Supplemental Information for

A 5′-to- 3′ Strand Exchange Polarity is Intrinsic to RecA Nucleoprotein Filaments in the Absence of ATP Hydrolysis

Yu-Hsuan Lin[1#], Chia-Chieh Chu[1#], Hsiu-Fang Fan[2#], Pang-Yen Wang[1#], Michael M. Cox[3]* and Hung-Wen Li[1]*

Sum of the two-dice model

In order to simulate the initial BM distribution, we utilized a "sum of the two-dice" model. In the model, we arbitrarily divide our surface-linked DNA substrate into 4 segments (x_1, $x_2 \ldots x_4$) that can contribute to BM, and the incoming DNA bound into a nucleoprotein filament into 10 segments (y_1, $y_2 \ldots y_{10}$) that can make a similar segmental contribution to BM. The 10:4 ratio reflects the length and mechanical properties conferred on the DNA when bound by RecA protein (1), which results in a larger Brownian motion. Similar results (Figure 2A) are obtained with any number of segments as long as the 2.5:1 ratio is maintained. The bead is attached to the incoming nucleoprotein filament. When tethered, the initial Brownian motion of the bead will reflect the point at which the filament first contacts the surface-linked DNA duplex, and will be the sum of the DNA segments between the contact point and the surface plus the nucleoprotein segments from the contact point to the end with the bead attached. The RecA filament is dice 1 and the surface-bound duplex DNA is dice 2. The initial contact point of RecA filament and surface-bound DNA decides the relative contribution of these two molecules to the initial BM value. Therefore, the distribution of initial BM value provides information on the initial contact point. Initial BM values are measured as in Figure S1A. We denote X as the outcome of N-faced dice 1 for RecA nucleoprotein filaments, and Y is the outcome of M-faced dice 2 for surface-bound DNA. The observed initial BM value is the sum of the two dice, ($X + Y$), with values ranging from 2 to 14. These values can be evenly distributed over the range of observed initial Brownian motion values (~10-80 nm; Figure S1B-D).

In the case of "random collision", collision between the RecA filament (N-faced dice 1) and surface-bound DNA (M-faced dice 2) can occur anywhere at random, i.e. dice 1 and 2 are both equally weighted. Random collision leads to a distribution of initial BM values that produces a trapezoidal shape due to N>M (there are four dice combinations that produce sums of 5, 6, 7, 8, 9, 10, and 11; fewer combinations that produce higher and lower sums; Figure S1B). This simple model describes the experimental results for the un-successful tethers very well (Figure 2B).

This simple random collision model suggests that invading events occur through random collisions between RecA filaments and surface-bound DNA (Figure 2B and 2E). In order to model the observed initial BM distribution for successful events (with a peak at lower BM), it is necessary to place a higher probability at the 5′-end segment of RecA nucleoprotein filaments with the probability gradually reduced as you progress away from the 5′-end (the unfair dice 1). The modeled data (as from Figure S1C) is shown in red triangles in Fig. 2. Similarly, in the case of a 3′-end preference simulation, a higher probability must be placed on the 3′-end proximal segment of the RecA filaments (as from Figure S1D). Results are presented with blue circles in Fig. 2. In our experimental results, the observations correspond strikingly to the 5′-end preference expectation based on this simple model.

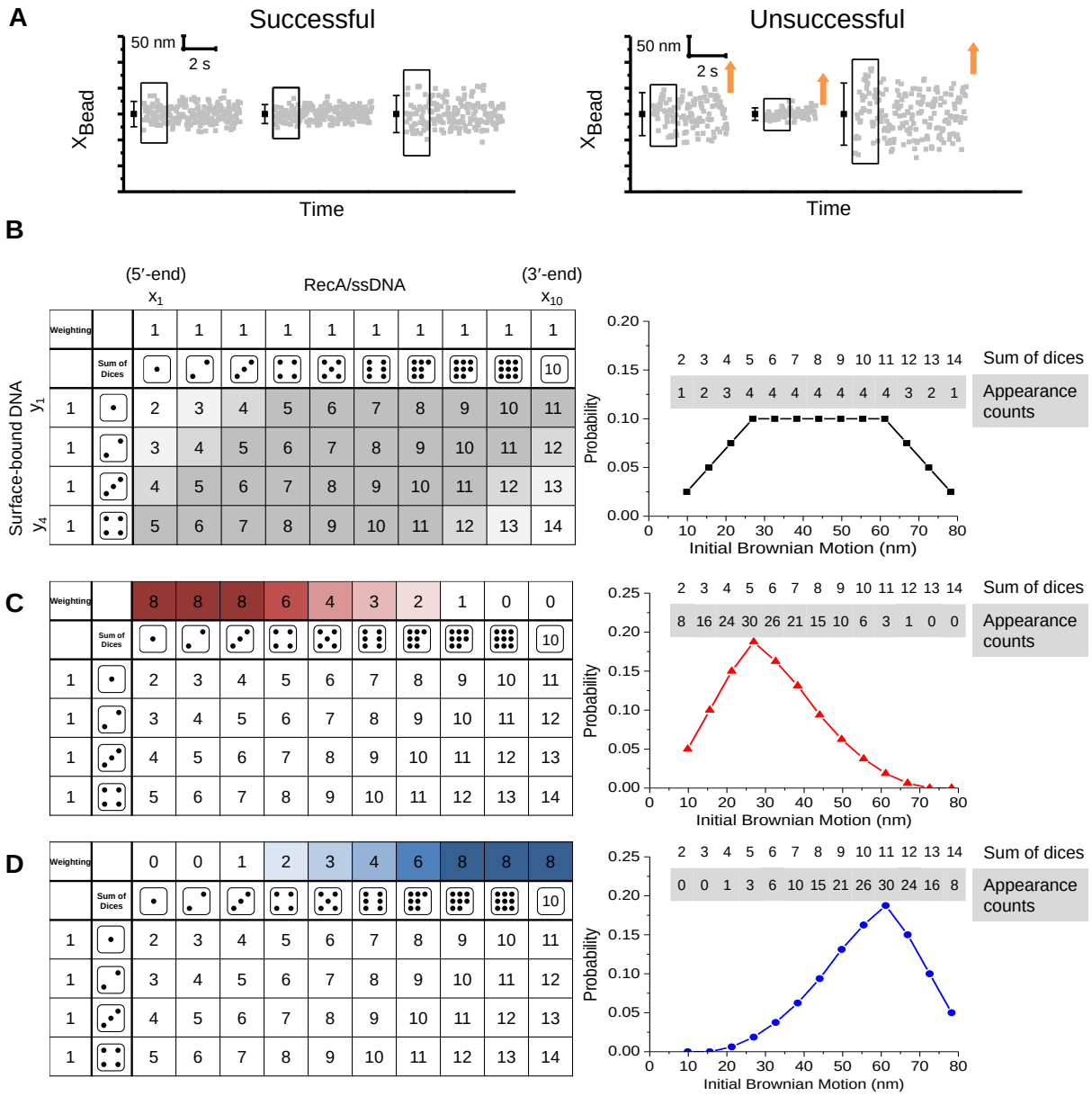

Figure S1. Initial Brownian motion (BM) determination and the sum of the two-dice model. **(A, left).** Exemplary time-courses of bead centroid position in x-axis (X_{Bead}) for the successful RecA-mediated strand exchange reactions in first several seconds. The BM value is defined as the standard deviation of bead centroid position (shown as the black bar range). Initial BM value is calculated from the standard deviation of centroid position during the first 1.3 seconds. Data used to define initial BM is boxed. **(A, right).** Exemplary bead centroid position time-courses for the unsuccessful strand exchange reactions. These transient tethers disappeared within 10 seconds of tethering (indicated by the upward arrows). **(B).** Results of the sum of the two-dice model for random collision. **(C).** Results of the sum of the two-dice model for a 5′-end preference. **(D).** Results of the sum of the two-dice model for a 3′-end preference.

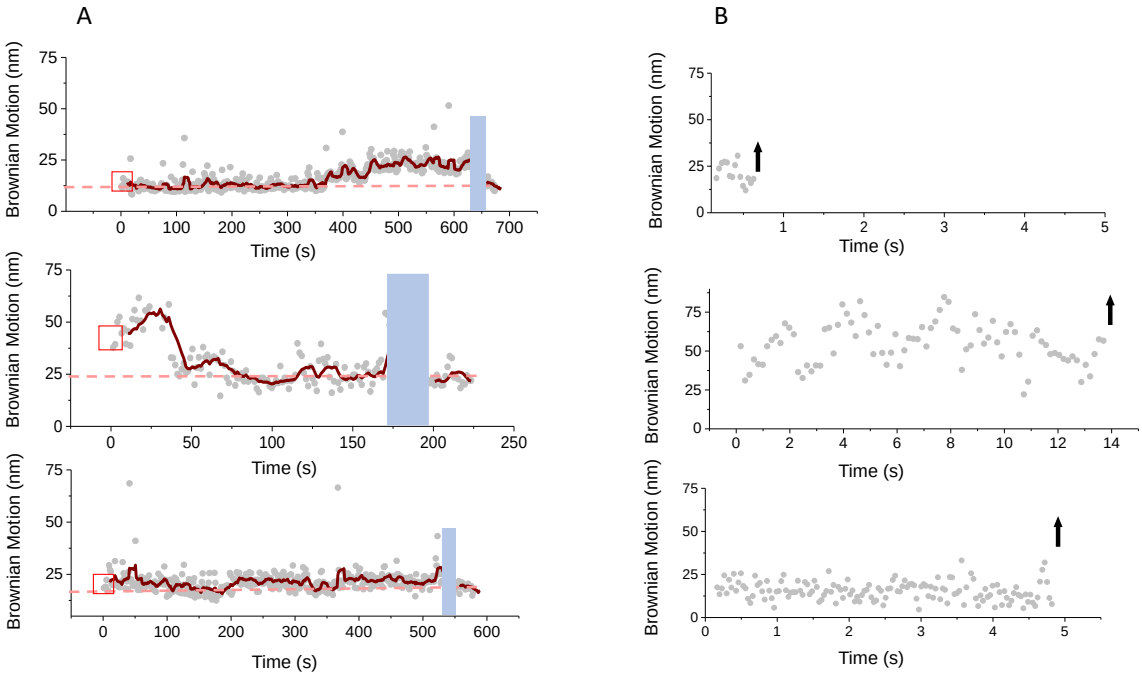

Figure S2. Exemplary time-courses for successful and unsuccessful strand exchange events. The reactions were carried out in 2 mM ATP with the presence of ATP regeneration system. Considering the nM-scale of DNA used and ~ 12 minutes observation window, the presence or absence of ATP regeneration system does not alter the reactions. All unsuccessful strand exchange events ended by the disappearance of tethered beads in the field-of-view (shown as an upward arrow).

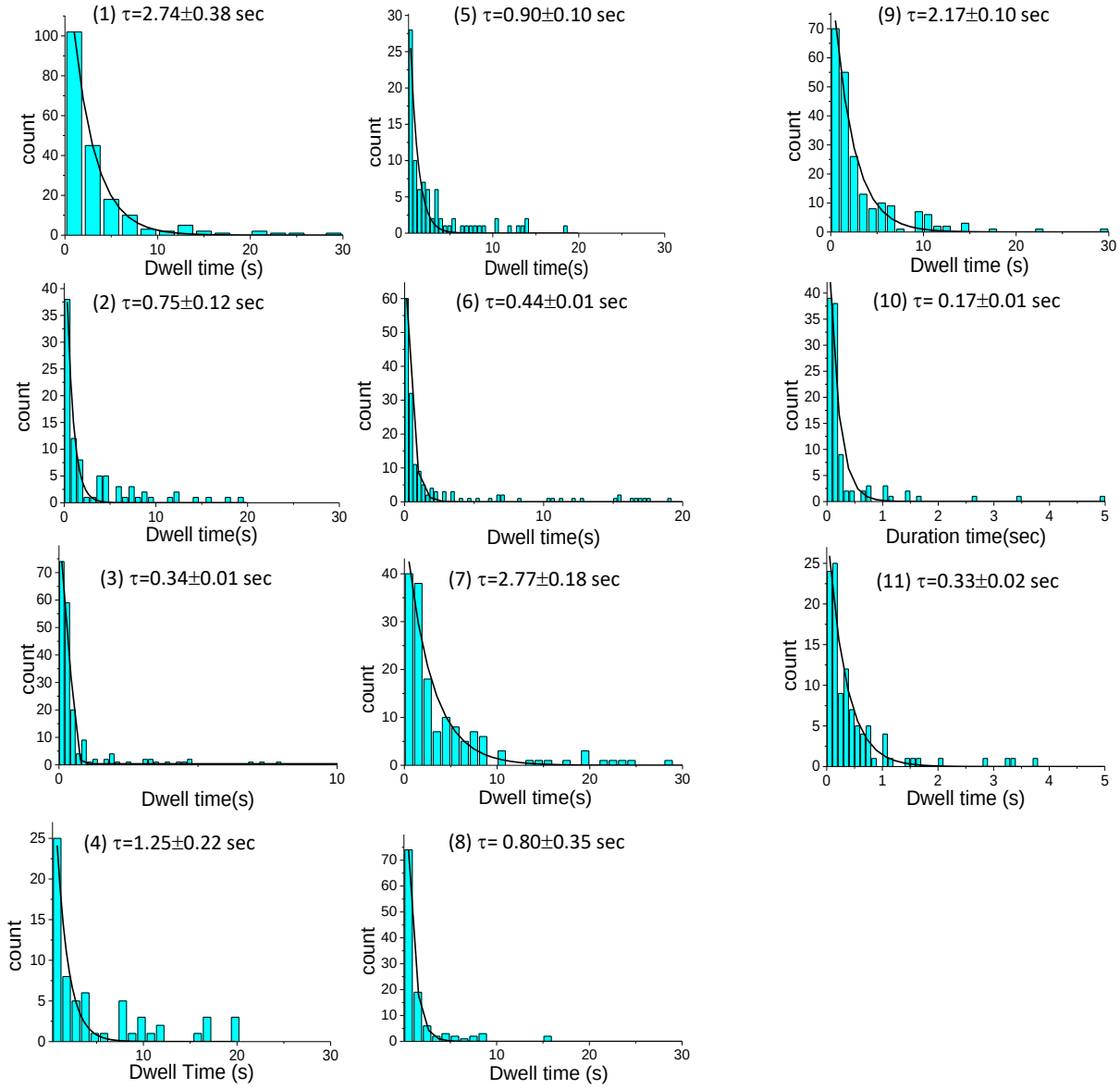

Figure S3. Dwell time histograms for unsuccessful strand exchange events. The reaction conditions of each histogram follow the numbering listed in Table 1.

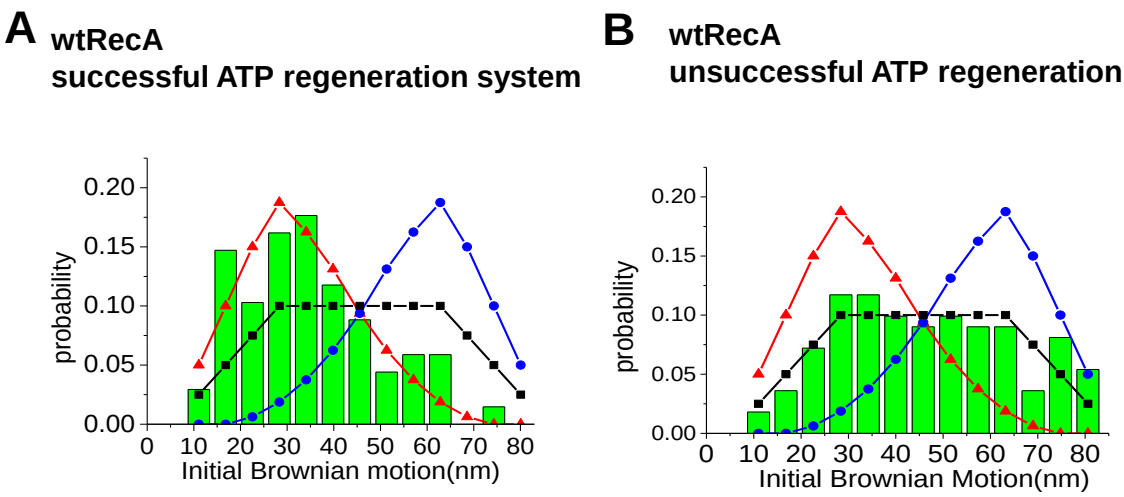

Figure S4. Initial BM distributions for successful and unsuccessful strand exchange of wild-type RecA carried out with ATP in the presence of ATP regeneration system (4 units/ml pyruvate kinase and 1 mM phosphoenolpyruvate, Sigma). (A) Initial BM histogram for successful events (N=69). (B) Initial BM histogram for unsuccessful events (N=111).

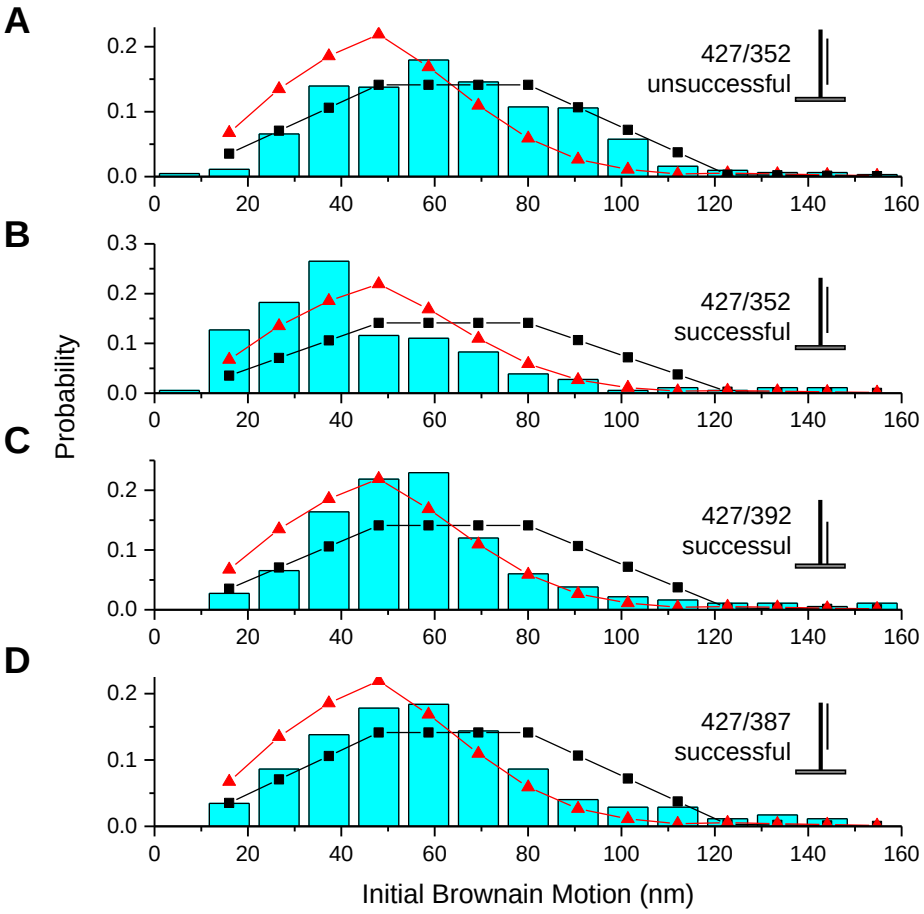

Figure S5. Sum of the two-dice model explains the observation of initial Brownian motion amplitude for strand exchange reactions with various different DNA substrates. **(A).** Initial Brownian motion amplitude probability distribution histogram of unsuccessful strand exchange of 427/352 substrates (surface bound hybrid DNA 427/352 with ssDNA gaps on both ends) with simulation result from the model (N=624). **(B).** Initial Brownian motion amplitude probability distribution histogram of successful strand exchange (427/352 substrate) (N=181). **(C).** Initial Brownian motion amplitude probability distribution histogram of successful strand exchange of 427/392 substrates (N=206). **(D).** Initial Brownian motion amplitude probability distribution histogram of successful strand exchange of 427/387 substrates (N=174). **(A-D).** All simulated curves have the same fixed parameters of M=4, N=10, and the weight in surface bound hybrid DNA, but the weight in RecA nucleoprotein filaments is different for successful (red) and unsuccessful (black) curves. Cyan bars represent the experimental results. Red triangle represents the simulation results with weights in $x_{i-1}>x_i$ (i=2 ~10) with secondary structure correction. Black square represents the simulation results with weights in x_i are the same with secondary structure correction.

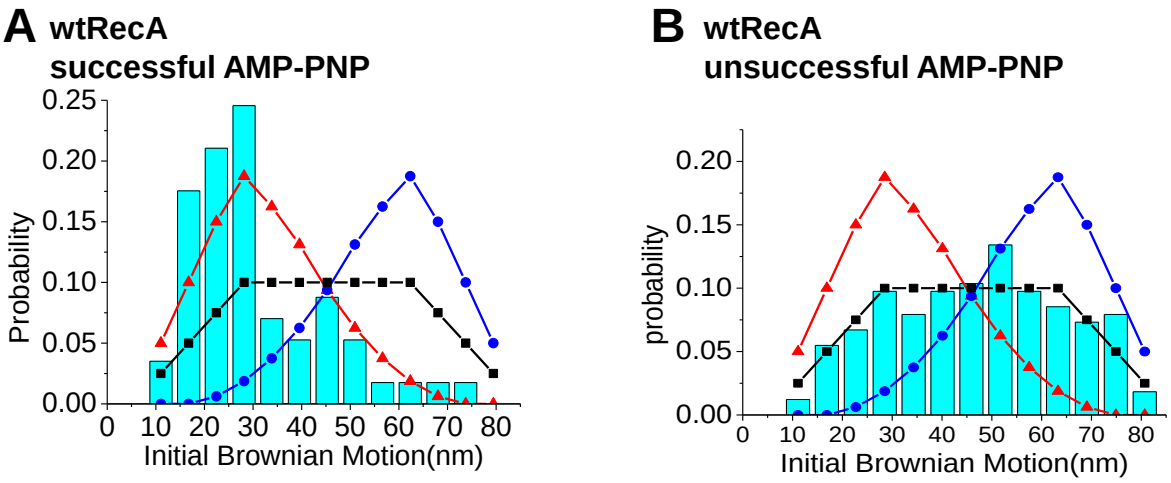

Figure S6. Initial BM distributions for successful and unsuccessful strand exchange of wild-type RecA carried out in the presence of 2 mM AMP-PNP. (A) Initial BM histogram for successful events (N=57). (B) Initial BM histogram for unsuccessful events (N=164).

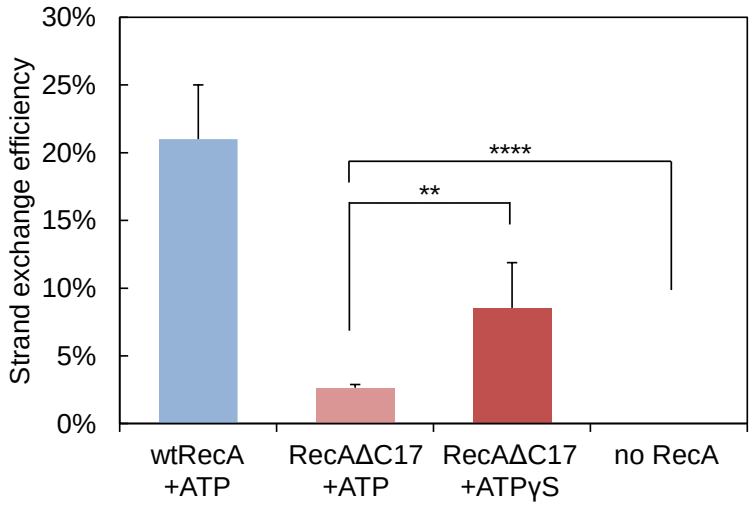

Figure S7. The strand exchange efficiency of wild-type RecA is 21.0 ± 4.0 % (modified from reference (2), N=7). The strand exchange efficiency of RecAΔC17 mediated reaction is lower than that of wild-type RecA. The efficiency of *RecAΔ*C17/ATP is significantly lower than *RecAΔ*C17/ATPγS (2.6 ± 0.3 % for 2 mM ATP (N=6), and 8.5 ± 3.3 % for ATPγS (N=6), p-value=0.0076). Invading experiments with no RecA, which are negative controls, with a 0 ± 0 % efficiency (N=5, all zero), is significantly lower than RecAΔC17/ATP reaction (p-value = 1.9×10^{-6}).

Table SI

Primers used for invading strand experiments

DNA substrate	template	Primer sequence
229/149 hybrid dsDNA	3χF3χH	5′-DigN-ACTACGATACGGGAGGGC 5′-Phos-CACCACGATGCCTGCAC 5′-OH-GCGCTGGCTGGTGGTCTAGA 5′-Phos-CCCAGTGCTGCAATGATACC
229 ssDNA	3χF3χH	5′-Phos-ACTACGATACGGGAGGGC 5′-Bio-CACCACGATGCCTGCAC
427/352 hybrid dsDNA	3χF3χH	5′-DigN-ACTACGATACGGGAGGGC 5′-Phos-CGGATGGCATGACAGTAAG 5′-OH-TGAGTGATAACACTGCGGC 5′-Phos-CCCAGTGCTGCAATGATACC
427 dsDNA	3χF3χH	5′-DigN-ACTACGATACGGGAGGGC 5′-Phos-CGGATGGCATGACAGTAAG
427 ssDNA	3χF3χH	5′-Phos-ACTACGATACGGGAGGGC 5′-Bio-CGGATGGCATGACAGTAAG
Nonhomologous 427 ssDNA	pBR322	5′-Bio-GATGGCGCCCAACAGTCCC 5′-Phos-CCTGGATGCTGTAGGCATAGG

DNA substrate	template	Primer sequence
231/167 hybrid dsDNA	3χF3χH	5′-Phos-GTGTCCACCAGCTCAGCATCGACCAC CAGCTCGAGTGCAGGCATCGTGGTG-DigN
		5′-Phos-GCGCTGGCTGGTGGTCTAGA 5′-OH-ACTACGATACGGGAGGGC
		5′-Phos-CCCAGTGCTGCAATGATACC 5′-OH-GCTGAGCTGGTGGACACG
231 ssDNA	3χF3χH	5′-Phos-CATTGCAGCACTGGGGCCAGATGGT AAGCCCTCCCGTATCGTAGT-Bio
		5′-Phos- ATACCGCGAGACCCACGC 5′-OH-CACCACGATGCCTGCAC
		5′-OH-GCCCCAGTGCTGCAATGATACCGCGA GACCCACGCT **(Oligo1)** 5′-OH-AGCGTGGGTCTCGCGGTATCATTGCA GCACTGGGGC **(Oligo2)**

Table SII

R^2 values of all invading data vs. different simulation

	Successful			Unsuccessful		
	5′ end preference	3′ end preference	No end preference	5′ end preference	3′ end preference	No end preference
wtRecA+ATP (5′-bead label, Figure 2B-C)	0.830	0.259	0.180	0.411	0.008	0.651
wtRecA+ATP (3′bead label, Figure 2E-F)	0.832	0.156	0.358	0.400	0.034	0.863
wtRecA+ATPγS (Figure 3A-B)	0.787	0.169	0.272	0.040	0.241	0.623
RecAΔC17+ATP (Figure 4A-B)	0.409	0.027	0.828	0.486	0.008	0.752
RecAΔC17+ATPγS (Figure 4C-D)	0.876	0.446	0.078	0.574	0.005	0.610

1. Hegner, M., Smith, S.B. and Bustamante, C. (1999) Polymerization and mechanical properties of single RecA-DNA filaments. *Proc. Natl. Acad. Sci. U. S. A.*, 96, 10109-10114.
2. Fan, H.F., Cox, M.M. and Li, H.W. (2011) Developing single-molecule TPM experiments for direct observation of successful RecA-mediated strand exchange reaction. *PLoS One*, 6, e21359.