

Supplemental Figures for

DNA sequence imperfections and base triplet recognition by the Rad51/RecA family of recombinases

Authors: Ja Yil Lee^{1,2}, Justin B. Steinfeld¹, Zhi Qi^{1,3}, YoungHo Kwon⁴, Patrick Sung⁴, and Eric C. Greene¹

Affiliations:

¹Department of Biochemistry & Molecular Biophysics, Columbia University, New York, NY, 10032, USA

²Present Address: School of Life Science, Ulsan National Institute of Science and Technology (UNIST), Ulsan, 689-798, Korea

³Present Address: Center for Quantitative Biology & Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, 100871, China

⁴Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT, 06510, USA

Fig. S1

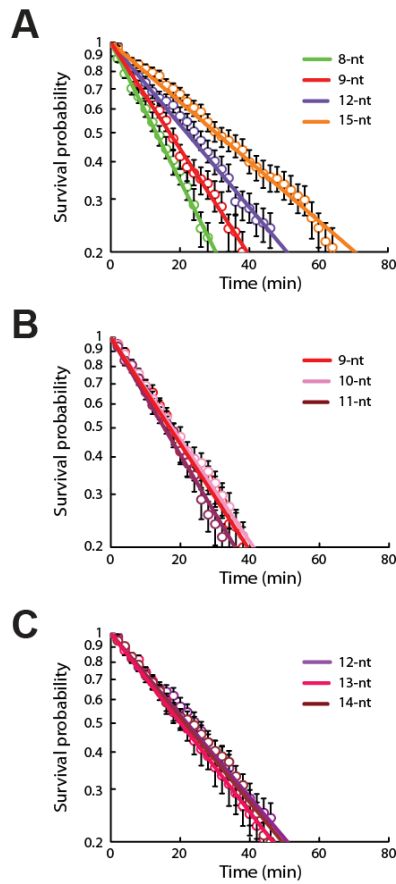


Fig. S1. Survival probability plots. (A) Examples of survival probability data for 70-bp oligonucleotide substrates with 8-, 9-, 12-, or 15-nt of microhomology, as indicated. (B) Survival probability data for the 9-, 10-, and 11-nt substrates and (C) 12-, 13-, and 14-nt substrates are presented separately, for clarity. All plots are for reactions with *S. cerevisiae* Rad51 in the presence of ATP. Error bars represent 70% confidence intervals obtained through bootstrap analysis, providing a close approximation to expectations for one standard deviation from the mean, as previously described (28,29).

Fig. S2

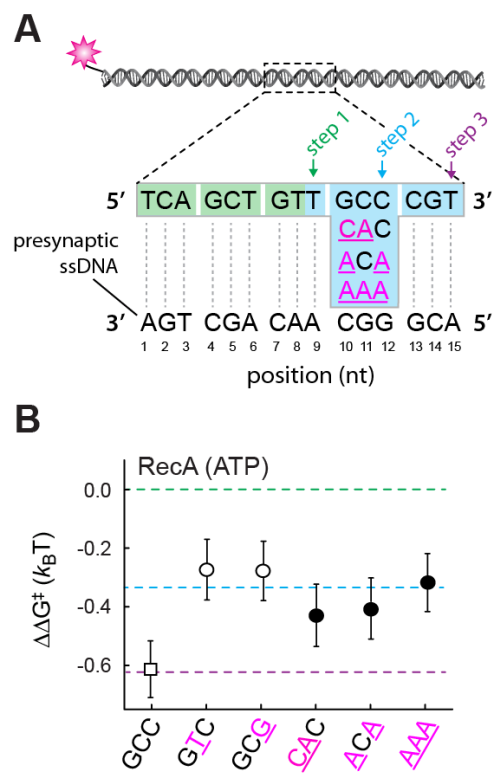


Fig. S2. RecA plus ATP with substrates bearing multiple mismatches. (A) Schematic illustration of the substrates bearing multiple mismatches in the penultimate triplet near the 3' end of the 15-nt tract of microhomology. The position and identity of the mismatches are highlighted in magenta underlined text. (B) Corresponding data for each of the different DNA substrates. Data for the complementary triplet is shown as a reference (open square), previously reported data for substrates bearing single mismatches are also shown (open circles) for comparison (28) along with the data for the multiple mismatch substrates (closed circles).

Fig. S3

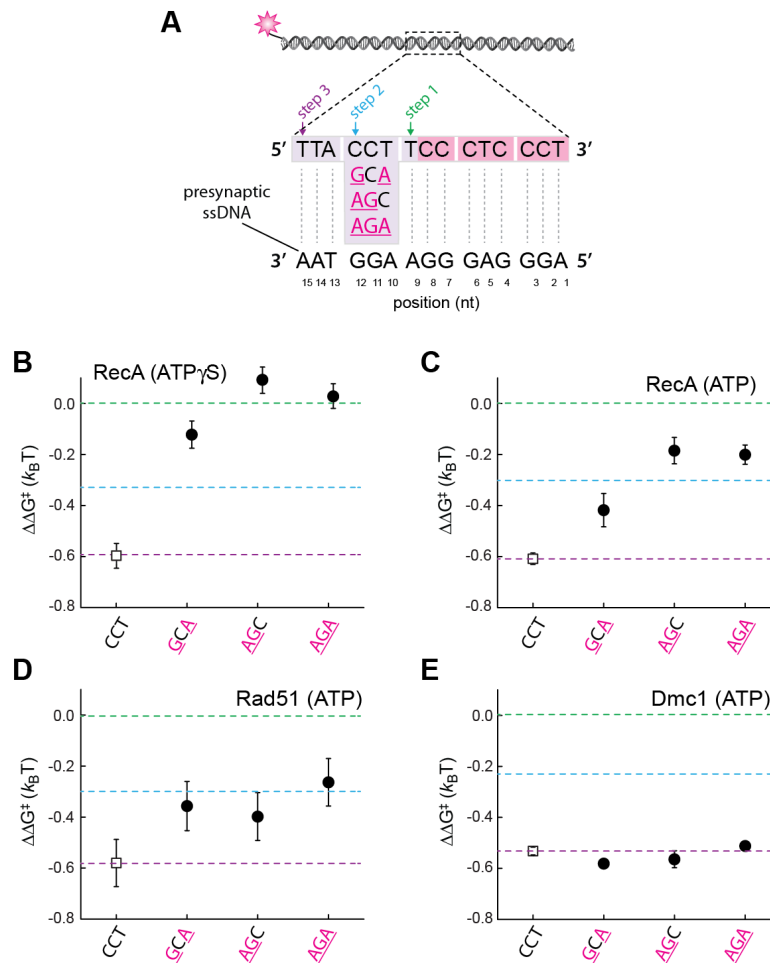


Fig. S3. Reactions with multiple mismatches near the 5' end of the microhomology. (A) Schematic illustration of the substrates bearing multiple mismatches in the penultimate triplet near the 5' end of the 15-nt tract of microhomology. The position and identity of the mismatches are highlighted in magenta underlined text. Corresponding data for each of the different DNA substrates with (B) RecA and ATP γ S, (C) RecA and ATP, (D) *S. cerevisiae* Rad51 and ATP, and (E) *S. cerevisiae* Dmc1 and ATP. Data for the complementary triplet is shown as a reference (open square), along with the data for the multiple mismatch substrates (closed circles).

Fig. S4

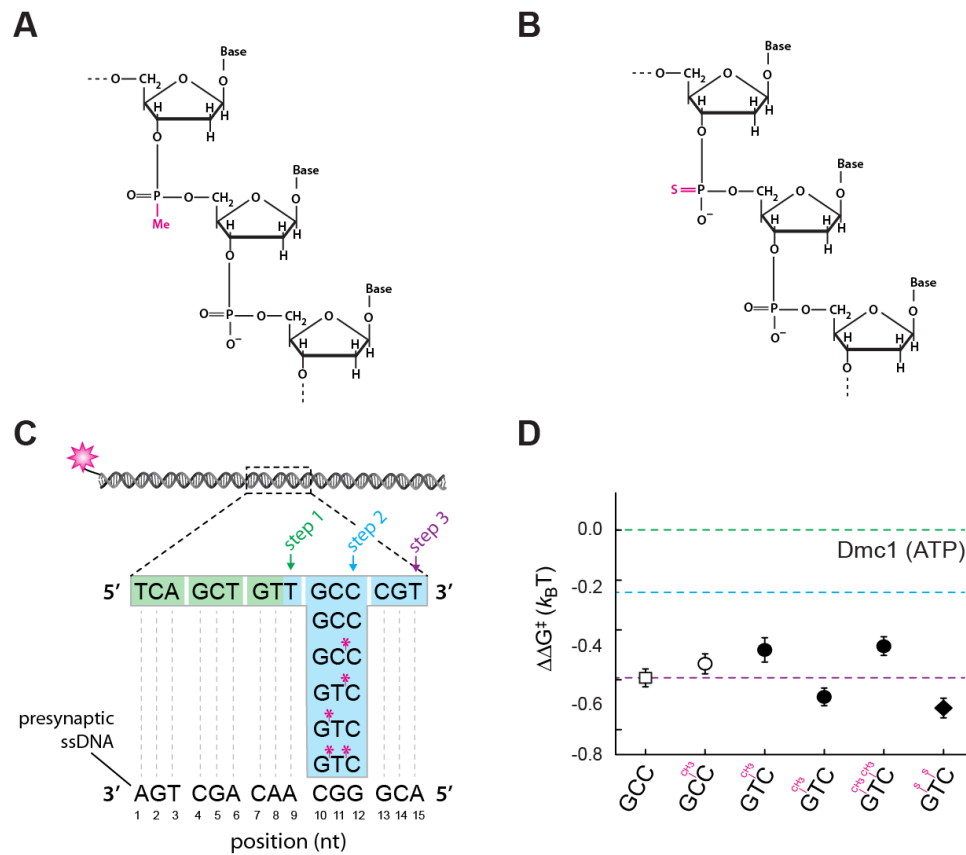


Fig. S4. Substitutions of the non-bridging oxygens do not affect mismatch stabilization by Dmc1. Illustrations of (A) methylphosphonate and (B) phosphorothioate backbone substitutions. (C) Schematic of dsDNA substrates bearing either methylphosphonate or phosphorothionate substitutions within the penultimate triplet at the 3' end of a 15-nt tract of microhomology. The locations of the backbone modifications are indicated with a magenta asterisk. (D) Binding data for the methyl phosphonate and phosphorothionate substituted substrates (as indicated) in reactions with Dmc1.

Table S1. Oligonucleotide sequences.

Table S1A. Lesions near the 3' end of the microhomology.

	Designation	Sequence [†]
1	15-nt "wild-type"	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-3'
2	Double mismatch #1	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT CAC CGT GGC TAG CTA GCT AGC TAG CTA GCT-3'
3	Double mismatch #2	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT ACA CGT GGC TAG CTA GCT AGC TAG CTA GCT-3'
4	Triple mismatch	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT AAA CGT GGC TAG CTA GCT AGC TAG CTA GCT-3'
5	Single-base insertion	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT GCC G CGT GGC TAG CTA GCT AGC TAG CTA GC-3'
6	Abasic #1	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT BCC CGT GGC TAG CTA GCT AGC TAG CTA GC-3'
7	Abasic #2	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT GBC CGT GGC TAG CTA GCT AGC TAG CTA GC-3'
8	Abasic #3	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCT TCA GCT GTT GCB CGT GGC TAG CTA GCT AGC TAG CTA GC-3'

Table S1B. 5' extension of the microhomology.

	Microhomology length	Sequence [†]
9	8-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA CCT GCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
10	9-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA CCA TCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
11	10-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA CAT TCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
12	11-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TTA ACT TCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
13	12-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG TAT CCT TCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
14	13-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG ACA CCT TCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
15	14-nt	5'-Atto565 C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG GTA CCT TCC CTC CCT AGC TAG CTA GCT AGC TAG CTA GCT-3'
16	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-3'

Table S1C. Lesions near the 5' end of the microhomology.

	Designation	Sequence [†]
17	15-nt “wild-type”	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA CCT</i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
18	Single Mismatch #1	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>ACT</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
19	Single Mismatch #2	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>CGT</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
20	Single Mismatch #3	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>CCG</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
21	Double mismatch #1	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>GCA</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
22	Double mismatch #2	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>AGT</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
23	Triple mismatch	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>AGA</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
24	Single-base insertion	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>G</u></i> <i>CCT TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
25	Abasic #1	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA</i> <i><u>B</u>CT TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
26	Abasic #2	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA C</i> <i><u>BT</u> TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'
27	Abasic #3	5'– <i>Atto565</i> C CGG AGG CCT TAG GCC TTA GGC CTT AGG CCG <i>TTA <u>CCB</u></i> <i>TCC CTC CCT</i> AGC TAG CTA GCT AGC TAG CTA GCT-3'

[†]The internal tracts of microhomology (≥ 8 -nt in length) targeted to the presynaptic ssDNA are highlighted as black italic text. Base mismatches, insertions, and abasic sites (shown here as “B”) are highlighted by underlined magenta text.