Structural mechanism of strand exchange by the RAD51 filament

Luay Joudeh[1,†], Robert E. Appleby[1,2,†], Joseph D. Maman[1], Luca Pellegrini[1,*]

[1]Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, UK

[2]Present address: Francis Crick Institute, London, London NW1 1AT, UK

†These authors contributed equally to this work.

*Corresponding author. Email: lp212@cam.ac.uk

SUMMARY

Homologous Recombination (HR) preserves genomic stability by repairing double-strand DNA breaks and ensuring efficient DNA replication. Central to HR is the strand-exchange reaction taking place within the three-stranded synapsis wherein a RAD51 nucleoprotein filament binds to a donor DNA. Here we present the cryoEM structure of a displacement loop of human RAD51 that captures the synaptic state when the filament has become tightly bound to the donor DNA. The structure elucidates the mechanism of strand exchange by RAD51, including the filament engagement with the donor DNA, the strand invasion and pairing with the complementary sequence of the donor DNA, the capture of the non-complementary strand and the polarity of the strand-exchange reaction. Our findings provide fundamental mechanistic insights into the biochemical reaction of eukaryotic HR.

INTRODUCTION

The transfer of genetic information between DNA molecules promoted by HR is essential in the repair of DNA double-strand breaks[1,2] and to overcome replicative stress during chromosomal DNA synthesis[3,4]. The molecular apparatus of HR is also responsible for the crossover reactions that generate genetic diversity in meiosis[5]. HR deficiency causes genetic instability that predisposes to cancer[4,6,7] and its synthetic lethality with other DNA repair pathways is exploited to induce cancer cell death in therapy[8,9].

In eukaryotic cells, the reaction of HR is performed by RAD51[10,11], a member of a highly-conserved family of strand-exchange ATPases that includes bacterial RecA and archaeal RadA[12,13]. RAD51 polymerises on the single-stranded (ss) DNA overhangs that form upon long-range resection of DNA ends[14]. The RAD51 nucleoprotein filament invades the donor DNA duplex and searches for sequence homology within a three-stranded synaptic filament intermediate[15,16]. When homology is found, the invading strand of the filament becomes stably paired to the complementary sequence of the donor DNA, yielding a structure known as a displacement loop (D-loop)[17–19] (Fig. 1A), which ultimately progresses to strand-exchange. Although cellular HR is controlled by a complex network of recombination mediators such as BRCA2[20], the RAD51 paralogues[21–23] and FIGNL1[24], RAD51 can catalyse the biochemical reaction of strand-exchange independently.

The RecA/RAD51/RadA protein family has been extensively investigated to elucidate the mechanistic basis for the reaction of strand exchange. Low-resolution electron microscopy (EM) and X-ray crystallography have demonstrated conservation in the helical parameters of the nucleoprotein filament[25,26], the structure of the ATPase domain[27,28] and the protomer self-association into a filament[29,30]. CryoEM analysis of the RAD51 nucleoprotein filament have revealed further conservation with RecA in the mode of interaction with DNA, such as the observation that DNA stretching results from its separation into B-form nucleotide triplets[31–34]. Furthermore, single-molecule experiments have determined the kinetic parameters of filament nucleation and growth[35,36], have shown that a minimum length of 8 homologous nucleotides is required for formation of stable synaptic joints[37,38] and that homologous pairing takes place in steps of three nucleotides[39,40].

High-resolution information about the RAD51 synaptic state is still missing, limiting our mechanistic understanding of the strand-exchange reaction. In this study, we have carried out the biochemical reconstitution, cryoEM structure and biophysical validation of a D-loop of

human RAD51. The structure elucidates the key steps of strand-exchange, including the mechanism of filament engagement with the donor DNA, strand invasion and pairing with the complementary strand, and capture of the exchanged strand. Comparison with the synaptic state of RecA filaments[41] reveals both conserved and distinct features of the strand-exchange reaction, shedding light on the molecular evolution of the biochemical reaction of HR.

RESULTS

Electron microscopy (EM) analysis of filamentous RAD51 relies normally on helical averaging, which is unsuitable for reconstruction of an aperiodic structure such as a D-loop. We therefore sought to prepare a D-loop sample of human RAD51 that is amenable to single-particle cryoEM. We designed a 32-nucleotide (nt) poly(dT) with a central, unique 10-nt long sequence complementary to a 50-nt donor double-stranded (ds) DNA; the complementary sequence of the donor DNA was incorporated in a 10-nt mismatch bubble to favour D-loop formation (Fig. S1A and Table S1). The length of the complementary sequence was decided based on evidence that RAD51 requires a homologous tract of at least 8 nucleotides for stable capture of the donor DNA[38]. As RAD51 can bind both ss- and dsDNA, we defined experimentally appropriate stoichiometry and buffer conditions for D-loop reconstitution. Electrophoretic mobility shift assay (EMSA) with fluorescently labelled oligos showed successful D-loop formation (Fig. S1B and Table S1).

We collected an initial cryoEM dataset of 2792 movies of the RAD51 D-loop sample and processed the data by single-particle analysis. 2D classification showed clear evidence of the presence of synaptic particles and yielded a D-loop reconstruction at 3.31 Å resolution (Fig. S2 and Table S2). To improve data homogeneity and resolution, we collected a larger dataset of 8960 movies of a D-loop sample with ss- and dsDNA that had been biotinylated at both ends and capped with monomeric streptavidin (Fig. S1C)[42], yielding a structure of the D-loop at 2.64 Å (Fig. S3 and Table S2). The final single-particle reconstruction based on 102,596 particles comprises a nucleoprotein filament of 9 RAD51 protomers and 26 nucleotides of single-stranded DNA, bound to a dsDNA containing the central bubble region flanked by 15- and 16-bp long segments of donor DNA. The cryoEM map ranges in resolution between 2.5 Å and 2.9 Å for the homologous pairing region of the D-loop and 3 Å to 6 Å for the flanking dsDNA arms (Fig. 1B).

D-loop structure

In the D-loop structure, RAD51 maintains the right-handed helical arrangement of protomers observed in pre- and post-synaptic filaments[31,32], with the same helical twist and a slightly compressed pitch relative to the pre-synaptic filament (Fig. 1C, Fig. S4A and Movie S1). Continuous density for the exchanged strand (es), that becomes excluded upon homologous pairing of the invading strand (is) with the complementary strand (cs) of the donor DNA, connects the 5'- and 3'-arms of the donor DNA (Fig. S4B). The filament-bound esDNA follows a peripheral path along the filament groove that runs parallel to the isDNA (Fig. 1D). The conformation of the duplex formed by pairing of the isDNA and csDNA strands resembles closely that of the dsDNA in the post-synaptic filament (Fig. S4C)[32,33], with sharp bends at the junctions where the donor DNA enters and exits the filament (Fig. 1C, D).

The arms of the donor DNA flanking the complementary region project away from the filament in opposite directions and with different tilt angles relative to the filament axis: the dsDNA arm closer to the 5'-end of the filament, hereinafter referred to as the 5'-arm, is almost perpendicular to the filament axis, with its filament-proximal end pointing slightly towards the 3'-end of the filament (3'-tilt), whereas the dsDNA arm closer to the 3'-end of the filament, hereinafter referred to as the 3'-arm, exhibits a more pronounced tilt in the opposite direction, towards the 5'-end of the ssDNA (5'-tilt) (Fig. 1D and Fig. S4D). The opposite tilt of the donor arms favours their connection via the esDNA strand in the D-loop structure.

Homologous pairing

A continuous span of 10 base-pair interactions is present between the isDNA and csDNA (Fig. 1E and Fig. S5A). Base pairing includes the entirety of the csDNA sequence designed to be homologous to the invading strand, with the exception of $G26_{cs}$ (Fig. 1E, F and Fig. S5A): its guanine base adopts a minor-occupancy conformation in the direction of C8 in the isDNA but too distant for Watson-Crick hydrogen bonding, while the predominant conformation is pointing away (Fig. S5B).

Base pairing is fully maintained in the transition between the csDNA : isDNA duplex and the flanking arms of the donor DNA (Fig. 1F, G), with the exception of the aforementioned $G26_{cs}$. Unexpectedly, an additional, mismatched base pair is formed at the junction of the csDNA : isDNA duplex with the 3'-arm of the donor DNA, where $T18_{is}$ becomes base paired with $G16_{cs}$ (Fig. 1E, G and Fig. S5A). Thus, the complementary pairing observed in the D-loop structure does not entirely follow our experimental design, with one unrealised $G26_{cs}$: $C8_{is}$ base pair and one mismatched $G16_{cs}$: $T18_{is}$ base pair (Fig. S5). This observation indicates that factors

additional to hydrogen-bonding potential contribute to determine pairing between the donor DNA and the invading strand of the filament.

Unwinding and strand invasion of donor DNA

Biochemical and structural analyses of pre- and post-synaptic filaments have revealed that RAD51 loops L1 and L2 mediate DNA binding, with prominent roles for amino acids L1 R235 and L2 V273 in partitioning filament DNA into nucleotide triplets[32,43]. However, L2 residues Q272 to P283 remained disordered in these filament structures (Fig. 2A).

In our D-loop structure, the L2 sequences of four contiguous RAD51 protomers – E, F, G and H (Fig. 2B) – have become ordered and traceable in the density map, folding in random-coil conformation with 3_{10} -helix elements. The structure shows that filamentous RAD51 achieves strand separation by sequential L2-loop insertion of adjacent RAD51 protomers between the strands of the donor DNA (Fig. 2C, D). Insertion of the RAD51-H L2 loop at the 5'-arm junction causes a drastic widening of the minor groove, progressing to full separation of the complementary and exchanged strands upon L2-loop insertion by RAD51-G (Fig. 2C, D, G, H). In total, donor DNA unwinding by RAD51-H and -G exposes a continuous csDNA stretch of five nucleotides for homologous pairing: A22 - C25 of csDNA base pair with G9 - T12 of isDNA, physically organised as a base-pair triplet of $(A22 - A24)_{cs}$ with $(T10 - T12)_{is}$ and one unstacked $C25_{cs}$: $G9_{is}$ base pair, while $G26_{cs}$ remains unpaired (Figs. 1F and 2C, D, G, H).

Successive L2-loop insertions extend the unwinding of donor DNA, with each insertion exposing a triplet of csDNA nucleotides for homology pairing with the invading strand (Fig. 2C). Thus, L2-loop insertions by RAD51-F and -E present $(A19 - C21)_{cs}$ for base pairing with $(G13 - T15)_{is}$, and $(G16 - T18)_{cs}$ for base pairing with $(A16 - T18)_{is}$, respectively (Fig. 2C-F). After completion of homologous pairing, the strands of the donor DNA rejoin in double-helical form at the 3'-arm junction (Figs. 1G and 2C).

The L2 loop acts as a physical spacer, directly separating the unwound strands of the donor DNA. Its sequence is dominated by hydrophobic amino acids (Fig. 2A) that make extensive contacts with the ribose-phosphate backbone of the esDNA and culminate in stacking of F279's phenylalanine side chain against the last base pair of the 5'-arm duplex (RAD51-H) or the aromatic bases of nucleotides in the exchanged strand (RAD51-G, F, E) (Fig. 2E-H). The L2 loop displays a considerable degree of conformational heterogeneity, likely because each loop engages in both common and unique contacts with the DNA. This is best exemplified by the L2 loops of RAD51-E and -H that interact with the duplex arms of the donor DNA in different

ways. RAD51-H L2 initiates donor DNA unwinding by buttressing the F279 side chain against $G27_{cs}$: $C15_{is}$, the last base pair of the 5'-arm duplex (Fig. 2H). Conversely, contact with RAD51-E L2 does not alter the conformation of the 3'-arm, as its F279 side chain stacks against C24 of the exchanged strand (Fig. 2E).

The asymmetric binding of RAD51-H and -E L2 loops at the boundaries between the arms of the donor DNA and the three-stranded synaptic DNA suggests that strand invasion and pairing follows a preferred direction. Thus, invasion by filament DNA likely begin at the 3'-end of the homology region in the complementary strand of the donor DNA and extends in the 5'-direction, as incrementally more nucleotides of csDNA are exposed by L2 loop insertions and become available for homologous pairing.

Capture of the exchanged strand

Strand invasion leads to displacement of the exchanged strand in the donor DNA. The structure of the D-loop shows that the esDNA is sequestered in a basic channel running parallel to the ssDNA-binding groove of the RAD51 nucleoprotein filament (Fig. 3A, B). One side of the channel is formed by RAD51 residues 303 to 313 (Fig. 3C), that fold in a beta hairpin on the inside of the filament groove, whereas the opposite side is formed by the C-terminal half of the L2-loop sequence (Fig. 3B, D). In the filament, the beta hairpins of successive protomers align to generate a continuous ssDNA-binding stripe (Fig. 3D). The esDNA is bound in the basic channel so that its aromatic bases remain largely exposed to solvent whereas the phosphodiester backbone faces the protein (Fig. 3B, D).

Three adjacent RAD51 protomers – F, G and H – engage the exchanged strand in the D-loop (Fig. 3D). The majority of contacts comprise charged interactions between four invariant RAD51 residues R303, K304, R306 and K313 and the phosphate groups of the esDNA (Fig. 3E-G). In addition to the electrostatic interactions, R306 is part of a GRG motif at the tip of the beta hairpin that comes into close contact with the phosphate-ribose backbone of the esDNA and makes stacking interactions with the aromatic bases of nucleotides $A19_{es}$ and $T25_{es}$ in RAD51-H and -G, respectively (Fig. 3E, F). The interaction of the 305-GRG-307 motif with esDNA resembles that of L1-loop 234-GRG-236 with csDNA in the homologous duplex of the D-loop.

Of the three RAD51 protomers F, G and H that bind the esDNA, RAD51-G makes the most extensive set of contacts, spanning five nucleotides, from $A21_{es}$ to $T25_{es}$, involving all four basic residues R303, K304, R306 and K313 (Fig. 3F). Binding of RAD51-G to the esDNA is

further strengthened by phosphate-group contacts of L2-loop K284 with $C22_{es}$ and of R130 with $C23_{es}$ and $T24_{es}$. RAD51-H contacts four nucleotides, from $A17_{es}$ to $A20_{es}$, with R303, K304 and 305-GRG-307 (Fig. 3E), whereas RAD51-F makes a single long-range interaction between K313 and $T25_{es}$ (Fig. 3G). In addition, R303 and R306 of RAD51-F engage the phosphate groups of T27, A28 and T29 of the 3'-arm duplex, possibly to guide the 3'-arm in the appropriate direction as it exits the filament groove, towards the RAD51-A NTD (Fig. 3G).

In total, the filament interacts with 9 of the 11 nucleotides of unpaired esDNA via contiguous interactions of the two RAD51 protomers G and H (5 and 4 nt, respectively). This compares with donor DNA opening by L2 loop insertion, which progresses by exposure of 3 nucleotides per RAD51 protomer.

Filament binding of donor DNA arms

The D-loop structure shows that the nucleoprotein filament interacts with the double-stranded arms of the donor DNA via the N-terminal domains of RAD51 protomers A and D (Fig. 4A, B). Although the position of the RAD51-A and -D NTDs relative to the donor DNA is broadly similar, their DNA-binding interfaces differ significantly in the extent and mode of their engagement with the 3'- and 5'-arm duplexes, respectively.

The RAD51-D NTD engages the minor groove of the 5'-arm DNA (Fig. 4C). The side chains of K39, K40 and K64 contact the phosphate groups of 10-GCTC-13 and 34-CG-35 in the upper and lower rungs of the minor groove respectively, while K70 and K73 contribute long-range electrostatic interactions to the binding (Fig. 4C, E). The association of RAD51-D NTD with 5'-arm is closer and more extensive than that of RAD51-A with the 3'-arm, and includes a direct hydrogen bond between the main-chain nitrogen of G65 and the phosphate group of T12 (Fig. 4C, E). In contrast, RAD51-A NTD binds in the major groove of the 3'-arm DNA, using electrostatic interactions mediated by the same set of lysine residues as in RAD51-D (Fig. 4D). However, the RAD51-A association with the 5'-arm involves a more limited set of direct contacts with the phosphate groups of C5 and 31-TTG-33 across the major groove (Fig. 4D, E).

The interaction of the NTD with the arms of the donor DNA highlights a lack of direct contacts with the bases and the predominantly electrostatic nature of the association. Such plasticity in the NTD - DNA interaction might be required for effective D-loop formation, as the direction at which the 3'-arm exits the filament upon D-loop formation is a function of the number of nucleotides involved in homologous pairing.

Structure-based mutagenesis of RAD51 residues important for D-loop formation

The atomic model of the RAD51 D-loop has revealed several new protein - DNA interfaces, including those involved in strand separation, capture of the exchanged strand and anchoring of the flanking DNA arms to the filament. To assess the relative importance and contribution of RAD51 residues involved in D-loop formation, we performed alanine mutagenesis of F279 (strand separation); R303, K304, R306, K313 (esDNA capture); K39, K40, K64, K70 and K73 (arm duplex binding) (Fig. 5A).

The designed mutations are distant from the binding site of ssDNA in filamentous RAD51, and therefore are not predicted to reduce the ability of the mutant protein to form nucleoprotein filaments on ssDNA. As expected, EMSA analysis showed that the F279 and NTD mutants retained wild-type ability to form filaments on ssDNA (Fig. S6A). However, RAD51 mutants R303A, K304A, R306A and K313A showed a clear decrease in ssDNA binding (Fig. S6A), suggesting that, in addition to their role in esDNA capture, they might mediate the initial contact of the incoming ssDNA during pre-synaptic filament nucleation. All RAD51 mutants showed a moderate decrease in affinity for dsDNA, with the exception of K304 (Fig. S6B).

We evaluated the ability of the alanine mutants to perform DNA-strand exchange using an *in vitro* assay that measures the increase in fluorescent signal when a fluorescently-labelled 60-nt oligo that is quenched in duplex form is strand-exchanged with the unlabelled sequence (Fig. 5B)[44]. The results of the strand-exchange assay revealed that all mutants decreased the efficiency of the reaction, but with varying degrees of impairment (Fig. 5B). The L2 loop F279A showed a modest impact, whereas the esDNA-capture mutants R303A, K304A, R306A and K313A displayed a more pronounced effect. Notably, the NTD mutations showed the largest negative impact on strand exchange: the double lysine-to-alanine substitutions of K39, K40 and K70, K73 reduced strand exchange to a fraction of the yield for the wild-type protein.

To provide further support for the role of the RAD51 NTD in D-loop formation, we developed a single-molecule D-loop assay using fluorescence imaging with optical tweezers[45]. In the assay, RAD51 filaments were reconstituted on an Alexa 488-labeled 49-nt ssDNA designed to anneal at position 10.6 kb of λ DNA (Fig. 5C). The filaments interacted with doubly-biotinylated λ DNA held between streptavidin-coated beads and D-loop formation was monitored by the presence of a fluorescent signal at the expected locus on λ DNA. Wild-type RAD51 filaments were capable of D-loop formation at the expected site, as assessed by kymograph and binding events analysis (Fig. 5D, E and Fig. S7). In contrast, filaments of the

KK39,40AA RAD51 double mutant failed to show measurable binding events; a 5-fold increase in the concentration of the double mutant protein showed non-specific binding and no stable D-loop formation at the expected λ DNA site (Fig. 5E and Fig. S7).

These findings validate our model for the D-loop structure and highlight the important role of the RAD51 NTD in the strand-exchange reaction, by anchoring the donor dsDNA to the nucleoprotein filament in the correct orientation for the strand invasion and homology search. The biochemical behaviour of our esDNA-capture site mutants agrees with the impaired D-loop formation of a yeast RAD51 protein bearing alanine mutations at amino acids corresponding to R130, K303 and K313 of human RAD51[46].

DISCUSSION

The architecture of the RAD51 D-loop structure reveals the mechanism of strand invasion and homology search within a synaptic RAD51 filament. The close engagement of the RAD51-D NTD with the 5'-arm duplex, coupled to minor-groove widening by the RAD51-H L2 loop that initiates DNA unwinding, indicates that synapsis formation starts by interaction of the filament with the 5'-arm of the donor DNA. In comparison, the interaction of the RAD51-A NTD with the 3'-arm is more limited and insertion of the L2 loop of RAD51-E does not remodel the proximal end of the 3'-arm duplex. Furthermore, the first isDNA : csDNA base pair formed at the 3'-arm junction is a T : G mismatch that is unlikely to stabilise the initial strand pairing.

Our findings form the basis for a mechanistic model of strand exchange by a RAD51 filament (Fig. 5F). In the model, RAD51-D NTD binding directs the donor duplex towards the filament groove, where L2-loop insertions by RAD51-H and -G initiate strand separation and capture of the exchanged strand. Such unwinding exposes 5 nucleotides of csDNA, of which 4 are available for pairing with the invading strand of the filament. Initial base pair formation in the presence of homology would drive further donor DNA unwinding by L2-loop insertions of RAD51 -F and -E; each insertion exposes a triplet of csDNA nucleotides, in agreement with experimental evidence that homologous pairing progresses in triplet steps[38,40]. Concomitant capture of the exchanged strand by RAD51-G after insertion of a third L2 loop by RAD51-F would lead to sequestration of nine esDNA nucleotides, contributing to the formation of a stable synaptic intermediate.

Successful homology pairing drives the engagement of the 3'-arm of the donor dsDNA with the NTD of the RAD51 molecule three protomer positions away towards the 3'-end of the filament, completing D-loop formation. The observed plasticity in the interaction of the NTD

with dsDNA would facilitate capture of the 3'-arm DNA emerging at slightly different angles from the filament axis, depending on the exact number of paired nucleotides. Overall, the structural features of the RAD51 D-loop provide a strong indication that strand pairing and exchange begins at the 3'-end of the complementary strand in the donor DNA and progresses with 3'-to-5' polarity (Fig. 5F).

The observation of an unrealised base pair involving $G26_{cs}$ at the 3'-end of the unwound csDNA, which was designed to base pair with $C8_{is}$, indicates that the mechanism of unwinding of the 5'-arm prevents base pairing of the first available csDNA nucleotide, likely due to physical constraints. Further unwinding of the 5'-arm to allow $G26_{cs}$ pairing was not observed, possibly because the presence of three consecutive G : C base pairs in the 5'-arm duplex acted as a GC clamp.

The presence of a $T18_{ss}$: $G16_{cs}$ mismatch as the last base pair at the 5'-end of the paired csDNA is intriguing (Fig. S5B). Of note, $T18_{ss}$ and $G16_{cs}$ do not form a 'wobble' base pair[47], indicating that fulfilment of hydrogen bonding potential is not critical. Rather, mismatch pairing that completes a base-pair triplet might occur regardless of perfect homology, to achieve a full complement of base stacking and protein - DNA interactions with RAD51 L1 and L2 loops. This observation provides a potential mechanism for how RAD51 can tolerate mismatches at certain positions of the homologous sequence during strand exchange.

The observed 5'-to-3' polarity of strand-exchange by RAD51 is opposite to the 3'-to-5' polarity of bacterial RecA (Fig. S8), that was determined based on cryoEM structures of RecA D-loops [41]. Comparison of RAD51 and RecA D-loop structures reveals that their different polarity follows from the opposite position occupied by their dsDNA-binding domains in the filament: in a vertically-oriented filament with the 5'-end of the ssDNA at the top and the 3'-end at the bottom, the RAD51 NTDs line the bottom of the filament groove whereas the RecA CTDs hang from the top (Fig. S9). The shared mode of minor-groove binding by RAD51-NTD and RecA-CTD and L2 loop insertion into the minor groove half a turn away leads naturally to opposite polarity in the progression of strand invasion and pairing with the complementary strand of the donor DNA.

Altogether, the available evidence suggests that the ability to perform DNA-strand exchange by members of the RecA-family of recombinases results from a process of convergent evolution (Fig. S10). Thus, an ancestral RecA-fold ATPase protein might first have evolved to bind ssDNA using the L1 and L2 loops, possibly as a means of protecting the exposed DNA

template during DNA replication. Independent acquisition of dsDNA-binding by RecA and RAD51 via incorporation of carboxy- and amino-terminal domains, respectively might have further endowed these proteins with the ability to engage simultaneously ss- and dsDNA and catalyse the exchange of homologous DNA strands.

Acknowledgments We would like to thank Chen Qian for early work on the project and the Biochemistry cryoEM facility team for assistance with data collection. This work was funded by a Wellcome Trust Investigator award to LP (221892/Z/20/Z) and a UKRI BBSRC PhD studentship award to LP and RA (grant number 2271086).

Author contributions LJ produced wild-type and mutant RAD51 proteins, performed the EMSA and strand-exchange assays, designed and analysed the optical-tweezers single-molecule imaging experiments, collected cryoEM data and supervised REA. REA performed the biochemical reconstitution of the D-loop, collected cryoEM data and performed an initial cryoEM analysis. JDM carried out early work on the RAD51 NTD mutants, supervised and assisted REA and LJ throughout the project. LP conceived and directed the project, solved the 3D structure of the D-loop and wrote the paper with input from all authors.

Competing interests The authors declare no competing financial interest.

Data and materials availability The coordinates and cryoEM map of the RAD51 D-loop structure have been deposited in the Protein Data Bank (accession code 9I62) and Electron Microscopy Data Bank (accession code 52646).

FIGURE LEGEND

Figure 1. CryoEM structure of the RAD51 D-loop. A Schematic drawing of RAD51 D-loop formation during double-strand DNA break repair by Homologous Recombination. **B** CryoEM map of the RAD51 D-loop, coloured according to local resolution. **C** Drawing of the RAD51 D-loop structure. The nine RAD51 protomers in the filament are shown as cylinder cartoon, coloured alternatively in lavender and pink. The DNA strands of the D-loop are represented as light yellow tubes for the phosphate ribose backbone and the nucleotides are drawn in stick representation with filled rings. **D** D-loop drawing that highlights the trajectories of the DNA strands of the donor DNA and the invading DNA. The complementary and exchanged strands of the donor DNA are coloured light green and orange; the invading strand of the filament is coloured as the exchanged strand to emphasise that they would share sequence homology *in*

vivo. The protein filament is drawn as a transparent surface in light blue. **E** Nucleotide sequences and observed base pairing of the D-loop DNA. Sequences are coloured as in panel D. **F** Cartoon drawing of the DNA strands at the 5'-arm junction. The phospho-ribose backbone is shown as narrow tube, the nucleotide bases as filled rings. DNA strands are coloured as in panel D. The Watson-Crick hydrogen bonds between base pairs are drawn as dashed purple lines. **G** Cartoon drawing of the DNA strands at the 3'-arm junction, coloured and annotated as in panel F.

Figure 2. Donor DNA unwinding. A Multiple sequence alignment of RAD51 L2 loop sequences, coloured according to conservation based on the Clustal scheme (*Hsa*: human, *Dre*: zebrafish, *Dme*: fly, *Ath*: arabidopsis, *Sce*: budding yeast, *Spo*: fission yeast). **B** Schematic drawing of the RAD51 D-loop. The N- and C-terminal domains of each RAD51 protomer are shown as spheres of size proportional to mass. RAD51 E, F, G and H are coloured in blue, the other protomers in pale blue. **C** Ribbon drawing of the L2 loop insertion into the donor DNA for RAD51-E, F, G and H. The L2 loops are drawn as thick blue tubes and the rest of the RAD51 protein is represented as transparent surface. The DNA strands are drawn as narrow tubes with the nucleotides shown as sticks; the DNA strands are coloured dark magenta except for the homologous sequences of the invading and complementary strands, coloured in pink. **D** Nucleotide sequences and observed base pairing of the D-loop DNA. Sequences are coloured in dark magenta and pink as in panel C. The approximate position of the L2 loop insertion in the unwound donor DNA is marked with the name of the RAD51 protomer. Vertical dashed lines mark the extent of donor DNA unwinding for each L2 insertion. **E-H** Cartoon drawings of the interface between the L2 loop of RAD51-E (panel E), RAD51-F (panel F), RAD51-G (panel G), RAD51-H (panel H) with the donor DNA. L2 loop amino acids and DNA-strand nucleotides are drawn as sticks and carbon-coloured as in panel C. The nucleotide bases are shown as filled rings and the Watson-Crick hydrogen bonds are drawn as dashed lines.

Figure 3. Capture of the exchanged strand. A Nucleotide sequences and observed base pairing of the D-loop DNA. The esDNA sequence is in yellow, the rest of the D-loop DNA in light brown. The span of esDNA bound by each of RAD51-H, -G and -F is underlined. **B** Drawing of esDNA capture in a basic channel of the filament, aligned with the filament axis. The DNA is shown as full-atom model with filled rings for the ribose and the bases, coloured in yellow for the captured esDNA nucleotides and light brown for the rest. Filament RAD51 is shown in surface representation, coloured according to electrostatic potential from blue (basic) to red (acidic). **C** Multiple sequence alignment of RAD51 esDNA-binding sequences, coloured

according to conservation based on the Clustal scheme (*Hsa*: human, *Dre*: zebrafish, *Dme*: fly, *Ath*: arabidopsis, *Sce*: budding yeast, *Spo*: fission yeast). **D** Ribbon drawing of the interface between the filament and the captured esDNA, highlighting the position of the esDNA-binding beta hairpins of RAD51-H, -G and -F in blue. The DNA is drawn as in panel B. The inset shows a schematic cartoon of the RAD51 D-loop, drawn as in figure 2B, to emphasise the position of RAD51-H, -G, -F in the filament. **E-G** Drawings of the interface between the beta-hairpin amino acids of RAD51-H (panel E), RAD51-G (panel F), RAD51-F (panel G) and the esDNA. Beta-hairpin residues and esDNA nucleotides are drawn as sticks; amino acids are coloured according to atom type, whereas esDNA nucleotides are coloured as in panel D.

Figure 4. Filament engagement of the donor DNA arms. A Drawing of the D-loop structure that highlights the interaction of RAD51-A and -D with the 3'- and 5'-arms of the duplex donor DNA, respectively. RAD51-A and -D are drawn as cylinder cartoons in pale blue with their NTDs in blue. The donor DNA arms are shown as molecular surfaces in light brown while the rest of the DNA is drawn as narrow tubes. The filament is shown as a transparent surface. The inset shows a schematic cartoon of the D-loop structure to highlight the position of the RAD51-A and -D protomers. **B** Multiple sequence alignment of RAD51 NTD sequences, coloured according to conservation based on the Clustal scheme (*Hsa*: human, *Dre*: zebrafish, *Dme*: fly, *Ath*: arabidopsis, *Sce*: budding yeast, *Spo*: fission yeast). **C, D** The interaction of RAD51-D and -A NTD with the 5'- and 3'-arms, respectively. The protein is drawn as a ribbon, with the side chains as sticks. The hydrogen bond between the main-chain nitrogen of G65 and the phosphate group of T12 is drawn as a dashed line. The DNA is in spacefill representation; the nucleotides contacted by the NTD are coloured yellow, and the rest of the DNA is light brown. **e** Nucleotide sequence of the donor DNA, coloured as in panels C and D.

Figure 5. Structure-based mutagenesis of the RAD51 D-loop. A Ribbon drawing of RAD51, highlighting residues targeted for alanine mutagenesis, colour-coded according to their role in D-loop formation (F279 in pink; R303, K304, R306, K313 in magenta; K39, K40, K64, K70, K73 in light blue). **B** Strand-exchange assay showing the change in fluorescence for each RAD51 mutant following injection of the complementary strand after 9 minutes, with traces coloured according to amino acid as in panel A. Error bars denote the standard deviation of three independent replicates. The upper panel shows a diagrammatic representation of the assay, with DNA shown as light or dark brown lines and the fluorophore and quencher represented as in the key. **C** Schematic drawing of the experimental setup for single molecule analysis of RAD51 D-loop formation. **D** Kymograph of a representative D-loop formation

event. The upper panel shows a time-lapse fluorescence intensity scan of the λ DNA, with a stable binding event at the expected target site (10.6 kb, indicated by the arrow). The white dash line marks the first 60-second interval that was used in frame acquisition. The lower panel shows a summed projection of the fluorescence signal over the acquisition time. **E** Quantification of RAD51 filament binding events. The left-side histogram displays the binding profile of wild-type RAD51 (0.5 μM), showing a Gaussian distribution of binding events centered on the expected position of the complementary sequence (~10.6 kb), consistent with specific ssDNA pairing and D-loop formation. The right-side histogram shows the binding profile of the KK39,40AA RAD51 double mutant (2.5 μM). **F** Mechanism of D-loop formation by RAD51 nucleoprotein filaments.

MATERIALS & METHODS

RAD51 expression and purification. Human RAD51 was expressed and purified as described[48]. Briefly, RAD51 was co-expressed with a HisTag - MBP (Maltose Binding Protein) - BRCA2 BRC4 fusion protein in Rosetta BL21(DE3) *E. coli* cells, and purified over successive steps of Ni^{2+} chromatography, Heparin chromatography and size-exclusion chromatography. The purified protein was concentrated and stored in small aliquots at -70°C. Structure-based RAD51 mutations were introduced according to standard protocols and the mutant RAD51 proteins were purified as for the wild-type protein.

DNA substrate preparation. DNA oligos were purchased from Integrated DNA Technologies (IDT). Unmodified or biotinylated oligos were supplied PAGE-purified. Fluorescently-labelled oligos were HPLC-purified. All oligos were resuspended to 100 μM in TE buffer (10 mM Tris pH 7.5, 1 mM EDTA). dsDNA reagents were prepared at a concentration of 10 μM duplex in TE buffer, by mixing equimolar concentrations of the constituent complementary ssDNA oligos and heating at 95°C for 5 minutes before slowly cooling to room temperature.

For the strand-exchange assay, a duplex DNA 60mer in which one strand was 5'-labelled with fluorescein and the complementary strand 3'-labelled with an Iowa Black quencher was prepared by mixing quencher strand to fluorescent strand in a 1.2 : 1 molar ratio to ensure full quenching, and purified on 10% native PAGE.

All DNA sequences can be found in Table S1.

Electrophoretic mobility shift assay (EMSA). EMSA reactions were prepared in buffer: 25 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT, 5 mM $CaCl_2$ and 2 mM ATP, by mixing

fluorescein-labelled ss- or dsDNA 60mer at 250 nM concentration with increasing concentrations of wild-type or mutant RAD51 (0.5, 1.0, 2.0, 4.0, 6.0, 8.0, and 10.0 μM) in a 20 μL volume, and incubated at room temperature for 10 min. 10 μL of 50% sucrose (w/v) was added to each reaction mixture prior to electrophoresis, and samples were resolved on a 0.5 % agarose gel in TB buffer (45 mM Tris, 45 mM boric acid) at 45 V for 2 hours at 4°C.

For D-loop reconstitution experiments, EMSA reactions were prepared in same buffer by mixing 500 nM of either Cy5-labelled ssDNA or doubly-biotinylated ssDNA with increasing concentrations of RAD51 (1.0, 2.5, 5.0, 10.0, 20.0, 30.0 μM), followed by incubation at room temperature for 10 minutes. 500 nM of either Cy3-labelled dsDNA or doubly-biotinylated Cy3-labelled DNA was then added to the reaction and incubated for a further 10 minutes before electrophoresis. In experiments with mono-streptavidin (mSA) capped DNA, mSA was first added to the DNA at 5x molar excess over biotin labels and incubated for 15 minutes at room temperature before adding it to EMSA reactions. SAvPhire™ monomeric streptavidin (mSA) was purchased from Sigma Aldrich and resuspended to a working stock of 4 mg/ml from lyophilised powder according to supplier instructions.

In experiments where unlabelled DNA was used, the gel was stained in a 2x SYBR Gold solution made up in milliQ water (10 min, 25°C) and then destained in milliQ water. Gels were visualised using a Typhoon FLA 9000 fluorescent gel imager, by excitation at 473 nm for fluorescein-labelled DNA or SYBR Gold staining, 532 nm excitation for Cy3-labelled DNA or 635 nm excitation for Cy5-labelled DNA.

DNA Strand Exchange (SE) Assay. Strand exchange assays were performed using a PHERAstar FS instrument (BMG Labtech) equipped with a fluorescence intensity optic module (excitation at 485 nm, emission at 520 nm, 10 nm bandwidth). The gain was calibrated to 80 % of the maximum fluorescence intensity using a 25 nM sample of fluorescein-labelled single-stranded DNA in strand-exchange buffer: 150 mM NaCl, 25 mM HEPES pH 7.5, 2 mM DTT, 2 mM $CaCl_2$, 2 mM ATP, 0.1 mg/ml BSA.

Reactions were prepared in 50 μL volumes of strand-exchange buffer containing 500 nM RAD51 and 50 nM ssDNA 60mer, corresponding to a 1 : 2 ratio of RAD51 to RAD51-binding sites in the DNA. RAD51 was incubated with ssDNA at 30°C for 9 minutes prior to initiation of the strand-exchange reaction by 50 μL injection of dsDNA labelled with fluorescein at the 5'-end of one strand and with the Iowa Black quencher at the 3'-end of the complementary strand, using the PHERAstar FI Slow Kinetics mode. Fluorescence measurements began 2

seconds after duplex injection, and data were collected at 1-minute intervals for an additional 30 minutes. As a specificity control for strand exchange, the reaction was performed with an oligo(dT) ssDNA 60mer. Blank measurements were performed by injecting dsDNA into strand-exchange buffer lacking both ssDNA and RAD51. All reactions were performed in triplicate.

CryoEM

Sample preparation. 5 μM RAD51 was first incubated with 250 nM ssDNA in buffer: 150 mM NaCl, 25 mM HEPES pH 7.5, 2 mM DTT, 2 mM ATP, 5 mM $CaCl_2$, for 10 minutes to assemble a nucleoprotein filament; 250 nM dsDNA was then added to the sample and incubated for a further 10 minutes. In experiments with mSA-capped DNA, mSA was first added to the DNA at 5x molar excess over biotin labels and incubated for 15 minutes before the addition of other components.

Vitrification protocol. UltrAuFoil R1.2/R1.3 300 mesh gold grids (Quantifoil) were glow-discharged twice for 1 minute using a PELCO easiGlow system (0.38 mBar, 30 mA, negative polarity). 3 μl of sample was applied to the mesh side of each grid before plunge-freezing in liquid ethane using a Vitrobot Mark IV robot (FEI), set to 100 % humidity, 4°C, 2 second blot time and -3 blot force.

Grid screening and data collection. Grids were screened using a Talos Arctica 200 keV transmission electron microscope fitted with a Falcon III direct electron detector (Thermo Fisher). Data were collected using a Titan Krios G3 300 keV transmission electron microscope (Thermo Fisher) fitted with a K3 (Gatan) direct electron detector, using the EPU package (Thermo Fisher). Two datasets were collected: a first dataset with free-end DNA (#1; 2792 movies in super-resolution mode at 0.326 Å/pixel) and a second dataset with mSA-capped ss- and dsDNA (#2, 8960 movies in super-resolution mode at 0.326 Å/pixel and 2xbinned). All data were recorded at the cryoEM facility of the University of Cambridge in the Department of Biochemistry. Data collection parameters are reported in Table S2.

Data processing. Data was processed using CryoSparc 4.4.1[49]. Micrographs were motion corrected with Patch Motion Correction with output F-crop factor set to ½. The Contrast Transfer Function was estimated using Patch CTF Estimation and exposures were manually curated to remove outliers.

Dataset #1 (Figs. S1, 2). 1477 particles were manually picked from 2711 micrographs and, after 2D classification, 578 particles were selected to generate filament templates for automated

picking. Template picking yielded 948,693 particles, which were reduced to 838,403 after outlier removal. 705,678 particles were extracted using a 400-pixel box. Multiple rounds of 2D classification led to the identification of a set of 2522 particles that yielded classes with clear evidence of synaptic filaments and was therefore used for Topaz training. The trained Topaz model picked 66,799 particles, which were manually curated to 64,942 particles; of these, 60,328 were extracted with a 416-pixel box. Ab Initio reconstruction of two 3D volumes followed by Heterogeneous Refinement yielded a 3D shape of a recognisable D-loop structure. The corresponding set of 44,864 particles was subject to 3D classification into 5 classes, three of which presented pronounced D-loop features with strong density for the synaptic dsDNA; these classes were pooled and the combined 29,398 particles refined by Homogenous Refinement to 4.43 Å. The resulting particle set was used for a second round of Topaz training and picking, which yielded 167,005 particles and from which 153,235 particles were extracted with a 416-pixel box. The particles were used directly for Ab Initio reconstruction and Heterogeneous Refinement, generating a D-loop structure from a set of 117,041 particles. These were fractionated into 10 classes by 3D classification, and the particles for four of these classes that presented the best D-loop shapes (50,363 particles) were pooled and subject to Homogenous Refinement to 3.62 Å, followed by Non-Uniform Refinement to yield a final D-loop reconstruction at 3.31 Å.

Dataset #2 (Figs. S1-3). The final 3D reconstruction obtained from dataset #1 processing was low-pass filtered at 8 Å and used to generate templates for particle picking. 2,940,045 particles were picked from 8834 micrographs, 2,680,398 particles were selected after manual curation and 2,245,379 particles extracted with a 208-pixel box. After multiple rounds of 2D classification, classes from a 19,116 particle set showed clear synaptic features and were used for Ab Initio reconstruction and Homogenous Refinement of a 3D volume from 13,626 particles, which presented the expected filament shape and D-loop features such as density attributable to the dsDNA arms of the target DNA. These particles were used for Topaz training and picking of 415,845 particles, of which 413,320 particles were extracted with a 208-pixel box. The extracted particles were fed directly into Ab Initio reconstruction and Heterogeneous Refinement, which yielded a 3D volume of a filament with recognisable D-loop features at lower map contouring for 328,724 extracted particles. 3D classification into 10 classes identified a single class of 43,824 particles of a complete D-loop structure, which refined in Non-Uniform Refinement to 2.96 Å. This particle set was used in a second round of Topaz training and picking that yielded 922,204 particles, of which 900,213 were extracted with a

208-pixel box and used directly in ab Initio reconstruction and Heterogenous Refinement of a filament volume of 711,480 particles with recognisable D-loop features at lower map contouring. 3D classification of this particles set over 10 classes identified one class of 74,483 particles corresponding to a high-quality D-loop structure, which was improved by further classification and refinement by Heterogenous Refinement. The resulting set of 102,679 particles was subject to Non-Uniform refinement to yield a 3D reconstruction at 2.72 Å, followed by Reference-Based Motion Correction of 102,596 particles and another step of Non-Uniform refinement, to yield a final reconstruction at 2.64 Å resolution.

Model building and refinement. An atomic model comprising nine RAD51-ATP protomers, 26 nucleotides of ssDNA and 41 nucleotides for each strand of the 'mismatch bubble' dsDNA was built into the 3D volume of the final dataset #2 reconstruction at 2.64 Å resolution using Coot[50]. Fitting of the model to the map was improved by real-space refinement in Phenix[51] and model rebuilding in Coot[50]. In the later stages of refinement, map fitting and model stereochemistry were improved using ISOLDE[52]. The register of the DNA sequence in the structure could be unambiguously assigned based on fitting of the nucleoside bases in the homologous dsDNA region, where the map resolution was highest.

Optical-tweezers based single-molecule fluorescence imaging. To reconstitute RAD51 D-loop formation at the single-molecule level, we used a C-Trap® (LUMICKS) instrument that combines optical tweezers and confocal fluorescence microscopy. The experimental setup featured a microfluidic chip with four distinct flow channels separated by laminar flow: one for streptavidin-coated polystyrene beads (Spherotech, 4.43 μm diameter), one for biotinylated λ DNA, and two for buffer exchange or RAD51 nucleoprotein filaments reservoirs.

Doubly-biotinylated λ DNA molecules (48.5 kbp) prepared as previously described[36] were tethered between two streptavidin-coated beads in the microfluidic flow cell. Successful capture of a single λ DNA was confirmed by measuring force-extension curves[53]. The optical traps were calibrated to a stiffness of 0.4 pN/nm, enabling stable manipulation of the DNA tether under tension[54,55]. Fluorescence imaging was performed using a single 488 nm laser for excitation with an illumination intensity of 2 μW. Fluorescence signals were detected on a highly sensitive single-photon counting detector. Confocal fluorescence scans were performed over a region of 300 × 50 pixels to visualize the DNA tether and the edges of the beads. Pixel width was set to 0.1μm and each pixel was illuminated for 0.1 ms. During imaging, DNA was held at a constant force of 20 pN, and the flow channels were closed.

RAD51 D-loop formation reactions were examined by incubating 20 μM RAD51 with 1 μM Alexa488-labelled ssDNA 49mer in buffer: 25 mM HEPES pH 7.5, 150 mM NaCl, 2 mM ATP, 2 mM $CaCl_2$, 2 mM DTT at room temperature for 20 minutes, 40-fold diluted in the same buffer and injected into the microfluidic flow cell for analysis of the interaction with the λ DNA. The 49 nt-long sequence of the ssDNA is complementary to nucleotides 10,617 to 10,666 of λ DNA, to promote D-loop formation at a single, specific site on the DNA tether. During imaging, the DNA tether tension was maintained at 20 pN, ensuring optimal conditions for detecting RAD51 interactions along the DNA substrate.

Data Extraction and Analysis. Two-dimensional (2D) fluorescence scans (300 × 50 pixels) of individual λ DNA molecules were recorded for a total acquisition time of 180 seconds at a temporal resolution of one frame every 2 seconds. Given that RAD51 binding induces DNA extension[56,57], leading to drift of the D-loop position under the constant applied force regime, only the initial 30 frames were considered for analysis to ensure accurate localization. The stack of frames was summed using the open-source image processing package FIJI (ImageJ), to enhance fluorescence signal detection by summing pixel intensity values across all frames in a stack[58]. Fluorescence intensity profiles were extracted by applying a linear scan along the DNA contour and bead peripheries, excluding regions near the bead peripheries to minimise fluorescence distortion, enabling precise quantification of binding events.

Due to fluorescence interference and reduced sensitivity near the beads, regions adjacent to each bead were excluded from analysis. The extent of this exclusion region was determined by performing a linear fit of the loss of fluorescence intensity measured at the junction areas versus the physical diameter of the beads and the contour length of λ DNA. Based on this analysis, the exclusion region was estimated to be approximately 0.75 kbp from each end of the DNA molecule.

To automate peak identification, fluorescence intensity profiles were plotted against position along λ-DNA. Using the peak fitting module in Origin Pro (OriginLab), fluorescence intensity peaks were identified and fitted with a 2D-Gaussian model which tracks each peak using a symmetrical bell-shaped curve characterized by amplitude, width (full width at half maximum) and center position (X-centre) to accurately locate and quantify binding events[59,60]. The X-centre values of the detected peaks were extracted to determine the precise binding positions along λ DNA. Binding event frequencies were binned at 1 kbp intervals along the DNA molecule length. Similar peak detection and Gaussian fitting were applied to the histogram of binding events, allowing for the identification of the focal binding site.

REFERENCES

1. Kowalczykowski SC. An Overview of the Molecular Mechanisms of Recombinational DNA Repair. Cold Spring Harbor Perspectives in Biology. 2015;7(11):a016410. https://doi.org/10.1101/cshperspect.a016410

2. Wright WD, Shah SS, Heyer W-D. Homologous recombination and the repair of DNA double-strand breaks. Journal of Biological Chemistry. 2018;293(27):10524–10535. https://doi.org/10.1074/jbc.tm118.000372

3. Chakraborty S, Schirmeisen K, Lambert SA. The multifaceted functions of homologous recombination in dealing with replication-associated DNA damages. DNA Repair. 2023;129:103548. https://doi.org/10.1016/j.dnarep.2023.103548

4. Triplett MK, Johnson MJ, Symington LS. Induction of homologous recombination by site-specific replication stress. DNA Repair. 2024;142:103753. https://doi.org/10.1016/j.dnarep.2024.103753

5. Brown MS, Bishop DK. DNA Strand Exchange and RecA Homologs in Meiosis. Cold Spring Harbor Perspectives in Biology. 2014;7(1):a016659. https://doi.org/10.1101/cshperspect.a016659

6. Prakash R, Zhang Y, Feng W, Jasin M. Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. Cold Spring Harbor Perspectives in Biology. 2015;7(4):a016600. https://doi.org/10.1101/cshperspect.a016600

7. Matos-Rodrigues G, Guirouilh-Barbat J, Martini E, Lopez BS. Homologous recombination, cancer and the 'RAD51 paradox.' NAR Cancer. 2021;3(2):zcab016. https://doi.org/10.1093/narcan/zcab016

8. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science. 2017;355(6330):1152–1158. https://doi.org/10.1126/science.aam7344

9. Dibitetto D, Widmer CA, Rottenberg S. PARPi, BRCA, and gaps: controversies and future research. Trends in Cancer. 2024;10(9):857–869. https://doi.org/10.1016/j.trecan.2024.06.008

10. Baumann P, Benson FE, West SC. Human Rad51 Protein Promotes ATP-Dependent Homologous Pairing and Strand Transfer Reactions In Vitro. Cell. 1996;87(4):757–766. https://doi.org/10.1016/s0092-8674(00)81394-x

11. Sung P. Catalysis of ATP-Dependent Homologous DNA Pairing and Strand Exchange by Yeast RAD51 Protein. Science. 1994;265(5176):1241–1243. http://science.sciencemag.org/content/265/5176/1241.long. https://doi.org/10.1126/science.8066464

12. Lin Z, Kong H, Nei M, Ma H. Origins and evolution of the recA/RAD51 gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. Proceedings of the

National Academy of Sciences. 2006;103(27):10328–10333. https://doi.org/10.1073/pnas.0604232103

13. Chintapalli SV et al. Reevaluation of the evolutionary events within recA/RAD51 phylogeny. BMC Genomics. 2013;14(1):240. https://doi.org/10.1186/1471-2164-14-240

14. Cejka P, Symington LS. DNA End Resection: Mechanism and Control. Annual Review of Genetics. 2021;55(1):285–307. https://doi.org/10.1146/annurev-genet-071719-020312

15. Renkawitz J, Lademann CA, Jentsch S. Mechanisms and principles of homology search during recombination. Nature Reviews Molecular Cell Biology. 2014;15(6):369–383. https://doi.org/10.1038/nrm3805

16. Greene EC. DNA Sequence Alignment during Homologous Recombination. Journal of Biological Chemistry. 2016;291(22):11572–11580. https://doi.org/10.1074/jbc.r116.724807

17. Shibata T, DasGupta C, Cunningham RP, Radding CM. Purified Escherichia coli recA protein catalyzes homologous pairing of superhelical DNA and single-stranded fragments. Proceedings of the National Academy of Sciences. 1979;76(4):1638–1642. https://doi.org/10.1073/pnas.76.4.1638

18. Jain SK, Cox MM, Inman RB. Occurrence of three-stranded DNA within a RecA protein filament. The Journal of biological chemistry. 1995;270(9):4943–4949.

19. Piazza A et al. Dynamic Processing of Displacement Loops during Recombinational DNA Repair. Molecular Cell. 2019;73(6):1255-1266.e4. https://doi.org/10.1016/j.molcel.2019.01.005

20. Bell JC et al. BRCA2 chaperones RAD51 to single molecules of RPA-coated ssDNA. Proceedings of the National Academy of Sciences. 2023;120(14):e2221971120. https://doi.org/10.1073/pnas.2221971120

21. Greenhough LA et al. Structure and function of the RAD51B–RAD51C–RAD51D–XRCC2 tumour suppressor. Nature. 2023:1–8. https://doi.org/10.1038/s41586-023-06179-1

22. Taylor MRG et al. Rad51 Paralogs Remodel Pre-synaptic Rad51 Filaments to Stimulate Homologous Recombination. Cell. 2015;162(2):271–286. https://doi.org/10.1016/j.cell.2015.06.015

23. Rawal Y et al. Structural insights into BCDX2 complex function in homologous recombination. Nature. 2023;619(7970):640–649. https://doi.org/10.1038/s41586-023-06219-w

24. Carver A et al. Molecular basis of FIGNL1 in dissociating RAD51 from DNA and chromatin. Science. 2024. https://doi.org/10.1126/science.adr7920

25. Ogawa T, Yu X, Shinohara A, Egelman EH. Similarity of the Yeast RAD51 Filament to the Bacterial RecA Filament. Science. 1993;259(5103):1896–1899. https://doi.org/10.1126/science.8456314

26. Yu X et al. Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA. Proceedings of the National Academy of Sciences. 2001;98(15):8419–8424. https://doi.org/10.1073/pnas.111005398

27. Story RM, Weber IT, Steitz TA. The structure of the E. coli recA protein monomer and polymer. Nature. 1992;355(6358):318–325. https://doi.org/10.1038/355318a0

28. Pellegrini L et al. Insights into DNA recombination from the structure of a RAD51–BRCA2 complex. Nature. 2002;420(6913):287–293. https://doi.org/10.1038/nature01230

29. Conway AB et al. Crystal structure of a Rad51 filament. Nature Structural & Molecular Biology. 2004;11(8):791–796. https://doi.org/10.1038/nsmb795

30. Wu Y et al. Crystal Structure of Archaeal Recombinase RadA A Snapshot of Its Extended Conformation. Molecular Cell. 2004;15(3):423–435. https://doi.org/10.1016/j.molcel.2004.07.014

31. Short JM et al. High-resolution structure of the presynaptic RAD51 filament on single-stranded DNA by electron cryo-microscopy. Nucleic acids research. 2016;44(19):9017–9030. https://doi.org/10.1093/nar/gkw783

32. Xu J et al. Cryo-EM structures of human RAD51 recombinase filaments during catalysis of DNA-strand exchange. Nature Structural & Molecular Biology. 2017;24(1):40–46. https://doi.org/10.1038/nsmb.3336

33. Appleby R et al. A metal ion-dependent mechanism of RAD51 nucleoprotein filament disassembly. iScience. 2023;26(5):106689. https://doi.org/10.1016/j.isci.2023.106689

34. Chen Z, Yang H, Pavletich NP. Mechanism of homologous recombination from the RecA–ssDNA/dsDNA structures. Nature. 2008;453(7194):489–494. https://doi.org/10.1038/nature06971

35. Galletto R, Amitani I, Baskin RJ, Kowalczykowski SC. Direct observation of individual RecA filaments assembling on single DNA molecules. Nature. 2006;443(7113):875–878. https://doi.org/10.1038/nature05197

36. Candelli A et al. Visualization and quantification of nascent RAD51 filament formation at single-monomer resolution. Proc Natl Acad Sci U S A. 2014;111(42):15090–15095. https://doi.org/10.1073/pnas.1307824111

37. Hsieh P, Camerini-Otero CS, Camerini-Otero RD. The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA. Proceedings of the National Academy of Sciences. 1992;89(14):6492–6496. https://doi.org/10.1073/pnas.89.14.6492

38. Qi Z et al. DNA Sequence Alignment by Microhomology Sampling during Homologous Recombination. Cell. 2015;160(5):856–869. https://doi.org/10.1016/j.cell.2015.01.029

39. Ragunathan K, Joo C, Ha T. Real-Time Observation of Strand Exchange Reaction with High Spatiotemporal Resolution. Structure. 2011;19(8):1064–1073. https://doi.org/10.1016/j.str.2011.06.009

40. Lee JY et al. Base triplet stepping by the Rad51/RecA family of recombinases. Science. 2015;349(6251):977–981. https://doi.org/10.1126/science.aab2666

41. Yang H, Zhou C, Dhar A, Pavletich NP. Mechanism of strand exchange from RecA–DNA synaptic and D-loop structures. Nature. 2020;586(7831):801–806. https://doi.org/10.1038/s41586-020-2820-9

42. Appleby R, Joudeh L, Cobbett K, Pellegrini L. Structural basis for stabilisation of the RAD51 nucleoprotein filament by BRCA2. Nature Communications. 2023;14(1):7003. https://doi.org/10.1038/s41467-023-42830-1

43. Matsuo Y et al. Roles of the human Rad51 L1 and L2 loops in DNA binding. The FEBS Journal. 2006;273(14):3148–3159. https://doi.org/10.1111/j.1742-4658.2006.05323.x

44. Huang F et al. Identification of Specific Inhibitors of Human RAD51 Recombinase Using High-Throughput Screening. ACS Chemical Biology. 2011;6(6):628–635. https://doi.org/10.1021/cb100428c

45. Heller I et al. Optical Tweezers Analysis of DNA–Protein Complexes. Chemical Reviews. 2014;114(6):3087–3119. https://doi.org/10.1021/cr4003006

46. Cloud V et al. Rad51 Is an Accessory Factor for Dmc1-Mediated Joint Molecule Formation During Meiosis. Science. 2012;337(6099):1222–1225. https://doi.org/10.1126/science.1219379

47. Hunter WN et al. The structure of guanosine-thymidine mismatches in B-DNA at 2.5-A resolution. Journal of Biological Chemistry. 1987;262(21):9962–9970. https://doi.org/10.1016/s0021-9258(18)61060-9

48. Brouwer I et al. Two distinct conformational states define the interaction of human RAD51-ATP with single-stranded DNA. The EMBO Journal. 2018;37(7):EMBJ201798162. https://doi.org/10.15252/embj.201798162

49. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290–296. https://www.nature.com/articles/nmeth.4169. https://doi.org/10.1038/nmeth.4169

50. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60:2126–2132.

51. Afonine PV et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallographica Section D. 2018;74(6):531–544. https://doi.org/10.1107/s2059798318006551

52. Croll TI. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallographica Section D. 2018;74(6):519–530. https://doi.org/10.1107/S2059798318002425. https://doi.org/10.1107/s2059798318002425

53. Smith SB, Cui Y, Bustamante C. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. Science. 1996;271(5250):795–799. https://doi.org/10.1126/science.271.5250.795

54. Rouzina I, Bloomfield VA. Force-Induced Melting of the DNA Double Helix. 2. Effect of Solution Conditions. Biophysical Journal. 2001;80(2):894–900. https://doi.org/10.1016/s0006-3495(01)76068-7

55. Rouzina I, Bloomfield VA. Force-Induced Melting of the DNA Double Helix 1. Thermodynamic Analysis. Biophysical Journal. 2001;80(2):882–893. https://doi.org/10.1016/s0006-3495(01)76067-5

56. Hilario J, Amitani I, Baskin RJ, Kowalczykowski SC. Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. Proceedings of the National Academy of Sciences. 2009;106(2):361–368. https://doi.org/10.1073/pnas.0811965106

57. Mameren J van et al. Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. Nature. 2009;457(7230):745–748. https://doi.org/10.1038/nature07581

58. Finkelstein IJ, Visnapuu M-L, Greene EC. Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. Nature. 2010;468(7326):983–987. https://doi.org/10.1038/nature09561

59. Morin JA et al. Sequence-dependent surface condensation of a pioneer transcription factor on DNA. Nature Physics. 2022;18(3):271–276. https://doi.org/10.1038/s41567-021-01462-2

60. Gorman J et al. Dynamic Basis for One-Dimensional DNA Scanning by the Mismatch Repair Complex Msh2-Msh6. Molecular Cell. 2007;28(3):359–370. https://doi.org/10.1016/j.molcel.2007.09.008

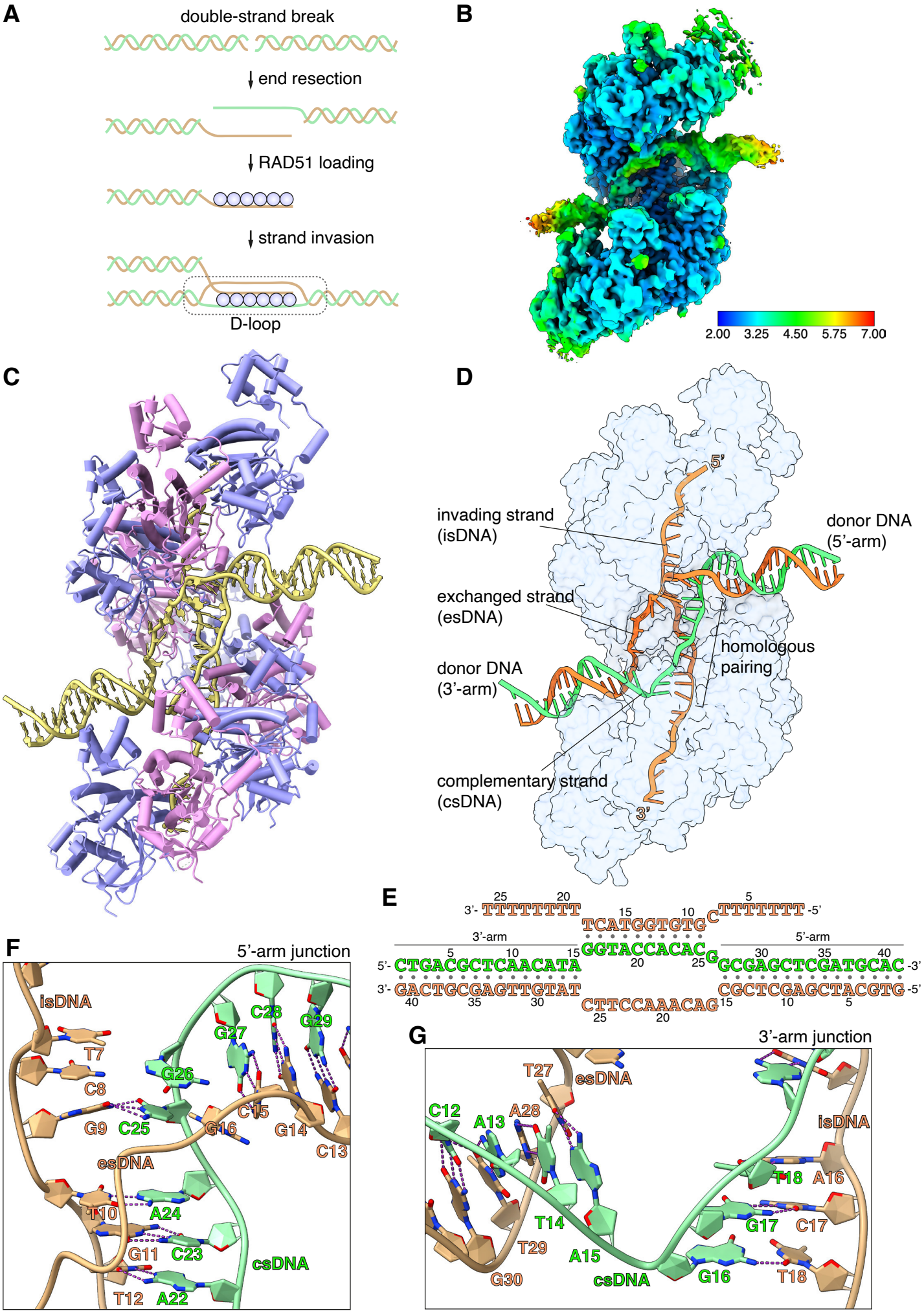
A
double-strand break
end resection
RAD51 loading
strand invasion
D-loop
B
2.00 3.25 4.50 5.75 7.00
C
D
5'
invading strand
(isDNA)
donor DNA
(5'-arm)
exchanged strand
(esDNA)
homologous
pairing
donor DNA
(3'-arm)
complementary strand
(csDNA)
3'
E
3'- TTTTTTTT TCATGGTGTG CTTTTTTT -5'
3'-arm
5'-arm
5'- CTGACGCTCAACATA GGTACCACAC G GCGAGCTCGATGCAC -3'
3'- GACTGCGAGTTGTAT CTTCCAAACAG CGCTCGAGCTACGTG -5'
F
5'-arm junction
isDNA
T7
C8
G9
C25
G26
G27
C28
G29
C15
G16
G14
C13
esDNA
T10
A24
G11
C23
csDNA
T12
A22
G
3'-arm junction
T27
esDNA
C12
A13
A28
isDNA
T18
A16
T14
G17
C17
T29
A15
G30
csDNA
G16
T18

Figure 1

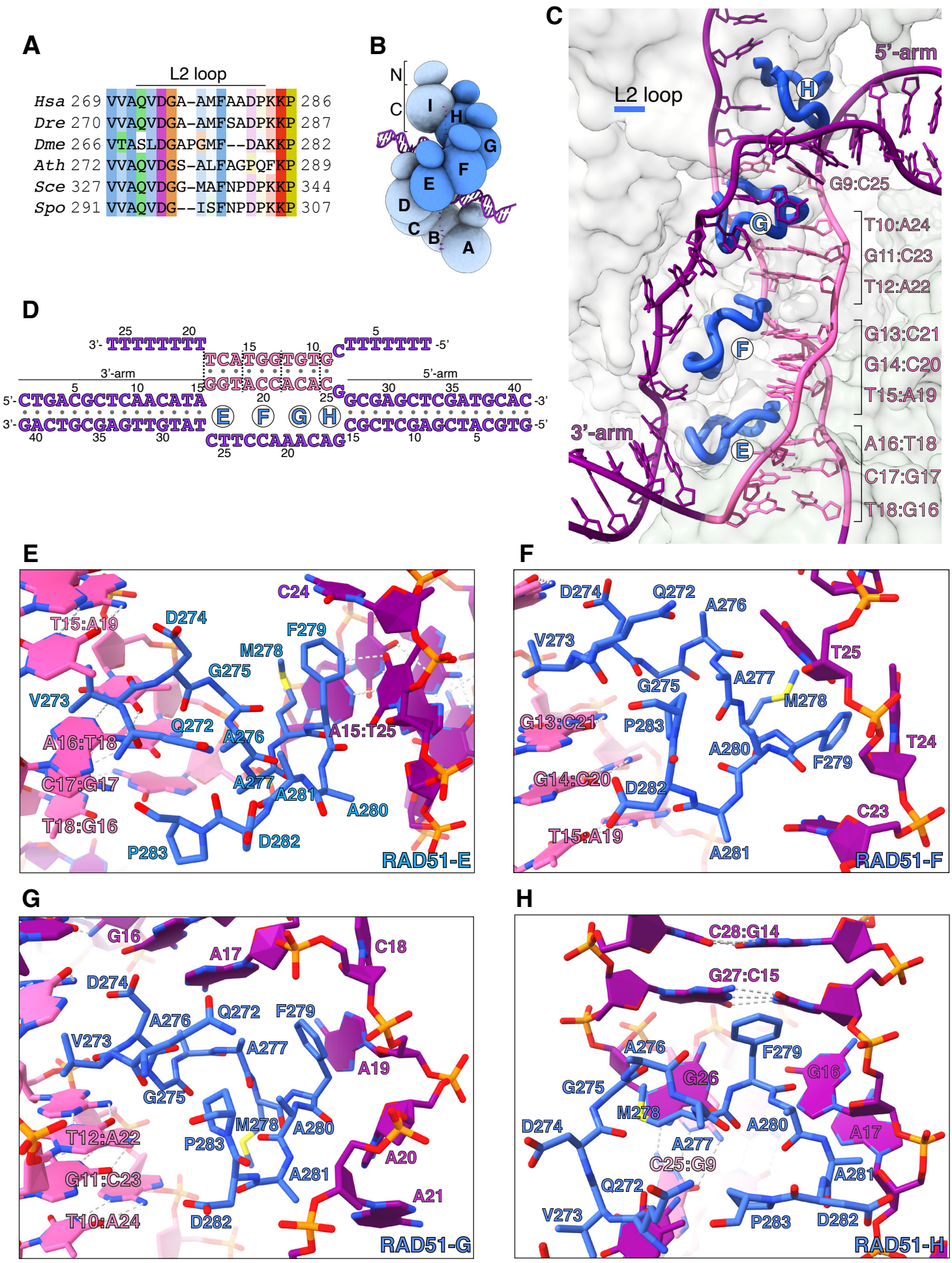
A
L2 loop
Hsa 269 VVAQVDGA-AMFAADPKKP 286
Dre 270 VVAQVDGA-AMFSADPKKP 287
Dme 266 VTASLDGAPGMF--DAKKP 282
Ath 272 VVAQVDGS-ALFAGPQFKP 289
Sce 327 VVAQVDGG-MAFNPDPKKP 344
Spo 291 VVAQVDG--ISFNPDPKKP 307
B
N
C
I
H
G
F
E
D
C
B
A
C
L2 loop
5'-arm
3'-arm
G9:C25
T10:A24
G11:C23
T12:A22
G13:C21
G14:C20
T15:A19
A16:T18
C17:G17
T18:G16
D
3'- TTTTTTTT
TCATGGTGTG
CTTTTTTT -5'
3'-arm
5'-arm
5'-CTGACGCTCAACATA
GGTACCACAC
GCGAGCTCGATGCAC -3'
3'-GACTGCGAGTTGTAT
CGCTCGAGCTACGTG -5'
CTTCCAAACAG
E
T15:A19
D274
C24
F279
M278
G275
V273
Q272
A276
A15:T25
A16:T18
C17:G17
A277
A281
A280
T18:G16
P283
D282
RAD51-E
F
D274
Q272
A276
V273
T25
G275
A277
M278
G13:C21
P283
A280
F279
T24
G14:C20
D282
C23
T15:A19
A281
RAD51-F
G
G16
A17
C18
D274
A276
Q272
F279
V273
A277
A19
G275
M278
A280
T12:A22
P283
A20
A281
G11:C23
A21
T10:A24
D282
RAD51-G
H
C28:G14
G27:C15
A276
F279
G26
G16
G275
M278
D274
A280
A17
A277
C25:G9
A281
Q272
V273
P283
D282
RAD51-H

Figure 2

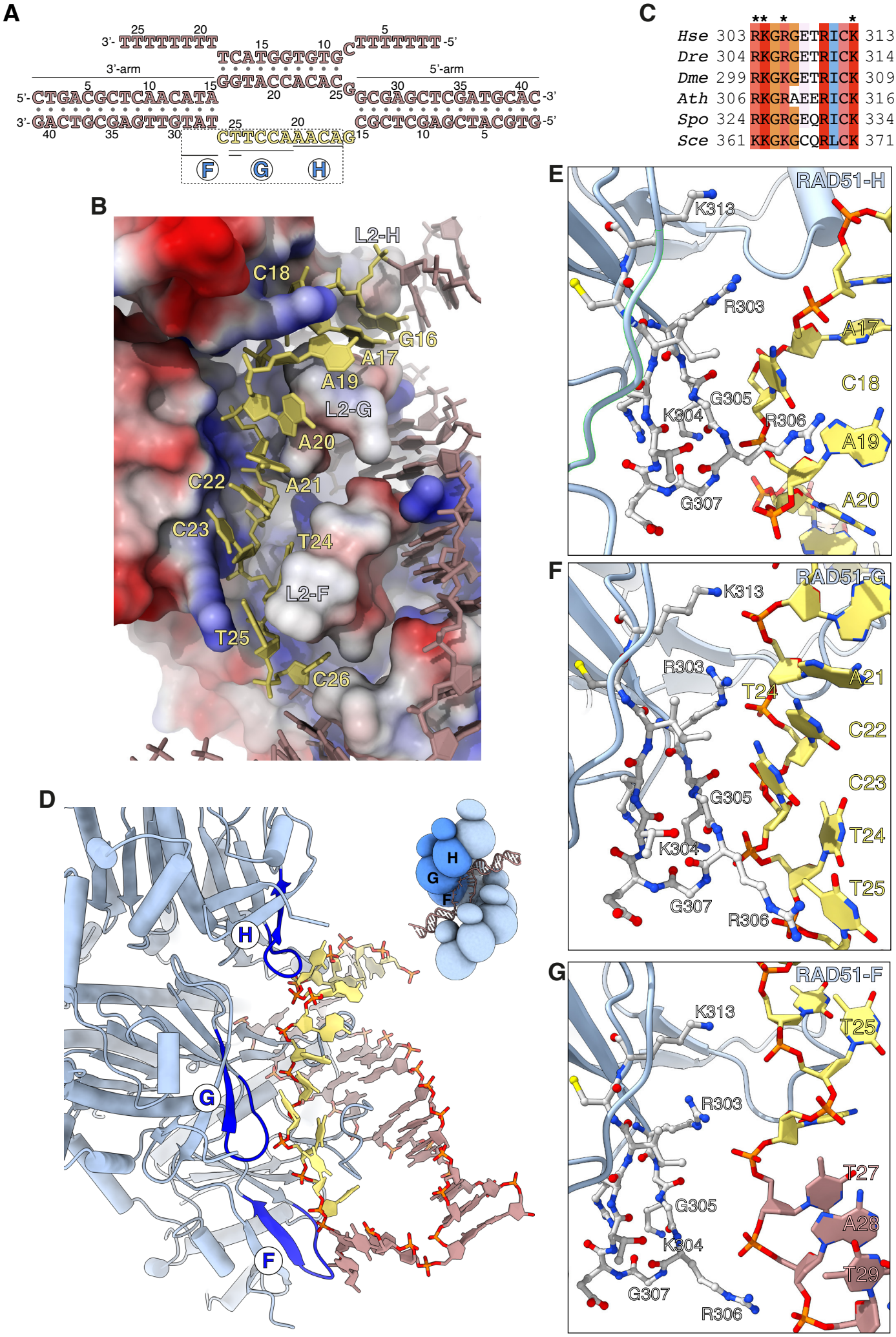
A
3'- TTTTTTTT
TCATGGTGTG
CTTTTTTT -5'
3'-arm
5'-arm
GGTACCACAC
5'-CTGACGCTCAACATA
GCGAGCTCGATGCAC -3'
3'-GACTGCGAGTTGTAT
CGCTCGAGCTACGTG -5'
CTTCCAAACAG
F
G
H
B
L2-H
C18
G16
A17
A19
L2-G
A20
C22
A21
C23
T24
L2-F
T25
C26
C
Hse 303 RKGRGETRICK 313
Dre 304 RKGRGETRICK 314
Dme 299 RKGKGETRICK 309
Ath 306 RKGRAEERICK 316
Spo 324 RKGRGEQRICK 334
Sce 361 KKGKGCQRLCK 371
D
H
G
F
E
RAD51-H
K313
R303
A17
C18
G305
K304
R306
A19
G307
A20
F
RAD51-G
K313
R303
A21
T24
C22
C23
G305
T24
K304
T25
G307
R306
G
RAD51-F
K313
T25
R303
T27
G305
A28
K304
T29
G307
R306

Figure 3

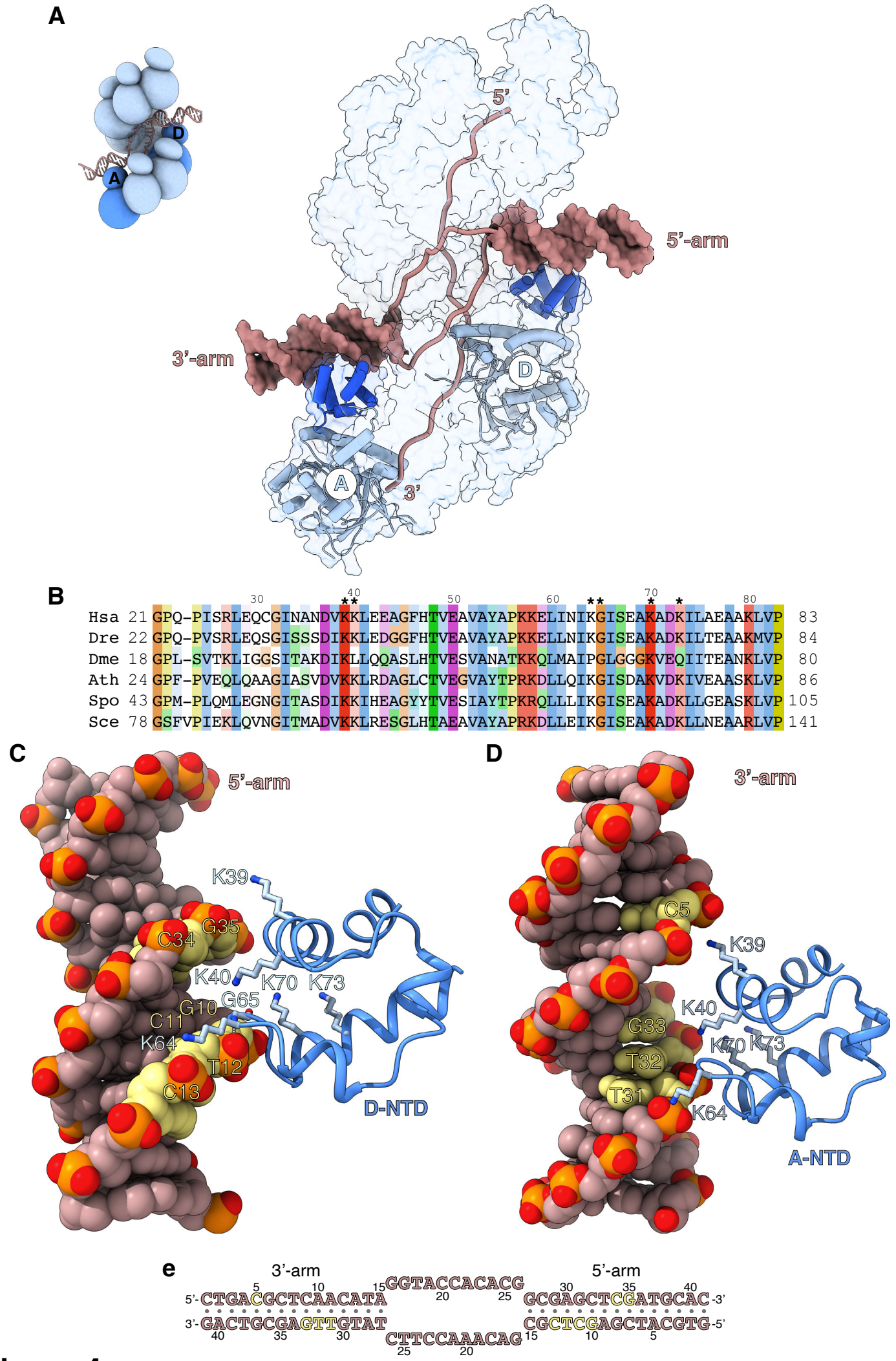
A
D
A
5'
5'-arm
3'-arm
D
A
3'
B
30 40 50 60 70 80
Hsa 21 GPQ-PISRLEQCGINANDVKKLEEAGFHTVEAVAYAPKKELINIKGISEAKADKILAEAAKLVP 83
Dre 22 GPQ-PVSRLEQSGISSSDIKKLEDGGFHTVEAVAYAPKKELLNIKGISEAKADKILTEAAKMVP 84
Dme 18 GPL-SVTKLIGGSITAKDIKLLQQASLHTVESVANATKKQLMAIPGLGGGKVEQIITEANKLVP 80
Ath 24 GPF-PVEQLQAAGIASVDVKKLRDAGLCTVEGVAYTPRKDLLQIKGISDAKVDKIVEAASKLVP 86
Spo 43 GPM-PLQMLEGNGITASDIKKIHEAGYYTVESIAYTPKRQLLLIKGISEAKADKLLGEASKLVP 105
Sce 78 GSFVPIEKLQVNGITMADVKKLRESGLHTAEAVAYAPRKDLLEIKGISEAKADKLLNEAARLVP 141
C
5'-arm
K39
G35
C34
K40
K70
K73
G65
G10
C11
K64
T12
C13
D-NTD
D
3'-arm
C5
K39
K40
G33
K70
K73
T32
T31
K64
A-NTD
e
3'-arm
5'-arm
5'-CTGACGCTCAACATA GGTACCACACG GCGAGCTCGATGCAC-3'
3'-GACTGCGAGTTGTAT CTTCCAAACAG CGCTCGAGCTACGTG-5'

Figure 4

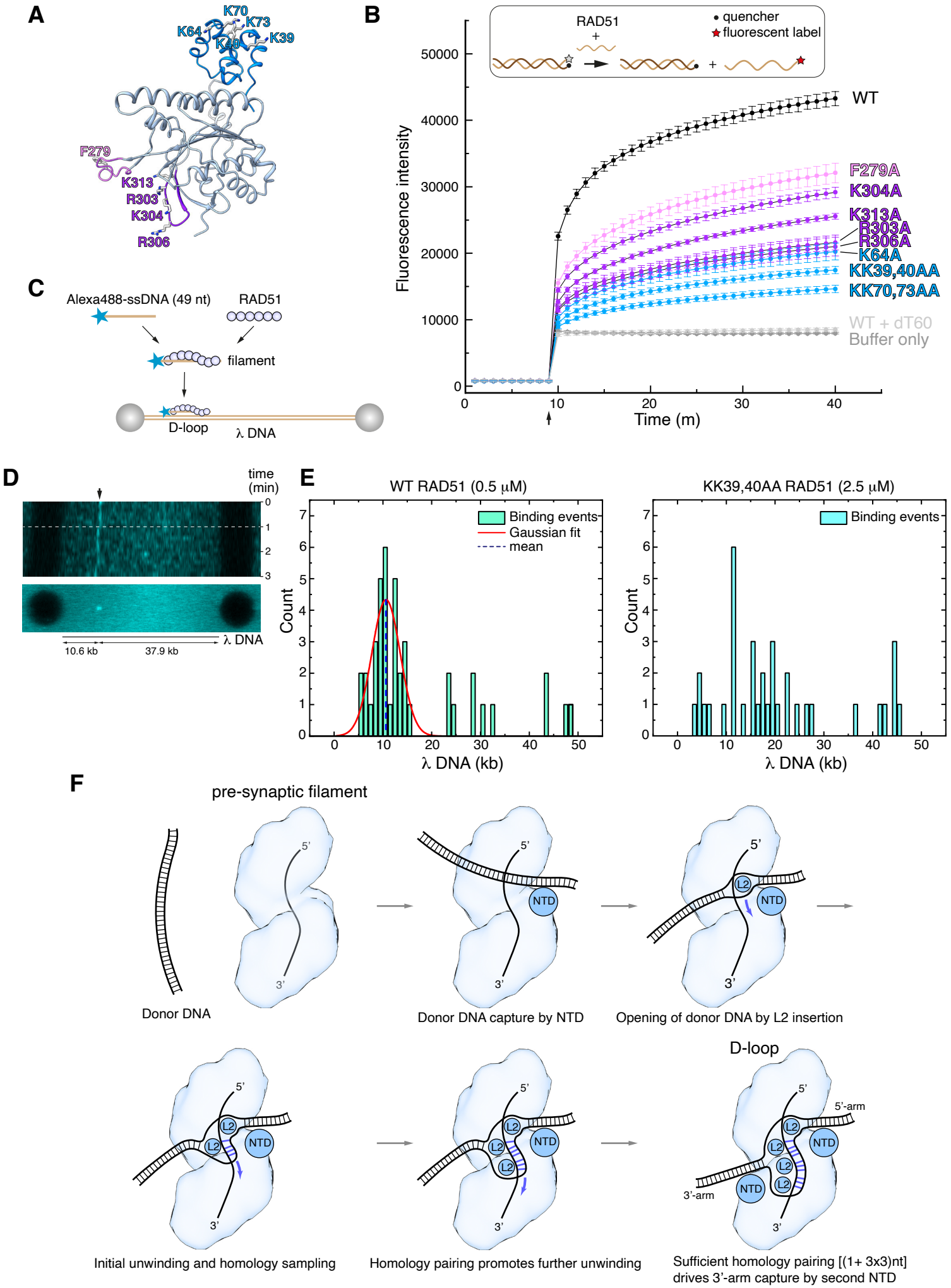
A
K70
K73
K64
K40
K39
F279
K313
R303
K304
R306
B
RAD51
+
quencher
fluorescent label
Fluorescence intensity
Time (m)
WT
F279A
K304A
K313A
R303A
R306A
K64A
KK39,40AA
KK70,73AA
WT + dT60
Buffer only
C
Alexa488-ssDNA (49 nt)
RAD51
filament
D-loop
λ DNA
D
time (min)
λ DNA
10.6 kb
37.9 kb
E
WT RAD51 (0.5 μM)
Binding events
Gaussian fit
mean
Count
λ DNA (kb)
KK39,40AA RAD51 (2.5 μM)
Binding events
Count
λ DNA (kb)
F
pre-synaptic filament
5'
3'
Donor DNA
NTD
Donor DNA capture by NTD
L2
Opening of donor DNA by L2 insertion
Initial unwinding and homology sampling
Homology pairing promotes further unwinding
D-loop
5'-arm
3'-arm
Sufficient homology pairing [(1+ 3x3)nt]
drives 3'-arm capture by second NTD

Figure 5