**ATP hydrolysis provides functions that promote rejection of pairings between different copies of long repeated sequences**

Claudia Danilowicz1, Laura Hermans1, Vincent Coljee1, Chantal Prévost2, Mara Prentiss1\*

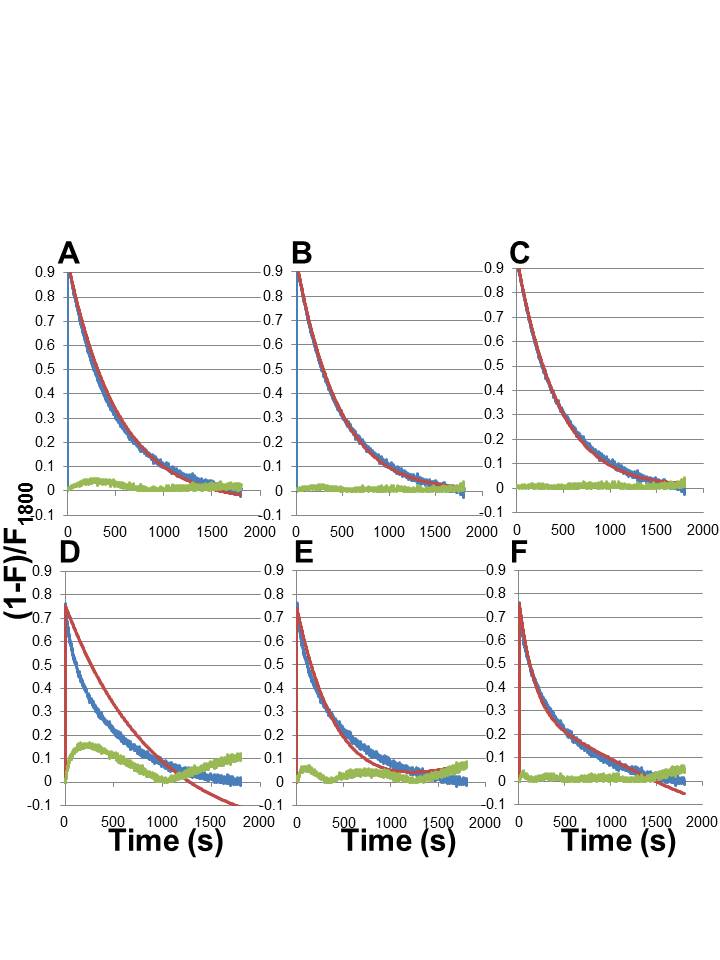
1Department of Physics, Harvard University, Cambridge, MA 02138, USA

2Laboratoire de Biochimie Théorique, CNRS UMR 9080, IBPC, Paris, France

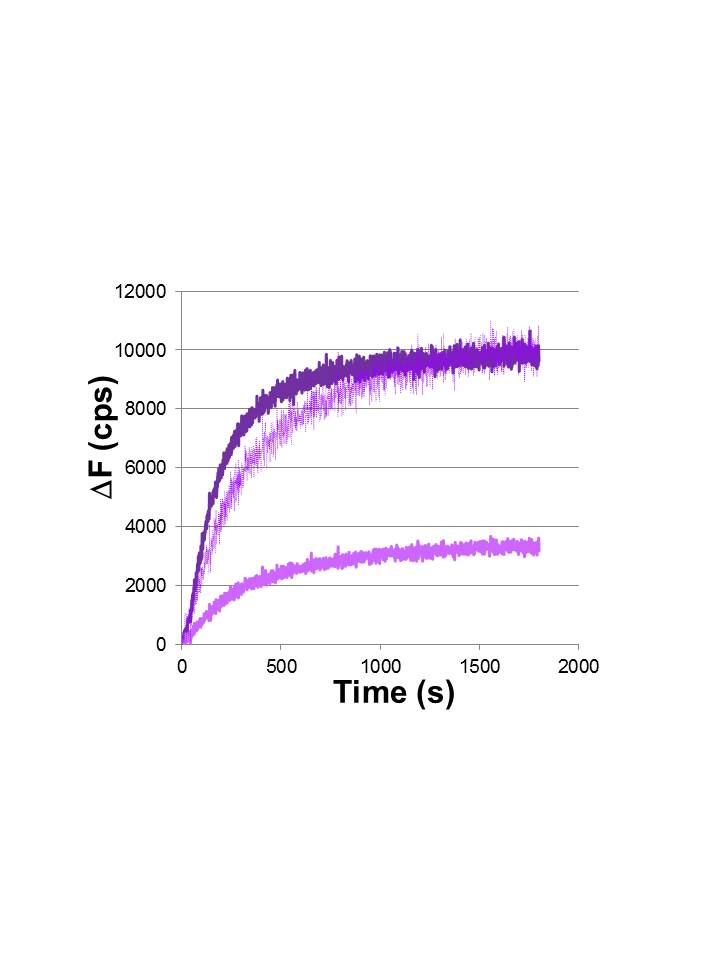
\*Corresponding Author: Mara Prentiss, 17 Oxford St, Cambridge, MA 02138. Telephone: (617) 495-2910. Email: prentiss@fas.harvard.edu



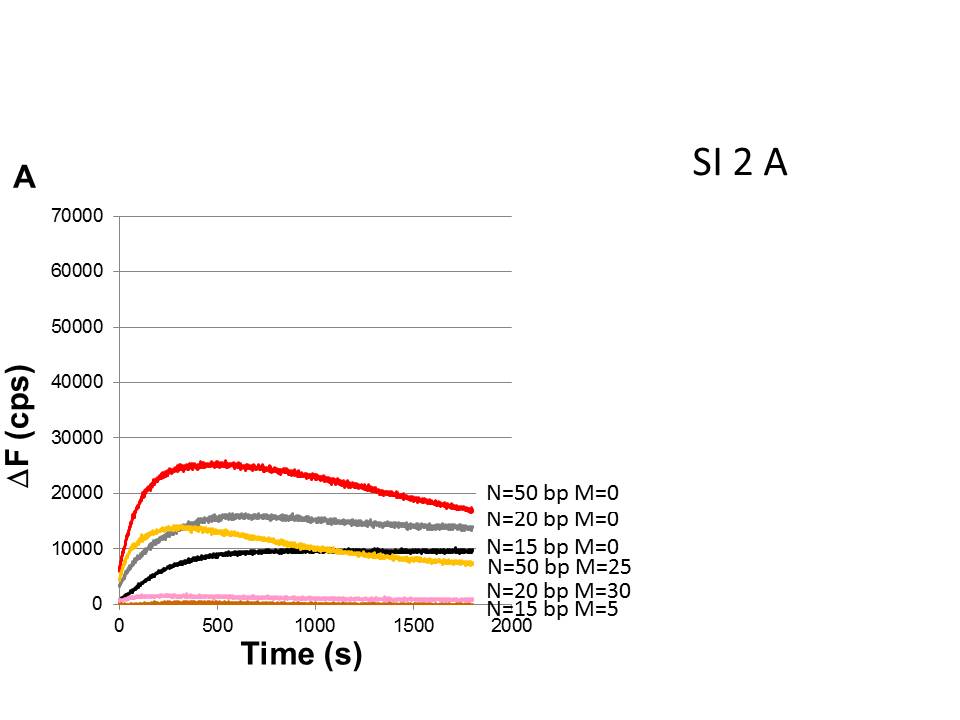
**Supplementary Figure S1. Progression of strand exchange for different N = L dsDNAs.** (A) Total fluorescence vs. time for 98 nt filaments and dsDNAs of increasing length N in ATPS; N=15 (black), 20 (gray), 50 (red), and 75 (blue) bp. (B) Same as (A) but in the presence of ATP. (C) Average F curves in ATP showing typical error bars in these FRET experiments for N = 15 bp (black), 20 bp (gray), 50 bp (red), and 75 bp (blue). Typical curves for the same conditions are shown in Figure 1C. (D) Total fluorescence vs. time for 98 nt filaments and dsDNAs of increasing length N in dATP; N=15 (black), 20 (gray), 50 (red), and 75 (blue) bp. (E) F vs. time curves in dATP show the change in cps from the initial value measured for a solution containing each dsDNA in the absence of ssDNA-RecA filaments. (F) Each value in (E) was subtracted to the final value averaged over the last ten seconds; the inset is the same data on a logarithmic y-scale.

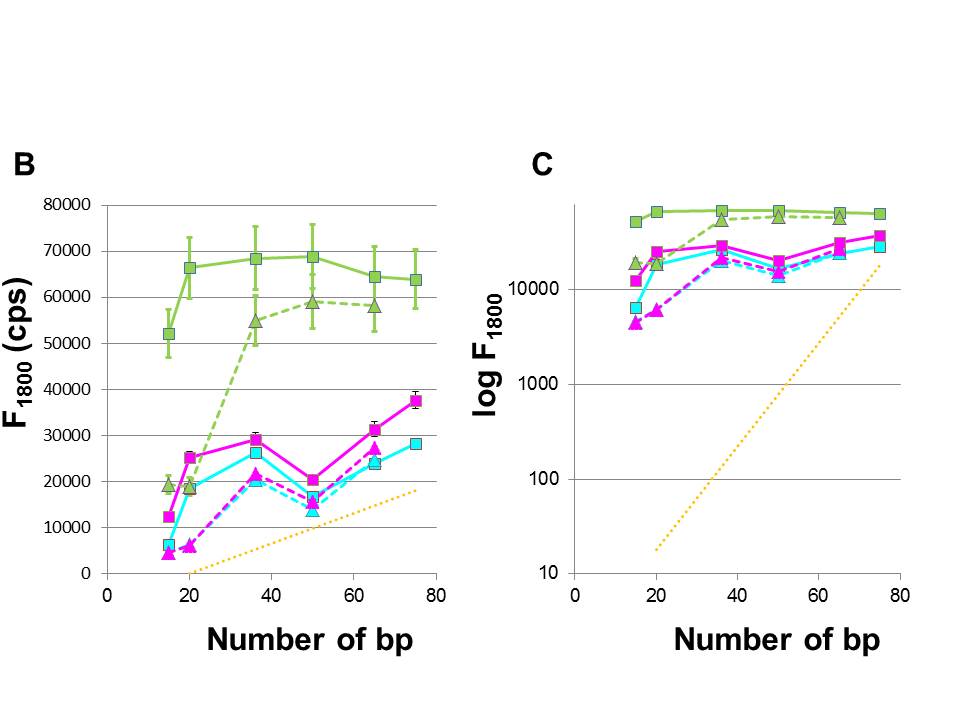


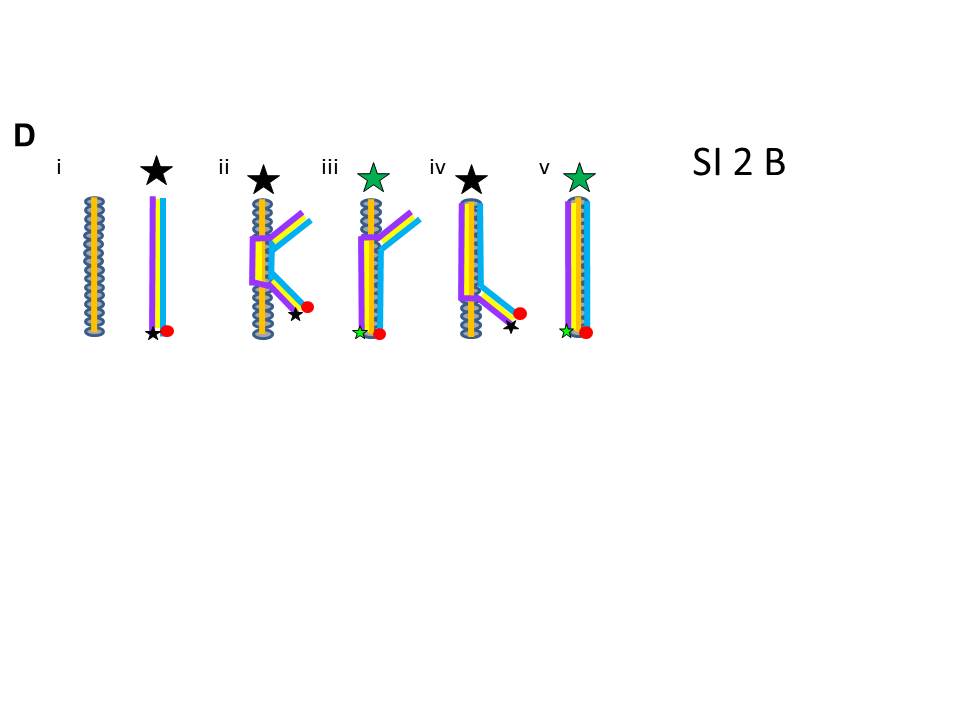
**Supplementary Figure S2. Exponential fits to typical strand exchange curves.** (A) Monoexponential fit (red) to strand exchange curve (blue) for 98 nt-RecA filament and 15 bp dsDNA; residual differences shown in green for all of the figures. (B) Biexponential fit (red) for 98 nt-RecA filament and 15 bp dsDNA data (blue). (C) Triexponential fit for 98 nt-RecA filament and 15 bp dsDNA. (D) Monoexponential fit (red) to strand exchange curve (blue) for 98 nt-RecA filament and 98 bp dsDNA. (E) Biexponential fit (red) for 98 nt-RecA filament and 98 bp dsDNA data (blue). (F) Triexponential fit (red) for 98 nt-RecA filament and 98 bp dsDNA data (blue).



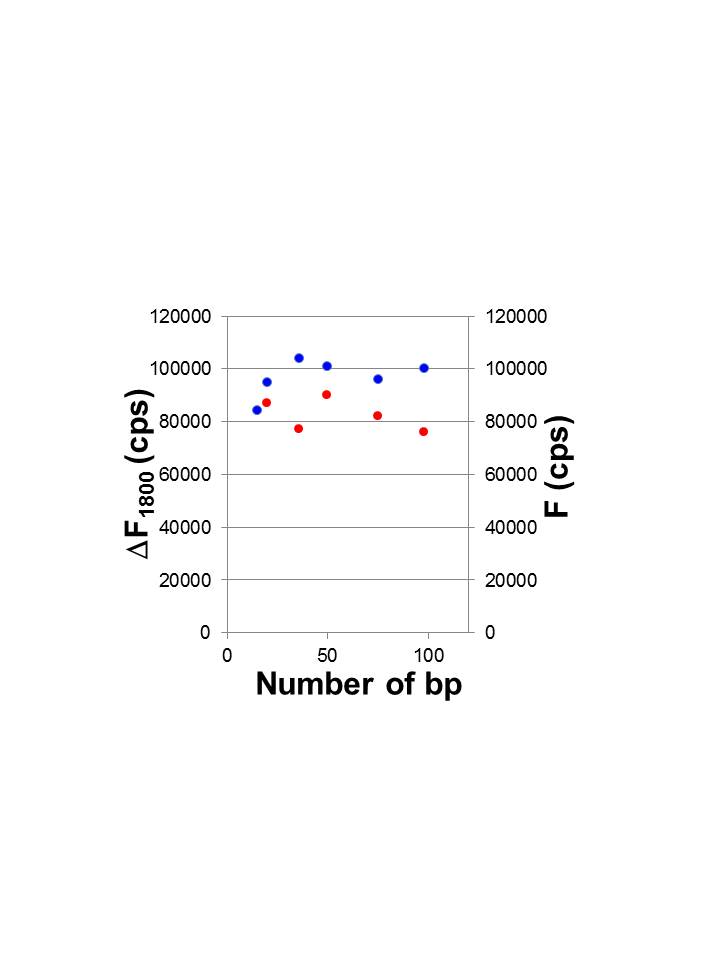
**Supplementary Figure S3. Effect of concentration on strand exchange.** Typical curves in ATP for 98 nt filament and 98 bp dsDNA: 0.06 M ssDNA (concentration in bases: 6 M), 0.06 M labeled dsDNA (concentration in bases: 12 M) (dark purple) and 0.02 M ssDNA (concentration in bases: 2 M), 0.02 M labeled dsDNA (concentration in bases: 4 M) (light purple); low concentration curve multiplied by 3 is shown by dotted line (light purple).



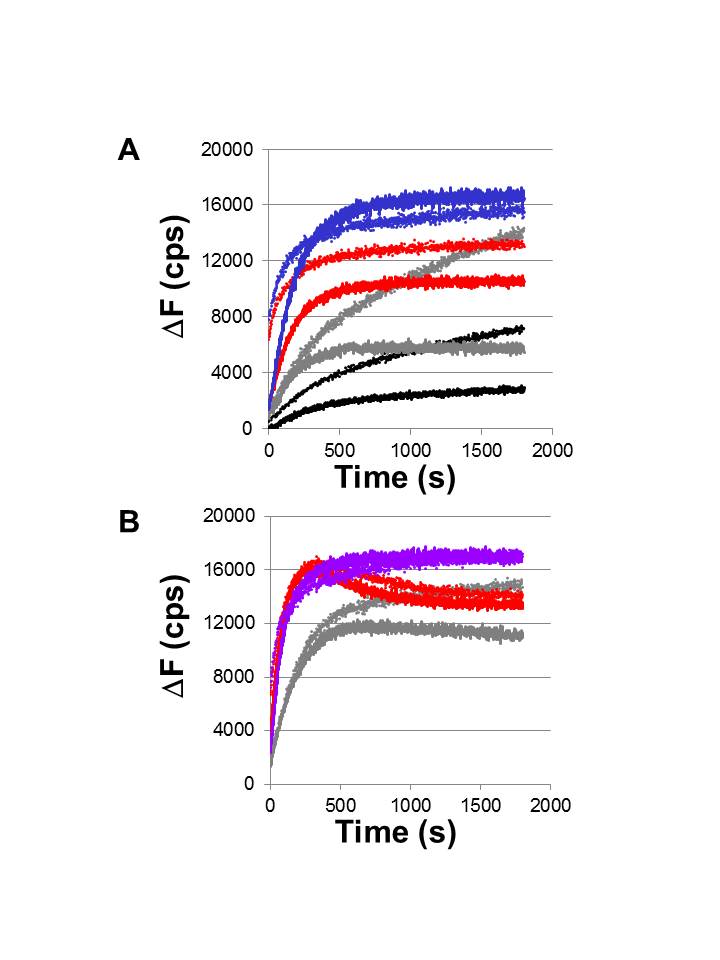




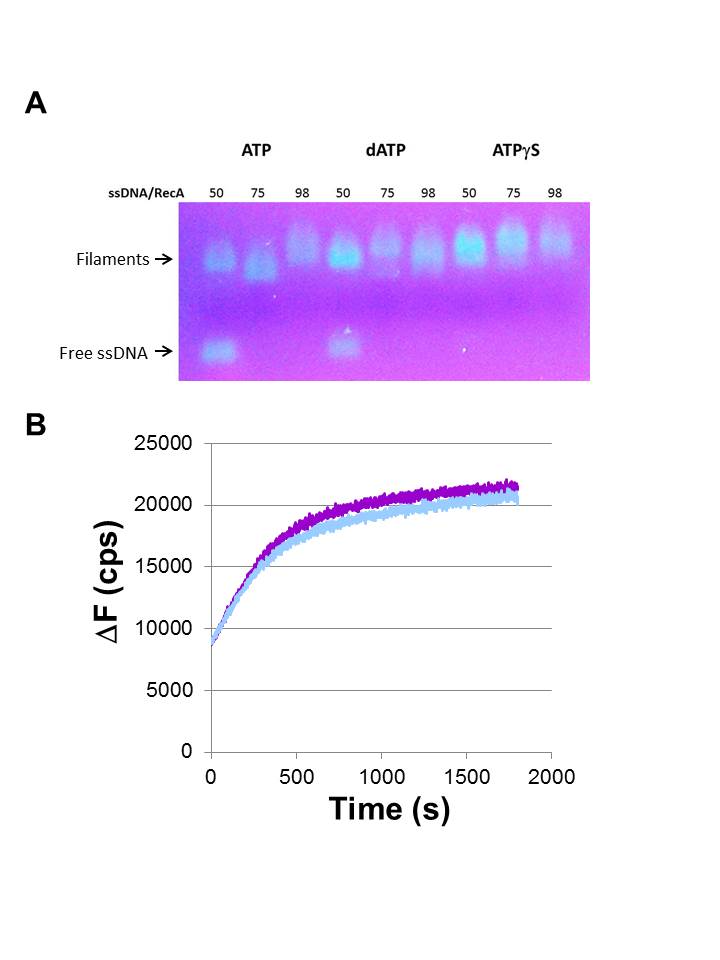
**Supplementary Figure S4. Strand exchange in the presence of heterologous tails.** (A) Effect of heterologous tails on strand exchange for 98 nt filaments and several dsDNAs in dATP shown as F vs. time for N=15 bp M=0 (black), N=15 + M=5 (brown), N=20 bp M=0 (gray), N=20 + M=30 (pink), N=50 + M=0 (red), and N=50 + M=25 (orange). (B) F1800 vs. N for 98 nt filaments in ATPS (green), ATP (cyan), and dATP (magenta), and Exp[0.13\*N](orange dotted line). The solid and dashed lines correspond to the same N values, but the dashed lines include tails with M >20 bp. (C) Same as (B) but the y-axis is in logarithmic scale. The dotted orange line shows the *in vivo* results scaled to match the F1800 values for *in vitro* results when N=75 bp. (D) Proposed bound structures during strand exchange events, some of which may not contribute to the fluorescent signal produced by labels at the 5´ end of the dsDNA.



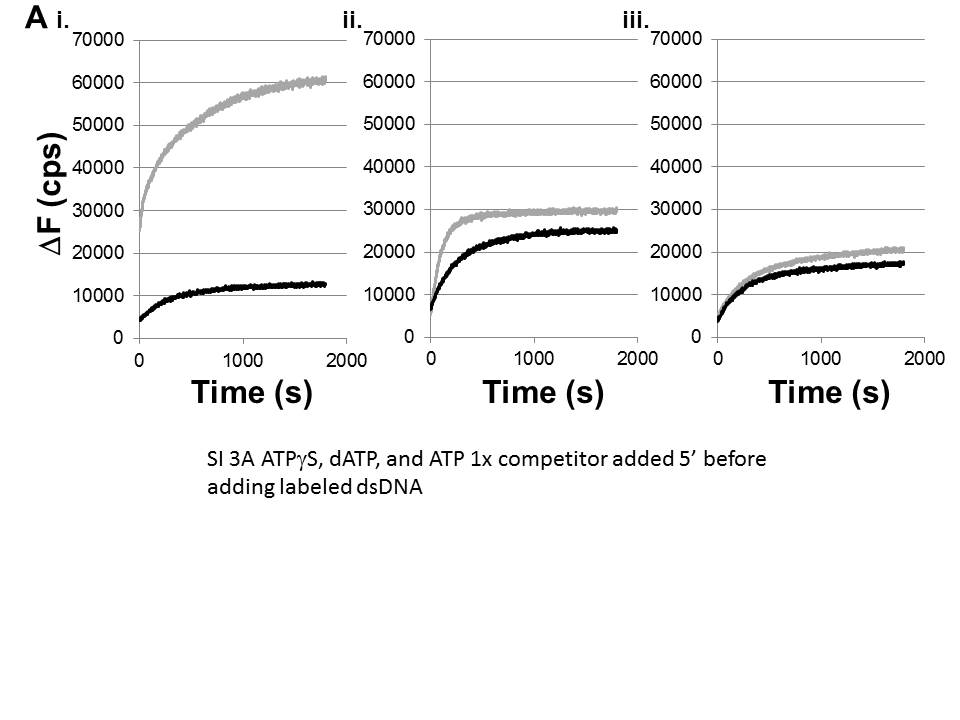
**Supplementary Figure S5. Evaluation of fluorescence signals of several dsDNA samples.** The change in fluorescence for ssDNA labeled with fluorescein after addition of complementary rhodamine oligonucleotide was measured at the same concentrations used in the strand exchange experiments: 0.06 M ssDNAs at 37°C (blue circles and left y-axis). Total fluorescence of 0.06 M dsDNA samples at 37°C containing fluorescein labeled oligonucleotide hybridized to unlabeled complementary oligonucleotide (red circles and right y-axis).

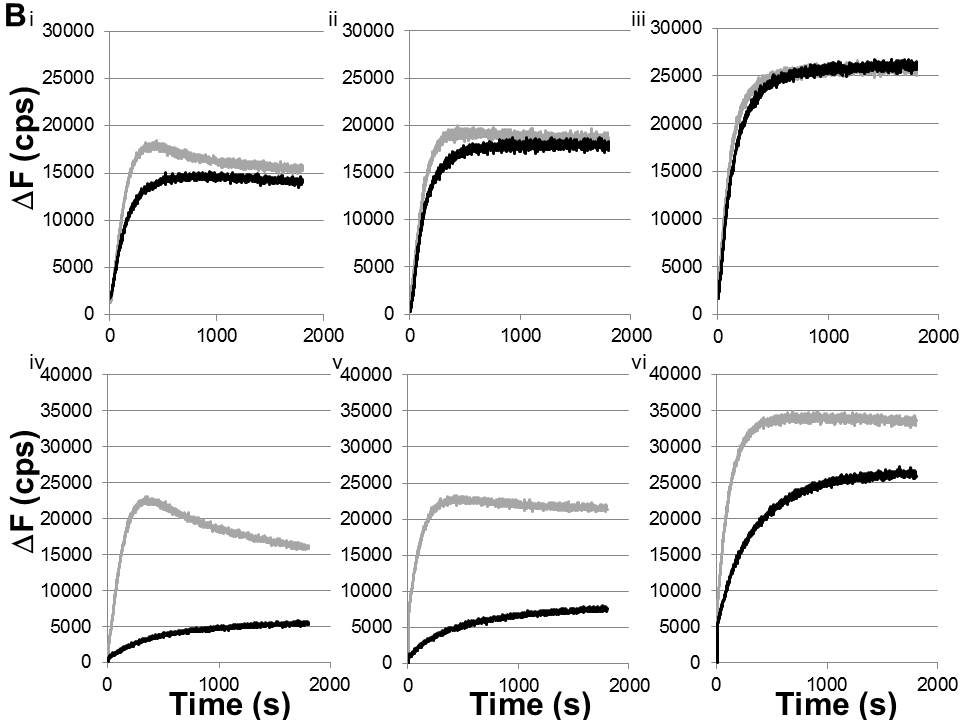
****

**Supplementary Figure S6. Effect of including an ATP regeneration system.** (A) Strand exchange reactions for 98 nt-RecA filaments in ATP vs. 15 bp (black), 20 bp (gray), 50 bp (red), and 75 bp (blue) dsDNAs in the absence (solid lines) and in the presence of an ATP regenerating system (dotted lines). (B) Strand exchange reactions for 98 nt-RecA filaments in dATP vs. 20 bp (gray), 50 bp (red), and 98 bp (purple) dsDNAs in the absence (solid lines) and in the presence of an ATP (dATP) regenerating system (dotted lines).

****

**Supplementary Figure S7. Study of filament quality in the presence of several co-substrates.** (A) Gel electrophoresis illustrating complete filament formation for 75 and 98 nt oligonucleotides in all three co-substrates and partial filament formation for 50 nt oligonucleotides in ATP and dATP. (B) Strand exchange reactions for 98 nt-RecA filaments vs. 98 bp dsDNA in the presence of 1 mM ATP (light blue) and 1 mM ATP /1 M ATPS (purple).





**Supplementary Figure S8. Competition experiments measuring F vs. time in the presence or absence of unlabeled dsDNA.** (A) Strand exchange of 98 nt ssDNA- RecA filaments with 98 bp dsDNA in ATPS (i), dATP (ii), and ATP (iii) in the absence (gray) and presence (black) of 1x concentration of identical unlabeled dsDNA; filament : total dsDNA ratio = 1:2; the unlabeled dsDNA was incubated with the filaments for five minutes before adding the labeled dsDNA. (B) Strand exchange of 98 nt ssDNA- RecA filaments with 50 bp (i), 75 bp (ii), and 98 bp (iii) dsDNA in dATP in the absence (gray) and presence (black) of 1x concentration of identical unlabeled dsDNA; filament : total dsDNA ratio = 1:1 and 50 bp (iv), 75 bp (v), and 98 bp (vi) dsDNA in dATP in the absence (gray) and presence (black) of 6x concentration of identical unlabeled dsDNA; filament : total dsDNA ratio = 1:7.

**Supplementary Data**

**Additional strand exchange reactions using an ATP or dATP regenerating system, 1 µM ATPγS added to 1 mM ATP filament solution, or lower filament concentration.**

Samples were prepared by mixing an aliquot of ssDNA 98 nt in length (final concentration 6 M in bases) with 2 M RecA (New England Biolabs) in the presence of a co-factor (1 mM ATP, dATP, or ATPS), and 0.2 M single-stranded binding protein (SSB) in RecA buffer (70 mM Tris-HCl, 10 mM MgCl2, and 5 mM dithiothreitol, pH 7.6) at 37°C for 10 minutes. An ATP or dATP regenerating system was included providing a final concentration of 10 U/ml of pyruvate kinase for ATP or 30 U/ml for dATP and 3 mM phosphoenolpyruvate (1-3). FRET experiments were subsequently performed after adding 0.06 M labeled dsDNA (final concentration total bases: 12 M for dsDNA 98 bp, 9 M for 75 bp, 6 M for 50 bp, etc.) and rapidly transferring the solution to a quartz cuvette.

Strand exchange reactions were also performed for ssDNA 98 nt in length (final concentration 6 M in bases) with 2 M RecA (New England Biolabs) in the presence of a co-factor mixture (1 mM ATP and 1 M ATPS) and 0.2 M single-stranded binding protein (SSB) in RecA buffer (70 mM Tris-HCl, 10 mM MgCl2, and 5 mM dithiothreitol, pH 7.6) at 37°C for 10 minutes. The low concentration of ATPS was added to prevent RecA unbinding from the filament 5´ end (4).

For strand exchange experiments with lower concentrations of filament and dsDNA, 0.02 M ssDNA (concentration in bases: 2 M), 0.66 M in RecA, 0.066 M SSB, and 1 mM co-substrate were incubated in RecA buffer at 37°C for 10 minutes. Finally the strand exchange reaction was started after adding 0.02 M labeled dsDNA (concentration in bases: 4 M).

**Gel experiments to evaluate filament formation.**

Oligonucleotides of increasing length and labeled with 3´-fluorescein (50, 75, and 98 nt; concentration in total bases 50, 75, and 100 M, respectively) were incubated with RecA protein (concentration ratio bases: RecA = 3:1), and SSB protein (RecA: SSB = 10:1 concentration ratio) in the presence of 1 mM co-substrate (ATP, dATP, or ATPS) and RecA buffer at 37°C for 10 minutes. Each solution was finally evaluated by gel electrophoresis in 1% agarose and 1X TBE (Tris/Borate/EDTA) buffer (5 V/cm for 2 hours). The gel was imaged with a midrange UV trans-illuminator.

**FRET measurements**

Oligonucleotides used for filament preparations were purchased from Integrated DNA Technologies (IDT). The dsDNA samples were prepared by melting the oligonucleotides at 90°C and slowly cooling down to 40°C. The sequences used for dsDNA samples are detailed below:

15 bp

5'/56-TAMN/CGG AAA AGT GCA TAT 3'

5'ATA TGC ACT TTT CCG /36-FAM/ 3'

(15+5) bp

5'/56-TAMN/CGG AAA AGT GCA TAT GGG CT3'

5' AGC CCA TAT GCA CTT TTC CG /36-FAM/ 3'

20 bp

5'/56-TAMN/CGG AAA AGT GCA TAT CCA GC 3'

5'GC TGG ATA TGC ACT TTT CCG /36-FAM/ 3'

(20+30) bp

5'- /56-TAMN/CGG AAA AGT GCA TAT CCA GCA TGA AAG AGA CGA CCA CTG CCA GGG ACG AA 3'

5' TTC GTC CCT GGC AGT GGT CGT CTC TTT CAT GCT GGA TAT GCA CTT TTC CG /36- FAM/ 3'

36 bp

5'/56-TAMN/CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT 3'

5'ATT TTC ATG ATG TTC TGC TGG ATA TGC ACT TTT CCG /36-FAM/ 3'

(36+25) bp

5'/56-TAMN/CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT A CGA ACG AGA CAA CGA ATG AAA CGG 3’

5' CCG TTT CAT TCG TTG TCT CGT TCG T ATT TTC ATG ATG TTC TGC TGG ATA TGC ACT TTT CCG /36-FAM/ 3'

50 bp

5'/56-TAMN/ CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT AAT GGG TAC TGT AA3'

5'TTA CAG TAC CCA TTA TTT TCA TGA TGT TCT GCT GGA TAT GCA CTT TTC CG /36-FAM/ 3'

(50+25) bp

5' /56-TAMN/ CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT AAT GGG TAC TGT AAA TGA GGG AGA CGA CGA CTG CCA CCG 3'

5'- CGG TGG CAG TCG TCG TCT CCC TCA TTT ACA GTA CCC ATT ATT TTC ATG ATG TTC TGC TGG ATA TGC ACT TTT CCG /36-FAM/ 3'

65 bp

5'/56-TAMN/ CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT AAT GGG TAC TGT AAA AGC GGT GCC AGT CG

5' CGA CTG GCA CCG CTT TTA CAG TAC CCA TTA TTT TCA TGA TGT TCT GCT GGA TAT GCA CTT TTC CG /36-FAM/ 3'

(65+25) bp

5'/56-TAMN/ CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT AAT GGG TAC TGT AAA AGC GGT GCC AGT CGA CGA ACG AGA CAA CGA ATG AAA CGG 3'

5'- CCG TTT CAT TCG TTG TCT CGT TCG T CGA CTG GCA CCG CTT TTA CAG TAC CCA TTA TTT TCA TGA TGT TCT GCT GGA TAT GCA CTT TTC CG /36-FAM/ 3'

75 bp

5' /56-TAMN/ CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT AAT GGG TAC TGT AA3'

5' TT ACA GTA CCC ATT ATT TTC ATG ATG TTC TGC TGG ATA TGC ACT TTT CCG /36-FAM/ 3'

Oligonucleotides used in the competition experiments

98 bp

5´/56-TAMN/ CGG AAA TCA C TC CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG CGT GGA GTG CAG GTA TAC AGA TT 3´

5´ AA TCT GTA TAC CTG CAC TCC ACG CCA CTG AGG TAT GCC GCA TTG CAC TTT CGT CCC TGG CAG TGG TCG TCT CTT TCA TAT ACC CGG GAG TGA TTT CCG /36-FAM/ 3´

75 bp

5´/56-TAMN/ CGG AAA TCA CTC CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG 3´

5´ CCA CTG AGG TAT GCC GCA TTG CAC TTT CGT CCC TGG CAG TGG TCG TCT CTT TCA TAT ACC CGG GAG TGA TTT CCG /36-FAM/ 3´

50 bp

5´/56-TAMN/ CGG AAA TCA CTC CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AA 3´

5´ TT CGT CCC TGG CAG TGG TCG TCT CTT TCA TAT ACC CGG GAG TGA TTT CCG /36-FAM/ 3´

Oligonucleotides containing internal labels for dsDNA:

d=10

5´ CGG AAA TCA C/iRho-T/C CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG CGT GGA GTG CAG GTA 3´

5´ TAC CTG CAC TCC ACG CCA CTG AGG TAT GCC GCA TTG CAC TTT CGT CCC TGG CAG TGG TCG TCT CTT TCA TAT ACC CGG GAG /iFluor-T/GA TTT CCG 3´

d=20

5´ CGG AAA TCA CTC CCG GGT A/iRho-T/A TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG CGT GGA GTG CAG GTA 3´

5´ TAC CTG CAC TCC ACG CCA CTG AGG TAT GCC GCA TTG CAC TTT CGT CCC TGG CAG TGG TCG TCT CTT TCA TA/iFluor-T/ ACC CGG GAG TGA TTT CCG 3´

d=36

5´ CGG AAA TCA CTC CCG GGT ATA TGA AAG AGA CGA CCA C/iRho-T/G CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG CGT GGA GTG CAG GTA 3´

5´ TAC CTG CAC TCC ACG CCA CTG AGG TAT GCC GCA TTG CAC TTT CGT CCC TGG CAG /iFluor-T/GG TCG TCT CTT TCA TAT ACC CGG GAG TGA TTT CCG 3´

d=56

5´ CGG AAA TCA CTC CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA /iRho-T/GC GGC ATA CCT CAG TGG CGT GGA GTG CAG GTA 3´

5´ TAC CTG CAC TCC ACG CCA CTG AGG TAT GCC GCA T/iFluor-T/G CAC TTT CGT CCC TGG CAG TGG TCG TCT CTT TCA TAT ACC CGG GAG TGA TTT CCG 3´

Oligonucleotides for filaments:

98 nt starting at position 10184-10281 in lambda phage sequence

5' CGG AAA AGT GCA TAT CCA GCA GAA CAT CAT GAA AAT AAT GGG TAC TGT AAA AGC GGT GCC AGT CGG CAT ACT CCG TGG ATG ACA TCC CGG CAA GCA TG 3'

98 nt: starting at position 19386-19483 in lambda phage sequence

5´ CGG AAA TCA CTC CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG CGT GGA GTG CAG GTA TAC AGA TT 3´

75 nt:

5´ CGG AAA TCA CTC CCG GGT ATA TGA AAG AGA CGA CCA CTG CCA GGG ACG AAA GTG CAA TGC GGC ATA CCT CAG TGG 3´

**Studies of filament quality**

One might be concerned that ATP hydrolysis depletes the concentration of ATP over time. Previous work has shown that this depletion can be removed by an ATP regeneration system (1-3). Thus, we conducted additional experiments in the presence of an ATP regeneration system. Supplementary Figure S6 shows that for dsDNA longer than ~ 20 bp, the F1800 values are insensitive to the presence of the regeneration system. Thus, in vitro artifacts associated with the depletion of ATP are unlikely to have played a significant role in the experimental results.

To further evaluate filament quality, we performed gel assays that measure filament formation (Supplementary Figure S7A). Consistent with previous work (5), we found that for filaments with lengths < 75 nt, filament formation was incomplete when ATP or dATP was the cofactor, though 50 nt filaments showed complete formation in ATPS. For filaments 75 nt or longer, filament formation was complete in all of the cofactors. In the presence of ATP, filaments contain regions to which no RecA is bound, binding of RecA to the filaments remains reversible, and the unbinding of RecA from ssDNA occurs most frequently at the 5´ ends of the filaments (6). Adding a small concentration of ATPS to the usual ATP concentration could stabilize the 5´ ends of filaments without compromising the hydrolysis efficiency of the filaments (4). Thus, we performed additional experiments in 1 mM ATP and 1 M ATPS (Supplementary Figure S7B). Again, the results were similar to those obtained with ATP as the only cofactor.

**Numerical simulations of fluorescence vs. time curves**

The Monte-Carlo simulations assumed that if the system can form irreversible products, pirr is the probability that a collision will lead to the formation of an irreversible product. Typical pirr values were ~ 0.2 to to 1.5 %. The model assumes that these products initiate at the 5´ end and achieve the maximum Lprod allowed by the interaction between the dsDNA and the filament. Thus, for these interactions, all labels report the same fluorescence.

If initiation does not occur at the 5´ end of the filament, then the model assumes that the interaction produces a product with a length of < 40 bp. If that product does not reach the 3´ end of the dsDNA or the filament, then the interaction does not contribute to the fluorescence because it is assumed to be too short lived. If the interaction does reach the 3´ end, all of the labels between the initiation point and the 3´ end contribute fluorescence for a time that depends on the length of the product. Various length dependences were considered. For the label interactions that initiate between s=36 and s=56 bp, reasonable fits were obtained with lifetimes between 15 and 70 seconds, depending on pirr and the lifetimes of the shorter lived states that initiate between the s=56 and the 3´ end. Once the product is reversed, no labels contribute to the fluorescence until a new product is initiated. Product initiations are separated by a characteristic time delay with an exponential distribution. Typical delay values were < 10 seconds. After the delay, a new product was formed. The process was repeated until the accumulated simulation time was equal to the observation time. Once the simulation time reached the observation time, we added up the fluorescence contributions due to each label at the end of the simulation to arrive at the total fluorescence that would be observed at that time. The calculation was repeated for each observation time t. The total number of realizations varied between 1000 and 100,000.

**Molecular dynamics protocol**

Two turns of RecA filament with mixed ATP and ADP interfaces were constructed using the molecular modeling library PTools (7) and its associated tools and the procedure described in (8). Each turn comprised five successive RecA-ADP.AlF4- monomers, where ADP.AlF4- is a non-hydrolyzable homolog of ATP, followed by one RecA-ADP monomer. The binding geometries of RecA-ATP.AlF4- monomers were taken from the two central monomers of the crystal structure with PDB code 3CMW (extended filament, (9)), while the interface between the central RecA-ADP monomer and its adjacent RecA-ADP.AlF4- monomer corresponds to the binding geometry in the compressed filament form (PDB code 2REB (10). This produced a dodecameric filament with all interfaces in the ATP geometry except between monomers 5 and 6. Trinucleotide DNA segments respectively taken from the crystal structures 3CMX for the heteroduplex in site I (strands 1 and 2) and from our model of DNA bound to the presynaptic filament for strand 3 in site II (11) were successively added to each filament monomer in such a way as to conserve their reference binding geometries with these monomers. Disruption of the DNA backbone continuity resulting from the filament perturbation near monomer 6 was corrected via energy minimization using the namd2.0 software (12). The resulting structure was solvated using a TIP3P water model in a 140x144x225 Å3 box, submitted to periodic boundary conditions. Electric neutrality and a physiological ionic concentration of 0.15 mol/l were obtained by the addition of 503 Na+ and 348 Cl− ions, leading to a total of 439,000 atoms.

The dodecameric filament bound to the three DNA strands was submitted to a 120 ns molecular dynamics simulation at 310 K, including 20 ns of equilibration and 100 ns of production, using the namd2.0 software (12) with the CHARMM 27 force field, including the CMAP correction (13). Long-range electrostatics were accounted for using the particle-mesh Ewald method and temperature and pressure were controlled along the simulation using Langevin dynamics and a Nose –Hoover–Langevin piston.

**Supplementary references**

1. Rehrauer, W.M. and Kowalczykowski, S.C. (1993) Alteration of the nucleoside triphosphate (NTP) catalytic domain within Escherichia coli recA protein attenuates NTP hydrolysis but not joint molecule formation. *J. Biol. Chem.* **268**, 1292-1297.
2. Menetski, J.P. and Kowalczykowski, S.C. (1989) Enhancement of Escherichia coli RecA protein enzymatic function by dATP. *Biochem*. **28**, 5871-5881.

# Kim, JI., Sharma, A.K., Abbott, S.N., Wood, E.A., Dwyer, D.W., Jambura, A., Minton, K.W., Inman, R..B, Daly, M.J. and Cox, M.M. (2002) RecA protein from the extremely radioresistant bacterium Deinococcus radiodurans: expression, purification, and characterization. *J.Bacteriol*. 184, 1649-1660.

# Lindsley, J.E. and Cox, M.M. (1989) Dissociation pathway for RecA nucleoprotein filaments formed on linear duplex DNA. *J. Mol. Biol.* 205, 695-711.

1. Bianco, P.R. and Weinstock, G.M. (1996) Interaction of the RecA protein of Escherichia coli with single-stranded oligodeoxyribonucleotides. *Nucleic Acids Res.,* **24**, 4933-4939
2. Lindsley, J.E. and Cox, M.M. (1990) Assembly and disassembly of RecA protein filaments occurs at opposite filament ends: Relationship to DNA strand exchange. *J. Biol. Chem*. **265**, 9043-9054.
3. Saladin, A., Fiorucci, S., Poulain, P., Prévost, C. and Zacharias, M. (2009) PTools: an open source molecular docking library. *BMC Struct. Biol*. **9**, 27.
4. Boyer, B., Ezelin, J., Poulain, P., Saladin, A., Zacharias, M., Robert, C.H. and Prévost, C. (2015) An integrative approach to the study of filamentous oligomeric assemblies, with application to RecA. *PLoS ONE*, **10**, e0116414.
5. Chen, Z., Yang, H. and Pavletich, N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature*, **453**, 489-494.
6. Story, R.M. and Steitz, T.A. (1992) Structure of the RecA Protein ADP Complex. *Nature*, **355**, 374-376.
7. Yang, D.R., Boyer, B., Prevost, C., Danilowicz, C. and Prentiss, M. (2015) Integrating multi-scale data on homologous recombination into a new recognition mechanism based on simulations of the RecA-ssDNA/dsDNA structure. *Nucleic Acids Res.*, **43**, 10251-10263.
8. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K. (2005) Scalable Molecular Dynamics with NAMD. *J. Comp. Chem.,* **26**, 1781-1802.
9. MacKerell, A.D., Feig, M. and Brooks, C.L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem.,* **25**, 1400-1415.