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

Received for publication, March 21, 2017, and in revised form, April 27, 2017. Published Papers in Press, May 5, 2017, DOI 10.1074/jbc.M117.787614

Ja Yil Lee1, Justin B. Steinfeld6, Zhi Qi2, YoungHo Kwon3, Patrick Sung6, and Eric C. Greene3

From the 1Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10023 and the 6Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Edited by Norma Allerwell

Homologous recombination plays key roles in double-strand break repair, rescue, and repair of stalled replication forks and meiosis. The broadly conserved Rad51/RecA family of recombinases catalyzes the DNA strand invasion reaction that takes place during homologous recombination. We have established single-stranded (ss)DNA curtain assays for measuring individual base triplet steps during the early stages of strand invasion. Here, we examined how base triplet stepping by RecA, Rad51, and Dmc1 is affected by DNA sequence imperfections, such as single and multiple mismatches, abasic sites, and single nucleotide insertions. Our work reveals features of base triplet stepping that are conserved among these three phylogenetic lineages of the Rad51/RecA family and also reveals lineage-specific behaviors reflecting properties that are unique to each recombinase. These findings suggest that Dmc1 is tolerant of single mismatches, multiple mismatches, and even abasic sites, whereas RecA and Rad51 are not. Interestingly, the presence of single nucleotide insertion abolishes recognition of an adjacent base triplet by all three recombinases. On the basis of these findings, we describe models for how sequence imperfections may affect base triplet recognition by Rad51/RecA family members, and we discuss how these models and our results may relate to the different biological roles of RecA, Rad51, and Dmc1.

Homologous recombination allows for the regulated exchange of genetic information between two different DNA molecules of identical or nearly identical sequence composition, and it is a driving force in evolution (1, 2). Homologous recombination contributes to double-strand DNA break (DSB) repair (3), the rescue of stalled or collapsed replication forks (4–6), chromosomal rearrangements (7–9), horizontal gene transfer (10), and meiosis (11–13). Defects in recombination compromise genome integrity and lead to the gross chromosomal rearrangements that are a hallmark of cancer (7–9). During recombination, single-stranded DNA (ssDNA), derived from the nucleolytic processing of a DSB or collapsed replication fork, is paired with the complementary strand of a homologous double-stranded DNA (dsDNA), resulting in the displacement of the non-complementary strand from the duplex to generate a D-loop (14, 15). This reaction is referred to as strand invasion, and the resulting intermediate can be channeled through a number of alternative pathways, any of which can allow for the repair of the originally broken DNA molecule using information derived from the template (3, 4).

The homology search and strand invasion reactions are catalyzed by the Rad51/RecA family DNA recombinases (14–16). These recombinases are among the most highly conserved of all DNA repair proteins (17), and prominent family members include bacterial RecA, the archaeal protein RadA, and the eukaryotic recombinase Rad51 (16, 18). In addition to Rad51, most eukaryotes also have the meiosis-specific recombinase Dmc1, although the reason most eukaryotes require a second recombinase for meiosis remains an enduring mystery (11, 19).

RecA is the archetypal recombinase originally identified in genetic screens for Escherichia coli mutants defective in recombination (20), and much of our current understanding of recombination mechanisms can be attributed to studies of this recombinase (15, 21). Rad51/RecA family members are all ATP-dependent DNA-binding proteins that form extended helical filaments on DNA (15, 16). The bound DNA is extended by ~50% relative to the contour length of B-form DNA (22–26), and crystal structures of RecA-ssDNA and RecA-dsDNA pre- and post-synaptic complexes reveal that the bound DNA is organized into near B-form base triplets separated by ~8 Å between adjacent triplets (Fig. 1A) (27). We have referred to this unique DNA architecture as recombinase-stretched-DNA (RS-DNA) to distinguish it from other forms of mechanically stretched DNA and as a reflection of its unique structural and mechanistic properties (28). To help better understand the mechanisms of genetic recombination, we have established single-molecule assays for studying how pre-synaptic complexes engage dsDNA during the early stages of homologous recombination (28–30). These experiments reveal that strand invasion...
Base triplet recognition

Figure 1. Base triplet stepping by the Rad51/RecA family of recombinases. A, crystal structure of the RecA-ssDNA pre-synaptic complex highlighting the base triplet organization of the bound pre-synaptic ssDNA. B, schematic illustration of base triplet stepping, in which each step coincides with complete pairing of all three bases within an RS-DNA triplet. C, experimental assay for quantitating the binding of Atto565-labeled dsDNA fragments to an unlabeled pre-synaptic complex. These panels were adapted with permission from Ref. 28.

Base triplet recognition

Assay for base triplet recognition with ssDNA curtains

We used ssDNA curtains and total internal reflection fluorescence microscopy to visualize RecA, Rad51, or Dmc1 presynaptic complexes. The ssDNA was generated using M13 as a template for rolling circle replication and then tethered to a lipid bilayer through a biotin-streptavidin linkage and aligned along chromium barriers, as described previously (28, 29, 31, 32). The ssDNA unravels when incubated with RPA-eGFP, and the downstream ends of the RPA-ssDNA are anchored to exposed chromium pedestals through non-specific surface adsorption. Addition of RecA, Rad51, or Dmc1, along with the required nucleotide cofactor, results in the displacement of the fluorescent RPA-eGFP from the ssDNA (28, 29). Once assembled, unbound proteins are flushed away, and dsDNA-binding activity of the pre-synaptic complexes is probed by the addition of short (70-bp) dsDNA oligonucleotides bearing a single tract of sequence microhomology 8–15 nucleotides (nt) in length that is complementary to a unique location on the M13 ssDNA (28, 29). Following a short incubation, the unbound dsDNA is flushed away, and the stability of the remaining dsDNA molecules is measured by survival probability analysis (Fig. 1C) (28, 29, 32).

Microhomology and base triplet recognition

We have previously reported assays in which the internal tract of microhomology was incrementally extended from 8- to 15-nt in the 5’ → 3’ direction relative to the complementary strand within the fluorescent dsDNA (28, 29). These assays, together with accompanying molecular dynamics simulations, revealed that strand invasion intermediates were stabilized in 3-nt increments (28, 29). Here, we sought to determine whether similar results would be obtained if we instead extended the tract of sequence microhomology 8–15 nucleotides (nt) in length that is complementary to a unique location on the M13 ssDNA (28, 29). Following a short incubation, the unbound dsDNA is flushed away, and the stability of the remaining dsDNA molecules is measured by survival probability analysis (Fig. 1C) (28, 29, 32).

Results

Assay for base triplet recognition with ssDNA curtains

We used ssDNA curtains and total internal reflection fluorescence microscopy to visualize RecA, Rad51, or Dmc1 presynaptic complexes. The ssDNA was generated using M13 as a template for rolling circle replication and then tethered to a lipid bilayer through a biotin-streptavidin linkage and aligned along chromium barriers, as described previously (28, 29, 31, 32). The ssDNA unravels when incubated with RPA-eGFP, and the downstream ends of the RPA-ssDNA are anchored to exposed chromium pedestals through non-specific surface adsorption. Addition of RecA, Rad51, or Dmc1, along with the required nucleotide cofactor, results in the displacement of the fluorescent RPA-eGFP from the ssDNA (28, 29). Once assembled, unbound proteins are flushed away, and dsDNA-binding activity of the pre-synaptic complexes is probed by the addition of short (70-bp) dsDNA oligonucleotides bearing a single tract of sequence microhomology 8–15 nucleotides (nt) in length that is complementary to a unique location on the M13 ssDNA (28, 29). Following a short incubation, the unbound dsDNA is flushed away, and the stability of the remaining dsDNA molecules is measured by survival probability analysis (Fig. 1C) (28, 29, 32).

Microhomology and base triplet recognition

We have previously reported assays in which the internal tract of microhomology was incrementally extended from 8- to 15-nt in the 5’ → 3’ direction relative to the complementary strand within the fluorescent dsDNA (28, 29). These assays, together with accompanying molecular dynamics simulations, revealed that strand invasion intermediates were stabilized in 3-nt increments (28, 29). Here, we sought to determine whether similar results would be obtained if we instead extended the tract of microhomology in the 3’ → 5’ direction (Fig. 2A). To address this question, we used a series of 70-bp oligonucleotides labeled at one end with a single Atto565 dye and bearing
an internal tract of microhomology targeted to a specific region on the pre-synaptic ssDNA. The internal tracts of microhomology were flanked by non-homologous sequences that lacked any microhomology exceeding 7 nt in length (supplemental Table S1). E. coli RecA, S. cerevisiae Rad51, or S. cerevisiae Dmc1 pre-synaptic complexes were assembled onto the ssDNA curtains. The fluorescent dsDNA substrates were then injected into the sample chamber and incubated for 10 min at 30 °C, and the unbound DNA was then flushed away while collecting data. Dissociation rates were then obtained from survival probability analysis of the bound dsDNA fragments (supplemental Fig. S1) (28, 29). These experiments revealed that extension of the internal tract of microhomology in the 3’ → 5’ direction yielded changes in dissociation rates for RecA, Rad51, and Dmc1, similar to our previously reported results for 5’ → 3’ extension (Fig. 2, B–D) (28, 29). In addition, each step coincided with an ~30% change in relative dissociation rates, corresponding to change in free energy (ΔΔG°) of ~0.3 k_BT (Fig. 2, B–D). From these results, we concluded that 3’ → 5’ extension of an internal tract of microhomology gave rise to periodic changes in dsDNA dissociation rates for RecA, Rad51, and Dmc1. These findings are consistent with our previous experimental observations using substrates in which we incrementally extended the microhomology length in the 5’ → 3’, as anticipated (28, 29). We were careful to note that although we obtained similar results for these two sets of substrates, the findings themselves do directly report upon the polarity of strand exchange because we are only assessing the stability of the reaction intermediates after they have already formed.

Mismatch discrimination at the 5’ terminus of microhomology

We next asked whether RecA, Rad51, and Dmc1 could discriminate against mismatched bases located within the 5’-terminal triplet of an embedded tract of microhomology. For these assays, a single mismatch was introduced at each of three possible positions at the 5’ end of a 12-nt tract of microhomology (Fig. 3A). These results revealed that a single mismatch at any position within the 5’-terminal triplet abolishes detectable binding of the mismatched triplet by the RecA, Rad51, and Dmc1 (Fig. 3, B–D). These findings are similar to our previous observations for mismatches located at the 3’ end of the internal tract of microhomology (28). Taken together, our results support the conclusion that RecA, Rad51, and Dmc1 are all intolerant of single nucleotide mismatches within terminal base triplets located at either the 5’ or 3’ end of an internal tract of sequence homology.
Internal triplets are destabilized within RecA and Rad51 complexes

RecA and Rad51 can step over mismatches located within the penultimate base triplet near the 3’ end of an internal tract of microhomology, but the reduced stability of the resulting intermediates suggests that the mismatched triplet may remain unpaired (28). Dmc1 can also step over internal triplets near the 3’ end of an internal tract of microhomology, but these mismatched triplets still contribute to binding free energy of the resulting intermediate (28). As an extension of these previous findings, we next asked whether RecA, Rad51, and Dmc1 could step over mismatches positioned within the penultimate base triplet near the 5’ end of an internal tract of microhomology (Fig. 3E). Here, a single mismatch was introduced at each of
three possible positions within the penultimate base triplet within a 15-nt tract of microhomology (Fig. 3E). For both RecA and Rad51, the binding stability of the resulting intermediate was most consistent with destabilization of the mismatch-bearing base triplet (Fig. 3, F and G). In contrast, the stability of the mismatched intermediates generated by Dmc1 was comparable with the fully paired intermediate (Fig. 3H). We conclude that Dmc1 can stabilize single mismatches located in the penultimate base triplet near either the 5’ or 3’ end of an embedded tract of microhomology. In contrast, RecA and Rad51 were unable to stabilize these mismatched triplets, but both were able to step over the mismatch-bearing triplets to stabilize an adjacent base triplet that is fully homologous to the pre-synaptic ssDNA.

**Dmc1 can stabilize triplets bearing multiple mismatches**

To explore the potential mechanism of mismatch stabilization by Dmc1, we next sought to determine whether Dmc1 could stabilize base triplets bearing more than one mismatch. For these assays, we utilized a 70-bp dsDNA substrate bearing a 15-nt tract of microhomology in which the penultimate triplet at the 3’ end of a 15-nt tract of microhomology. Corresponding binding data for RecA (Fig. 4B, RAD1, and Dmc1 (D) as indicated. Green, blue, and purple dashed lines correspond to the averaged binding stability for data obtained from the non-mismatched 9–11-nt substrates; the 12–14-nt substrates; and 15-nt substrates, respectively, reflecting steps 1–3 in A. Previously reported data for substrates bearing single mismatches are shown (open circles) for comparison (28).

**Figure 4. Impact of multiple mismatches on base triplet recognition.** A, schematic illustration of dsDNA substrates bearing either two or three nucleotide mismatches within the penultimate triplet at the 3’ end of a 15-nt tract of microhomology. Corresponding binding data for RecA (B), RAD1, and Dmc1 (D) are as indicated. Green, blue, and purple dashed lines correspond to the averaged binding stability for data obtained from the non-mismatched 9–11-nt substrates; the 12–14-nt substrates; and 15-nt substrates, respectively, reflecting steps 1–3 in A. Previously reported data for substrates bearing single mismatches are shown (open circles) for comparison (28).

Matches (Fig. 4, B and C, and supplemental Fig. S2). Based upon the observed changes in binding free energy, RecA appears to be incapable of stabilizing a homologous triplet that lies beyond a triplet bearing either two or three mismatches in reactions using ATPγS. This differential response to substrates harboring multiple mismatches is the only substantive difference we have found between RecA and Rad51 in response to base triplet imperfections. Bulk biochemical assays have revealed that RecA is unable to bypass longer (>50 bp) heterologous insertions during strand invasion when ATPγS is used as the nucleotide cofactor, but it can bypass these insertions in reactions with ATP (33, 34). Therefore, we sought to determine whether RecA might be able to stabilize the substrates with multiple mismatches in the presence of ATP. These experiments revealed that in the presence of ATP, RecA is capable of stabilizing a homologous triplet that lies beyond an internal triplet bearing multiple mismatches (supplemental Fig. S3). We conclude that in the presence of ATP, RecA and Rad51 respond similarly to internal base triplets bearing multiple mismatches.

**Dmc1 can stabilize triplets bearing abasic sites**

We next tested substrates bearing a single abasic site at each of the three possible positions within the penultimate base triplet near the 3’ or 5’ end of the microhomology (Fig. 5). Neither RecA nor Rad51 was able to stabilize internal triplets bearing an abasic site, consistent with their intolerance for single base mismatches (Fig. 5, B, C, F, and G). In contrast, Dmc1 was able to stabilize these substrates regardless of the location of the abasic site within the triplet, and similar findings were obtained for abasic sites located near the 3’ or 5’ ends of the internal tract of microhomology (Fig. 5, D and H). These findings provide further support for the hypothesis that Dmc1 does not stabilize mismatched intermediates through a mechanism involving protein-enhanced pairing of non-Watson-Crick interactions, because abasic sites cannot form any type of non-canonical pairing interaction. The finding that Dmc1 can stabilize substrates bearing abasic sites also suggests that the stabilization mechanism likely does not involve direct protein contacts with the nucleotide bases of the suboptimal triplet, but instead it may involve contacts with the ribose-phosphate backbone of the DNA.

**Non-bridging oxygen modifications do not prevent mismatch stabilization by Dmc1**

We next tested whether Dmc1-mediated mismatch stabilization might involve recombinase contacts with the phosphate backbone. To address this possibility, we conducted assays with Dmc1 using mismatched substrates harboring either methylphosphonate or phosphorothioate substitutions within the phosphate backbone of the mismatched triplet (supplemental Fig. S4). Surprisingly, the methylphosphonate substitutions had no appreciable impact upon the ability of Dmc1 to stabilize the mismatched base triplets. These findings suggest that if Dmc1 contacts the phosphate backbone of incoming complementary strand during mismatch stabilization, then these protein–DNA contacts may not involve the non-bridging oxygen atoms, but it would more likely involve other contacts with the ribose-phos-
phate backbone. Future work will be essential to determine the structural basis for mismatch stabilization by Dmc1.

**Single nucleotide insertions disrupt triplet stabilization**

We next sought to establish whether RecA, Rad51, or Dmc1 could tolerate single nucleotide insertions. For this purpose, we designed a dsDNA substrate harboring a single base insertion between the penultimate and terminal triplet at either the 3’ or 5’ end of an internal tract of microhomology, which is anticipated to position the terminal base triplets just 1-nucleotide out of register with the pre-synaptic ssDNA (Fig. 6, A and C). Surprisingly, the results revealed that a single base insertion results
in a change in binding free energy consistent with the loss of stabilizing interactions involving the terminal base triplets for each of the three recombinases (Fig. 6, B and D). We conclude that RecA, Rad51, and Dmc1 cannot stabilize a homologous 3′- or 5′-terminal base triplet that is moved out of register by the presence of a single nucleotide insertion.

Discussion

We have established that the eukaryotic meiosis-specific recombinase Dmc1 is more tolerant of imperfect base triplet-pairing interactions relative to Rad51 and RecA. As discussed below, our findings suggest a model in which Dmc1 stabilizes imperfectly paired triplets through a mechanism involving protein-mediated contacts that compensate for the binding energy that would otherwise be lost due to the unpaired triplet.

Strand exchange takes place in 3-nt steps

Our work suggests that incrementally increasing the length of homology in either the 5′ → 3′ or 3′ → 5′ around a central 8-nt tract of microhomology gives rise to stepwise changes in binding free energy that follow a highly predictable base triplet pattern. In addition, the free energy changes associated with base triplet stabilization appear to be indistinguishable for base triplets located at either end of the internal tract of microhomology. These observations, together with the finding that the recombinases can all step over mismatches or abasic sites positioned within the penultimate triplet at either the 5′ or 3′-ends of the internal tract of microhomology, support a model in which strand exchange might proceed in either the 5′ → 3′ or 3′ → 5′ directions, and they suggest that bidirectional strand exchange may be conserved across the Rad51/RecA family of recombinases (Fig. 7A). Several studies have suggested that strand exchange can take place with a defined polarity (35–38), whereas others have reported that strand exchange can occur in either direction (39, 40). We note that our assays do not directly observe strand exchange steps as they are taking place, but rather assess the stability of these early recombination intermediates only after they have already formed. Therefore, we do not rule out the possibility that there may exist polarity-dependent differences in the rates of strand exchange that cannot be assessed in our experimental system.

Responses of RecA, Rad51, and Dmc1 to sequence imperfections

Single mismatches within the terminal base triplet at either the 5′ or 3′ end of the microhomology prevent stabilization of these terminal triplets by RecA, Rad51, and Dmc1, suggesting that this may be a universal response of Rad51/RecA family members to imperfectly paired terminal base triplets (Fig. 7A). This result is surprising for Dmc1 because this recombinase can stabilize all other tested imperfections (i.e. single, double, and triple mismatches, and abasic sites), with the exception of single nucleotide insertions, so long as the imperfections are flanked by at least one fully homologous base triplet (Fig. 7B). These findings raise the question of whether the inability of Rad51/RecA recombinases to stabilize imperfect terminal triplets
provides some specific biological advantage for homologous recombination or whether it might be a reflection of the thermodynamic properties of RS-DNA or both. One possibility is that imperfect terminal triplets play an important role in ensuring accurate recombination, perhaps by providing a kinetic barrier to slow the progression of strand invasion, and providing a decision point for the pre-synaptic complex (and associated factors) to either continue strand invasion or dissociate to sample other regions of the genome for a more perfectly matched donor sequence.

**Triplet pairing requires perfect alignment**

Single base insertions abolish recognition of the adjacent triplet by all three recombinases, suggesting that this may be a conserved feature of base triplet recognition by the Rad51/RecA family of DNA recombinases (Fig. 7A). This finding highlights intolerance of the pre-synaptic complex to either continue strand invasion or dissociate to sample other regions of the genome for a more perfectly matched donor sequence.

Interestingly, RecA and Rad51 can both step over triplets bearing single mismatches at either the 5' or 3' ends of the microhomology, but neither Rad51 nor RecA can stabilize the resulting mismatched triplets (Fig. 7B). This finding contrasts with results for Dmc1, which can both step over and stabilize single mismatches at either the 5' or 3' ends of the microhomology (Fig. 7B). RecA and Rad51 can also step over abasic sites, but cannot stabilize the resulting triplets bearing the abasic residues. This finding is again in striking contrast with Dmc1, which can step over and stabilize base triplets bearing abasic sites regardless of the location of the abasic site within the base triplet (Fig. 7C). These results have important implications for understanding the mechanistic basis by which Dmc1 stabilizes substrates bearing imperfect triplets (see below).

**Effects of multiple mismatches**

Experiments looking at multiple mismatches yielded several important findings. First, RecA can step over triplets bearing a single mismatched nucleotide but not two or three mismatches in reactions with ATPγS (Fig. 7D). This contrasts with the findings for Rad51, which is able to step over base triplets bearing two or even three mismatches (Fig. 7D). We speculate that in the presence of ATPγS, RecA might be capable of stepping over

---

**Figure 7. Effects of lesions on base triplet recognition by RecA, Rad51, and Dmc1.** Details of these models are presented under “Discussion.”
these tandem mismatches if there were additional sequence homology beyond the tandem mismatches, perhaps by initiating an independent strand invasion reaction downstream of the mismatches. Interestingly, RecA can step over multiple mismatches in reactions with ATP (Fig. 7D), which is consistent with prior biochemical assays demonstrating that RecA can bypass larger sequence heterologies only in the presence of ATP (33, 34). This requirement for ATP has led to the hypothesis that RecA may have an ATP-dependent motor activity that is coupled to heterology bypass (41). An alternative possibility is that ATP hydrolysis-dependent protein turnover may be necessary to reduce the stiffness of the RecA filament, allowing a separate segment of the pre-synaptic complex to more readily re-engage the dsDNA substrate beyond tandem mismatches.

These experiments also revealed that Dmc1 can stabilize base triplets bearing either one, two, or three mismatches, so long as those triplets are flanked by at least one homologous triplet (Fig. 7D). This finding highlights the marked tolerance of Dmc1 to sequence imperfections at the level of a single base triplet, and it also has crucial implications for understanding how Dmc1 stabilizes imperfectly matched RS-DNA triplets (see below).

Models for base triplet stability

Our results can be considered within the context of two contrasting models describing the physical stability of mismatched RS-DNA base triplets. The first model is that imperfectly matched RS-DNA base triplets are themselves inherently unstable (Fig. 7E). This unstable triplet model would account for all of our observations with RecA and Rad51, indicating that the inability of these two proteins to stabilize imperfect triplets stems from the lesion-bearing triplets themselves being unstable relative to a perfectly paired triplet. The unstable triplet model is also supported by the previous reports that RecA makes very limited contact with the incoming complementary DNA strand, which is instead held in place primarily by Watson-Crick hydrogen bond pairing interactions with the pre-synaptic ssDNA (27). However, if the unstable triplet model is correct, then Dmc1 must somehow stabilize the imperfect triplet, or otherwise compensate for the loss of binding free energy that would be expected if the imperfect triplets were not paired within the Dmc1–dsDNA postsynaptic complex.

An alternative model, which we do not favor, is that perfect RS-DNA base triplets are themselves inherently stable (Fig. 7E). This model could explain all of our results with Dmc1, but it would then require that RecA and Rad51 have built-in mechanisms for sensing and actively destabilizing imperfect base triplets. Moreover, it is difficult to envision how an RS-DNA base triplet would tolerate single, double, and triple mismatches and even abasic sites without any loss of triplet pairing stability. Indeed, molecular dynamics (MD) simulations of either RS-DNA or isolated base triplets reveal that RS-DNA is highly dynamic relative to B-DNA, and single base mismatches cause rapid destabilization of triplet pairing interactions.5Recent reports suggest that non-complementary sequences are stable within RecA filaments during a 10-ns MD simulation (42), so future work will be required to more fully understand how mismatched sequences behave during recombination. However, our experimental results seem to argue against a model where mismatched triplets are inherently stable and must be actively destabilized by RecA and Rad51. Specifically, Dmc1 can stabilize base triplets bearing single abasic sites at any of the three possible positions, and it is again difficult to envision that a base triplet harboring an abasic site could be inherently stable within an RS-DNA triplet. Together, we consider our findings to be most consistent with the unstable triplet model, suggesting that RecA and Rad51 need not actively destabilize imperfect RS-DNA base triplets, but instead Dmc1 must somehow compensate for the loss of Watson-Crick binding free energy due to mismatched base triplets.

Possible implications for homologous recombination

Our results suggest that RecA and Rad51 require perfect Watson-Crick pairing to allow stabilization of RS-DNA base triplets, whereas Dmc1 can stabilize non-complementary triplets embedded with longer tracts of homology. We emphasize that we do not yet know the biological implications of these findings. One possibility is that this biophysical difference may reflect the biological specialization of the meiosis-specific recombinase Dmc1, namely the requirement for Dmc1 to promote recombination between polymorphic alleles of different parental origins (see below). Future work will therefore be essential to determine whether or not the differences we observe between Rad51 and Dmc1 in response to mismatches and abasic sites reflect some broader biological difference between these two proteins.

Importantly, our work only reflects the effects of mutations within individual base triplets. The overall pairing interactions are highly stable, with half-lives still exceeding tens of minutes, even when mismatches or basic sites are present. Furthermore, we anticipate that the relative differences in stability for longer strand-exchange products will be very small. A single unpaired triplet within a paired intermediate that is tens or perhaps hundreds of base triplets in length will not in and of itself destabilize pairing interactions. Indeed, experimental and theoretical studies of RecA have recently shown that after initial recognition of an 8-nt homologous sequence, the presence of mismatches outside this region has little impact on overall stability (42). The previous work with RecA by Prentiss and co-workers (42) is in good general agreement with our findings. At this stage, we can only infer that the differential stabilization of imperfectly matched RS-DNA triplets might reflect some broader difference in the overall stability of a complete strand-exchange product, which in turn may manifest as a recombinase- and mismatch-dependent difference in recombination efficiency. Future work will be essential to continue testing for potential relationships between triplet pairing, or mispairing, and recombination efficiencies.

Finally, the results reported here reflect the basal activities of E. coli RecA, S. cerevisiae Rad51, and S. cerevisiae Dmc1. These recombinases acts in concert with many other proteins; for example, a total of ~20 proteins or protein complex participate in recombination in E. coli (14, 43, 44), and ~45 distinct proteins or protein complexes participate in recombination in E. coli...
Base triplet recognition

*S. cerevisiae* (4, 45). Interestingly, recent in vivo studies have shown that Rad51-mediated recombination is unexpectedly tolerant of mismatches (46). Therefore, the crucial next step will be to start understanding how recombination accessory proteins augment the basal base triplet recognition activities of the recombinases during strand invasion.

Experimental procedures

DNA curtains

Single-stranded DNA substrates were generated using M13mp18 (7,249-nt; New England Biolabs) as a template for rolling circle replication (RCR), as described (29, 31). In brief, a biotinylated primer was annealed to the M13mp18 template (~30 nm) at a 1:1.2 molar ratio (primer, M13mp18) in buffer containing 40 mM Tris-HCl (pH 8.0), 50 mM NaCl, and 10 mM MgCl2. The annealing mixture was heated for 5 min at ~90 °C and then cooled slowly to room temperature. Annealed products were diluted to a final concentration of 15 nm and stored at 4 °C until use. Replication reactions were prepared in RCR buffer (50 mM Tris-HCl (pH 7.5), 4 mM DTT, 10 mM ammonium sulfate, and 10 mM MgCl2), containing annealed primer/M13mp18 mixture (350 pm), 150 nm ph29 DNA polymerase, and dNTP mixture (200 μM each). Reactions were incubated at 30 °C for 25 min and then diluted 10-fold into buffer containing 40 mM Tris-HCl (pH 7.5), 2 mM MgCl2, 50 mM NaCl, 1 mM DTT, and 0.2 mg/ml BSA prior to use.

Flow cells were made by depositing chromium barriers onto the surface of a fused silica slide by electron beam lithography and then assembled into a sample chamber using double-sided tape and a borosilicate cover slide, as described (47). Bilayers were prepared with 91.5% 1,2-dioleoyl-sn-glycero-3-phosphocholine, 0.5% biotinylated 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(biotinyl), and 8% mPEG 2000–1,2-dioleoyl-sn-glycero-3-phosphoethanolamine, and the biotinylated ssDNA molecules were attached to the bilayer through a biotin-streptavidin linkage. Anchored ssDNA molecules were then aligned at the barriers by application of flow in buffer containing 40 mM Tris-HCl (pH 7.5), 2 mM MgCl2, 50 mM NaCl, 1 mM DTT, 0.2 mg/ml BSA, and 0.1 mM RPA-eGFP for 15 min at 1 ml/min. All experiments were conducted with a custom-built prism-type total internal reflection fluorescence microscope equipped with a 488-nm blue laser (Coherent Sapphire, 200 milliwatt) and a 561-nm yellow laser (Coherent Sapphire, 200 milliwatt).

Reaction conditions and data analysis

*E. coli* RecA, *S. cerevisiae* Rad51, and *S. cerevisiae* Dmc1 were prepared as described (28, 29). Rad51/RecA pre-synaptic complexes were assembled as described previously (28, 29), using the following reaction buffers: *E. coli* RecA buffer, 25 mM Tris acetate (pH 7.5), 4 mM magnesium acetate, 10 mM sodium acetate, 1 mM ATP, or 2.5 mM ATP plus an ATP-regenerating system (20 mM creatine and 0.04 mg/ml creatine kinase), 1 mM DTT, and 0.2 mg/ml BSA; *S. cerevisiae* Rad51 buffer, 30 mM Tris acetate (pH 7.5), 20 mM magnesium acetate, 50 mM KCl, 1 mM DTT, 2.5 mM ATP, and 0.2 mg/ml BSA; *S. cerevisiae* Dmc1, 40 mM Tris-HCl (pH 7.5), 2 mM MgCl2, 1.5 mM CaCl2, 100 mM KCl, 2.5 mM ATP, 1 mM DTT, and 0.2 mg/ml BSA. Note that the identity of the nucleotide cofactor used in each different experiment is specified in each of the corresponding figure panels. All data were collected at 30 °C.

All dsDNA binding data were collected and analyzed as described previously (28, 29). In brief, Atto565-tagged dsDNA oligonucleotides (2–10 nm; supplemental Table S1) were injected into the sample chamber in the same buffers (see above), and reactions were incubated for 10 min at 30 °C in the absence of buffer flow. Flow cells were quickly flushed (40 s 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 remove dissociated dsDNA and to replenish free nucleotide cofactor. Data were obtained by acquiring 100-ms frames at 30- or 60-s intervals; 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 generated using Fiji and were analyzed by measuring the time that each Atto565-dsDNA molecule remained bound to the pre-synaptic complexes after flushing unbound DNA from the sample chamber. The probability that a bound molecule survived to a particular time point (ε) 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 data points were calculated from an average of ~200 molecules (n = 150–250). The dissociation kinetics for all dsDNA substrates were well described by single exponential fits to the survival probability data (supplemental Fig. S1).

Free energy calculations were performed as described previously (28, 29). In brief, for dsDNA substrates harboring ≥8 nt of microhomology, the free energy barrier, ΔG‡, for escape from a potential well can be related to the rate as shown in Equation 1,

\[
\Delta G^+ = A e^{-\frac{k_B T}{k_d} G^N} \quad \text{(Eq. 1)}
\]

where A is a jump frequency, k_B is the Boltzmann constant, and T is the temperature. The difference between the barrier heights between two different escape processes can be compared, leading to Equation 2.

\[
\Delta \Delta G^+ = \Delta G^N_2 - \Delta G^N_1 = k_B T \ln \frac{k_d^{-1}}{k_d^1} \quad \text{(Eq. 2)}
\]

All reported ΔΔG‡ values were normalized such that ΔG‡ for the dsDNA containing a single 8-nt tract of microhomology is zero, and the experimentally measured data used to calculate ΔΔG‡ were the k_d values for each different substrate obtained from survival probability data.

Author contributions—J. Y. L designed and conducted the single molecule experiments and data analysis. J. B. S. and Z. Q. assisted with single molecule experiments, data analysis, and experimental design. Y. K. expressed and purified Rad51 and Dmc1. E. C. G. supervised the project and wrote the manuscript with input from J. Y. L., J. B. S., Z. Q., and P. S.

Acknowledgments—We thank Kyle Kaniecek, Corentin Moevus, Luisina de Tullio, Tsuyoshi Terakawa, Daniel Duzdevich, and Brooks Crickard for comments on the manuscript.
