Supporting Information

Pike et al. 10.1073/pnas.1417594112

SI Materials and Methods

Protein Purification. RECQ1^{FL} proteins (WT or mutated) were expressed and purified from baculovirus-infected insect cells as described (1). RECQ1^{T1} proteins (spanning amino acids 49–619) were expressed and purified from recombinant *E. coli* as described (2). The isolated WH domain (amino acids 481–624) and a fragment containing the D2, Zn, and WH domain (amino acids 282–624) were cloned in the vector pNIC28-Bsa4 (3), expressed in *E. coli*, and purified as described (2).

Tailed Duplex. The proper combination of DNA oligonucleotides was determined after screening >50 combinations for cocrystallization. Diffracting crystals were obtained using the combination of top: 5′-CGGTATTGGATCTCGACGCTCTCCCTT-3′ and bottom: 5′-AGCGTCGAGATCC-3′. The two oligonucleotides were mixed at a final concentration of 100 μM each in 10 mM Hepes (pH 7.5) and 50 mM NaCl, and were annealed by heating to 95 °C and gradual cooling to 15 °C. To form protein/DNA complexes, RECQ1^{T1} protein and the annealed DNA were combined and diluted to concentrations of 3.3 and 4.0 μM, respectively, in a buffer consisting of 10 mM Hepes (pH 7.5) and 10 mM NaCl. The complexes were then concentrated by ultrafiltration (Centricon) to ∼140 μM protein. Crystals were grown by vapor diffusion from nanoliter sitting drops at 4 °C. The best crystals were obtained from drops comprising 150 nL of protein/DNA solution and 50 nL of reservoir solution [0.2 M sodium sulfate, 20% (wt/vol) PEG 3350, 10% (vol/vol) ethylene glycol] equilibrated against reservoir solution. Crystals were soaked in reservoir solution supplemented with an additional 15% (vol/vol) ethylene glycol before flash-cooling in liquid nitrogen.

ssDNA Complex. RECQ1^{T1} protein was buffer-exchanged into 10 mM Hepes (pH 7.5), 10 mM NaCl, 5% (vol/vol) glycerol, and 1 mM DTT; diluted to 15 μM; and combined with 20 μM single-stranded oligonucleotide (5′-GGATCTCGACGCTCTCCCTT-3′). The complex was then concentrated by ultrafiltration (Centricon) to 150 μM protein (10 mg/mL). Crystals were grown at 4 °C by vapor diffusion from nanoliter sitting drops against a reservoir solution containing 17% (wt/vol) PEG 3350, 5% (vol/vol) ethylene glycol, 0.18 M potassium citrate, and 0.1M Bis-Tris propane (pH 6.5). Crystals were soaked in reservoir solution supplemented with an additional 20% (vol/vol) ethylene glycol before flash-cooling in liquid nitrogen.

Data Collection. Diffraction data were collected on beamline I04 at the Diamond Light Source (Harwell, United Kingdom). Data were processed with XDS (4) (tailed duplex), MOSLFM (5) (ssDNA complex), and SCALA (6) (statistics are provided in Table S2).

Structure Determination (Tailed Duplex). The structure was solved using the coordinates for the DNA-free enzyme [PDB ID code 2V1X (2)] as a search model in PHASER (7). The search model was split into two domains comprising either the N-terminal RecA-like domain (D1, residues 63–281) or residues 282–592 (RecA D2 C-terminal domain), and these models were used to locate two copies of RECQ1^{T1} in the asymmetrical unit. Initial electron density maps clearly showed the presence of a single-tailed DNA duplex associated with each of the two protein chains. The structure was built using COOT (8) and refined initially with REFMAC (9) and then with PHENIX (10) using appropriate noncrystallographic symmetry (NCS) restraints, a single B-factor per residue, and two translation/libration/screw (TLS) groups for each protein chain plus an additional group for the associated DNA duplex (statistics are provided in Table S2). The final dsDNA model contains a RECQ1^{T1} dimer comprising residues 63–592 (residues 64–593 for chain B). Each chain also contains a zinc ion, two sulfates, and an ethylene glycol molecule. A single-stranded tailed duplex is associated with each chain, with the shorter oligo being fully resolved (13 bases) and the longer 27-nt strand being partially ordered (6–26 bases resolved). In addition, the 5′ end of the longer strand (G^3) from an adjacent molecule in the crystal interacts with the C^{21} at the hairpin junction.

Structure Determination (ssDNA Complex). The apparent symmetry of the ssDNA complex was orthorhombic, but both intensity moments and L-tests indicated that the data were twinned. Data were processed assuming monoclinic symmetry with a β-angle close to 90° with MOSFLM and SCALA. Free flags were generated by PHENIX, taking into consideration the known pseudosymmetry. The structure was solved using the coordinates of DNA-free enzyme (PDB ID code 2V1X) as a search model in PHASER. This procedure correctly placed three of the four protein chains present in the asymmetrical unit. The fourth protein chain was placed manually using knowledge of the RECQ1's known dimeric arrangement, followed by rigid-body refinement. Initial maps indicated the presence of two pseudotailed DNA duplexes that are sandwiched between the WH domain's β-hairpins of adjacent molecules in the crystal. The single-stranded oligonucleotides were initially modeled as poly-T, and the sequence was then tentatively assigned based on comparison with the dsDNA tailed-duplex complex and the visible start point of the longer single strand that extends between the RecA domains (assigned as the first nucleotide of the sequence because no additional electron density was visible in the 5′ direction). Twinned refinement was carried out with PHENIX, with additional torsional restraints using the high-resolution apo enzyme coordinates (PDB ID code 2V1X) as a reference molecule, NCS restraints, and grouped B-factor and TLS refinement. The final ssDNA model comprises four RECQ1 chains and four single-stranded oligonucleotides. The pseudoduplex (assigned chains P/Q) associated with protein chains A and B is better resolved than the pseudoduplex associated with the other two chains (statistics are provided in Table S2).

DNA Substrate Preparation. All of the oligonucleotides used in this study were chemically synthesized and HPLC-purified by IDT Technologies, and suspended in Milli-Q water (Millipore) (Table S3). A single nucleotide of each substrate was 5′ end-labeled with [γ^{32}P]ATP (3,000 Ci/mmol) for 45 min using T_4 polynucleotidyl kinase (New England Biolabs), and the reaction was terminated by heat-inactivating the enzyme at 95 °C for 6 min. The unincorporated [γ^{32}P]ATP was removed using MircoBio-spin 30 columns (Bio-Rad). To prepare the forked duplex substrate, the labeled oligonucleotide B was annealed with a 1.4 excess of unlabeled complementary strand A in annealing buffer [10 mM Tris·HCl (pH 7.5), 50 mM NaCl], followed by heating at 95 °C for 6 min and then slowly cooling to room temperature. For the annealing assays, the labeled strand A was used as a substrate.

To prepare the HJ (1) substrate, labeled (oligonucleotide C) and a 1.5-fold excess of unlabeled DNA intermediates (oligonucleotides D, E, and F) were incubated in annealing buffer supplemented with 5 mM $MgCl_2$ for 30 min at 37 °, and then for an additional 30 min at room temperature. The HJ was purified in a Sepharose 4B (5 mL) column, and the collected fractions

were analyzed on a 10% PAGE gel. The fractions were selected based on their purity. The 6-fluorescein amidite–labeled (6-FAM) HJ substrate for the analytical ultracentrifugation (AUC) experiments was prepared using the oligonucleotides G, H, I, and L incubated in annealing buffer at 96 °C for 5 min, followed by gradual cooling to room temperature. The same procedure was used for preparation of the 5′-FAM–labeled, blunt-ended dsDNA substrates listed in Table S3 (oligonucleotides M, N, O, P, Q, and R).

DNA Unwinding Assays. All unwinding assays were performed as described (2), using a fork duplex substrate with a duplex region of 20 bp (oligonucleotides A and B in Table S3). Briefly, we prepared 20 μL of reaction mix containing ^{32}P-labeled DNA substrate (2 nM) in the helicase buffer [20 mM Tris·HCl (pH 7.5), 8 mM DTT, 5 mM $MgCl_2$, 10 mM KCl, 80 mg/mL BSA, 10% glycerol] with 5 mM ATP. The reaction was initiated by the addition of the indicated RECQ1 concentration, and the mixture was incubated at 37 °C for 20 min. The reactions were terminated by the addition of 20 μL of quenching solution [0.4 M EDTA (pH 8.0), 10% (vol/vol) glycerol, 1% SDS]. The reaction products were resolved on native 10% PAGE and visualized by autoradiography.

DNA Strand-Annealing Assays. Strand-annealing reactions (20 μL), containing one DNA strand and the indicated concentration of protein, were carried out at 37 °C in the helicase buffer in the absence of ATP and were initiated by adding the cDNA strand. Reactions were terminated by the addition of 20 μL of quenching solution [0.4 M EDTA (pH 8.0), 10% (vol/vol) glycerol, 1% SDS]. Reaction products were subsequently resolved on native 10% PAGE and visualized by autoradiography. The concentrations of DNA in the reactions were 1.1 nM and 2.2 nM for the labeled and unlabeled strands, respectively. Experiments were set up by adding either the labeled or unlabeled strand first, with similar results. In addition, we found that the optimal temperature for observation of strand annealing depended on the length of the duplex: 37 °C was used for long oligonucleotides, whereas 25 °C gave better signal-to-background ratio for oligos shorter than 20 nt. For duplexes of 20–30 nt, the results at either temperature were similar.

The gels were exposed to a phosphorimaging screen (Kodak K) overnight and then analyzed on a Bio-Rad Personal Molecular Imager using the analysis software provided (Quantity One). For each reaction, the fraction of annealed oligonucleotide was calculated and the value for the enzyme-free control was subtracted. Experiments were performed in three independent replicates. Similar results were obtained with duplex lengths of 15, 20, 25, and 30 nt, with or without unpaired tails on either end.

Branch Migration Assays. Branch migration assays were performed in a 20-μL volume containing 2 nM HJ (1) and the indicated protein concentrations in branch migration buffer [35 mM Tris·HCl (pH 7.5), 20 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA, 2 mM DTT, 15 mM phosphocreatine, 30 U/mL creatine phosphokinase, 5% glycerol] at 37 °C for 20 min. The reaction was initiated by the addition of 2 mM ATP. DNA substrates were deproteinized by adding 3× stop buffer [1.2% SDS, 30% glycerol supplemented with proteinase K (3 mg/mL)], followed by incubation at room temperature for 10 min before being resolved on 8% native PAGE. Labeled DNA fragments were detected by autoradiography.

Analytical Ultracentrifugation. Sedimentation velocity analytical ultracentrifugation was carried out in an Optima XL-A analytical ultracentrifuge (Beckman Coulter) and Epon charcoal-filled, double-sector centerpieces at 40 krpm with an An-60 Ti rotor (Beckman Coulter). The experiments were carried out in AUC buffer [10 mM Hepes (pH 7.5), 150 mM KCl, 5% glycerol, 0.5 mM tris (2-carboxyethyl) phosphine (TCEP), 5 mM $MgCl_2$] at 8 °C. For RECQ1 alone, samples were scanned at 280 nm every 3′ at a spacing of 0.003 cm, whereas for the FAM-labeled HJ alone, samples were scanned at 495 nm. For experiments with RECQ1 in the presence of FAM-labeled HJ, the samples were scanned at 495 nm (fluorescein), where protein absorbance does not contribute to the observed signal, and at 280 nm. Sedimentation profiles were processed and analyzed using SedFit (developed by Peter Schuck, NIH/National Institute of Biomedical Imaging and Bioengineering), and the apparent sedimentation coefficient (S) was corrected for buffer composition and temperature using SEDNTERP (developed by D. Hayes, University of New Hampshire and J. Philo, Alliance Protein Laboratories) to obtain the sedimentation coefficient in water at 20 °C ($s_{20,w}$) values.

Sedimentation equilibrium (SE) data were acquired in six-channel centerpieces filled with 110-μL samples at three different rotor speeds (7, 10, and 13 krpm) with an An-60 Ti rotor (Beckman Coulter) in the absorbance and interference modes starting with the lowest rotor speed and finishing with the highest rotor speed. At each speed, samples were sedimented to equilibrium, when successive scans at 2-h intervals were superimposable. The SE data were analyzed using SEDPHAT v6.21 software (11) with the "Species analysis" model to obtain the buoyant molecular mass corrected for the buoyancy factor by the partial specific volume calculated by SEDNTERP. Hydrodynamic properties of the RECQ1/HJ complexes were calculated using the following equation (12):

$$\nu_{pHJ} = \left(nM_p\nu_p + M_{HJ}\nu_{HJ}\right)/\left(nM_p + M_{HJ}\right)$$

where n is the number of RECQ1 molecules in the complex, M_P and M_{HJ} are the molecular weights of RECQ1 and the HJ, respectively, ν_p is the partial specific volume of RECQ1, and ν_{HJ} is the partial specific volume of the HJ.

Fluorescence Anisotropy Assays. Proteins were dialyzed from storage buffer against assay buffer [20 mM Tris (pH 7.5), 50 mM KCl, 2% glycerol, 5 mM $MgCl_2$, 0.5 mM TCEP]. Stocks of fluorescein-labeled DNA in Tris-EDTA buffer were diluted to the appropriate concentration using the assay buffer. Fluorescence anisotropy assays were performed in 96-well, low-binding, black plates (Corning). An increasing amount of protein was added to the assay buffer containing the indicated concentration of ssDNA, dsDNA, or HJ, as mentioned in the main text. The reaction mixture containing the protein and DNA was incubated at 37 °C for 15 min before data acquisition. The reaction volume of all binding assays was 100 μL. All reactions were run in triplicate. Fluorescent assays were read at 485/528 nm using a BioTek Synergy 2 plate reader. Fluorescent polarization anisotropy values were normalized using the following equation:

$$A = (A_i - A_0)/A_0,$$

where A_0 and A_i are fluorescence anisotropy values for free substrate and for each titration point, respectively. Binding constants were determined with GraphPad Prism software using one-site specific binding with a Hill slope equation, and the graphs were plotted using Prism.

GraFix Sample Preparation for EM Imaging. Glycerol and glutaraldehyde gradient sedimentation of full-length RECQ1 was performed according to the GraFix protocol (13, 14). Briefly, a 5-mL 20–45% glycerol gradient was poured, which contained 0.1% glutaraldehyde in the heavy solution. A 150-μL, 3.7 μM aliquot from the S200 tetrameric RECQ1 peak was dialyzed into 500 mM KCl and, subsequently, to 50 mM KCl. This low-salt fraction was loaded on top of the gradient and centrifuged at 50,000 rpm at 4 °C in a SW55Ti ultracentrifuge rotor (Beckman Coulter) for 15 h. Two hundred-microliter fractions were collected manually

from the top of the gradient and analyzed by SDS/PAGE and silver staining, using a 4–12% gradient gel (Biorad) in 3-(N-morpholino)propansulfonic acid buffer at room temperature.

Negative Stain Grid Preparation. Negative stain grids were prepared using 400-mesh copper grids (Agar Scientific). Carbon was evaporated onto freshly cleaved mica with a Q150TE coater (Quorum Technologies) and baked overnight at 50 °C before floating. Dried carbon grids were glow-discharged for 30 s at 45 mA using a 100× glow discharger (Electron Microscopy Sciences). Before staining, peak fractions from GraFix-treated samples were buffer-exchanged with a PD MiniTrap G-25 column (GE Healthcare) to remove glycerol. Freshly glow-discharged grids were laid on top of a 50-μL drop for 4 h at 4 °C in a custom-made humidity chamber. Subsequently, grids were sequentially laid on top of four distinct 50-μL drops of 2% (wt/vol) uranyl formate solution and stirred for 10 s each. The staining solution was then blotted dry, and the grids were stored at room temperature before imaging.

EM Data Collection. Negative stain analyses of all complexes were performed using a JEM-2100 LaB6 electron microscope (JEOL) operated at 120 kV. Images were recorded at a nominal magnification of 40,000× on an Ultrascan 4k × 4k CCD camera (Gatan), resulting in a 2.73-Å pixel size at the specimen level. A total of 156 micrographs were collected with a defocus of 0.5–1.5 μm and using 35 $e^{-}/Å^{2}$.

Image Processing. Contrast transfer function estimation was performed on the raw micrographs using CTFFIND3 (15), within the RELION image processing suite, version 1.3 (16). RECQ1 particles were selected semiautomatically using the e2boxer function in EMAN2, version 2.1beta1 (17). All further image processing was performed using RELION. An initial dataset of 16,489 particles was extracted, normalized, and subjected to six iterative rounds of 2D classification. This effort yielded a final dataset of 5,749 nicely averaging particles of recognizable stoichiometry.

1. Lucic B, et al. (2011) A prominent β-hairpin structure in the winged-helix domain of RECQ1 is required for DNA unwinding and oligomer formation. *Nucleic Acids Res* 39(5):1703–1717.
2. Pike AC, et al. (2009) Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci USA* 106(4):1039–1044.
3. Savitsky P, et al. (2010) High-throughput production of human proteins for crystallization: The SGC experience. *J Struct Biol* 172(1):3–13.
4. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
5. Leslie AGW (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):48–57.
6. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72–82.
7. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
8. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
9. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3): 240–255.
10. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
11. Vistica J, et al. (2004) Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem* 326(2):234–256.
12. Barranco-Medina S, Galletto R (2010) DNA binding induces dimerization of Saccharomyces cerevisiae Pif1. *Biochemistry* 49(39):8445–8454.
13. Kastner B, et al. (2008) GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat Methods* 5(1):53–55.
14. Simon AC, et al. (2014) A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature* 510(7504):293–297.
15. Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334–347.
16. Scheres SH (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180(3):519–530.
17. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38–46.
18. Manthei KA, Hill MC, Burke JE, Butcher SE, Keck JL (2014) Structural mechanisms of DNA binding and unwinding in bacterial RecQ helicases. *Proc Natl Acad Sci USA*, in press.

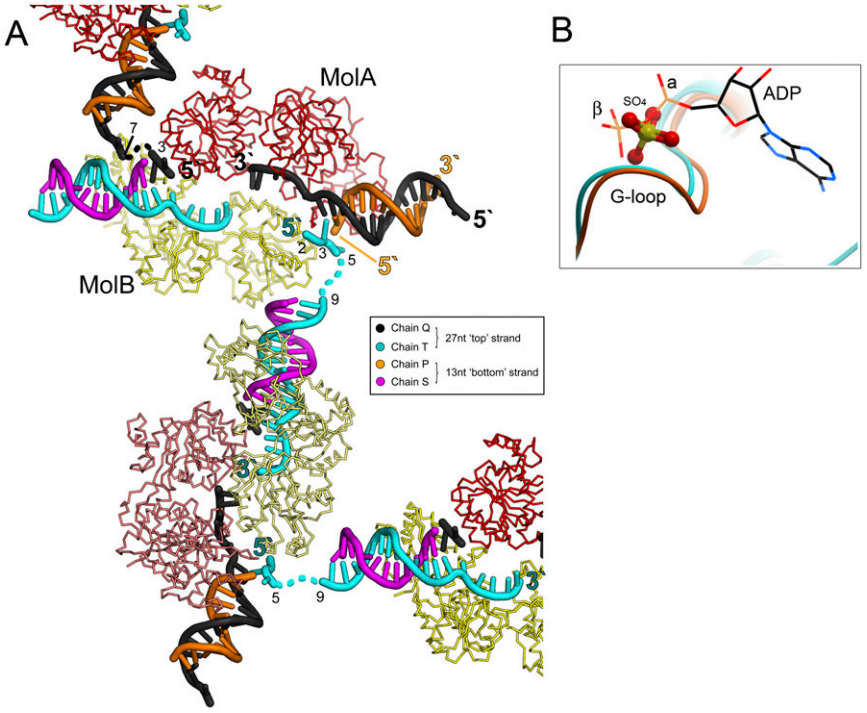


Fig. S1. Structure of the RECQ1/DNA complex. (*A*) Molecular packing in RECQ1/dsDNA crystals. The two protein chains in each RECQ1 dimer are colored red and yellow. The tailed DNA duplexes are colored black/orange and cyan/magenta (the shorter, 13-nt "bottom" strands are in orange/magenta). The single-stranded 5′ end of the long "top" strand of an adjacent molecule (G^3) in the crystal (cyan) is base-paired with the free C^{21} nucleotide at the hairpin branch point. No electron density is observed for the intervening four nucleotides (nucleotides 5–8) in the long strand, but the distance between the resolved bases (18 Å) is entirely compatible with such a connection. In the crystal structure (PDB ID code 2WWY), this short interacting region has been designated as chain O; however, in reality, it corresponds to the extreme 5′ end of a symmetry-related chain T. (*Upper Left*) Similarly, the corresponding region in the tailed duplex bound to molB (chain R) derives from the 5′ terminus of an adjacent top (black) strand (chain Q). (*B*) Overlay of the current structure (PDB ID code 2WWY, orange) with the structure of a RECQ1/ADP complex (PDB ID code 2V1X, cyan) shows a sulfate ion occupying a position similar to the position of the β-phosphate of ADP. A portion of the G-rich loop (G116-G118) of RECQ1 is shown.

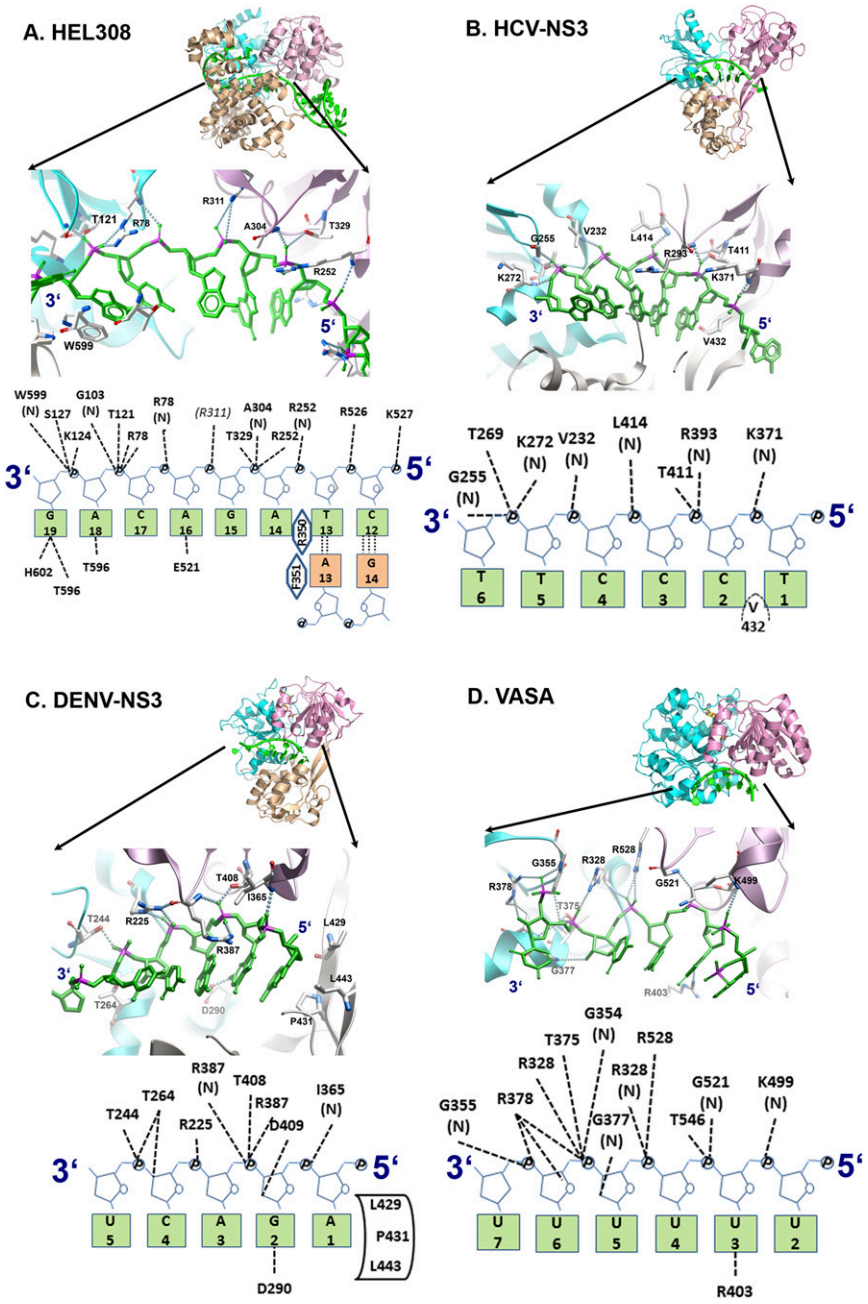


Fig. S2. Protein/DNA contacts in SF2 helicase structures. The diagrams are based on PDB ID codes 2P6R (HEL308; panel *A*), 3KQH (HCV-NS3; panel *B*), 2JLV (DEN4-NS3; panel *C*), and 2DB3 (VASA; panel *D*). Details and references are provided in the main text. Each panel shows an overview of the protein/nucleic acid complex, a close-up view of the ssDNA/RNA interaction with the core helicase domain, and a schematic summary of the specific contacts. The corresponding data for RECQ1 are shown in Fig. 1. The chains are color coded: cyan, RecA-like domain 1; pink, RecA domain 2; gold, additional domains; green, nucleic acid.

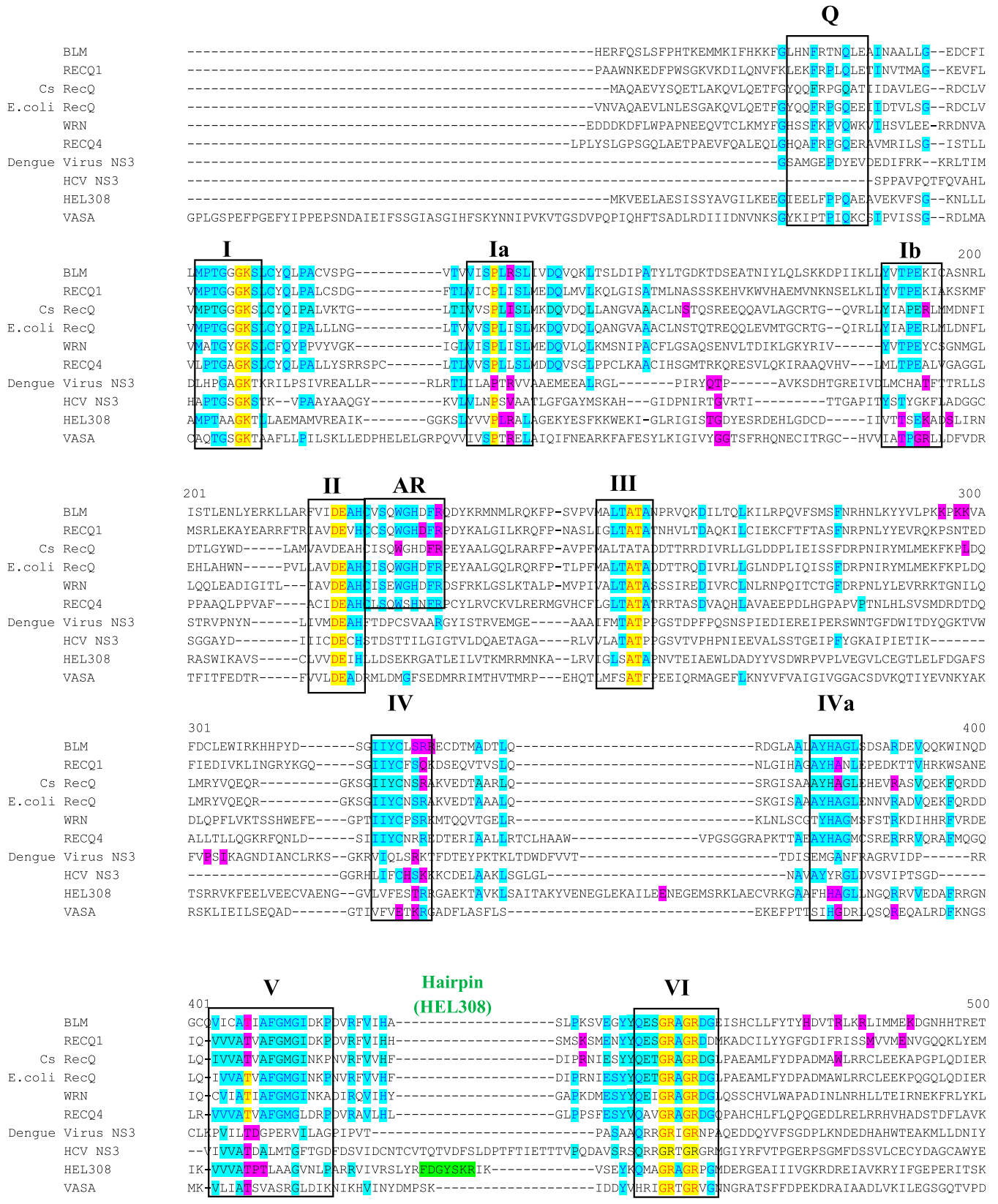


Fig. S3. Structure-based sequence alignment of the RecA-like domains of RecQ helicases and other SF2 helicases. The boxes indicate the helicase motifs Q, I–VI, which are present in most SF2 helicases; AR is the aromatic-rich motif that is characteristic of RecQ helicases. Residues in cyan and yellow represent conserved and invariant residues. Purple denotes residues shown to contact DNA in crystal structures, derived from the following PDB entries: PDB ID codes 2P6R (HEL308), 3KQH (HCV-NS3), 2JLV (DEN4-NS3), 2DB3 (VASA), 2CGZ (BLM), and 4TMU for *C. sakazakii* RecQ (CsRecQ) (18).

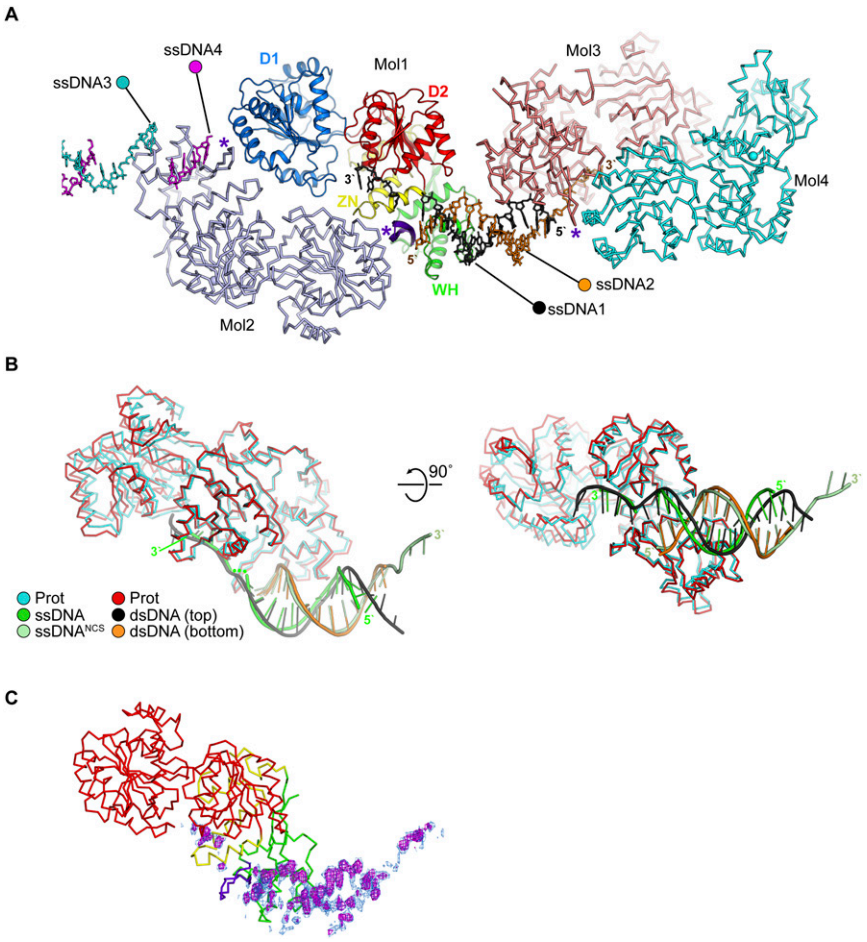


Fig. S4. Structure of the ssDNA complex. (*A*) Arrangement of molecules with the asymmetrical unit of the ssDNA complex is shown. Two RECQ1 dimers are shown, with each constituent monomer bound to a single ssDNA oligonucleotide. The individual domains of a single monomer are color-coded as in Fig. 1*A*. The 5′ end of the ssDNA spans WH domains from two adjacent monomers, with the 3′ end extending between the RecA domains. A pseudoduplex is formed between nucleotides 8 and 20 of one strand and nucleotides 1 and 13 of the paired ssDNA oligonucleotide. The pseudoduplex between Mol1/Mol3 is better resolved than the pseudoduplex between Mol2/symmetry-related copy of Mol4. (*B*) Comparison of the RECQ1 monomer in complex with either single-stranded DNA or tailed duplex. Two perpendicular views are shown. The two DNA strands in the tailed duplex are colored black/orange. The ssDNA is shown in green. An additional noncrystallographic copy of the ssDNA is shown to illustrate the pseudoduplex formed between the WH domains of adjacent monomers in the ssDNA crystals. The RECQ1 monomers have been superposed using the RecA D2, ZN, and WH domains only (residues 280–592). Prot, protein. (*C*) Omit electron density for the pseudoduplex region in the ssDNA complex. The omit electron density, calculated before inclusion of the DNA in the model, is shown, along with a RECQ1 monomer. The sugar phosphate backbone is clearly visible in the pseudoduplex region, whereas the 3′ extensions and the nucleotide bases are less well resolved. The 2Fo–Fc electron density (blue mesh) and Fo–Fc electron density (magenta) maps, calculated by PHENIX, are contoured at 1.2σ and 2.5σ, respectively.

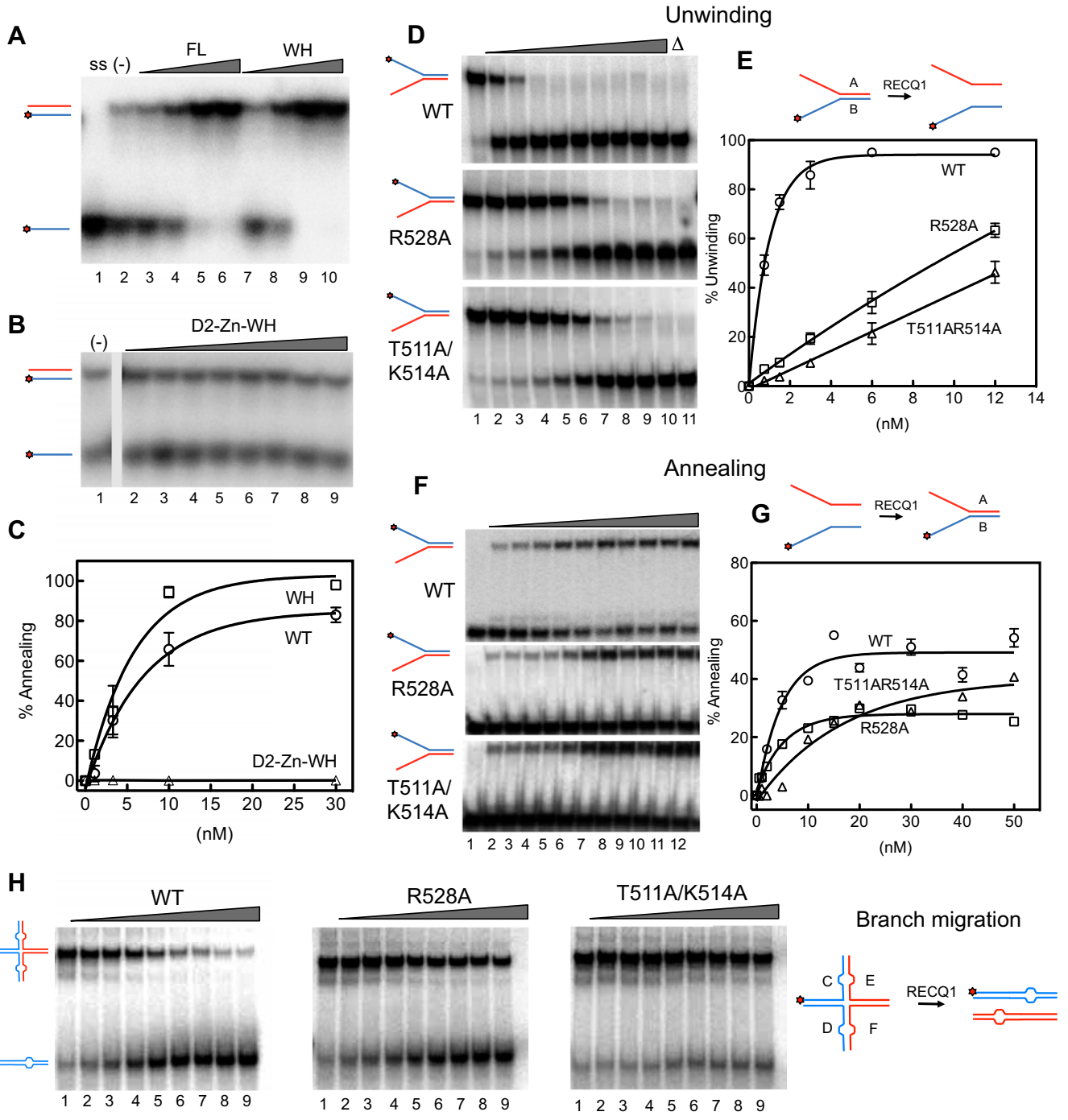


Fig. S5. Analysis of the annealing, unwinding, and branch migration activity of RECQ1 mutants. (*A*) Isolated WH domain of RECQ1 can drive DNA strand annealing. Representative gel image. Lane 1: labeled ssDNA alone. Lane 2: labeled ssDNA in the presence of the complementary unlabeled strand. Lanes 3–6: annealing assays performed using increasing concentrations of WT RECQ1 (FL) (1.1, 3.3, 10, and 30 nM). Lanes 7–10: annealing assays performed using increasing concentrations of the isolated WH domain (WH) (1.1, 3.3, 10, and 30 nM). (*B*) D2-Zn-WH RECQ1 fragment completely lacks annealing activity. Lane 1: labeled ssDNA in the presence of the complementary unlabeled strand. Lanes 2–10: annealing assays performed using increasing duplicate concentrations of the D2-Zn-WH protein fragment (1.1, 1.1, 3.3, 3.3, 10, 10, 30, and 30 nM). All annealing assays in *A* and *B* were performed using the oligonucleotides S and T listed in Table S3. (*C*) Plots of the annealing assays performed in the presence of WT RECQ1 or the isolated WH domain. The data points represent the mean of five independent experiments. Error bars indicate SEM. (*D*) Unwinding assays using RECQ1^{FL} and mutants R528A and T511A/K514A. Lanes 1–10: unwinding assays performed using increasing RECQ1 concentrations (0, 0.75, 1.5, 3, 6, 12.5, 25, 50, 100, and 200 nM) and a fixed concentration of forked duplex substrate (2 nM). Lane 11: heat-denatured substrate control. All of the reactions were stopped after 20 min. (*E*) Plots of the unwinding assays performed in the presence of WT RECQ1, R528A, or T511/K514A. The data points represent the mean of three independent experiments. Error bars indicate SEM. (*F*) Annealing assays using RECQ1^{FL} and mutants R528A and T511A/K514A. Lane 1: labeled ssDNA in the absence of the complementary unlabeled strand. Lanes 2–12: annealing assays performed at increasing RECQ1 concentrations (0, 0.5, 1, 2, 5, 10, 15, 20, 30, 40, and 50 nM). (*G*) Plots of the annealing assays performed in the presence of WT RECQ1, R528A, or T511/K514A. The data points represent the mean of three independent experiments. Error bars indicate SEM. (*H*) Analysis of the branch migration activity of RECQ1 mutants, R528A, and T511A/K514A using an HJ substrate with a heterology region of one base [HJ (1)]. Lanes 1–9: branch migration assays at increasing RECQ1 concentrations (0, 6.25, 12.5, 25, 50, 75, 100, 150, and 200 nM). The HJ (1) concentration is 2 nM.

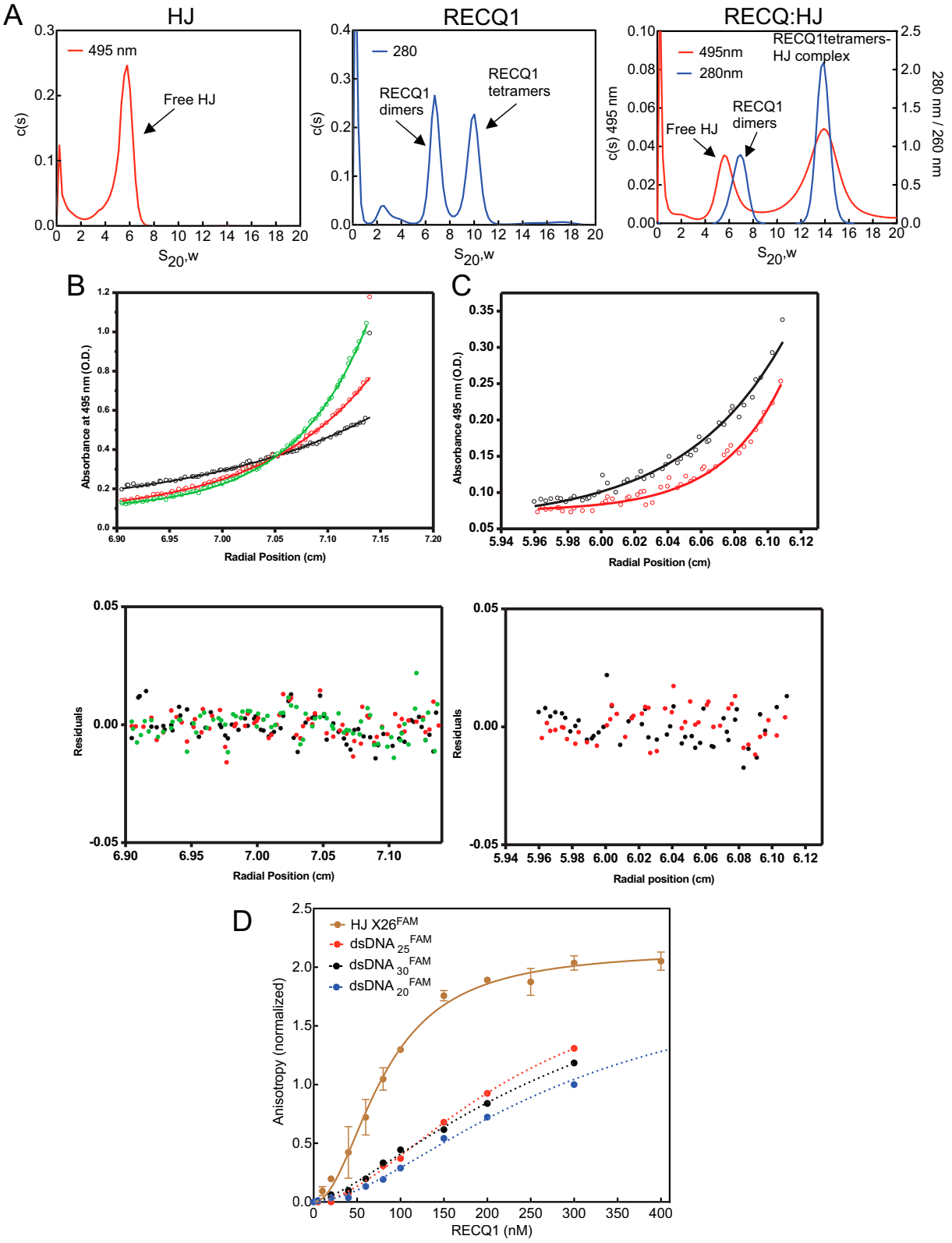


Fig. S6. Biochemical and biophysical analysis of the RECQ1 + HJ complex. (*A*) Sedimentation velocity AUC studies of RECQ1 in complex with FAM-labeled HJ. Scans were acquired at 495 nm for FAM detection (red) and 280 nm for protein detection (blue). (*Left*) FAM-HJ alone (3 μM). FAM-HJ sediments are shown as a single species with $s_{20,w}$ = 5.8 S. (*Center*) RECQ1 alone (6 μM). RECQ1 sediments are shown as two species with $s_{20,w}$ = 6.8 and $s_{20,w}$ = 9.9 S. These two species have an estimated molecular mass of 130 kDa (dimers) and 290 kDa (tetramers), respectively, in agreement with previous findings (1). (*Right*) Molar ratio of 1:6 of FAM-HJ (3 μM)/RECQ1 (18 μM). The RECQ1/HJ complex sediments at $s_{20,w}$ = 13.8 S, with an estimated molecular mass of ∼352 kDa. (*B* and *C*) Sedimentation equilibrium analytical ultracentrifugation analysis with HJ alone or HJ + RECQ1 complex. Experiments were carried out at 8 °C in six-channel centerpieces, as

Legend continued on following page

described in *SI Materials and Methods*, and analyzed using SEDPHAT. The solid lines are the global analysis of each set of data with the "species analysis" model. (*B*, *Upper*) Sedimentation equilibrium profiles for the HJ alone (3 μM) at 7,000, 10,000, and 13,000 rpm (An-60 Ti rotor from Beckman Coulter). (*C*, *Upper*) Sedimentation equilibrium profiles for HJ (3 μM) in the presence of RECQ1 (30 μM) at 7,000 and 10,000 rpm (An-60 Ti rotor from Beckman Coulter). (*B* and *C*, *Lower*) Graphs show the residuals for the given fits. (*D*) Fluorescence anisotropy assays using 56-FAM–labeled dsDNA and HJ substrates. The FAM-labeled substrates (10 nM) were titrated with an increasing concentration of RECQ1 as described in *SI Materials and Methods*. The solid lines represent the best fits of the data to "one-site binding with Hills equation." The values of the equilibrium dissociation constants (K_d) and the Hill coefficient (h) measured for the 56-FAM–labeled HJ substrate are 80.3 ± 3.5 nM and 2.0 ± 0.1, respectively. The value of the Hill coefficient is consistent with a model where two tight-binding dimers of RECQ1 form a tetrameric complex on the four-way junction substrate. RECQ1 binds linear duplexes of 20, 25, and 30 bp with much lower affinity compared with four-way junction substrates: K_d > 250 nM.

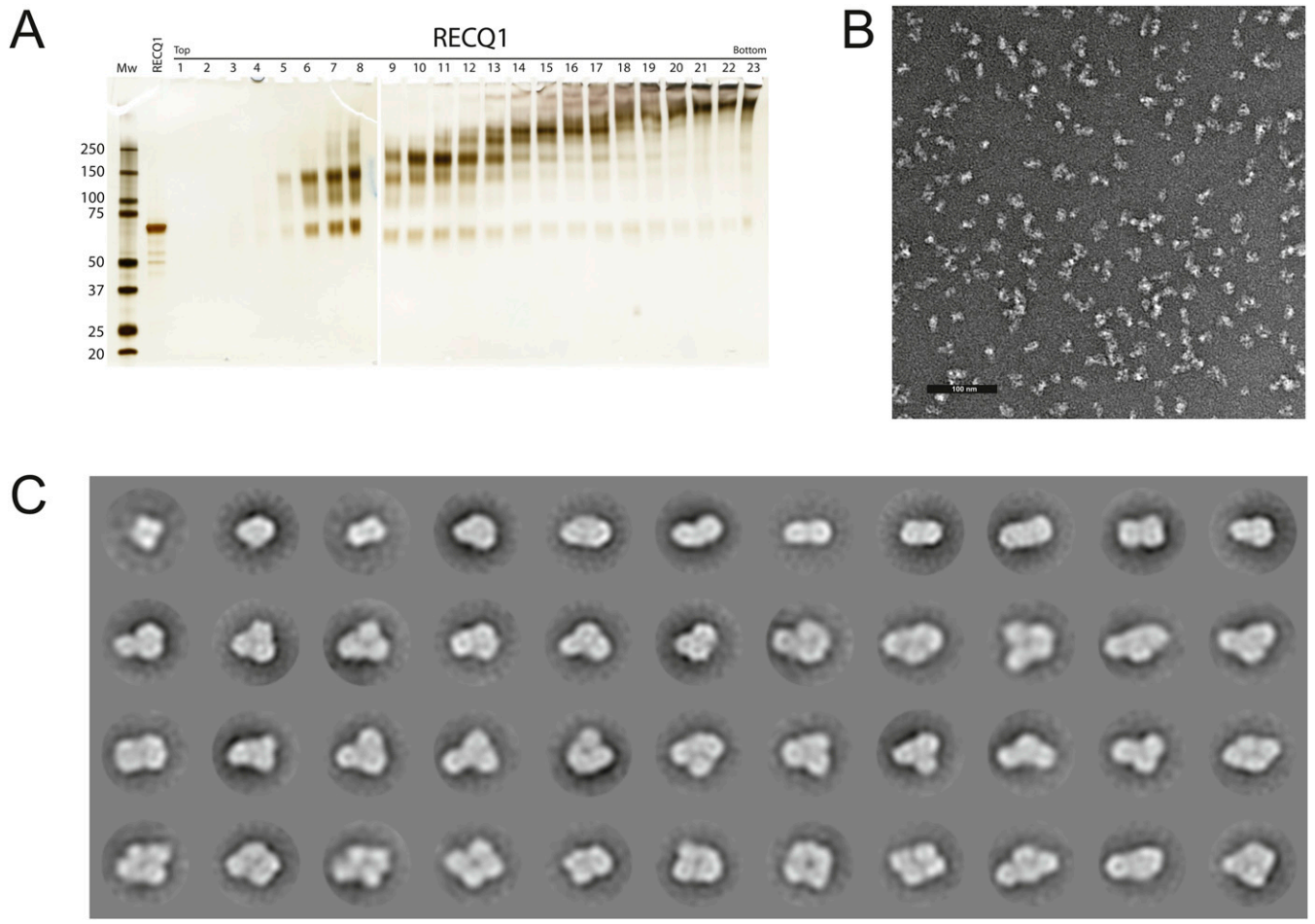


Fig. S7. EM analysis of RECQ1. (*A*) GraFix (20–45% glycerol and 0–0.1% glutaraldehyde) gradient ultracentrifugation of RECQ1, analyzed by SDS/PAGE and silver staining. Supershifted bands correspond to higher order assemblies, indicating that RecQ1 exists as homo-oligomers of varying stoichiometry. (*B*) Characteristic micrograph of fraction 16 of the GraFix gradient. (*C*) RecQ1 reference-free class averages. RECQ1 forms apparent monomeric (number of particles = 116, 2% of the "clean" dataset), dimeric (number of particles = 1,854, 33% of the clean dataset), trimeric (number of particles = 1,916, 35% of the clean dataset), or tetrameric (number of particles = 1,593, 29% of the clean dataset) assemblies. Although we cannot rule out that apparent dimeric and trimeric views are side or tilted views of RECQ1 tetramers, the presence of a flat, four-membered RECQ1 assembly provides compelling evidence for a tetrameric planar configuration. Box size, 168 pixels or 459 Å.

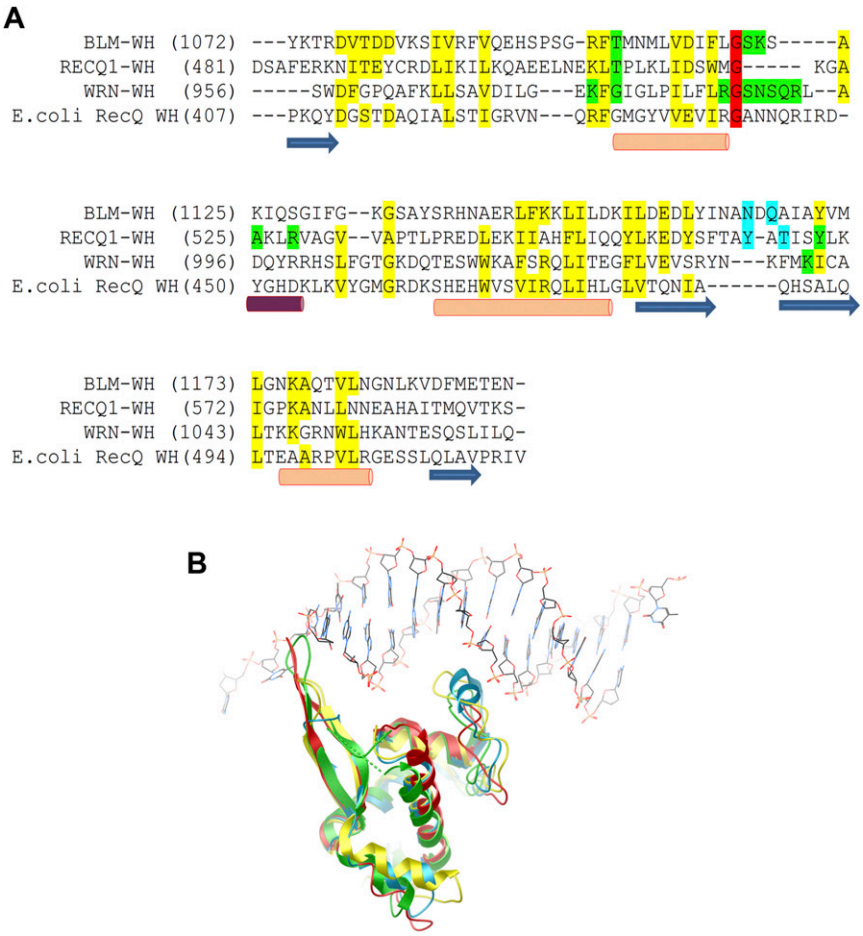


Fig. S8. Structure-based alignment of RecQ and other SF2 helicases. (*A*) Structure-based sequence alignment of the WH domains of human BLM, RECQ1, WRN, and bacterial RecQ. Residue numbers are listed to the left of each row. Yellow boxes denote similar residues, red is the (only) residue conserved in all of the proteins, green boxes denote residues forming H-bonds with the dsDNA backbone, and cyan denotes contacts with DNA bases. The secondary structure elements of RECQ1 are indicated under the alignment (pink cylinders, α-helices; purple cylinder, 3_{10} helix; blue arrows, β-strands). The secondary structure elements are generally aligned in all structures. (*B*) Structural superposition of the WH domains of REQ1 (red, PDB ID code 2WWY), WRN (yellow, PDB ID code 3AAF), BLM (green, PDB ID code 4O3M) and *E. coli* RecQ (cyan, PDB ID code 1OYW), together with the DNA from the RecQ1 complex (PDB ID code 2WWY).

Table S1. Fork-duplex unwinding, single-strand annealing, HJ resolution activity, and oligomeric properties of the RECQ1^{T1} and RECQ1^{FL} mutants

	RECQ1^{T1}				RECQ1^{FL}			
Phenotype	Fork unwinding	Single-strand annealing	HJ resolution	Oligomeric state	Fork unwinding	Single-strand annealing	HJ resolution	Oligomeric state
Hairpin mutations								
WT	Yes	No	No	Dimer	Yes	Yes	Yes	Dimer + tetramer
Y564A	No	Yes	No	Dimer	Partial (50%)	Yes	Partial (<10%)	Dimer + tetramer
Y564A shortened (Δ2, Δ3, Δ8)	No	Yes	No	Monomer	No	Yes	No	Dimer*
WH mutations								
R528A					Partial (<20%)	Partial (<60%)	Partial (<50%)	Dimer + tetramer
T511A-K514A					Partial (<15%)	Partial (<60%)	Partial (<20%)	Dimer + tetramer
Isolated WH					No	Yes	No	
D2-Zn-WH					No	No	No	

*Different kind of dimer held together by the N terminus of RECQ1, as described (1). The Δ2, Δ3, and Δ8 are deletion mutants lacking, respectively, two, three, and eight residues on both strands of the β-hairpin, as described (1).

Table S2. Data collection, phasing and refinement statistics

	dsDNA	ssDNA
PDB ID code	2WWY	4U7D
Data collection		
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c*; Å	58.3, 174.2, 100.2	55.38, 138.20, 207.60
α, β, γ; °	90, 93.31, 90	90, 90.02, 90
Resolution,* Å	2.9 (2.9–3.06)*	3.4 (3.40–3.58)*
Resolution limits,† Å	2.90, 3.32, 2.90 (2.90, 3.18, 2.90)	3.4, 3.40, 3.90 (3.40, 3.40, 3.40)
Total reflections	324,606	87,431
Unique reflections	44,075	40,962
Completeness, %	100 (100)*	95.3 (96.5)*
R_{meas}	0.111 (1.06)*	0.135 (0.477)*
R_{pim}	0.041 (0.386)*	0.089 (0.311)
CC1/2	0.994 (0.716)	0.975 (0.704)
$I/\sigma I$	14.8 (2.2)*	5.7 (2.0)*
Redundancy	7.4 (7.5)*	2.1 (2.1)*
Refinement		
Resolution, Å	43.38–2.90	40–3.4
No. of reflections (free)	44,020 (2,217)	40,919 (1,951)
R_{work}/R_{free}	19.8/23.6	20.51/24.86
No. of atoms		
Protein	8,241	16,240
DNA	1,385	980
Other	107	4
B-factors, Å^2		
Protein	79	94
DNA	141	112
Other	76	98
rmsd		
Bond lengths, Å	0.011	0.007
Bond angles, °	1.441	1.140

CC1/2, Pearson correlation coefficient (CC) for averaged intensities in a random half-dataset analysis as output by AIMLESS; $I/\sigma I$, signal-to-noise ratio, mean intensity/standard deviation, average <merged $<I_h>$/sd($<I_h>$)>;

R_{meas}, $\sum_{hkl}\left[\frac{N}{(N-1)}\right]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where N is the redundancy, $I_i(hkl)$ is the intensity measurement for reflection i, and $<I(hkl)>$ is the mean intensity for multiple recorded reflections;

R_{pim}, $\sum_{hkl}\left[\frac{1}{(N-1)}\right]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$;

R_{work}, $\sum_{hkl}\frac{|Fobs(hkl)-Fcalc(hkl)|}{\sum_{hkl}Fobs(hkl)}$;

R_{free}, as R_{work} but calculated with a 5% subset of observed structure factors that were excluded from the refinement process.

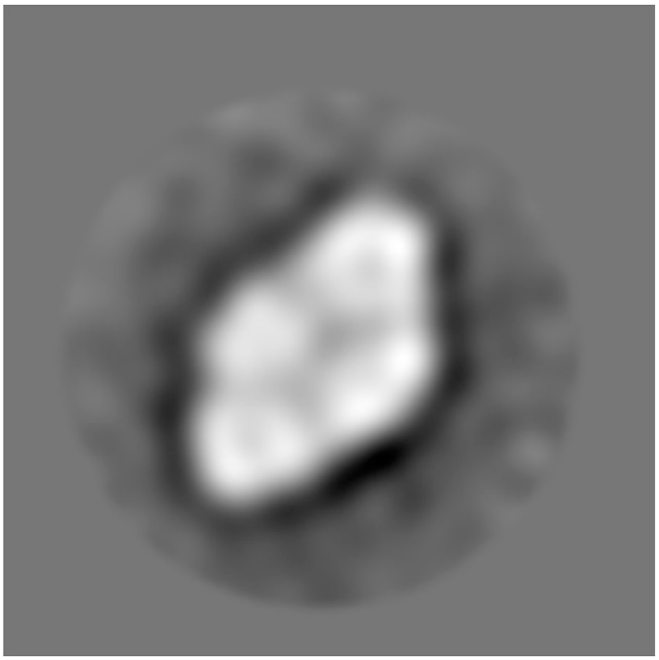
*Values in parentheses are statistics for the highest resolution shell.

†Anisotropic resolution limits as defined by AIMLESS based on Mn [I/sd(I)] > 2. Values in parentheses are resolution limits based on half-dataset correlation >50%.

Table S3. Sequences of the oligonucleotides used in the study of forked duplex unwinding and annealing, and HJ branch migration

Name	Length, nt	Sequence 5′→3′
A	50	GAA CGA ACA CAT CGG GTA CGT TTT TTT TTT TTT TTT TTT TTT TTT TT
B	50	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT CGT ACC CGA TGT GTT CGT TC
C	60	CAC TGT GAT GCA CGA TGA TTG ACG ACA GTA GTC AGT GCT GCA GTG GTC AGG TGT CAT CAC
D	60	CCT GCA TAC AGA TGT TGA CCC AGC ACT GAC TAC TGT CGT CAA TCA TCG TGC ATC ACA GTG
E	60	GTG ATG ACA CCT GAC CAC TGC AGC ACT GAC TAC TGT CGT C**G**A TCA TCG TGC ATC ACA GTG
F	60	CAC TGT GAT GCA CGA TGA T**C**G ACG ACA GTA GTC AGT GCT GGG TCA ACA TCT GTA TGC AGG
G (X26-1)	60	/56-FAM/CCG CTA CCA GTG ATC ACC AAT GGA TTG CTA GGA CAT CTT TGC CCA CCT GCA GGT TCA CCC
H (X26-2)	60	TGG GTG AAC CTG CAG GTG GGC AAA GAT GTC CTA GCA ATC CAT TGT CTA TGA CGT CAA GCT
I (X26-3)	60	GAG CTT GAC GTC ATA GAC AAT GGA TTG CTA GGA CAT CTT TGC CGT CTT GTC AAT ATC GGC
L (X26-4)	61	TGC CGA TAT TGA CAA GAC GGC AAA GAT GTC CTA GCA ATC CAT TGG TGA TCA CTG GTA GCG G
M	20	/56-FAM/CCG CTA CCA GTG ATC ACC AA
N	20	TTG GTG ATC ACT GGT AGC GG
O	25	/56-FAM/CCG CTA CCA GTG ATC ACC AAT GGA T
P	25	ATC CAT TGG TGA TCA CTG GTA GCG G
Q	30	/56-FAM/CCG CTA CCA GTG ATC ACC AAT GGA TTG CTA
R	30	TAG CAA TCC ATT GGT GAT CAC TGG TAG CGG
S	30	GGC AAG AAC GAA CAC ATC GGG TAC GGC CTC
T	30	GAG GCC GTA CCC GAT GTG TTC GTT CTT GCC

Bold red letters indicate the nucleotides that form mismatched pairs in the branch migration products.



Movie S1. Reference-free 2D averages of negatively stained, single-particle EM images. RECQ1 exists as a heterogeneous population that contains apparent dimers, trimers, and tetramers.

Movie S1