Complementary strand relocation may play vital roles in RecA-based homology recognition

Published online 31 August 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 20* ***10441–10451***
doi:10.1093/nar/gks769

Complementary strand relocation may play vital roles in RecA-based homology recognition

Alexandra Peacock-Villada, Darren Yang, Claudia Danilowicz, Efraim Feinstein, Nolan Pollock, Sarah McShan, Vincent Coljee and Mara Prentiss*

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

Received June 14, 2012; Revised July 19, 2012; Accepted July 20, 2012

ABSTRACT

RecA-family proteins mediate homologous recombination and recombinational DNA repair through homology search and strand exchange. Initially, the protein forms a filament with the incoming single-stranded DNA (ssDNA) bound in site I. The RecA–ssDNA filament then binds double-stranded DNA (dsDNA) in site II. Non-homologous dsDNA rapidly unbinds, whereas homologous dsDNA undergoes strand exchange yielding heteroduplex dsDNA in site I and the leftover outgoing strand in site II. We show that applying force to the ends of the complementary strand significantly retards strand exchange, whereas applying the same force to the outgoing strand does not. We also show that crystallographically determined binding site locations require an intermediate structure in addition to the initial and final structures. Furthermore, we demonstrate that the characteristic dsDNA extension rates due to strand exchange and free RecA binding are the same, suggesting that relocation of the complementary strand from its position in the intermediate structure to its position in the final structure limits both rates. Finally, we propose that homology recognition is governed by transitions to and from the intermediate structure, where the transitions depend on differential extension in the dsDNA. This differential extension drives strand exchange forward for homologs and increases the free energy penalty for strand exchange of non-homologs.

INTRODUCTION

Homologous recombination (HR) plays an important role in meiosis and DNA damage repair (1–3). During HR, an incoming single-stranded DNA (ssDNA) molecule bound to recombinase protein molecules, such as RecA, searches for a sequence matched double-stranded DNA (dsDNA) molecule (4,5). Once a sequence-matched dsDNA is found, the incoming ssDNA displaces one of the strands in the dsDNA (outgoing strand) and Watson–Crick pairs with the other strand (complementary strand), yielding heteroduplex dsDNA in site I (6,7) and an unpaired outgoing strand. The crystal structure of the searching state with ssDNA in site I is known (8). The final structure of the heteroduplex dsDNA in site I and some residues associated with the final position of the outgoing strand are also known (8). Figure 1 illustrates the structure of the heteroduplex in the final post-strand exchange state and the binding sites associated with the final position of the outgoing strand (9). The structure of the incoming ssDNA in the searching state is almost identical to the structure of the dsDNA in the final post-strand exchange state. In both cases, the DNA consists of base triplets with nearly B-form spacing, where the base triplets are separated by backbone extensions that are much larger than the equilibrium extension for B-form dsDNA (8). This non-uniform spacing results in an average extension along the filament axis of $\sim 1.5\times$ the B-form extension.

An outstanding question has been what drives strand exchange forward given that the Watson–Crick pairing is the same before and after strand exchange, yet the process must be free energetically favorable since it occurs in the absence of hydrolysis. Another outstanding question is whether strand exchange is simply a two-step process in which the Watson–Crick pairing of the complementary strand bases is transferred from the outgoing strand to the incoming strand by base flipping, or whether additional steps occur. Many experiments have probed for additional intermediate states (10–14), but the existence, number, nature and roles of intermediate states remain controversial. The existence of an additional step involving a structure other than the known initial and final states might allow strand exchange to result in a

*To whom correspondence should be addressed. Tel: +1 617 495 4483; Fax: +1 617 495 0416; Email: prentiss@fas.harvard.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

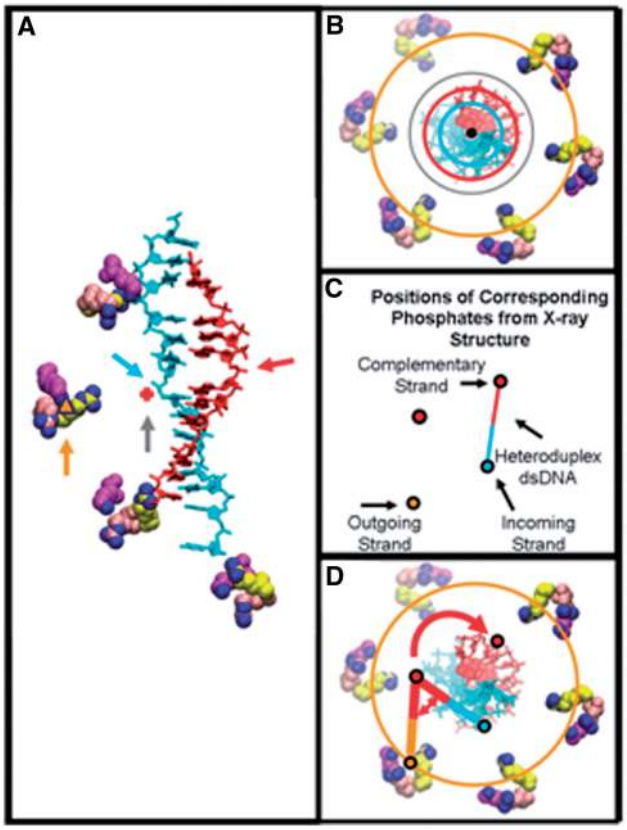

Figure 1. Proposed schematics of the strand exchange process. (**A**) Representation of the side view of the X-ray structure of the dsDNA heteroduplex in the final post-strand exchange state with incoming and complementary strands indicated by the cyan and red stick renderings. Orange, cyan and red arrows indicate the positions of the corresponding phosphates on the outgoing, incoming, and complementary strands in the final post-strand exchange structure. The VMD (9) renderings of RecA crystal structure 3CMX (8) show site II residues Arg226 (pink), Arg243 (yellow) and Lys245 (magenta) with charged nitrogen atoms (blue). The outgoing strand position was calculated by minimizing the energy of the interaction with those residues. The grey arrow points at the red plus sign indicating the proposed position of the corresponding complementary strand phosphate in the proposed intermediate structure. (**B**) Bottom view of the same structure. Circles correspond to the radii occupied by the phosphates. (**C**) Information as in Figure 1B, but only one set of the corresponding phosphates is shown. Final state phosphate positions for the incoming, complementary, and outgoing strands are shown with filled colored circles. Paired bases shown as lines. The complementary strand position in the proposed intermediate structure is shown by the grey circle. (**D**) Same as Figure 1B with circles and base pairs added to show radial positions of the corresponding phosphate groups and repositioning of the complementary strand.

significant change in the DNA–protein interaction that could produce the free energy reduction required to drive strand exchange forward.

The X-ray structure of the filament shows that the incoming strand backbone is located ~1 nm from the center of the helical RecA filament, whereas the binding sites associated with the outgoing strand are ~2 nm from the center of the RecA filament, as shown in Figure 1. As a result of this difference in separation from the center of the RecA filament, the total linear extension of the outgoing strand backbone is significantly longer than the total linear extension of the incoming strand. If the base pairs in the outgoing strand are organized in B-form triplets separated by rises, and strand exchange is accomplished by base flipping, then the bases in the outgoing strand must also be arranged in nearly B-form triplets separated by large rises; however, in the rises the linear extension of the outgoing strand backbone must exceed the linear extension of the incoming strand backbone since the outgoing strand is located farther from the center of the RecA protein helix, as shown in Figure 1. A similar effect appears in a spiral staircase with a hollow center: the railing on the outside of the staircase has a longer linear extension than the railing on the inside, even though the steps connecting them are perpendicular to the axis of the helix. In metal spiral staircases, the longer outer railing simply includes more metal than the inner railing. In contrast, in DNA molecules all three backbones contain the same number of phosphates, and all of the molecules are extended beyond their B-form lengths; therefore, larger extensions can result in larger mechanical stress because the bonds in the more extended backbones are more deformed. For backbones with strong direct interactions with the protein, the protein–DNA interactions may be so free energetically favorable that they more than compensate for stress in the rises; however, if direct DNA–protein interactions are weak, mechanical stress due to large rises may be free energetically unfavorable.

In the final post-strand exchange state, the complementary strand is attached to the RecA–ssDNA filament dominantly via its Watson–Crick pairing with the incoming strand (8); therefore, large rises between bases in the incoming strand would place great mechanical stress on the complementary strand bases unless the protein structure provides additional mechanical support for those rises. If the complementary strand is paired with the even more highly extended outgoing strand, then the stress on the complementary strand will be even larger than the stress on the complementary strand paired with the incoming strand. Thus, strand exchange may reduce the stress on the complementary strand bases and the complementary strand backbone, which would reduce the mechanical energy of the system. In sum, strand exchange might be free energetically favorable because it reduces the tension on the complementary strand by transferring the base pair interaction from the very highly extended outgoing strand backbone to the somewhat less extended incoming strand backbone.

Seminal strand exchange models proposed that the transition between the homology searching state and the final strand exchange state could be achieved by flipping the bases in the complementary strand while the positions of all three DNA backbones remained approximately fixed (4); however, the X-ray structures show that in the final post-strand exchange state, the separation between the outgoing and complementary strands is too large to allow for Watson–Crick pairing (Figures 1A–C and 2A). Earlier FRET results had suggested that the complementary strand backbones reposition after the complementary strand base pairing is transferred from the outgoing strand to the incoming strand (10). The suggestion that the incoming and outgoing strands do not reposition during

strand exchange is sensible since the incoming and outgoing strands are attached strongly through direct contacts between the backbone and the protein (8) (Supplementary Figure S1). Thus, it is reasonable to assume that the complete strand exchange process involves the relocation of the complementary strand backbone from a position where its bases can flip between pairing with the outgoing strand and pairing with the incoming strand to the final post-strand exchange position, known from the X-ray structure. Since changing the position of the complementary strand backbone from the proposed initial position to its known final position requires such a large relocation of the backbone, it is plausible that the time required to relocate the backbone might represent the rate-limiting step in strand exchange. In the final post-strand exchange structure portions of the L1 and L2 loops are located in the rises (8). Thus, in the final post-strand exchange state, the protein may provide mechanical support for the rises as well as additional free energetically favorable interactions (8), which make the final repositioning of the complementary strand free energetically favorable.

Earlier experiments had suggested that the structure of dsDNA in the final post-strand exchange state is similar to the structure of dsDNA in a RecA filament created by the dsDNA binding to RecA that is free in solution (15). If the incoming strand backbone maintains its position throughout strand exchange, then the X-ray structure suggests that the binding of free RecA to dsDNA would occur in a conformation similar to the conformation of the dsDNA in the strand exchange process immediately after the pairing interaction of the complementary strand bases has shifted from the outgoing strand to the incoming strand. If this were true, the binding of free RecA to dsDNA would involve the same complementary strand relocation that occurs during strand exchange after the Watson–Crick pairing of the complementary strand has been transferred from the outgoing strand to the incoming strand. Thus, if free RecA binding is not diffusion limited, then the extension rates due to free RecA binding should be the same as the extension rate due to strand exchange because both would be limited by the time required to relocate the complementary strand backbone.

In this work we investigate the strand exchange process in which the base pairing of the complementary strand is transferred from the outgoing strand to the incoming strand by considering the interaction between RecA–ssDNA filaments and homologous dsDNA held under tension by an external force. Since the dsDNA is homologous to the ssDNA, the Watson–Crick pairing of the complementary strand can be transferred from the outgoing strand to the incoming strand. Thus, we study the possibility that strand exchange is free energetically favorable because it reduces the mechanical stress on the complementary strand bases due to the transfer of Watson–Crick pairing from the highly extended outgoing strand to the less extended incoming strand. Such a reduction would occur if the complementary strand were bound to the RecA filament dominantly through its Watson–Crick pairing since the bases that connect the complementary strand would experience great mechanical stress from either of its pairing partners, which have extensions that greatly exceed the B-form length.

We probe for effects of tension associated with the differential extension of the complementary strand and its Watson–Crick pairing partners by applying force to the ends of homologous dsDNA during strand exchange. If a reduction in the differential extension between the complementary strand and its pairing partners drives strand exchange forward, then using an external force to reduce that differential extension should slow strand exchange. Thus, pulling on the 3′5′-ends of the complementary strand would slow strand exchange because it decreases the differential extension between the complementary strand and its pairing partners, whereas pulling on the 3′5′-ends of the outgoing strand should not affect the strand exchange rate because it does not reduce the differential extension. Finally, we performed experiments where we investigated both the extension rates during strand exchange and during free RecA binding to dsDNA as a function of the force applied to the ends of the dsDNA to determine whether the relocation of the complementary strand could represent the rate-limiting step in strand exchange.

MATERIALS AND METHODS

Sample preparations

Double-stranded λ DNA molecules (New England Biolabs) were modified by hybridizing and ligating biotinylated oligonucleotides yielding 3′3′-labeled dsDNA. Ligation steps were done in the presence of a thermostable DNA Ligase (Ampligase, Epicentre, Madison, WI, USA). The oligonucleotides at both ends included a ssDNA tail [$(dT)_7$–(biotin-$dT)_6$] to allow free rotation of the bonds. After each modification step was completed, the dsDNA sample was washed three times using Amicon YM-100 filters (Millipore, USA) and 70 mM Tris buffer pH 7.6. The final concentration was determined by the absorbance at 260 nm.

The RecA–ssDNA filaments were prepared by mixing 3 μM ssDNA (5 kb) with 1 μM RecA, 1 mM ATPγS and 0.2 μM SSB in RecA buffer (70 mM Tris–HCl, 10 mM $MgCl_2$, and 5 mM dithiothreitol, pH 7.6). The ssDNA was previously prepared by amplifying a 5-kb fragment using λ-phage as a template where one of the primers was 5′-phosphorylated, and purified using a Macherey–Nagel kit. The dsDNA PCR fragment was subsequently incubated with λ-exonuclease enzyme (NEB) at 37°C for 30 min, and the resulting ssDNA was further purified using a Qiagen kit.

Experimental conditions and data analysis

For experiments where free RecA binds to dsDNA, an aliquot of dsDNA in RecA buffer, 1 μM RecA (New England Biolabs), 1 mM ATPγS and the beads were placed in a square micro-cell. Initially, a force of ~70 pN was briefly applied to achieve overstretching. After overstretching, the force was quickly reduced to between 15 and 36 pN.

Strand exchange experiments were performed with and without initial overstretching, and the results in both cases were the same. In strand exchange experiments, free RecA was replaced by a 5 µl aliquot of the filament preparation in a 50 µl final volume.

The experiments that presented different concentrations of Mg(II) were done by initially incubating all the reagents in 10 mM Mg(II) and finally diluting to each Mg(II) concentration to allow for the initial binding of the filaments in the homolog region and subsequently following strand exchange at several Mg(II) concentrations.

The temperature of the square capillary was varied from 22 to 37°C using a thermoelectric cooler placed on top of the aluminum mount holding the square capillary. A temperature sensor close to the capillary channel provides feedback for the stabilization loop controlling the thermoelectric cooler.

Experiments were carried out using our magnetic tweezers set up (16). Data analysis was performed using scripts custom-written in Matlab (www.mathworks.com) as described in our previous work (17).

RESULTS

We perform single-molecule experiments to measure the changes in dsDNA extension as a function of time in the presence of either free RecA or RecA–ssDNA filaments (Supplementary Figure S2). The total extension of each dsDNA molecule is followed at various constant forces, with values from 15 to 36 pN. Given that the binding of dsDNA to RecA extends the average length of the dsDNA by 0.51 nm per bound base pair triplet, measurements of changes in the dsDNA extension can be used to calculate the binding rate. A detailed description of the algorithms used to determine the binding rates was presented elsewhere (17).

For strand exchange experiments, we mix and briefly incubate the RecA–ssDNA filaments with dsDNA at zero force and subsequently apply a constant force to monitor the change in dsDNA extension due to strand exchange. Initial force-induced nucleation is not required for experiments using RecA–ssDNA filaments since these filaments spontaneously pair with homologous dsDNA. In contrast, for the experiments where free RecA binds to dsDNA, nucleation is required before RecA can begin to polymerize along a dsDNA molecule. Since dsDNA in the post-strand exchange complex is bound in site I of RecA, properties of this complex are often probed via measurements of direct binding of free RecA to dsDNA, which is believed to result in dsDNA bound to site I (15). Thus, in free RecA-binding experiments nucleation is achieved by briefly applying a force of ~70 pN to each dsDNA molecule before starting the extension measurement (17–20).

Histograms of the slopes of the extension curves for primary site binding experiments usually show clear periodic peaks corresponding to integer multiples of a characteristic elongation rate (17), whereas histograms for the strand exchange experiment usually show one dominant peak that may sometimes be accompanied by a few additional peaks that are integer multiples of the dominant peak (Figure 3A). In the free RecA experiments the integer multiples may correspond to multiple nucleation sites (17).

In the experiments that studied whether or not strand exchange was affected by applying a force to single-stranded tails at the ends of ~50-kb λ-phage dsDNA molecules during strand exchange (Figure 3B), we used three different dsDNA constructs and two different ssDNA filaments (Figure 3C). The sequence of λ-phage is more GC-rich at one end than the other, so the two strands of λ-phage dsDNA can be unambiguously identified by specifying the direction with respect to the GC-rich end (indicated in Figure 3C by the lavender band). In Figure 3C, the black dsDNA backbone corresponds to the 5′3′-strand starting from the GC-rich end and the grey dsDNA backbone indicates the 3′5′-strand starting from the GC-rich end. The two ssDNA filaments were prepared from a 5-kb fragment amplified from the GC-rich end of λ-phage. After treatment with λ-exonuclease, ssDNA is obtained. Depending on the strand chosen to interact with the exonuclease, the resulting ssDNA is complementary to either the 5′3′-strand or the 3′5′-(Supplementary Figure S3). The filament complementary to the 5′3′-strand is shown in grey, and the filament complementary to the 3′5′-strand is shown in black. These schematics illustrate that pulling on the 3′5′-ends of the complementary strand reduces the differential tension between the complementary and outgoing strands. In contrast, pulling on the 3′3′-ends or pulling on the 3′5′-ends of the outgoing strand does not reduce the differential tension between the complementary and outgoing strands.

Figure 3B shows the rate at which dsDNA is extended by strand exchange. The solid line indicates the strand exchange rate observed in previous work when no force was applied to the ends of the dsDNA (21). The symbols in the figure correspond to results where the dsDNA is pulled along a single backbone, as well as results where dsDNA was pulled from the 3′3′-ends. The blue circles show the results when the force is applied to the outgoing strand, as determined by the filaments shown in grey, whereas the red squares show the results when the force is applied to the complementary strand, as determined by the filaments shown in grey. The red diamond corresponds to the case when the strand pulled is complementary to the RecA–ssDNA black filament (Figure 3Ciii). Thus, though Figures 3Ci and 3Ciii both correspond to pulling on the strand that is complementary to the ssDNA in the filament, they represent opposite strands of the dsDNA molecule. Similarly, though Figures 3Cii and 3Ciii correspond to pulling on the same physical strand of the dsDNA, they correspond, respectively, to the outgoing and complementary strands for the RecA–ssDNA filaments used in each of those experiments. The figure also shows that when pulling on the 3′3′-ends (grey triangles) or pulling on the 3′5′-ends of the outgoing strand, the observed strand exchange rate is the same as the rate that is observed in the absence of force. In contrast, pulling on the 3′5′-ends of the complementary strand significantly reduces the observed strand

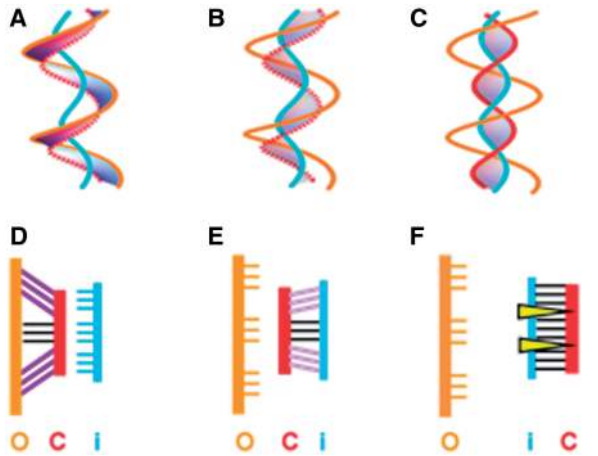

Figure 2. Illustrations of the RecA–DNA structures. (**A–C**) show the approximate positions of the outgoing (orange), complementary (red) and incoming (cyan) backbones, where the base pairs are located within the shaded regions. (A), (B) and (C) show the searching state, the intermediate post-strand exchange state, and the final post-strand exchange state respectively. (**D–F**) have the same backbone color code as (A). They show a schematic representation of the stress on the base pairs, where purple indicates high stress, lavender indicates moderate stress, and black indicates low stress. (D), (E) and (F) show the searching state, the intermediate post-strand exchange state, and the final post-strand exchange state, respectively.

exchange rate. This result is true whether the filament is complementary to the sequence at the 3′-end (black) or the 5′-end (grey) of the dsDNA.

We conducted these experiments to allow us to discriminate between the following possibilities: (1) the results depend on whether the pulled strand was complementary to the RecA–ssDNA filament, (2) the results depend on the physical strand being pulled, (3) the results depend on the sequence in the filament and (4) the results depend on whether the pulled dsDNA end nearest the filament was 3′ or 5′. If (1) were true, then cases represented in Figure 3C(i) and (iii) would be the same, whereas 3C(ii), (iv) and (v) would be different. If (2) were true, then 3C(ii) and (iii) would be the same, whereas 3C(i) would be different. If (3) were true, then 3C(i) and (ii) would be the same, whereas 3C(iii) would be different. Finally if (4) were true, then 3C(ii), (iii), (iv) and (v) would be the same and 3C(i) would be different.

In sum, the results show that the strand exchange rate is reduced if and only if a force between 15 and 36 pN is applied preferentially to the dsDNA that is complementary to the ssDNA in the searching filament. In all other cases, the strand exchange rates are independent of force for forces <36 pN. At sufficiently low forces, the strand exchange rate for 3′5′ pulling on the complementary strand must return to the rate observed in the absence of force, but forces <5 pN were inaccessible to us in our experiment because the low dsDNA tension increased the uncertainty in the bead position due to Brownian motion making the extension rate determinations too difficult.

Furthermore, we compared the strand exchange rates and the rates for free RecA binding. The red plus signs and blue × symbols in Figure 3B show the rate at which dsDNA is extended by the binding of free RecA. The force-independent extension rates for free RecA binding are independent of which strand is being pulled and similar to the force independent extension rates obtained during strand exchange when the 3′3′-ends were pulled, shown as grey triangles. The extension rates observed when pulling on the 3′5′-ends of the outgoing strand are also similar.

In addition, Figure 4A shows an Arrhenius log plot of the characteristic elongation rates as a function of 1000/Temperature for dsDNA during strand exchange and during free RecA binding. The red circles show new results for free RecA binding to dsDNA under tension in a buffer containing ATPγS. New results for strand exchange obtained while force was applied to the 3′3′-ends of dsDNA molecules in the presence of either ATPγS or a mixture of ATPγS and ADP are shown by the black outlined red triangles and squares, respectively. For comparison, previous strand exchange rates obtained earlier by other groups are also shown in the figure. The line corresponds to the best fit results from earlier bulk experiments in the absence of force (21). The rest of the points correspond to single-molecule experiments (22–24). The grey diamonds show zero force results in a buffer containing ATP or ATPγS (23); the grey squares show results in a buffer containing either ATP or ATPγS when ~20–1000 bp were bound to the filament and a 0.5-pN force was applied to both termini at both ends of the dsDNA (24). Finally, the grey triangle corresponds to single-molecule experiments measured in a buffer containing ATP, where the strand exchange rate was determined by measuring the decrease in extension due to the unbinding of heteroduplex dsDNA at the lagging end of the strand exchange window (22).

Figure 4A shows that the strand exchange rates are insensitive to hydrolysis, applied force and the amount of dsDNA bound to the filament. Earlier work has shown that the elongation rate due to the binding of free RecA is insensitive to the applied force, free RecA concentration, hydrolysis and ATP concentration (17). Thus, the process or processes that limit the rates must not depend strongly on any of these factors.

We further studied the effect of salt concentration dependence of the rate at which dsDNA is extended by strand exchange and the rate at which dsDNA is extended by the binding of free RecA for the cases where the dsDNA is pulled from the 3′3′-ends (Figure 4B). Both rates increase as a function of $MgCl_2$ concentration. Similar results were obtained at 24°C. These data provide additional support for the contention that the rate-limiting step in strand exchange is the same as the rate-limiting step in the extension of dsDNA due to free RecA binding from solution. Though increasing $MgCl_2$ concentration increases the observed extension rate, earlier results have shown that increasing the NaCl concentration decreases the observed extension rate for free RecA binding (25) and strand exchange (26). Together these results suggest that the characteristic extension rate is not a simple function of Debye screening, but rather

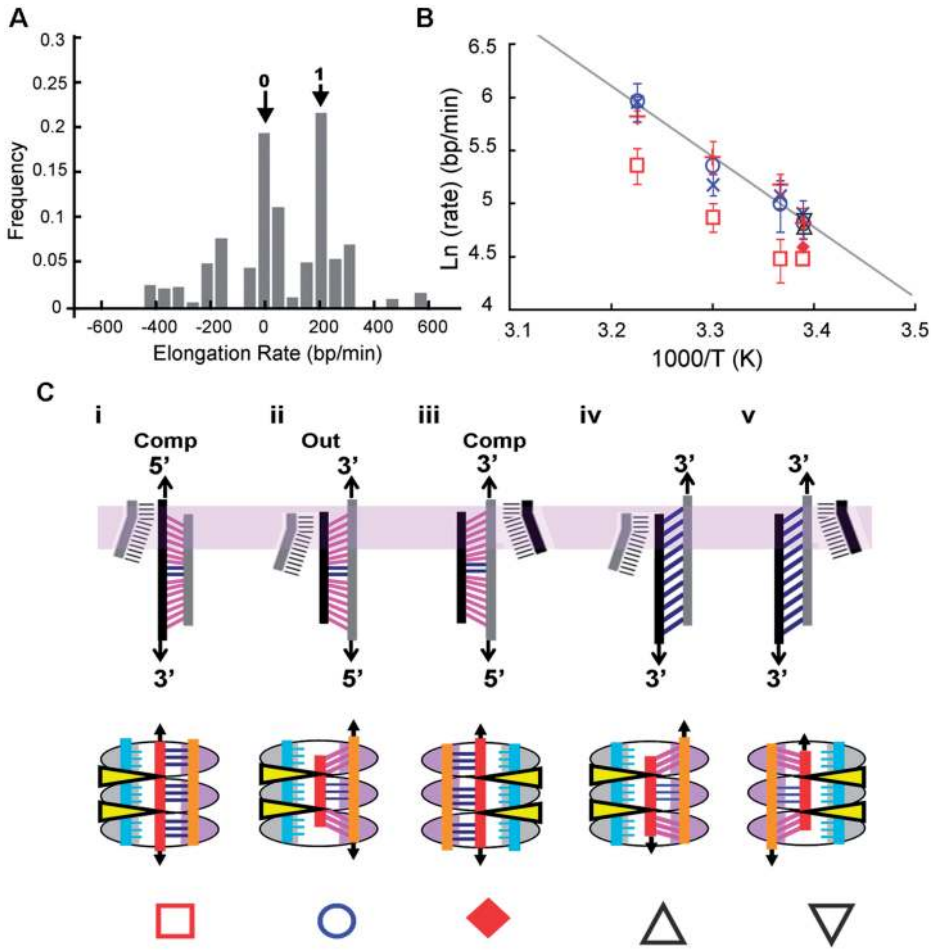

Figure 3. Effect of force applied to different ends of dsDNA λ-phage during strand exchange and free RecA binding. (**A**) Elongation rate histogram for strand exchange in ATPγS at 30°C; peak 1 shows the characteristic rate of 211.8 bp/min (0.60 nm/s) whereas peak 0 corresponds to molecules that were followed but showed no change in extension. (**B**) Extension rates as a function of temperature for different 3′5′- and 3′3′-pulling techniques. Arrhenius plot of single-molecule extension rates as a function of temperature for free RecA binding and strand exchange in ATPγS and 10 mM $MgCl_2$ with 3′5′ and 3′3′ pulling techniques. Strand exchange rates in bulk experiments and no external force (21) (grey line); free RecA binding while pulling 3′5′ from the complementary strand (red plus signs); free RecA binding while pulling 3′5′ from the outgoing strand (blue × symbols). Strand exchange rate while pulling 3′5′ from the complementary strand (red squares) and while pulling the other 3′5′ strand with the alternative filament complementary to the pulled strand (red diamond). Strand exchange rate while pulling 3′5′ from the outgoing strand (blue circles); strand exchange rate while pulling 3′3′ from the outgoing strand nearest the filament (grey triangle) and strand exchange rate while pulling 3′3′ from the complementary strand nearest the filament (grey upside-down triangle). Error bars: confidence intervals. (**C**) Schematic representation of the effect of force applied to different ends of the dsDNA constructs during strand exchange experiments. (**i**, **ii** and **iii**) dsDNA pulled from 3′5′-ends with stressed and unstressed base pairs shown in magenta and blue, respectively. (**iv** and **v**) dsDNA pulled from the 3′3′-ends with stressed and unstressed base pairs shown in magenta and blue, respectively. The grey and black ssDNA correspond to filaments complementary to opposite strands of the dsDNA. In the representation of strand exchange in the first row, the RecA molecules were omitted for better clarity. The lavender band indicates the GC-rich end in λ-phage dsDNA. The ellipses in the second row indicate RecA monomers with Site I and Site II shown in grey and purple, respectively. The outgoing, complementary, and incoming strands are shown in orange, red and cyan, respectively. The effect of pulling the 3′5′-ends of the complementary strand, 3′5′-ends of the outgoing strand, and 3′3′-ends of the dsDNA is represented. The symbols under each figure correspond to the symbols in Figure 3B.

depends on details of the interaction between the salts and the RecA–DNA complex.

DISCUSSION

When the dsDNA is in the intermediate state, pairing the complementary strand bases with either the outgoing strand or the incoming strand results in significant tension on the base pairs due to the differential extension of the complementary strand and its pairing partners. One would expect that pulling on the complementary strand would make strand exchange less favorable because it would reduce the free energy advantage gained by strand exchange. In contrast, pulling on the outgoing strand would not make the transition less favorable. Similarly, applying a uniform tension on both strands should not alter the strand exchange rate. Finally, pulling on any set of dsDNA ends should have no effect on the binding of free RecA since that process does not involve strand exchange.

The results shown in Figure 3B are completely consistent with the proposal that the differential extension

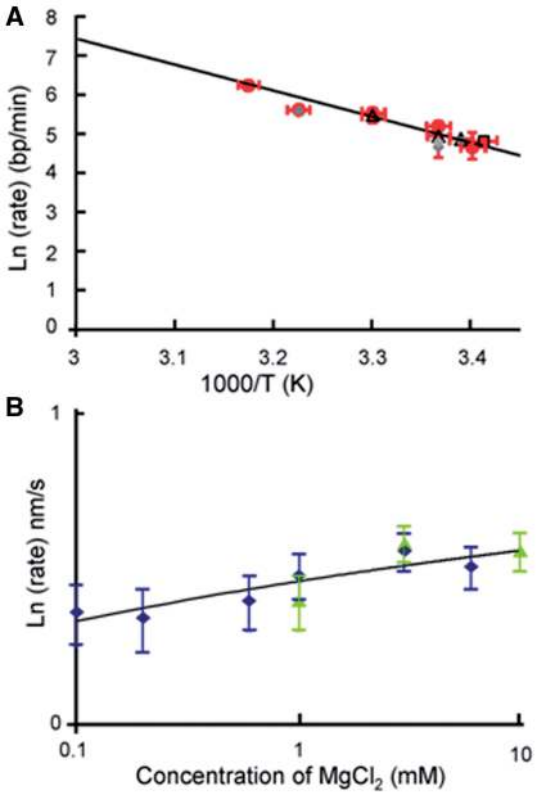

Figure 4. Comparison of the measured dsDNA elongation rates due to strand exchange and the binding of free RecA to dsDNA. (**A**) Arrhenius plot: strand exchange rates in bulk experiments and no external force (21) (black line); single-molecule strand exchange rates in ATP or ATPγS: grey diamonds (23) and grey squares (24); dsDNA release rate at the back of the strand exchange window in ATP: grey triangle (22); new results for strand exchange in ATPγS (red triangles), ATPγS–ADP (outlined-red square), and free RecA binding (red circles). Error bars: confidence intervals. (**B**) Extension rates as a function of $MgCl_2$ concentration. Log of single-molecule dsDNA extension rates (nm/s) as a function of $MgCl_2$ concentration (mM) for both free RecA binding (navy diamonds) and strand exchange (green triangles) in ATPγS at 22°C. $MgCl_2$ concentrations varied from 0.1 to 10 mM. The dependence of the rates of free RecA binding on $MgCl_2$ concentration is fit by a logarithmic trend-line (black line).

between the incoming strand and the outgoing strand drives strand exchange forward when the dsDNA is homologous to the ssDNA in the RecA–ssDNA filament. In particular, Figure 3B shows that the extension rates due to free RecA binding are independent of force. Similarly, the extension rates due to strand exchange are independent of force, except for the case where a force in excess of 15 pN is applied to the ends of the complementary strand. In that case, the strand exchange rate is significantly reduced. This reduction occurs whether the ssDNA is homologous to the 3′- or the 5′-end. Similarly, the results do not depend on which dsDNA strand was physically pulled, but on whether or not the pulled strand is complementary to the ssDNA filaments used in the experiments. Thus, the consistent differences shown in Figure 3B must not be the result of some artifact associated with the preparation of the dsDNA construct, or with some peculiarity of one of the two filaments. These results are also consistent with earlier experimental results suggesting that the differential tension between the complementary and outgoing strands drives the unbinding of non-homologous dsDNA in the homology searching state (16).

The experimental results shown in Figure 3 suggest that the samples do not contain significant internal nicks; otherwise, the observed results would be independent of the ends to which the force is applied because the nicks would redistribute the stress between the two strands (see 'Supplementary Discussion' section). Similarly, the finite stiffness of the backbone might redistribute the tension between the dsDNA strands so that the tension on both backbones was the same except at short regions near the end of the molecules (see 'Supplementary Discussion' section and Supplementary Figure S4). In this case, force-dependent effects would depend only on the terminus pulled at the end nearest the filament since the terminus pulled at the other end would have no effect on the force distribution near the filament. In this case, the situation illustrated in Figure 3C(ii–v) would produce the same result, whereas 3C(i) would produce a different result. In contrast, the results show that force-dependent effects are not determined by the choice of terminus at the filament end, rather they depend on whether or not the force is applied to both ends of the strand that is complementary to the ssDNA in the filament. Thus, the results show that when the dsDNA is pulled from the 3′3′-ends, the tension on both backbones extends fairly uniformly along the length of the molecule, whereas for 3′5′ pulling the tension over the entire length of the molecule is largely confined to the single backbone being pulled, as illustrated in Figure 3C.

We propose that the differential extension between the three DNA strands not only drives strand exchange forward for perfect homologs but also plays a vital role in decreasing the probability that partially homologous dsDNA will remain in the post-strand exchange state of the intermediate structure. Theory suggests that the differential extension in the searching state is so large that thermal energy fluctuations are insufficient to bind $>\sim$15 bp in the searching state, which represents less than one helical turn of the filament (27). Thus, the state shown in Figure 2A will not occur. For perfect homologs, strand exchange reduces the stress on the dsDNA and allows more base pairs to bind, but non-homologs simply unbind. In the intermediate state, the differential extension of the complementary and incoming strands results in a non-linearity in the free energy as a function of the number of bound triplets because the complementary strand backbone physically connects neighboring triplets. The physically connected backbone redistributes stress along all of the base pairs in the filament. As a result, the stress on any given base pair depends on the binding of all of the other base pairs in the filament (27); consequently, a mismatched triplet in the strand exchange state is not only unfavorable because of the loss of its Watson–Crick pairing but also unfavorable because the mismatch does not relieve any of the stress on the other bases in the strand exchange state. In contrast, the strand exchange of a sequence-matched triplet reduces the stress on other strand exchanged triplets, which substantially

reduces the total free energy of the system (27). Thus, in the intermediate strand exchange state the free energy difference between a homologous triplet and a non-homologous triplet can be much larger than the Watson–Crick pairing difference because the strand exchange of matched triplets reduces the stress on other homologous triplets in the intermediate strand exchange state.

A possible sequence of steps in the strand exchange of a perfect homolog is shown in Supplementary Figure S5. For sequences <30 bp, the stress due to the differential extension is distributed across all of the base pair triplets. Under these conditions, the presence of a mismatch in the final state increases the stress in all the other triplets. Thus, the strand exchange of subsequent triplets beyond a mismatched triplet can be quite free energetically unfavorable, even if the subsequent triplet is perfectly sequence-matched to the corresponding triplets in the incoming strand. We note that previous experiments have shown that extending past a mismatch can be unfavorable (28).

The discussion above assumed that the base flipping that tests for homology recognition occurs in an intermediate state where the position of the complementary strand differs from its know position in the final post-strand exchange structure. This proposition is supported by the similarity between the dsDNA extension rate due to strand exchange and the dsDNA extension rate due to the binding of free RecA to dsDNA, as we discuss below.

Figure 4A shows that for a range of temperatures, the dsDNA extension rates due to free RecA binding from solution are the same as the extension rates due to strand exchange, while pulling on the outgoing strand. Figure 4B shows that the rates as a function of $MgCl_2$ concentration are also similar suggesting that the two processes have a common rate-limiting step (Figure 2B). Given that binding of RecA to dsDNA increases dsDNA extension in both strand exchange and free RecA polymerization along dsDNA, one might suggest that dsDNA extension is the common rate-limiting step; however, we propose that this is unlikely for the following reasons: (i) Earlier measurements that divided the strand exchange process into steps have suggested that the sequence independent dsDNA extension occurs much more rapidly than subsequent sequence dependent steps (13). (ii) Each homology search attempt requires extension (13,16), and the measured *in vitro* homology searching rate exceeds ~50 bp/s (29) or 100 bp/s (30). (iii) *In vivo,* the search rate must also exceed 100 bp/s in order for the homology search to occur on a biologically relevant timescale. (iv) The observed rates for free RecA binding and strand exchange are both insensitive to force, but theory suggests that if the rate depended on extending the dsDNA (18,19), then the observed extension rates should depend on force, like the force dependence of the nucleation rate for free RecA (18).

Since transition steps that do not occur in both strand exchange and free RecA binding cannot provide the common rate-limiting step, we can eliminate several processes that might have limited either individual process. Unlike strand exchange, the binding of free RecA from solution does not involve base flipping, homology recognition or repositioning of the outgoing strand. Similarly, unlike free RecA binding, extension of the strand exchange product does not involve free protein diffusion or assembly of protein monomer interfaces from interactions with a free monomer.

Additional support for the contention that the strand exchange rate is not limited by the homology search process is provided by earlier experimental work that studied strand exchange in bulk reactions (31). These bulk reaction experiments showed that the process, which limits the strand exchange rate, occurs after the homology searching process is complete where the search process includes the dsDNA binding, extending and the base flipping to test for homology.

Having ruled out all of the possible rate-limiting factors detailed above, we propose that the repositioning of the complementary strand is the rate-limiting step in both free RecA binding and strand exchange. If in free RecA binding experiments one strand of the dsDNA initially binds to the strong contacts in site I while the other strand is pulled toward the positive charges in the direction of site II, then the pull orients the bases. The interaction between the L2 loop and the bases may also position Phe203 in the rise between the triplets in the intermediate state, providing a functional role for Phe203, which does not appear to have a functional role in the initial or final states. The resulting RecA–dsDNA structure would be similar to the post-strand exchange intermediate (Figure 5A, 'Supplementary Discussion' section, Supplementary Figures S6 and S7). If this is true, the repositioning of the complementary strand could be the rate-limiting step for free RecA binding and strand exchange (Figure 5B).

In sum, we propose that the final step in strand exchange is the transition from the intermediate state to the final state, which requires the repositioning of the complementary strand. In the final state, the differential extension between the incoming and complementary strands is slightly lower than the differential extension in the intermediate state. In addition, for perfect homologs, interactions with the L1 and L2 loops may provide additional mechanical support for the rises which is absent in the intermediate state (see 'Supplementary Discussion' section, Supplementary Figures S6, S7 and S8). If the linear term in the free energy is favorable for the intermediate post-strand exchange state, but the non-linear term is favorable for the final state because of the reduced stress on the base pairs due dominantly to support from the protein, then the transition from the intermediate state to the final state will not become free energetically favorable until a sufficient number of contiguous homologous bases have undergone strand exchange (27). If the required number were less than six triplets or 18 bases (32), then the non-linear difference in the free energy due to the differential extension would provide yet another significant enhancement in sequence stringency in comparison with a system that relied only on the Watson–Crick pairing energy of each individual base, since the correct sequence matching of all of the

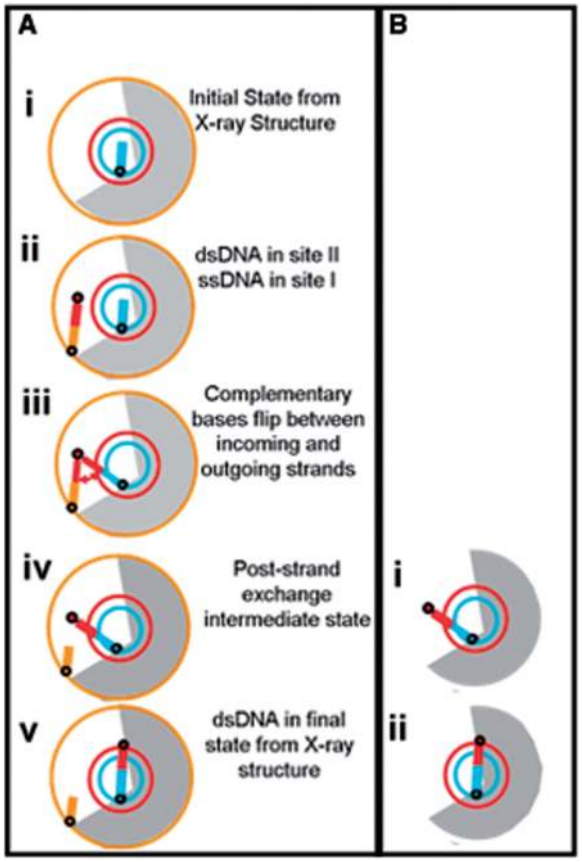

Figure 5. Proposed steps in strand exchange and free RecA binding. (**A**) Schematic of the steps involved in the proposed strand exchange process. Grey areas show regions occupied by the protein, excluding the L2 loop: (**i**) incoming strand (cyan), in the post-strand exchange state. (**ii**) dsDNA outgoing strand (orange circle) bound into Site II; complementary strand (red filled circle) in proposed intermediate position. (**iii**) Complementary strand bases and incoming strand bases rotate in search of homology, where an L2 loop may rotate with the incoming strand. (**iv**) Strand exchanged state that may benefit from an interaction with the L2 loop if the bases are homologous. (**v**) The heteroduplex dsDNA rotates to the final X-ray structure position, possibly accompanied by the L2 loop. (**B**) Schematic of possible steps in the binding of free RecA to dsDNA: (**i**) an additional free RecA binds to the dsDNA in the intermediate state, possibly accompanied by an interaction with the L2 loop. (**ii**) dsDNA (red circle) rotates to the final state shown by the x-ray structure, possibly accompanied by the L2 loop.

contiguous bases is required to move forward with strand exchange.

In conclusion, we speculate that the intermediate structure represented in Figure 5A(iv) may play the following roles in homology recognition: (i) The existence of an intermediate structure increases homology stringency because the total probability that the system will make a transition from the initial state to the final state is the product of each of the intermediate transition probabilities (27). (ii) If in the intermediate structure, the DNA tension is between the tension in the initial and final states, then dsDNA tension in the searching and intermediate states can be large enough to enforce accurate sequential kinetic proofreading, while the final state tension is low enough that adding dsDNA to site II remains free energetically favorable as long as almost all of the bound base pairs are in the final state and the total number of bound base pairs $<\sim$80 (27). (iii) For homologous triplets, strand exchange is driven forward by the reduction in dsDNA tension that occurs as a result of the differential extension of the backbones. (iv) The extension difference between the complementary strand and its pairing partners amplifies the free energy penalty due to the strand exchange of non-homologous bases so the penalty may be much larger than the Watson–Crick pairing penalty. (v) If incorrectly paired heteroduplex bases cannot stack with minimal free energy, then the transition from the intermediate state to the final state may be more favorable for homologs than for heterologs because homologs have less stress on the base pairs and more amino acid residue interactions. (vi) The free energy penalty for a sequence mismatch may be large enough to make the transition from the intermediate state to the final state so unfavorable that it does not occur before strand exchange spontaneously reverses and the dsDNA unbinds. This may explain why if $<\sim$21 bp are exchanged, encountering a base pair mismatch allows strand exchange to be fully reversed. In contrast, if $>$30 bp are exchanged, strand exchange cannot be reversed (28,32–34). A detailed discussion of these points is presented in 'Supplementary Discussion' section.

The type of kinetic proofreading system described in the previous paragraph represents a departure from Hopfield's original proposal for kinetic proofreading (28,35–37). In that original proposal, heterologs initially unbind faster than homologs, and the system makes a sequence independent irreversible transition to a tightly bound state after a time T. This article, however, considers a system where the initial unbinding rates for homologs and heterologs are the same, but a series of fully reversible sequence dependent intermediate transitions to more tightly bound states provides homology recognition. Thus, this article may provide insights about intermediate structures in the homology recognition–strand exchange process and their possible roles in fully reversible kinetic proofreading. Similar models may apply to other natural systems. Finally, since RecA family protein-based homology recognition is fully reversible it may provide a new paradigm for artificial self-assembly of nanoscale systems.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Discussion, Supplementary Figures 1–8, and Supplementary References [8,16,27, 38–44].

ACKNOWLEDGEMENTS

The authors thank C. Prévost, M. M. Cox and D. Bishop for useful discussions and A. Conover for manuscript corrections. The authors thank P. Rice and A. Rodriguez for useful suggestions.

FUNDING

NSF Graduate Research Fellowship (to D.Y.); National Institutes of Health [GM025326 to C.D.]; Harvard University (to M.P.). Funding for open access charge: Harvard University.

Conflict of interest statement. None declared.

REFERENCES

1. Kowalczykowski,S.C. and Eggleston,A.K. (1994) Homologous pairing and DNA strand-exchange. *Annu. Rev. Biochem.*, **63**, 991–1043.
2. Roca,A.I. and Cox,M.M. (1990) The RecA protein: structure and function. *Crit. Rev. Biochem. Mol. Biol.*, **25**, 415–456.
3. West,S.C. (2003) Molecular views of recombination and their control. *Nat. Rev. Mol. Cell Biol.*, **4**, 435–445.
4. Honigberg,S.M., Rao,B.J. and Radding,C.M. (1986) Ability of RecA protein to promote a search for rare sequences in duplex DNA. *Proc. Natl Acad. Sci. USA*, **83**, 9586–9590.
5. Müller,B., Koller,T. and Stasiak,A. (1990) Characterization of the DNA binding activity of stable RecA-DNA complexes. Interaction between the two DNA binding sites within RecA helical filaments. *J. Mol. Biol.*, **212**, 97–112.
6. Takahashi,M., Kubista,M. and Nordén,B. (1989) Binding stoichiometry and structure of RecA-DNA complexes studied by flow linear dichroism and fluorescence spectroscopy. Evidence for multiple heterogeneous DNA co-ordination. *J. Mol. Biol.*, **205**, 137–147.
7. Mazin,A.V. and Kowalczykowski,S.C. (1998) The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. *EMBO J.*, **17**, 1161–1168.
8. Chen,Z., Yang,H. and Pavletich,N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature*, **453**, 489–494.
9. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J Mol. Graphics*, **14**, 27–38.
10. Xiao,J. and Singleton,S.F. (2002) Elucidating a key intermediate in homologous DNA strand exchange: structural characterization of the RecA–triple-stranded DNA complex using Fluorescence Resonance Energy Transfer. *J. Mol. Biol.*, **320**, 529–558.
11. Bazemore,L.R., Takahashi,M. and Radding,C.M. (1997) Kinetic analysis of pairing and strand exchange catalyzed by RecA. Detection by fluorescence energy transfer. *J. Biol. Chem.*, **272**, 14672–14682.
12. Gumbs,O.H. and Shaner,S.L. (1998) Three mechanistic steps detected by FRET after presynaptic filament formation in homologous recombination. ATP hydrolysis required for release of oligonucleotide heteroduplex product from RecA. *Biochemistry*, **37**, 11692– 11706.
13. Xiao,J., Lee,A.M. and Singleton,S.F. (2006) Direct evaluation of a kinetic Model for RecA-mediated DNA-strand exchange: the importance of nucleic acid dynamics and entropy during homologous genetic recombination. *ChemBioChem*, **7**, 1265–1278.
14. Kaushik,R., Joo,C. and Ha,T. (2011) Real-time observation of strand exchange reaction with high spatiotemporal resolution. *Structure*, **19**, 1064–1073.
15. Pugh,B.F. and Cox,M.M. (1987) Stable binding of RecA protein to duplex DNA. Unraveling a paradox. *J. Biol. Chem.*, **262**, 1326–1336.
16. Danilowicz,C., Feinstein,E., Conover,A., Gunaratne,R., Kleckner,N. and Prentiss,M. (2012) RecA homology search is promoted by mechanical stress along the scanned duplex DNA. *Nucleic Acids Res*, **40**, 1717–1727.
17. Feinstein,E., Danilowicz,C., Conover,A., Gunaratne,R., Kleckner,N. and Prentiss,M. (2011) Single molecule studies of the stringency factors and rates governing the polymerization of RecA on double stranded DNA. *Nucleic Acids Res.*, **39**, 3781–3791.
18. Leger,J.F., Robert,J., Bourdieu,L., Chatenay,D. and Marko,J.F. (1998) RecA binding to a single double-stranded DNA molecule: a possible role of DNA conformational fluctuations. *Proc. Natl Acad. Sci. USA*, **95**, 12295–12299.
19. Hegner,M., Smith,S.B. and Bustamante,C. (1999) Polymerization and mechanical properties of single RecA-DNA filaments. *Proc. Natl Acad. Sci. USA*, **96**, 10109–10114.
20. Shivashankar,G.V., Feingold,M., Krichevsky,O. and Libchaber,A. (1999) RecA polymerization on double-stranded DNA by using single-molecule manipulation: the role of ATP hydrolysis. *Proc. Natl Acad. Sci. USA*, **96**, 7916–7921.
21. Bedale,W.A. and Cox,M.M. (1996) Evidence for the coupling of ATP hydrolysis to the final (extension) phase of RecA protein-mediated DNA strand exchange. *J. Biol. Chem.*, **271**, 5725–5732.
22. van der Heijden,T., Modesti,M., Hage,S., Kanaar,R., Wyman,C. and Dekker,C. (2008) Homologous recombination in real time: DNA strand exchange by RecA. *Mol. Cell*, **30**, 530–538.
23. 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.
24. Fulconis,R., Mine,J., Bancaud,A., Dutreix,M. and Viovy,J.L. (2006) Mechanism of RecA-mediated homologous recombination revisited by single molecule nanomanipulation. *EMBO J*, **25**, 4293–4304.
25. Galletto,R., Amitani,I., Baskin,R.J. and Kowalczykowski,S.C. (2006) Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature*, **443**, 875–878.
26. Cox,M.M. and Lehman,I.R. (1982) RecA protein promoted DNA strand exchange. *J Biol Chem*, **257**, 8523–8532.
27. Feinstein,E. and Prentiss,M. (2012) The tension on dsDNA bound to ssDNA/RecA filaments may play an important role in driving efficient and accurate homology recognition and strand exchange. arXiv: 1108.4936v2.
28. Sagi,D., Tlusty,T. and Stavans,J. (2006) High fidelity of RecA-catalyzed recombination: a watchdog of genetic diversity. *Nucleic Acids Res.*, **34**, 5021–5031.
29. Forget,A.L. and Kowalczykowski,S.C. (2012) Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature*, **482**, 423–427.
30. Honigberg,S.M., Rao,B.J. and Radding,C.M. (1986) Ability of RecA protein to promote a search for rare sequences in duplex DNA. *Proc. Natl Acad. Sci. USA*, **83**, 9586–9590.
31. Yancey-Wrona,Y.E. and Camerini-Otero,R.D. (1995) The search for DNA homology does not limit stable homologous pairing promoted by RecA protein. *Current Biol.*, **5**, 1149–1158.
32. Hsieh,P., Camerini-Otero,C.S. and Camerini-Otero,R.D. (1992) The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA. *Proc. Natl Acad. Sci. USA*, **89**, 6492–6496.
33. Mani,A., Braslavsky,I., Arbel-Goren,R. and Stavans,J. (2010) Caught in the act: the lifetime of synaptic intermediates during the search for homology on DNA. *Nucleic Acids Res.*, **38**, 2036–2043.
34. Shen,P. and Huang,H.V. (1986) Homologous recombination in Escherichi coli: dependence on substrate length and homology. *Genetics*, **112**, 441–457.
35. Hopfield,J.J. (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl Acad. Sci. USA*, **71**, 4135–4139.
36. Savir,Y. and Tlusty,T. (2010) RecA-mediated homology search as a nearly optimal detection system. *Mol. Cell*, **40**, 388–396.
37. Ninio,J. (1975) Kinetic amplification of enzyme discrimination. *Biochimie*, **57**, 587–595.
38. Datta,S., Ganesh,N., Chandra,N.R., Muniyappa,K. and Vijayan,M. (2003) Structural studies on MtRecA-nucleotide complexes: insights into DNA and nucleotide binding and the structural signature of NTP recognition. *Proteins*, **50**, 474–485.
39. Datta,S., Krishna,R., Ganesh,N., Chandra,N.R., Muniyappa,K. and Vijayan,M. (2003) Crystal structures of Mycobacterium smegmatis RecA and its nucleotide complexes. *J. Bacteriol.*, **185**, 4280–4284.

40. Danilowicz,C., Limouse,C., Hatch,K., Conover,A., Coljee,V.W., Kleckner,N. and Prentiss,M. (2009) Demonstration that the force versus extension curves for overstretched DNA depend on which ends are pulled. *Proc. Natl Acad. Sci. USA*, **106**, 13196–13201.
41. Danilowicz,C., Hatch,K., Conover,A., Ducas,T., Gunaratne,R., Coljee,V. and Prentiss,M. (2010) Study of force-induced melting of dsDNA as a function of length and conformation. *J. Phys.: Condens. Matter*, **22**, 414106.
42. de Gennes,P.G. (2001) Maximum pull out force on DNA hybrids. *Comp. Rendus de l'Acad. des Sci. Ser. IV Phys.*, **2**, 1505–1508.
43. Hatch,K., Danilowicz,C., Coljee,V. and Prentiss,M. (2008) Demonstration that the shear force required to separate short double-stranded DNA does not increase significantly with sequence length for sequences longer than 25 base pairs. *Phys. Rev. E*, **78**, 011920 1–4.
44. Baker,N.A., Sept,D., Joseph,S., Holst,M.J. and McCammon,J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.