\_2/RAD52, n = 8; siHELQ\_3/RAD52, n = 3 independent experiments ± S.E.M, compared to siNT. **h**, Detection of formation of yH2AX, RPA2, and RAD51 foci after 10 Gy IR treatment in U2OS WT and *HELQ*<sup>-/-</sup> cells. **i**, Quantification of RPA2, RAD51 and yH2AX foci in experiments as shown in **h**. n = 3 independent experiments ± S.E.M, compared to siNT. **j**, Detection of formation of yH2AX, RPA2, pRPA2 S4/8 and RAD51 foci after treatment with 4 μM CPT (Camptothecin) of U2OS WT and *HELQ*<sup>-/-</sup> cells. **k**, Quantification of yH2AX, pRPA S4/8 foci and RAD51 foci in experiments as shown in **j**. n = 3 independent experiments ± S.E.M, compared to siNT. **l**, Quantification of formation of pRPA2 S4/8 and RAD51 foci after 6 h CPT treatment in U2OS cells treated with indicated siRNA. pRPA2 S4/8 n = 5, RAD51 n = 3 independent experiments ± S.E.M, compared to siNT. siNT, siRNA not-targeted. The statistical significance was determined by using two-tailed paired t test.



**Extended Data Fig. 10 | A possible mechanism of HELQ function in HR and SSA.** **HR.** In HR, following DSB, 5' DNA ends are resected by nucleases, which generates RPA-coated 3' overhangs. RPA is displaced by RAD51 giving rise to presynaptic nucleoprotein filament formation. This RAD51-filament invades homologous DNA duplex through its 3' end. DNA invasion result in D-loop formation, which is extended by DNA synthesis. If the SDSA pathway is invoked, HELQ can disrupt the D-loop to displace and/or anneal the extended strand with the broken duplex via its helicase and strand annealing activities. Alternatively, in DSBR, the D-loop can be greatly extended and RPA-coated DNA

strands in D-loop and second 3' overhang can be captured by HELQ. During this second-end capture step, HELQ strips RPA from ssDNA and anneals complementary DNA strands together to prime DNA synthesis, restoring the broken DNA and resulting in either crossover or non-crossover products. **SSA**, similar to HR, longer RPA-coated 3' overhangs can be captured by HELQ. Post capture, HELQ strips RPA from both ssDNAs and actively anneals the DNA strands together using base-pairing. The remaining flaps are cleaved, and nicks are ligated by specific nuclease and ligases completing the DSB repair reaction.

## 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

We used commercial software available with the respective instruments for data collection. These include Chemidoc MP Image Lab Touch Software, Biorad (version 2.2.0.08) for gel imaging, Clariostar BML Labtech (version 5.20 R5) for Quenching-based kinetic assay for DNA unwinding, Bluelake software for single-molecule imaging (LUMICKS), NimOS software (ONI) Version 1.16.4.13788 for smFRET, BD FACSDiva software (v8.0.1) was used with BD Biosciences LSR Fortessa analyzer for flow cytometry data acquisition; Zen 2.3 SP1 FP3 (black) v14.0.18.201 was used for confocal microscopy image acquisition.

Data analysis

We used combination of custom made, commercial or publicly available software used for data analysis. These include ImageJ (NIH Version 1.52k) for quantifying gel based assays, Mars Data analysis software (BML Labtech version 3.10 R6), Pylake software (Lumicks) for single-molecule imaging by optical tweezer, Custom script made for some of DNA unwinding/translocation SMI assay (<https://github.com/singlemoleculerroup>), iSMS software (open source) for smFRET analysis (Preus et al. iSMS: single-molecule FRET microscopy software. Nat Methods 12, 593–594 (2015)), Image J (NIH, Version 1.53e) for the analysis of microscopy data; graphs and numerical data (including statistics/error bars) was analyzed and plotted by Prism (GraphPad, Version 8.2.1 and 8.4.2); flow cytometry data was analyzed using FlowJo v10.6.2; QuantStudio Design and Analysis Software v2 was used with QuantStudio 6 Pro real-time PCR instrument for relative gene expression analysis; Cas9 DSB repair assay sequencing data was analyzed as in Hussain et al, Nucleic Acids Research 2021, using PEAR software for read stitching, BLOSUM62 for alignment, and code for microhomology/deletion analysis available on Github <https://github.com/cjsifuen/delmh>.

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.J.B and S.P on reasonable request. For gel source data, see source data file. All data are archived at the Francis Crick Institute or Sloan Kettering Institute. The custom script made for determining the DNA unwinding/translocation by SMI is available at link "https://github.com/singlemoleculergroup".

## Field-specific reporting

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

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

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

## Life sciences study design

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

Sample size	The sample size was determined according to standard practices of field of biochemistry and Cell biology and what was pragmatic number to testing to carry out. Wherever quantification is provided, minimum 3 independent experiments were carried out to perform statistical analysis, As per field practices, for all experiments showing quantification bar/graphs, experiments were repeated at 3 (or more) times.
Data exclusions	In general, no data was excluded except where experiments failed due to failed control and technical problems.
Replication	To ensure the replication, each experiment was performed multiple times in exact condition on different days. For protein-protein interaction, interaction assay was repeated thrice with slightly different conditions and all showed the same result indicating the robustness of the result. Some gels were not included in the study due to technical failure. Any result, which did not replicate was not included in the study.
Randomization	Randomization is not relevant to the kind of experiments performed in this study because randomization of samples is not applicable to cell lines and in vitro studies. On the contrary, researcher needs to know every sample and assay to carry out these kinds of experiments.
Blinding	Blinding is not relevant to the kind of experiments performed in this study because of the same above reason. For example, a researcher exactly needs to know which protein he/she is working with to carry out relevant assays to understand their functions. The same is applicable to cell lines and thus cell biology data. Just like randomization, this is not at all applicable to these kinds of experiments.

## 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	Involved 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"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

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

## Antibodies

### Antibodies used

For Western blot, we used the following antibodies. anti-HEL308 (2406C1a, Santa Cruz Biotechnology, sc-81095, lot#J1217, 1:200 dilution), anti-RAD52 (F-7, Santa Cruz Biotechnology, sc-365341, lot#I2616, dilution 1:500), anti-BRCA2 Ab-1 (clone 2B, Millipore Sigma, OP95, lot#3011827, dilution 1:300), anti-SMC1 (Bethyl laboratories, A300-055A, lot#6, dilution1:1000). For immunoprecipitation of HELQ, we used Hel308 (Novus Biologicals NBP1-91842, lot#A91998). For IF studies, Cells were stained with



the following antibodies: anti-phospho-histone H2A.X (Ser139) (clone JBW301, 05-636, Millipore Sigma, lot#3292608, dilution 1:500), anti-RAD51 (Ab-1, PC130 Calbiochem Millipore Sigma, lot#3092494, dilution 1:500), anti-RPA32 (4E4, #2208, Cell Signaling Technology, lot#5, dilution 1:400), anti-phospho-RPA32 (S4/S8) (ab87277, Abcam, lot#GR3182765-31, dilution 1:200). AlexaFlour 488-labeled goat anti-rabbit IgG (Invitrogen, A11008, dilution 1:1000), AlexaFlour 568-labeled donkey anti-mouse IgG (Invitrogen, A10037, dilution 1:1000) AlexaFlour 568-labeled goat anti-rat (Invitrogen, A11077, dilution 1:1000).

#### Validation

The specificity of anti-HEL308 (Santa Cruz Biotechnology, sc-81095) was validated in Liu, DN. et al. 2017. *Oncol. Rep.* 37: 1107-1113. It was also validated by manufacturer in human cells for WB (<https://datasheets.scbt.com/sc-81095.pdf>). anti-RAD52 (Santa Cruz Biotechnology, sc-365341) specificity was determined previously in multiple studies including most recent research articles Kilgas et al. 2021. *Cell Rep.* 35: 109153 and Zhu et al. 2021. *NAR Cancer.* 3: zcab010. It was also validated by manufacturer in human cells for WB (<https://datasheets.scbt.com/sc-365341.pdf>). anti-BRCA2 (Millipore Sigma, OP95) has validated in numerous studies including the most recent studies Ghoul et al. *Nat Commun* 12, 4605 (2021). It was also validated by manufacturer in human cells for WB. anti-SMC1 was used only as loading control in western blot and has been validated by Bethyl Laboratories for WB (<https://www.bethyl.com/product/A300-055A?referrer=search>). anti-HEL308 (Novus Biologicals, NBP1-91842) specificity was determined the manufacturer where they tested the specificity by protein array containing target protein and other 383 non-specific proteins. anti-phospho-histone H2A.X (Ser139) (Millipore Sigma, 05-636) was validated in Xe, X., et al. (2015) *Nat. Cell Biol.* 20 (3): 320-331. It was also validated by manufacturer in human cells for IF (<https://www.sigmaaldrich.com/GB/en/product/mm/05636i>). anti-RAD51 (Millipore Sigma, PC130) was validated in Brendel, V., et al. 1997. *J. Mol. Evol.* 44, 528 and Boulikas, T., et al. 1997. *Anticancer Res.* 17, 843. It was also validated by manufacturer in human cells for IF ([https://www.merckmillipore.com/GB/en/product/Anti-Rad51-Ab-1-Rabbit-pAb,EMD\\_BIO-PC130#documentation](https://www.merckmillipore.com/GB/en/product/Anti-Rad51-Ab-1-Rabbit-pAb,EMD_BIO-PC130#documentation)). anti-RPA32 (Cell Signaling technology, 2208) was validated in Wang et al. *Cell Death Dis* 9, 923 (2018) and Dharm et al. *J Cell Biol* 6 November 2017; 216 (11): 3521–3534. It was also validated by manufacturer in human cells for IF (<https://www.cellsignal.co.uk/datasheet.jsp?productId=2208&images=1>). anti-phospho-RPA32 (S4/S8) (Abcam, ab87277) was validated in Shengqin et al. *Nucleic Acids Research*, Volume 40, Issue 21, 1 November 2012, Pages 10780–10794. It was also validated by manufacturer in human cells for IHC (<https://www.abcam.com/rpa32rpa2-phospho-s4-s8-antibody-ab87277.html>). To validate Hel308 (Novus Biologicals NBP1-91842, lot#A91998) specificity, western blot was performed in wild type, HELQ knockout and HELQ siRNA treated cells where single band was detected in only wild type cells.

## Eukaryotic cell lines

### Policy information about cell lines

#### Cell line source(s)

We used sf9 insect cells for expression of recombinant proteins available from The Francis Crick Institute, London, UK and these cells are available on request. We used U2OS cells integrated with either DRGFP or SAGFP reporters gifted by Dr. Maria Jasin from Memorial Sloan Kettering Cancer Center New York, NY USA. U2OS cells with integrated SCR reporter were gifted by Dr. Ralph Scully from Harvard Medical School Boston, MA USA. U2OS-DRGFP, U2OS-SA: generated from U2OS cell line ATCC HTB-96. Specifically, Gunn A, Stark JM (2012) I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol Biol* 920: 379–391. pmid:22941618. U2OS-EJDR: generated from U2OS-DRGFP cell line Bindra RS, Goglia AG, Jasin M, Powell SN. Development of an assay to measure mutagenic non-homologous end-joining repair activity in mammalian cells. *Nucleic Acids Res.* 2013 Jun;41(11):e115. doi: 10.1093/nar/gkt255. Epub 2013 Apr 12. PMID: 23585275; PMCID: PMC3675474. U2OS-SCR: generated from U2OS cell line ATCC HTB-96 Chandramouly G, Kwok A, Huang B, Willis NA, Xie A, Scully R. BRCA1 and CtIP suppress long-tract gene conversion between sister chromatids. *Nat Commun.* 2013;4:2404. doi: 10.1038/ncomms3404. PMID: 23994874; PMCID: PMC3838905

#### Authentication

Cell lines were not authenticated.

#### Mycoplasma contamination

All human cell lines tested negative for mycoplasma contamination. sf9 cells were not tested for the contamination.

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

No commonly misidentified lines were used.