www.sciencemag.org/content/349/6251/977/suppl/DC1

Supplementary Materials for

Base triplet stepping by the Rad51/RecA family of recombinases

Ja Yil Lee, Tsuyoshi Terakawa, Zhi Qi, Justin B. Steinfeld, Sy Redding, YoungHo Kwon, William A. Gaines, Weixing Zhao, Patrick Sung, Eric C. Greene*

*Corresponding author. E-mail: ecg2108@cumc.columbia.edu

Published 28 August 2015, *Science* **349**, 977 (2015)
DOI: 10.1126/science.aab2666

This PDF file includes:

Materials and Methods

Fluorescent dsDNA substrates.

Oligonucleotides were obtained from Integrated DNA Technologies (IDT) and purified by HLPC reverse phase chromatography. Stock solutions were quantified by UV absorbance and diluted to 100 μM oligonucleotide using 10 mM Tris–HCl [pH 8.5]. Fluorescently–tagged and untagged complementary ssDNA oligonucleotides were mixed in a 1:1.2 ratio and annealed in buffer containing 40 mM Tris–HCl [pH 8.0], 50 mM NaCl, and 10 mM $MgCl_2$. Reaction mixes were heated 5 minutes at 90˚C, then cooled slowly to room temperature over a period of 4 hours. Annealed oligonucleotides were purified using a QIAquick Gel Extraction Kit (#28704), and the purity of the annealed substrates was verified by electrophoresis on 12% DNA polyacrylamide gels. Fluorescent substrates were protected from light during preparation and storage.

All dsDNA sequences (table S1) were analyzed for microhomology to ensure that they were targeted to a single unique location on the M13mp18 presynaptic ssDNA using a MatLab algorithm (MathWorks, Inc., Natick, MA), as described (*11*). This analysis was essential to ensure that the survival probability data reflected the dissociation of Atto565–dsDNAs bound to identical locations on the presynaptic ssDNA. In brief, the dsDNA strands (top and bottom stands) were scanned in 3– to 10–nt increments along the M13mp18 ssDNA sequence to identify all corresponding tracts of microhomology. This process was also repeated for all mismatched substrates (tables S3 & S4) to ensure that introduction of the mismatched bases did not also result in new ≥8–nt tracts of microhomology. If a desired mismatch resulted in a new tract of microhomology, then we first attempted to modify nucleotides in the flanking sequences to remove the undesired microhomology. If it was not possible to ensure that a particular mismatched dsDNA substrate was targeted to a single region on M13mp18, then that substrate was omitted from the study.

DNA binding experiments and data analysis.

Flowcells, presynaptic ssDNA substrates, and ssDNA curtains were prepared essentially as described (*20, 21*). Rad51/RecA presynaptic complexes were assembled as previously described (*11*), using reaction buffers adapted from prior studies of EcRecA (*22-24*), ScRad51 (*25*), hRad51 (*26*), and ScDmc1 (*27*), as summarized in table S2. Atto565–tagged dsDNA oligonucleotides (2–10 nM) were then injected into the sample chamber in the same buffers, and reactions were incubated for a period of 10 minutes at 30˚C in the absence of buffer flow. Flowcells were quickly flushed (40 sec at 1.0 ml min^{-1}) with fresh buffer to remove unbound Atto565–dsDNA, and the flow rate was then reduced (0.2 ml min^{-1}) to ensure removal of dissociated dsDNA and replenishment of free nucleotide cofactor while images were being collected. Data were obtained by acquiring single 100–msec frames at 30– or 60–second intervals. The data collection intervals were optimized relative to the overall lifetime of each dsDNA substrate and the laser was shuttered between acquired images to minimize photobleaching. Kymographs were then generated from the resulting movies using Fiji. Survival probabilities were determined from analysis of the resulting kymographs by measuring the time (dwell time) that each

molecule of Atto565–DNA remained bound to the presynaptic complexes after flushing the unbound DNA from the sample chamber. The probability that a bound molecule survived to a particular time point (*t*) was determined as the fraction of Atto565–dsDNA molecules that remained bound at time *t*, and survival probability graphs were constructed from the resulting data. All reported data points were calculated from an average of 150 different molecules (N=70–250). In each instance, the dissociation kinetics for the different dsDNA substrates were well described by single exponential fits to the survival probability data (fig. S3) (*11*). Error bars for the survival probability measurements and binding distributions represent 70% confidence intervals obtained through bootstrap analysis, providing a close approximation of expectations for one standard deviation from the mean (*11*). All reported $\Delta\Delta G^{\ddagger}$ values were calculated from the dissociation rate data for the Atto565–dsDNA substrates, as previously described (*11*), and the shared energetic signature (mean ± s.d.) reported in the main text for EcRecA, ScRad51, hRad51, and ScDmc1 reflects the combined average of the first, second and third triplet steps observed for all four recombinases using both sets of oligonucleotide substrates (see Fig. 2 & fig. S4).

Simulations.

Simulations were performed using the Academic Center for Computing and Media Studies (ACCMS) Cray XE6 supercomputer at Kyoto University. Monte Carlo and molecular dynamics simulations were performed using a course–grained oxDNA model (*18, 28*). The RS–DNA model was prepared by mapping the center of mass (COM) of the sugar atoms in each base in the crystal structure of RecA–dsDNA postsynaptic filament (PDB ID: 3CMX) (*9*) to the COM of each base in the coarse–grained representation of a DNA molecule 15 base pairs in length. We performed a short low temperature (0˚C) simulation with the COM of each base restrained in space to relax positions of base and backbone beads. The structure of B–DNA was prepared using generate–sa.py provided by the oxDNA software package.

In all of the simulations, one of the two DNA strands was constrained in space to mimic the protein–bound presynaptic ssDNA. We refer to this constrained strand as the presynaptic strand, which is colored red in Fig. 4 and fig. S9. We refer to the other strand as the complementary strand, which is colored blue in Fig. 4 and fig. S9. For both the B–DNA and RS–DNA models the first three base pairs were also constrained in space by imposing a harmonic potential (Stiffness parameter k = 34.3 kJ mol^{-1} $Å^{-2}$) on the COM of each of these three base pairs. This constraint was applied so that the initial structural models within the simulations represented a partially paired recombination intermediate. Two additional constraints were applied for the RS–DNA simulations. First, distances between adjacent base triplets the complementary strand were restrained by a harmonic potential. Second, the end–to–end distance of the complementary strand was also restrained by a harmonic potential (Stiffness parameter k = 3.43 kJ mol^{-1} $Å^{-2}$). These constraints were intended to emulate the expectation that Rad51/RecA would need to locally stretch the incoming complementary strand to more closely match the extended configuration of the presynaptic ssDNA (*9, 10, 19, 29-31*). Monte Carlo simulations (see below) confirmed that RS–DNA duplex formation was strongly disfavored in the absence of these constraints on the complementary strand (fig. S9A).

Using these initial structures and constraints, we performed virtual move Monte Carlo (MC) simulations to obtain canonical NVT ensemble at 27˚C (*32*), and this ensemble was used to elucidate the free energy profiles of duplex formation for RS–DNA (Fig. S9A). Umbrella sampling, where the number of base pairs with a negative hydrogen–bonding energy was used as an order parameter, was used to enhance conformational sampling (*33*). Weight for each biasing potential was calibrated manually by iterating relatively short MC simulations (10^8 steps). Each simulation was then performed for 10^9 steps and the conformations were collected every 10^4 steps. The free energy profile was calculated from the canonical conformational ensemble using the equation, $-k_B T \log(P)$, where k_B is the Boltzmann constant, T is temperature, and P is the probability of each conformational state.

Molecular dynamics (MD) simulations were performed by integrating a Langevin equation using a diffusion coefficient of 12 Å^2 picosecond^{-1} and an integral step size of 15 femtoseconds. The initial structure for the annealing simulations was prepared by performing a short (10^6 steps) high temperature (100˚C) simulation to melt the duplex. The RS–DNA and B–DNA annealing simulations were conducted at 27˚C, RS–DNA melting simulations were performed at 80˚C, and B–DNA melting simulations were performed at 100˚C. The temperature was kept constant using an Anderson–like algorithm. Each simulation was performed for 10^7 steps, and conformations were collected every 10^4 steps.

The melting and annealing trajectories of B–DNA and RS–DNA were then analyzed to determine which bases were paired (Fig. 4B,E) as well as the duration (lifetime) of these pairing interactions (Fig. 4C,F & fig. S9B). Within the simulations, a base was considered paired if the beads representing the two bases on complementary strands were within 8.5 Å of one another, and the bases were considered unpaired if this distance exceeded 8.5 Å. We define the "n bases paired state" as when the n^{th} base pair forms (where $n = 1\ to\ 15$) and the $n + k^{th}$ base pair does not form (where $k\ =\ 1\ to\ 15 - n$); examples of these states are shown as snapshots in Fig. 4A,D and graphical representation of the data are presented in Fig. 4B,D where the y–axes reflect the position of the most distal base pair relative to the pre–annealed triplet. We then calculated the lifetime (in simulation steps) of each paired state. For annealing simulations, we report the lifetimes of each n bases paired state (Fig. 4C,F). For the melting simulations, we report the lifetimes as the average value for each nucleotide position within the base triplets (fig. S9D). In both cases, lifetime data for the terminal base triplets are excluded from the reported values.

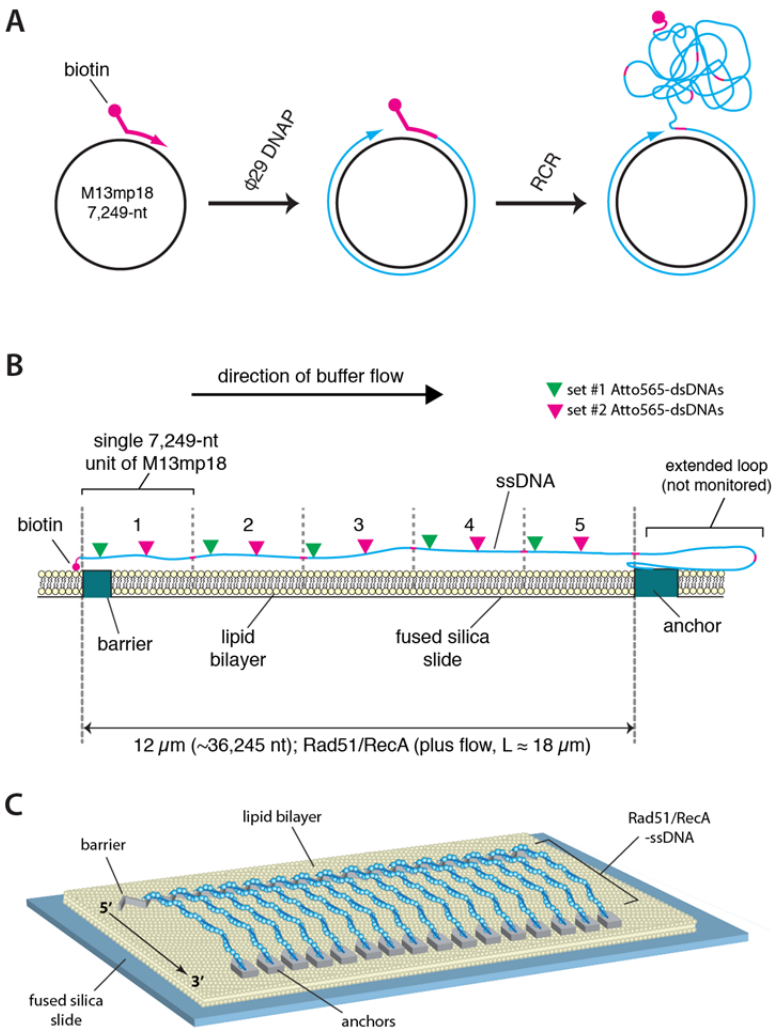

Fig. S1. Preparation of ssDNA curtains by rolling circle replication. (**A**) Schematic showing the production of the presynaptic ssDNA by rolling circle replication (RCR) using an M13mp18 circular ssDNA template. Each substrate is biotinylated at the 5' end and is comprised of multiple units of M13mp18. (**B**) Schematic showing the ssDNA anchored to the surface of the microfluidic sample chamber. Each anchored substrate reflects 5 units of M13mp18 (numbered 1 through 5); any ssDNA longer than the 5 units extends beyond the anchor and is not monitored in the experiments (as indicated). Green and magenta arrowheads indicate the relative locations of the two sets of dsDNA oligonucleotides used in the binding experiments (see tables S1,S3,and S4). (**C**) Schematic of an ssDNA curtain bound by Rad51.

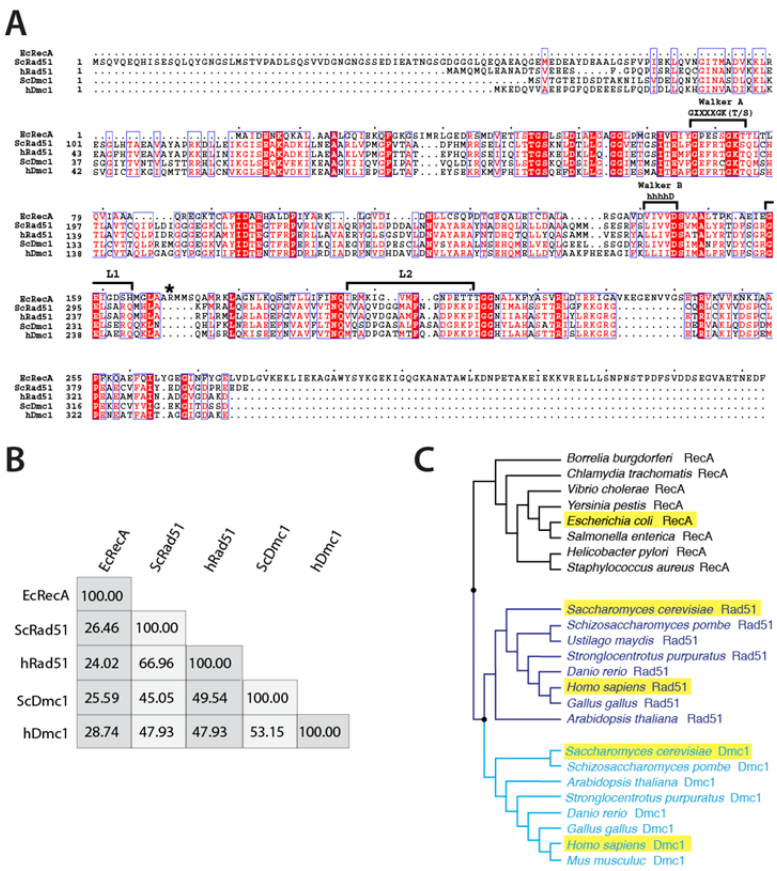

Fig. S2. Relationship among Rad51/RecA proteins used in this study. (**A**) Sequence alignment of EcRecA, ScRad51, hRad51, ScDmc1 and hDmc1. Identical amino acids are highlighted by a red background, similar amino acids are highlighted in red text and the locations of the Walker A and B motifs are indicated, as are the L1 and L2 loops, which contact the presynaptic ssDNA. Arginine 169 in EcRecA, which has been proposed to sense the minor groove geometry of the postsynaptic dsDNA (9), is marked with an asterisk. The numbering scheme for EcRecA begins at residue 2 because the N–terminal methionine is removed *in vivo*. (**B**) Percent identity between EcRecA, ScRad51, hRad51, ScDmc1, and hDmc1. (**C**) Phylogenetic relationship of RecA, Rad51 and Dmc1; the three different lineages are color–coded, and the five recombinases used in this study are highlighted in yellow.

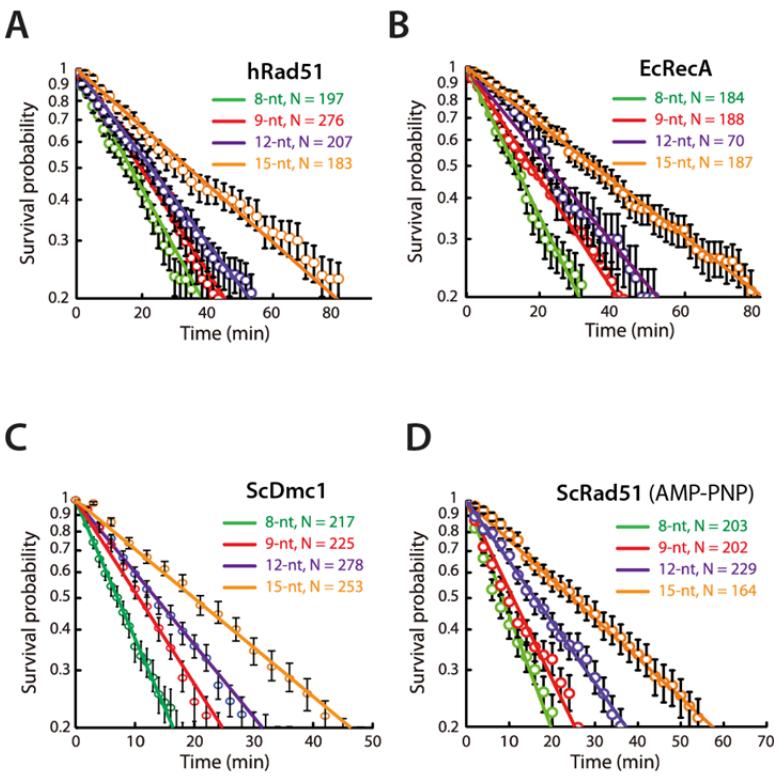

Fig. S3. Representative survival probability curves. Survival probability data for (**A**) hRad51, (**B**) EcRecA, (**C**) ScDmc1, and (**D**) ScRad51 (plus AMP–PNP) using set#1 substrates with 8– to 15–nt of microhomology, as indicated; survival probability curves for the 10–, 11–, 13–, and 14–nt substrates superimpose with the 9–nt and 12–nt data sets and are omitted for clarity. Data were analyzed by measuring the amount of time (dwell time) that each Atto565–DNA molecule remained bound to the Rad51–ssDNA presynaptic complexes after flushing unbound DNA from the sample chamber. The probability that a bound molecule survived up to time point (*t*) was determined as the fraction of Atto565–DNA molecules that remained bound at time *t*, and survival probability graphs for the different substrates were constructed from the resulting data. N corresponds to the total number of Atto565–DNA molecules measured. Error bars for survival probability plots represent 70% confidence intervals obtained through bootstrap analysis; our choice of 70% confidence intervals for the bootstrapped data provides a close approximation to expectations for one standard deviation from the mean.

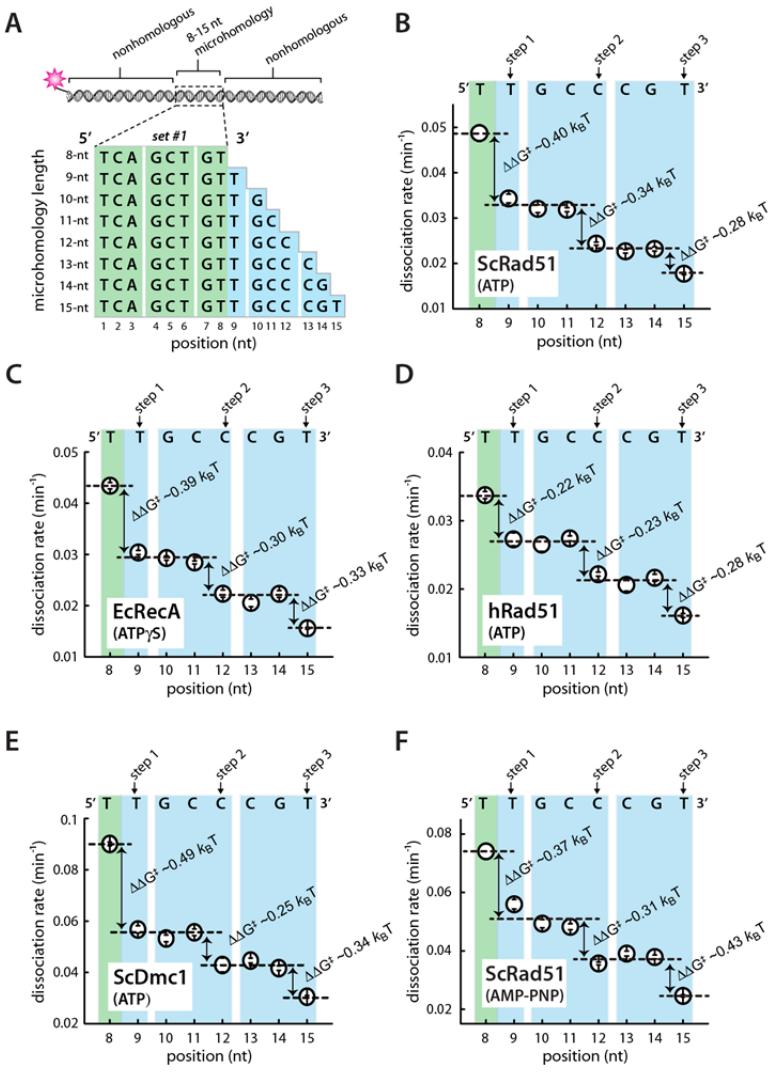

Fig. S4. Conservation of base triplet stepping. (**A**) Schematic of the 70–bp dsDNA substrates (set #1; see table S1). All substrates contain an internal 8– to 15–nt tract of microhomology (as indicated) flanked by nonhomologous sequence. (**B**) Atto565–dsDNA dissociation rates (mean ± s.d.) for ScRad51 plus ATP. (**C**) EcRecA plus ATPγS. (**D**) hRad51 plus ATP. (**E**) ScDmc1 plus ATP. (**F**) ScRad51 plus AMP–PNP. In (**B-F**), the color–coded shading highlights each base triplet, arrows indicate stepwise reductions in dissociation rates coincident with recognition of the 3rd base of each triplet, and the free energy changes ($\Delta\Delta G^{\ddagger}$) associated with each step are indicated.

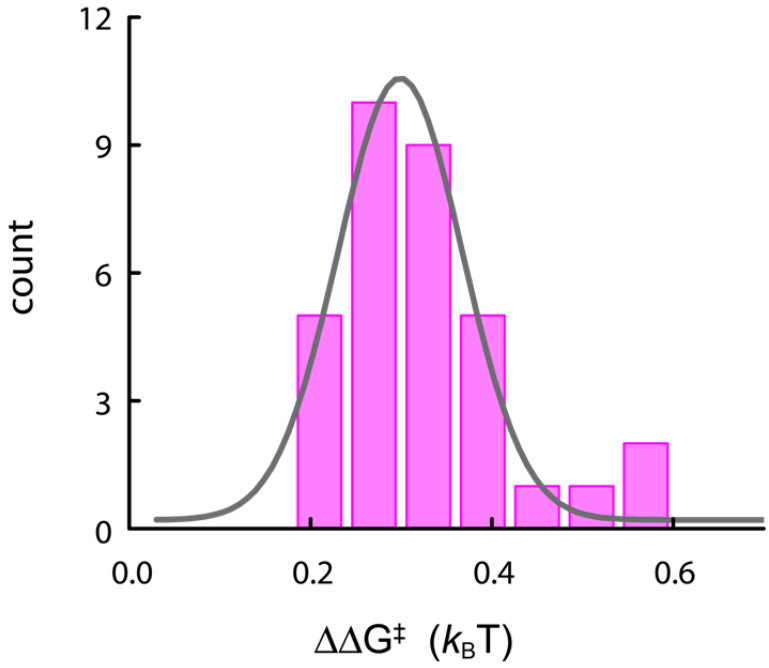

Fig. S5. Shared energetic signature for triplet stepping. Histogram showing the free energy changes associated with base triplet recognition for all individual steps (N=33 total) measured for EcRecA, ScRad51, hRad51, ScDmc1 and hDmc1 (see Fig. 2, fig. S4 & fig. S8). A Gaussian fit to the distribution yields a value of 0.3 ± 0.14 k_BT (mean ± s.d.) for the free energy change that takes place during 3–nt strand exchange steps.

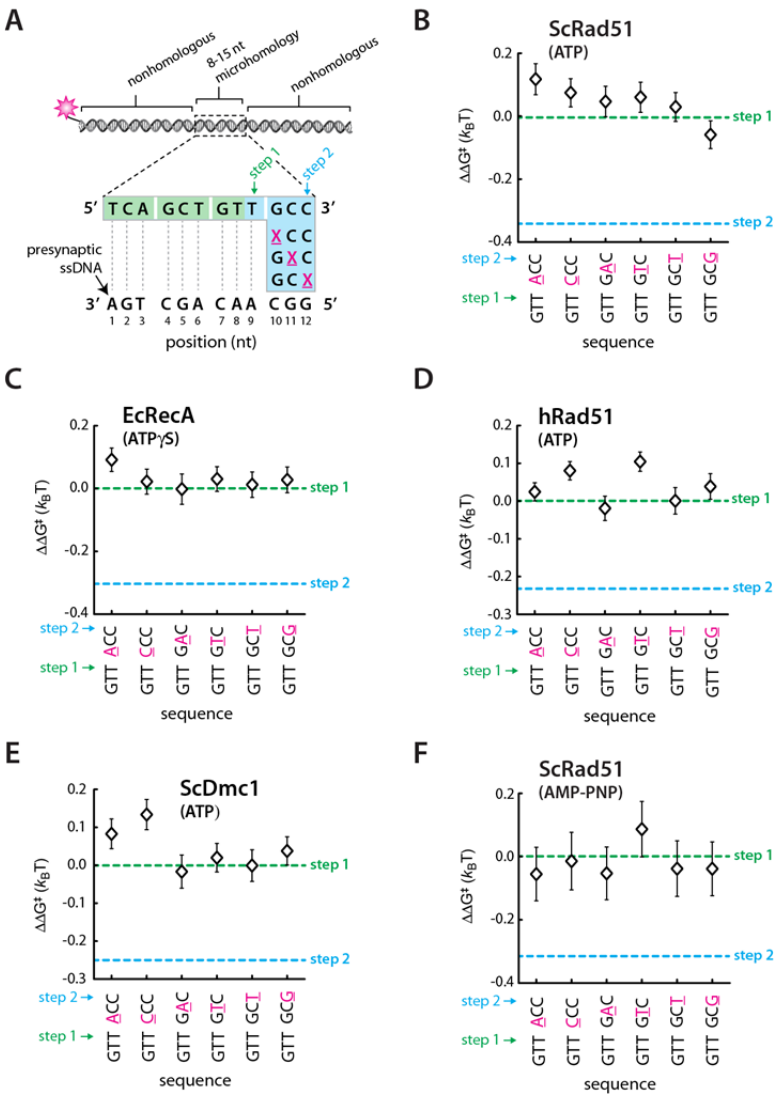

Fig. S6. Effects of mismatches on base triplet recognition. (**A**) Schematic illustration of mismatched set #1 dsDNA substrates. The location of each mismatch within the fourth base triplet is indicated as an underlined, magenta "X". The corresponding sequence of the M13mp18 presynaptic ssDNA is shown to highlight the identity of the mismatched pair (Also see table S3). (**B**) Changes in binding free energy for mismatch substrates in reactions with ScRad51, EcRecA, hRad51, and ScDmc1, as indicated. The green and light blue dashed lines correspond to step 1 and step 2, respectively, obtained for non–mismatched substrates with each of the different recombinases (see fig. S4).

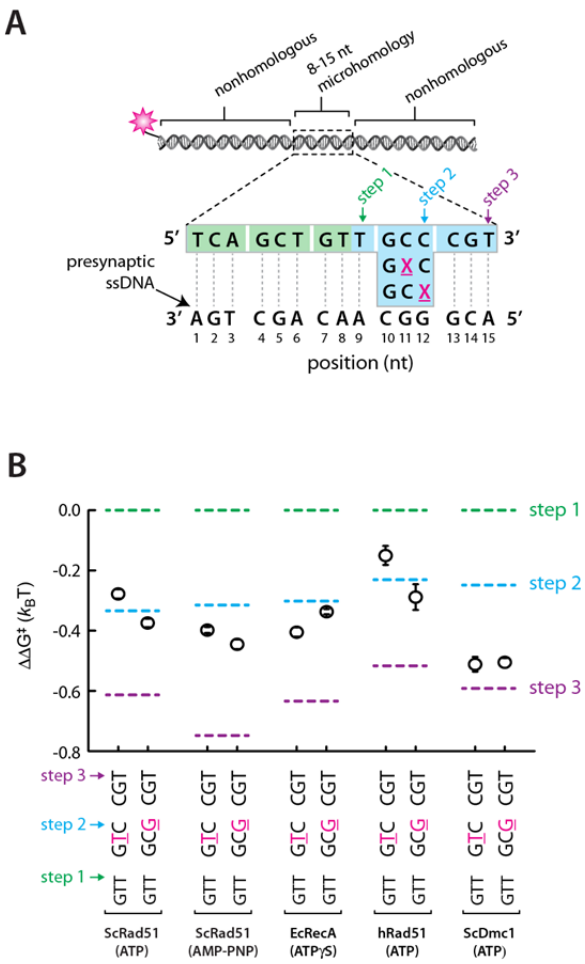

Fig. S7. RecA, Rad51 and Dmc1 can step over mismatched triplets. (**A**) Substrate design based on oligonucleotide set #1 (table S4) for testing whether Rad51/RecA is capable of stepping over a mismatched triplet; note that position 10 could not be altered without introducing new tracts of microhomology, so mutations at this position were not tested. (**B**) Triplet binding data illustrating how base mismatches within the 4th triplet affect recognition of the 5th triplet.

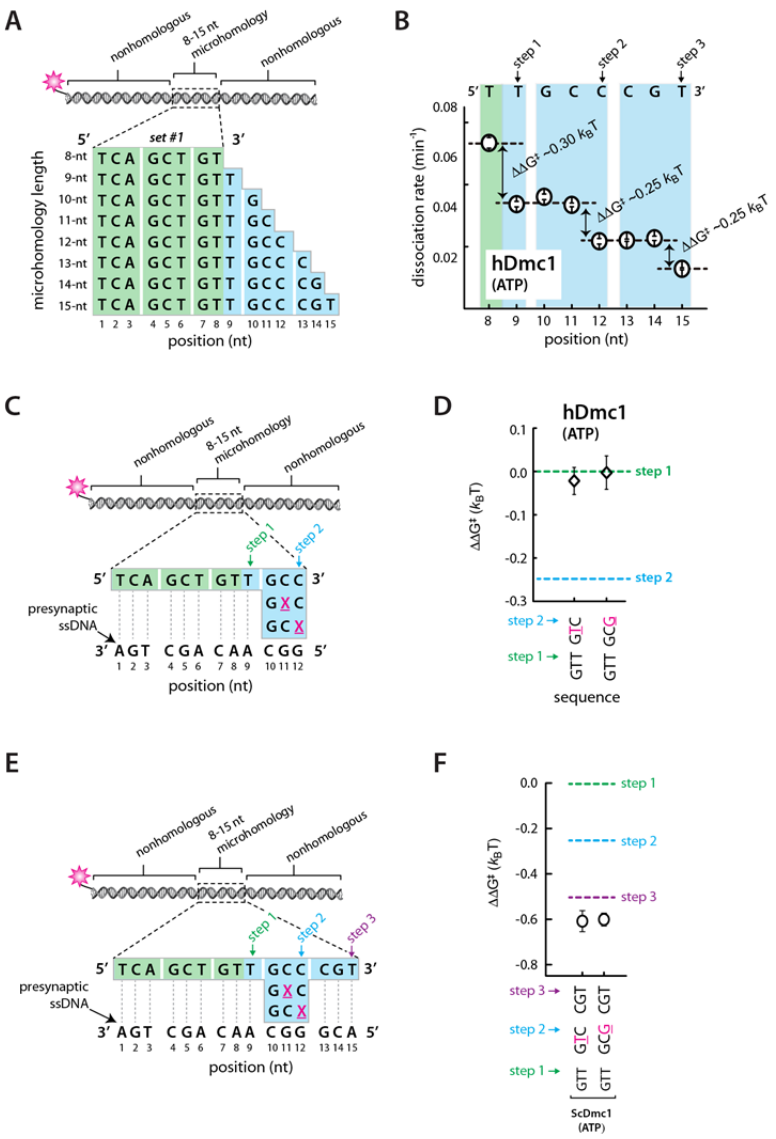

Fig. S8. Strand exchange characteristics of human Dmc1. (**A**) Schematic of the 70–bp dsDNA substrates (set #1). All substrates contain an internal 8– to 15–nt tract of microhomology (as indicated) flanked by nonhomologous sequences. (**B**) Atto565–dsDNA dissociation rates (mean ± s.d.) for hDmc1 plus ATP. (**C**) Schematic of mismatched substrates. (**D**) Mismatch substrate binding for hDmc1. All indicated $\Delta\Delta G^{\ddagger}$ values are relative to the step 1 binding data for the substrate bearing 9–nt of microhomology. (**E**) Substrate design for testing mismatch bypass. (**F**) hDmc1 binding data illustrating how base mismatches within the 4th triplet affect recognition of the 5th triplet.

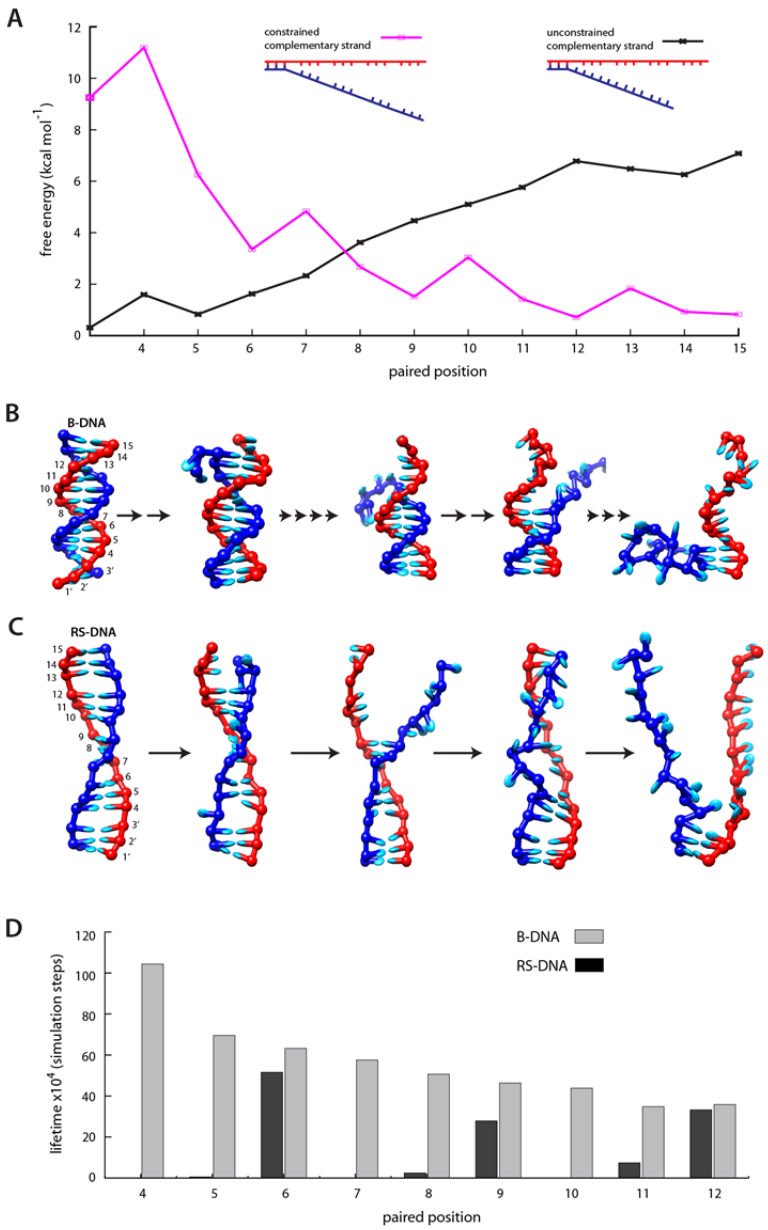

Fig. S9. RS–DNA melting occurs in 3–nt steps. (**A**) Free energy profile of RS–DNA annealing generated by Monte Carlo simulation. (**B**) Molecular dynamics snapshots of B–DNA melting intermediates. (**C**) Molecular dynamics snapshots of RS–DNA melting intermediates. (**D**) Lifetimes of B–DNA and RS–DNA intermediates calculated from 50 separate simulation runs ($5x10^8$ total simulation steps). In (**B**) and (**D**), the number of arrowheads between each image corresponds to the number of intermediates necessary to reach the depicted state. The data in (**D**) do not include the terminal triplets.

Table S1. Sequences of 70–bp oligonucleotides with 8–nt to 15–nt tracts of microhomology.

Set	Microhomology length	Top strand sequence[†]
#1	8–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GT***A GGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	9–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT*** AGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	10–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT G***AC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	11–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GC***G GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	12–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	13–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC C***AT AGC TAG CTA GCT AGC TAG CTA GCT
#1	14–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CG***G AGC TAG CTA GCT AGC TAG CTA GCT
#1	15–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CGT*** GGC TAG CTA GCT AGC TAG CTA GCT
#2	8–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TC***A GGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	9–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC*** AGC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	10–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC C***GC CTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	11–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CT***A ATT AGC TAG CTA GCT AGC TAG CTA GCT
#2	12–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC*** ACT AGC TAG CTA GCT AGC TAG CTA GCT
#2	13–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC C***TT AGC TAG CTA GCT AGC TAG CTA GCT
#2	14–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CC***A AGC TAG CTA GCT AGC TAG CTA GCT
#2	15–nt	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT

[†] Only the Atto565 labeled top strands of each oligonucleotide are shown; the tracts of microhomology complementary to the M13mp18 ssDNA substrate are designated in underlined bold italics. The corresponding bottom strands are complementary to the top strand sequences, but lack a fluorescent label (not shown). Microhomology analysis confirmed that the sequences were targeted to only a single region on M13mp18, corresponding the complement of the underlined sequences.

Table S2. Reaction conditions required for each of the recombinase.

Protein	Reaction buffer
EcRecA	25 mM Tris–Acetate [pH 7.5], 4 mM Mg–Acetate, 10 mM Na–Acetate, 1 mM ATPγS, 1 mM DTT, and 0.2 mg ml^{-1} BSA
hRad51	30 mM Tris–Acetate [pH 7.5], 1 mM $MgCl_2$, 5 mM $CaCl_2$, 100 mM KCl, 1 mM ATP, 1 mM DTT, and 0.2 mg ml^{-1} BSA
ScRad51	30 mM Tris–Acetate [pH 7.5], 20 mM Mg–Acetate, 50 mM KCl, 1 mM DTT, 2.5 mM ATP or 1 mM AMP–PNP (as indicated), and 0.2 mg ml^{-1} BSA
ScDmc1 and hDmc1	40 mM Tris–HCl [pH 7.5], 2 mM $MgCl_2$, 1.5 mM $CaCl_2$, 100 mM KCl, 2.5 mM ATP, 1 mM DTT, and 0.2 mg ml^{-1} BSA

Table S3. DNA sequences for 70–bp oligonucleotides with 12–nt microhomology tracts bearing mismatched triplets.

Set	4th Triplet‡	Top strand sequence†
#1	GCC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	XCC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT ACC*** CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	XCC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT CCC*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GAC*** CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GTC*** CTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GCX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCT*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#1	GCX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCG*** GTT AGC TAG CTA GCT AGC TAG CTA GCT
#2	CTC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC*** ACT AGC TAG CTA GCT AGC TAG CTA GCT
#2	XTC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC GTC*** TCT AGC TAG CTA GCT AGC TAG CTA GCT
#2	CXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CGC*** ACT AGC TAG CTA GCT AGC TAG CTA GCT
#2	CTX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTA*** ACT AGC TAG CTA GCT AGC TAG CTA GCT

† Only the Atto565 labeled top strands of each oligonucleotide are shown; the tracts of microhomology complementary to M13mp18 ssDNA designated in underlined bold italics. The corresponding bottom strands are complementary to top strand sequences, but lack fluorescent label.

‡ "X" corresponds to the mismatched position, and the identity and location of the mismatched bases in the DNA sequences are highlighted in magenta. All oligonucleotide sequences were analyzed for microhomology to ensure that the introduction of the mismatched bases did not result in new regions of microhomology; mismatches that resulted in new regions of microhomology were excluded from this study.

Table S4. DNA sequences for 70–bp oligonucleotides with 15–nt microhomology tracts bearing mismatched triplets.

Set	4th Triplet‡	Top strand sequence†
#1	GCC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCC CGT*** AGC TAG CTA GCT AGC TAG CTA GCT
#1	GXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GTC CGT*** GGC TAG CTA GCT AGC TAG CTA GCT
#1	GCX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT ***TCA GCT GTT GCG CGT*** GGC TAG CTA GCT AGC TAG CTA GCT
#2	CTC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT
#2	XTC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC GTC CCT*** AGC TAG CTA GCT AGC TAG CTA GCT
#2	CXC	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CGC CCT*** CGC TAG CTA GCT AGC TAG CTA GCT
#2	CTX	*5'–Atto565* C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ***TTA CCT TCC CTA CCT*** AGC TAG CTA GCT AGC TAG CTA GCT

† Only the Atto565 labeled top strands of each oligonucleotide are shown; the tracts of microhomology complementary to M13mp18 ssDNA designated in underlined bold italics. The corresponding bottom strands are complementary to top strand sequences, but lack fluorescent label.

‡ "X" corresponds to the mismatched position, and the identity and location of the mismatched bases in the DNA sequences are highlighted in magenta. All oligonucleotide sequences were analyzed for microhomology to ensure that the introduction of the mismatched bases did not result in new regions of microhomology; mismatches that resulted in new regions of microhomology were excluded from this study.