

# Supporting Information

Manthei et al. 10.1073/pnas.1416746112

## SI Materials and Methods

**Protein Purification.** The open-reading frame of the *recQ* gene was PCR amplified from *C. sakazakii* genomic DNA (ATCC BAA-894D-5) and was subcloned into pET28b (pKH0130) for over-expression. The catalytic core was purified from full-length CsRecQ in pET28b through extensive thrombin cleavage, which removes the helicase and HRDC domain. To avoid removal of the HRDC from full-length protein, a second plasmid (pKH0134) was made by subcloning a sequence encoding for a tobacco etch virus (TEV) cleavage site between the His-tag and CsRecQ. Rosetta 2 (DE3) *E. coli* transformed with pLysS (Novagen) and either CsRecQ-encoding plasmid were grown at 37 °C in LB supplemented with 50 µg/mL kanamycin and 50 µg/mL chloramphenicol. At an OD<sub>600</sub> of ~0.6, cells were induced to overexpress protein by adding 1 mM isopropyl β-D-thiogalactopyranoside at ~37 °C for 3.5 h. Protein was purified as described previously for EcRecQ (1), with changes in His-tag removal: For the catalytic core, thrombin cleavage was allowed to proceed overnight to remove the HRDC, and for the full-length CsRecQ, TEV was used overnight instead of thrombin to retain the HRDC. The CsRecQ proteins were purified further on a Sephacryl S-100 column (GE Healthcare) to ensure protein homogeneity.

EcRecQ full-length proteins were purified essentially as previously described, with short thrombin cleavage reactions to ensure the HRDC would not be removed (1). A HiPrep QFF ion exchange column (GE Healthcare) was substituted for MonoQ.

## CsRecQ/DNA Complex Crystallization and Structure Determination.

CsRecQ catalytic core at 7 g/L in minimal buffer [10 mM Tris-HCl (pH 8), 1 M ammonium acetate] was mixed with the self-annealed hairpin DNA with the sequence 5'-CGG TAG AAT TGT AAA ATT CTA CCG GTG CCT TAC T-3' in 10 mM Tris-HCl (pH 8), at a protein:DNA ratio of ~1:1.2. This DNA forms a hairpin structure with a GTAA tetraloop that stabilizes hairpin formation (2). Previous structural studies have shown that the first and last nucleotides of the tetraloop form a sheared G-A base pair (3). The complex was combined at a 1:1 (vol/vol) ratio with mother liquor [70 mM sodium acetate, 24% (vol/vol) glycerol, 11% (wt/vol) PEG 4000], and crystals were formed by hanging-drop vapor diffusion. Crystals formed only in the presence of the DNA substrate. Crystals were flash-frozen in liquid nitrogen. Crystallization trials also were performed with an oligonucleotide from which the last four bases were removed (5'-CGG TAG AAT TGT AAA ATT CTA CCG GTG CCT-3') and with A32 exchanged with T34 (5'-CGG TAG AAT TGT AAA ATT CTA CCG GTG CCT TTC A-3'). No crystals formed with either oligonucleotide.

X-ray diffraction data were collected at the Advanced Photon Source (LS-CAT beam line 21ID-F) and were indexed and scaled using HKL2000 (4). The structure of the CsRecQ/DNA complex was determined by molecular replacement using the apo RecQ structure (5) as a search model in the program Phaser (6) followed by rounds of manual fitting using Coot (7) and refinement using REFMAC5 (8) and PHENIX (9). Coordinate and structure factor files have been deposited in the Protein Data Bank (PDB ID code 4TMU).

**SAXS.** SAXS data were collected at the National Magnetic Resonance Facility at Madison (NMRFAM) using a Bruker NanoStar instrument at a wavelength of 1.54 Å and a sample-detector distance of 67.7 cm, resulting in a momentum transfer range of  $0.01 < q < 0.40 \text{ \AA}^{-1}$ . Data were collected for buffer and

protein at 25 °C for 2–5 h (NMRFAM) in 20 mM Tris-HCl (pH 8), 300 mM NaCl, 5% glycerol, and 1 mM β-mercaptoethanol. Buffer subtraction was adjusted for the excluded volume of the protein. The predicted scattering intensity at  $q = 0 \text{ \AA}^{-1}$  and the  $R_g$  were determined by Guinier analysis (10) and were compared between two concentrations (2.5–5 g/L CsRecQ catalytic core) to detect interparticle interference.

SAXS data were processed using Primus and Gnom (ATSAS, European Molecular Biology Laboratory) (11) to determine the  $R_g$  and maximum dimension ( $D_{\text{max}}$ ) of the protein and to obtain regularized scattering amplitudes extrapolated to  $q = 0 \text{ \AA}^{-1}$  and the pair distance distribution function (PDDF) for each sample. The latter provides a profile of all pair distances within the molecule and therefore is descriptive of the overall shape. When computing the PDDF,  $D_{\text{max}}$  was adjusted by increments of 1–2 Å and was optimized based on the following criteria: (i) a smooth drop-off to zero probability at  $D_{\text{max}}$ ; (ii) agreement between the  $R_g$  measured using the Guinier transform and the  $R_g$  extracted from the PDDF; and (iii) agreement between the regularized scattering determined from the PDDF and the experimental data. Furthermore, the regularization parameter, alpha, was kept below 5 to avoid oversmoothing. The resulting regularized scattering amplitudes were used to calculate 10 replicate ab initio dummy atom models using DAMMIF (12). The quality and uniqueness of the results were assessed further by agreement between replicate models as quantitated by the mean normalized spatial discrepancy (13). Models were superimposed with one another using the Supcomb20 algorithm (14). Existing crystal structures (PDB ID code 1OYW and the protein portion of the CsRecQ/DNA complex) were docked manually into the dummy atom models, and the scattering of the resulting all-atom model was predicted using the FOXS web server (15).

**DNA-Binding Assay.** 3'-Fluorescein-conjugated 30-base ssDNA (5'-GCG TGG GTA ATT GTG CTT CAA TGG ACT GAC-3') was annealed to an 18-base ssDNA (5'-AAG CAC AAT TAC CCA CGC-3') to create an 18-bp duplex with a 12-base 3' overhang (annealed substrate: F-3'overhang). RecQ proteins were diluted serially from 1,000–0.01 nM in 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 4% (vol/vol) glycerol, 1 mM β-mercaptoethanol, and 0.1 g/L BSA. Dilutions of RecQ proteins were incubated for 20 min at room temperature with 500 pM F-3'overhang in a total volume of 100 µL. The fluorescence polarization of each sample was measured at 25 °C with a Beacon 2000 fluorescence polarization system. Measurements are reported in triplicate, and error bars/uncertainties represent 1 SD of the mean. Uncertainties reported with  $K_{\text{d, app}}$  values are 1 SD of the mean.

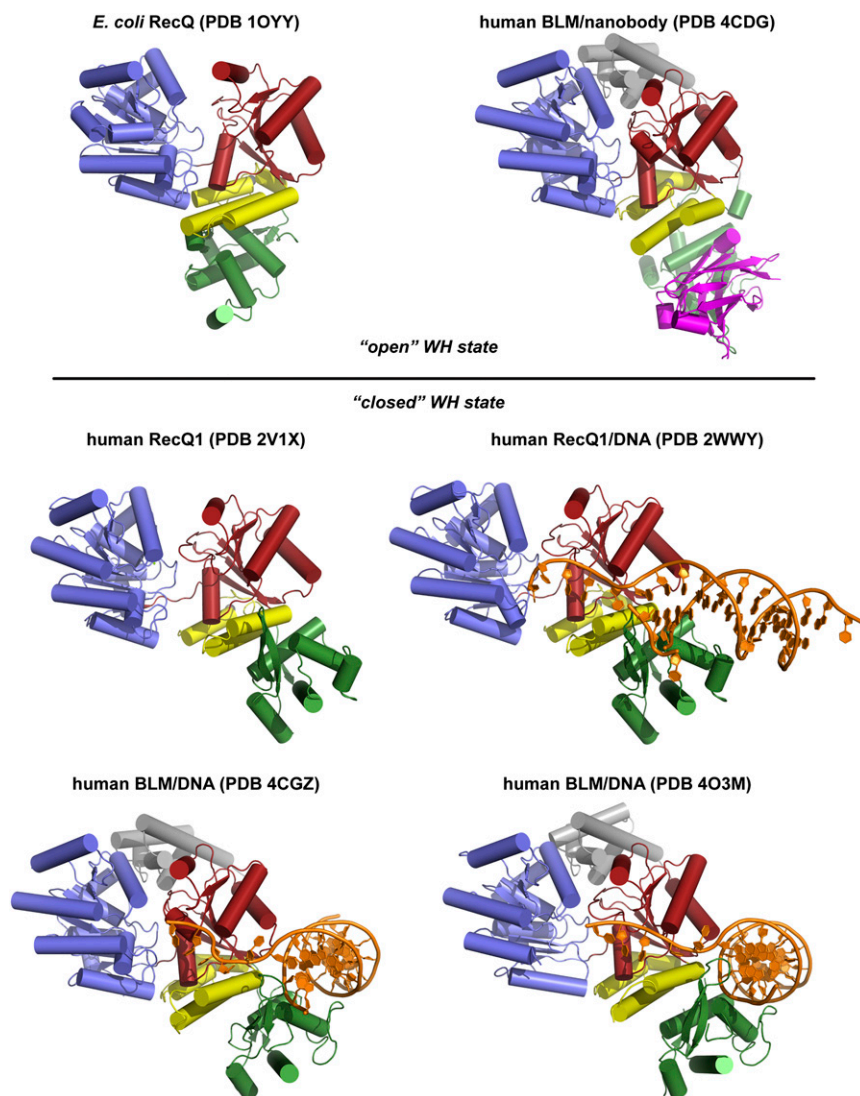
**ATPase Assays.** To compare DNA-dependent ATP hydrolysis profiles, purified EcRecQ variants were mixed with 0–1,000 nM dT<sub>28</sub> in 20 mM Hepes-HCl (pH 8.0), 50 mM NaCl, 1 mM β-mercaptoethanol, 2 mM MgCl<sub>2</sub>, 0.1 g/L BSA, and 1 mM ATP at 25 °C. Homopolymeric ssDNA (dT<sub>28</sub>) was used as the DNA cofactor, because ssDNA and dsDNA differ in their ability to stimulate ATP hydrolysis by RecQ (16) and because DNA unwinding alters the concentrations of ssDNA and dsDNA during measurement. Wild-type and variant EcRecQ protein concentrations were 1, 5, or 50 nM. The assay also included an ATP regeneration system that converts ADP to ATP in a reaction that is coupled to the conversion of NADH to NAD<sup>+</sup> (17). This coupled reaction can be detected spectrophotometrically by observing the decrease of A<sub>340nm</sub> caused by NADH oxidation. A Synergy 2 plate reader

(BioTek) was used to monitor  $A_{340\text{nm}}$  using automated path-length correction. The  $\Delta A_{340\text{nm}}/\Delta t$  rates were measured and converted to  $\Delta[\text{ATP}]/\Delta t$  to determine the ATPase rates and were normalized to the concentration of RecQ present in each reaction. The intrinsic ATPase rate is reflected by the constant  $k_{\text{min}}$  and is measured as the rate of ATP hydrolysis in the absence of the DNA cofactor. The data were graphed using Prism (GraphPad), with  $K_{1/2}$  reflecting the concentration of RecQ required to reach 50% of the maximal ATPase activity and the maximal ATPase rate represented by the constant  $k_{\text{max}}$ . Measurements are reported in triplicate, and error bars and uncertainties represent 1 SD of the mean.

**Helicase Assays.** A fluorescence helicase assay was adapted from ref. 18, in which a quencher is placed opposite a fluorophore, and unwinding leads to an increase in fluorescence. In this assay, the same 3' overhang substrate was used as in the DNA-binding assay, with the exception that 1 bp was swapped to avoid putting a guanine next to the Tetrachlorofluorescein (TET) fluorophore. A 5'-TET-conjugated 30-base ssDNA (5'-CCG TGG GTA ATT GTG CTT CAA TGG ACT GAC-3') was annealed to a 3' black hole quencher-1 (BHQ)-conjugated 18-base ssDNA (5'-AAG CAC AAT TAC CCA CGG-3') to create the 3' overhang heli-

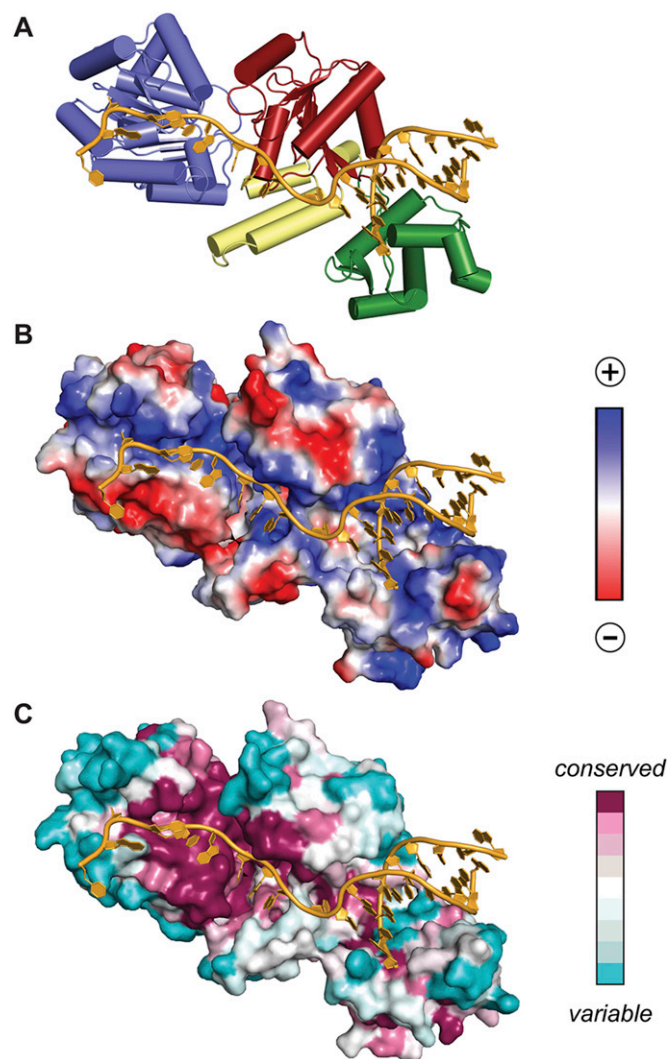
case substrate. The Synergy 2 plate reader (BioTek) was used to measure emission at a wavelength of 528 nm (excitation at 485 nm) with reactions in 96-well black plates with clear bottoms (Costar 3603; Fisher Scientific). Purified RecQ variants were incubated with 100 nM substrate in 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM  $\beta$ -mercaptoethanol, 1 mM  $\text{MgCl}_2$ , 1 mM ATP, 0.1 g/L BSA, and 4% (vol/vol) glycerol for 5 min at 25 °C with protein concentrations between 0.001 and 100 nM in a reaction volume of 100  $\mu\text{L}$ . Reactions were terminated by the addition of 0.5% SDS (to denature RecQ) and 1  $\mu\text{M}$  unlabeled oligo 3 (to prevent quenching by reannealing of the unwound TET-conjugated DNA). For wild-type RecQ and variants with helicase activity near wild-type levels, the highest concentration of RecQ tested was 300 nM, because at higher concentrations all available ATP was consumed before the 5-min reaction was quenched. The percent of DNA unwound was determined by normalizing each reaction between negative and positive controls using Prism (GraphPad), where the negative control contained no RecQ, and the positive control was boiled for 5 min in the presence of the quench and then was placed on ice. Measurements are reported in triplicate; error bars/uncertainties represent 1 SD of the mean.

- Zittel MC, Keck JL (2005) Coupling DNA-binding and ATP hydrolysis in *Escherichia coli* RecQ: Role of a highly conserved aromatic-rich sequence. *Nucleic Acids Res* 33(22):6982–6991.
- Antao VP, Tinoco I, Jr (1992) Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res* 20(4):819–824.
- van Dongen MJ, et al. (1997) Structural features of the DNA hairpin d(ATCCTA-GTTA-TAGGAT): Formation of a G-A base pair in the loop. *Nucleic Acids Res* 25(8):1537–1547.
- Otwinowski Z, Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*, eds Carter, CW Jr, Sweet RM (Academic, New York), Vol 276, pp 307–326.
- Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.
- McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
- Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
- Winn MD, Isupov MN, Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1):122–133.
- 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.
- Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* 40(3):191–285.
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst* 36(5):1277–1282.
- Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Cryst* 42(2):342–346.
- Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Cryst* 34(1):33–41.
- Volkov VV, Svergun DI (2003) Uniqueness of ab-initio shape determination in small-angle scattering. *J Appl Cryst* 36(3):860–864.
- Schneidman-Duhovny D, Hammel M, Sali A (2010) FoXS: A web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* 38(web server issue):W540–4.
- Umezū K, Nakayama K, Nakayama H (1990) *Escherichia coli* RecQ protein is a DNA helicase. *Proc Natl Acad Sci USA* 87(14):5363–5367.
- Morrison SW, Lee J, Cox MM (1986) Continuous association of *Escherichia coli* single-stranded DNA binding protein with stable complexes of recA protein and single-stranded DNA. *Biochemistry* 25(7):1482–1494.
- Manhart CM, McHenry CS (2013) The PriA replication restart protein blocks replicase access prior to helicase assembly and directs template specificity through its ATPase activity. *J Biol Chem* 288(6):3989–3999.



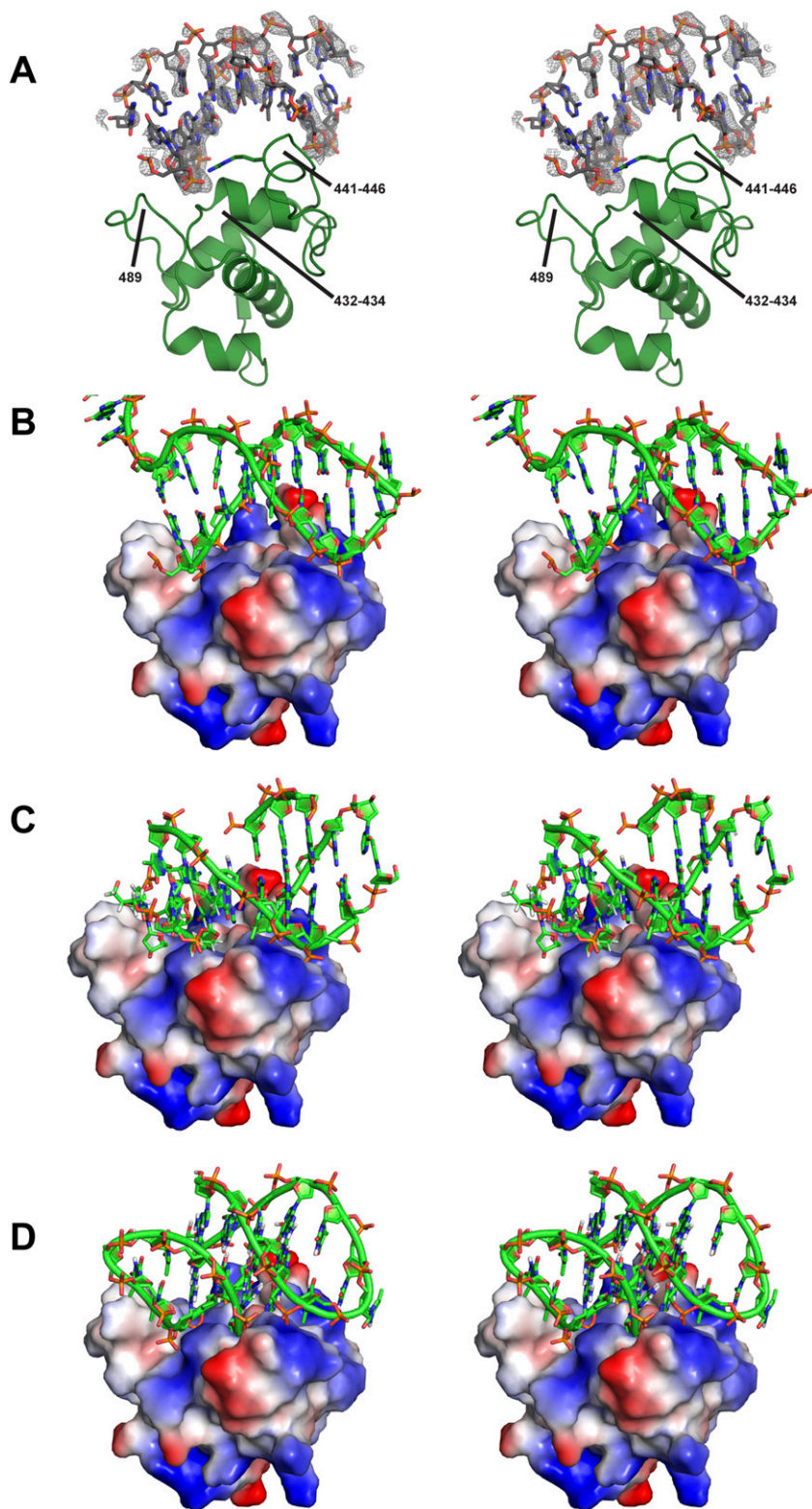
**Fig. S1.** Comparison of RecQ structures reveals differences in the position of the WH domain. (*Upper*) Ribbon diagrams of RecQ helicases with an open WH state [EcRecQ catalytic core (PDB ID code 1OYY) (1)] and an antibody-stabilized form of human BLM (unpublished structure available through the Protein Data Bank, PDB ID code 4CDG). In these structures the WH domain (green) is directly below the helicase (blue and red) and Zn<sup>2+</sup>-binding (yellow) domains. The nanobody from the BLM structure (magenta) binds in the cleft that is open in this form. (*Lower*) Ribbon diagrams of RecQ helicases with a closed WH state [human RecQ1 catalytic core, PDB ID code 2V1X (2)], a partial-duplex DNA-bound human RecQ1 catalytic core domain (unpublished structure available through the Protein Data Bank, PDB ID code 2WWY) and two independent structures of partial-duplex DNA-bound human BLM catalytic core with its HRDC domain [PDB ID codes 4O3M (3) and 4CGZ]. In these structures the WH domain (green) is adjacent to the helicase (blue and red) and Zn<sup>2+</sup>-binding (yellow) domains. DNA (orange) binds in the cleft formed between the WH and helicase domains.

- Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.
- 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.
- Swan MK, et al. (2014) Structure of human Bloom's syndrome helicase in complex with ADP and duplex DNA. *Acta Crystallogr D Biol Crystallogr* 70(Pt 5):1465–1475.



**Fig. S2.** Electrostatics and conservation of the DNA-binding surface in CsRecQ. (A) Ribbon diagram of the CsRecQ/DNA crystal structure. Domains and DNA are colored as in Fig. 1C in the main text. (B) Electrostatic surface features of CsRecQ (blue, electropositive; red, electronegative; white, neutral). (C) Evolutionary conservation of CsRecQ (the conservation scale, from variable to invariant RecQ protein sequences, is shown).

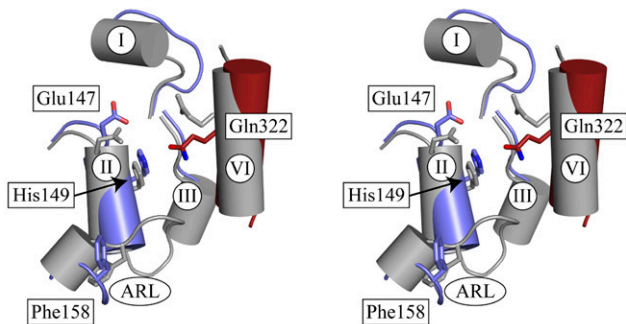




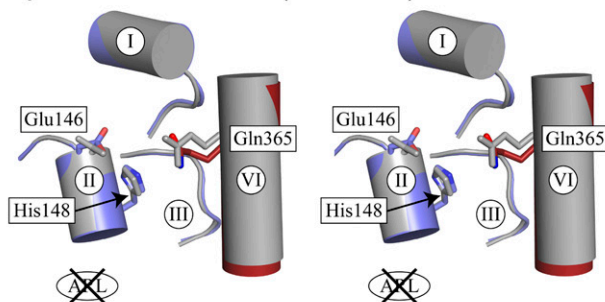
**Fig. 54.** Structural models of the RecQ WH domain bound to duplex, triplex, and quadruplex DNA. (A) Stereo diagram of the CsRecQ WH/dsDNA interface. An  $F_o - F_c$  DNA omit electron density map (contoured to  $2\sigma$ ) is shown in gray with the refined DNA model overlaid. Loops from the WH domain that contact dsDNA are labeled. (B) Stereo diagram of the CsRecQ WH domain/DNA interface. The WH domain surface is stained to highlight electropositive (blue), electronegative (red), and neutral (white) surfaces. (C) Model of triplex DNA bound to the duplex DNA-binding site on CsRecQ. Triplex DNA [PDB ID code 134D (1)] was fit onto the duplex region in the CsRecQ/DNA structure. (D) Model of quadruplex DNA bound to the duplex DNA-binding site on CsRecQ. Quadruplex DNA [PDB ID code 1D59 (2)] was fit onto the duplex region in the CsRecQ/DNA structure.

1. Radhakrishnan I, Patel DJ (1993) Solution structure of a purine.purine.pyrimidine DNA triplex containing G.GC and T.AT triples. *Structure* 1(2):135–152.
2. Kang C, Zhang X, Ratliff R, Moyzis R, Rich A (1992) Crystal structure of four-stranded Oxytricha telomeric DNA. *Nature* 356(6365):126–131.

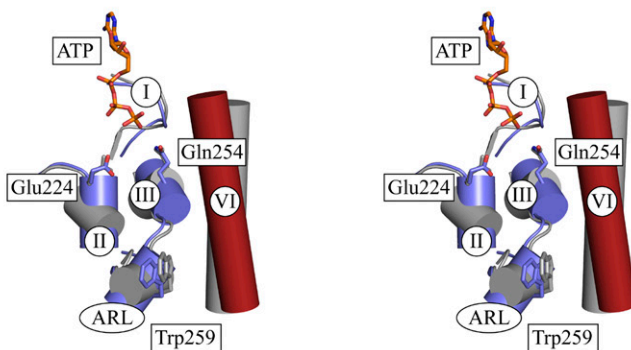
(A) Apo and DNA-bound RecQ (SF2 helicase)



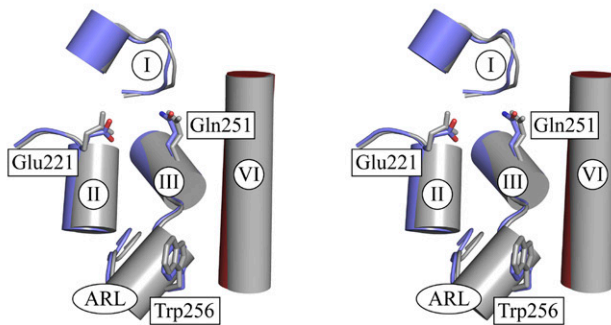
(B) Apo and DNA-bound Hel308 (SF2 helicase)



(C) Apo and DNA-bound PcrA (SF1 helicase)



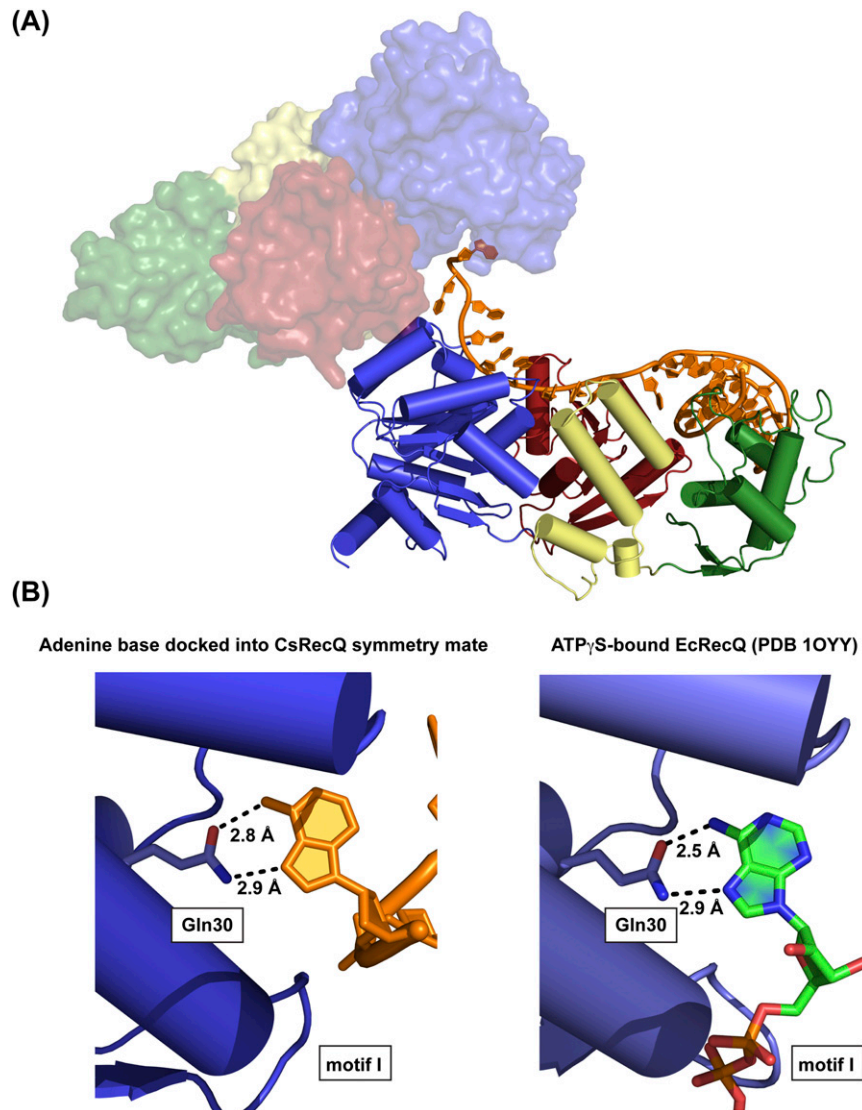
(D) Apo and DNA-bound UvrD (SF1 helicase)



**Fig. S5.** Comparison of helicase ATPase active sites in the absence and presence of DNA. Stereo ribbon diagrams of SF2 helicases RecQ (1) (A) and Hel308 (2) (B) and single-subunit SF1 helicases PcrA (3, 4) (C) and UvrD (5) (D) are shown with the apo structures displayed in gray and the DNA-bound structure displayed with the helicase domains shown in blue and red. Only structures in which both apo- and helicase substrate-bound DNA substrates (not ssDNA) were available are shown. BLM and RecQ1 structures were not included because the ssDNA portions of the structures do not extend to the ARL. Selected helicase motifs and the ARL are labeled. Key ATPase active site and base-stacking residues are shown in sticks and are labeled. ATP, which was included in the helicase substrate-bound structure, is shown with PcrA.

1. Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E.coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.  
 2. Büttner K, Nehring S, Hopfner KP (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol* 14(7):647–652.

3. Subramanya HS, Bird LE, Brannigan JA, Wigley DB (1996) Crystal structure of a DExx box DNA helicase. *Nature* 384(6607):379–383.
4. Velankar SS, Soutanas P, Dillingham MS, Subramanya HS, Wigley DB (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97(1):75–84.
5. Lee JY, Yang W (2006) UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* 127(7):1349–1360.



**Fig. S6.** DNA links symmetrically related CsRecQ monomers within the crystal lattice. (A) The partial duplex DNA (orange) in the crystal lattice is bound primarily to one protein (shown in ribbons), but the 3' extension docks into the ATP-binding site of an adjacent symmetrically related protein (shown in surface representation). (B) An adenine base on the DNA used in crystallization binds in place of the adenine in ATP. (Left) Distances between Gln30 of CsRecQ and atoms in the adenine from the 3' extension of the helicase substrates are shown. (Right) Distances between Gln30 of EcRecQ and atoms in the adenine from ATP<sub>γ</sub>S (1) are shown.

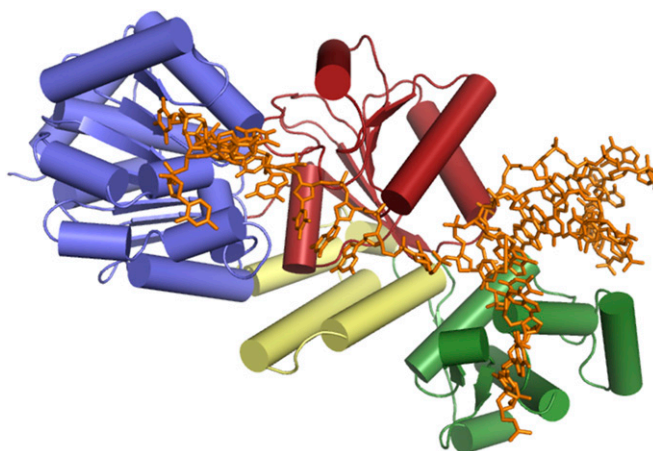
1. Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E.coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.







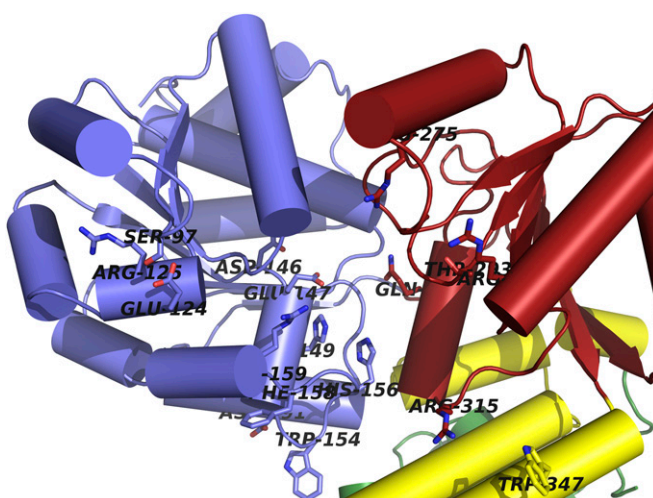




**Movie S1.** Apparent RecQ domain dynamics associated with binding to partial-duplex DNA. Chimera (1) was used to interpolate a trajectory for RecQ and DNA structural changes by comparing the structures of the apo EcRecQ (2) and DNA-bound CsRecQ catalytic core domains. A modified model of DNA showing the fully base-paired duplex portion was added to the free form to allow interpolation. The movie was rendered using PyMol (3).

#### [Movie S1](#)

1. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
2. Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E.coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.
3. Delano WL (2002) *The PyMol Molecular Graphics System* (DeLano Scientific, San Carlos, CA).



**Movie S2.** Close up of RecQ dynamics associated with binding to partial-duplex DNA. Chimera (1) was used to interpolate a trajectory for RecQ and DNA structural changes by comparing the structures of the apo EcRecQ (2) and DNA-bound CsRecQ catalytic core domains. The DNA is not shown in the movie to allow visualization of several residues. The movie was rendered using PyMol (3).

#### [Movie S2](#)

1. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
2. Bernstein DA, Zittel MC, Keck JL (2003) High-resolution structure of the *E.coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910–4921.
3. Delano WL (2002) *The PyMol Molecular Graphics System* (DeLano Scientific, San Carlos, CA).