# Mismatch sensing by nucleofilament deciphers mechanism of RecA-mediated homologous recombination

Xingyuan Huang[a,b,1], Ying Lu[a,c,1], Shuang Wang[a,b,1], Mingyu Sui[d], Jinghua Li[a], Jianbing Ma[a,b], Dongfei Ma[a,b], Qi Jia[a,b], Shuxin Hu[a,b], Chunhua Xu[a,b,2], and Ming Li[a,b,c,2]

[a]Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, 100190 Beijing, China; [b]School of Physics, University of Chinese Academy of Sciences, 100049 Beijing, China; [c]Songshan Lake Materials Laboratory, 523808 Dongguan, Guangdong, China; and [d]School of Physical Science and Technology, Lanzhou University, 730000 Lanzhou, Gansu, China

Edited by David Baker, University of Washington, Seattle, WA, and approved July 16, 2020 (received for review November 18, 2019)

**Recombinases polymerize along single-stranded DNA (ssDNA) at the end of a broken DNA to form a helical nucleofilament with a periodicity of ∼18 bases. The filament catalyzes the search and checking for homologous sequences and promotes strand exchange with a donor duplex during homologous recombination (HR), the mechanism of which has remained mysterious since its discovery. Here, by inserting mismatched segments into donor duplexes and using single-molecule techniques to catch transient intermediates in HR, we found that, even though 3 base pairs (bp) is still the basic unit, both the homology checking and the strand exchange may proceed in multiple steps at a time, resulting in ∼9-bp large steps on average. More interestingly, the strand exchange is blocked remotely by the mismatched segment, terminating at positions ∼9 bp before the match–mismatch joint. The homology checking and the strand exchange are thus separated in space, with the strand exchange lagging behind. Our data suggest that the strand exchange progresses like a traveling wave in which the donor DNA is incorporated successively into the ssDNA–RecA filament to check homology in ∼9-bp steps in the frontier, followed by a hypothetical transitional segment and then the post-strand-exchanged duplex.**

homologous recombination | RecA | strand exchange | molecular mechanism | dynamic configuration

Homologous recombination (HR) is a major pathway for repair of double-stranded DNA (dsDNA) break (1, 2). For instance, RecA polymerizes along single-stranded DNA (ssDNA) at the end of a broken DNA to form a presynaptic filament, which is a right-handed helix with approximately six monomers per turn and three bases per monomer (3). It catalyzes the search for homologous sequences and promotes strand exchange with a donor duplex composed a complementary strand (C strand) and an outgoing strand (O strand) (4). The strand exchange occurs likely through base flipping of the C strand from paring with the O strand to pairing with the incoming strand (I strand) (3, 5–8). DNA bases in the nucleofilament are grouped into nearly B-form triplets separated by rises of ∼0.78 nm (3). It is logical and actually widely accepted that the strand exchange advances in units of 3 base pairs (bp) because each monomeric RecA is associated with a base triplet in the filament (9–13). Recently, single-molecule assays showed that HR may involve approximately eight contiguous bases for initial homology testing (9, 10). We further found that ssDNA–RecA filaments interrogate the double-stranded donor DNA in ∼9-bp steps during strand exchange (14). These results suggest that, to facilitate the search for homology, the presynaptic filament checks the homology of base-pair units with a length of ∼9 bp. Despite the progresses, the molecular mechanism of strand exchange and the structure–function relationship in the presynaptic filament are still unclear. The lack of understanding can be traced to the transient and heterogeneous nature of the underlying intermediates, which cannot be accessed by typical experimental methodologies (15).

In an effort to solve the challenge to catch the transient states, we realized that the question of how strand exchange progresses is actually equivalent to how it is blocked. We therefore designed a series of mismatch-containing donor DNAs to assess the dynamics of HR at the single-molecule level. We found, surprisingly, that the strand exchange was blocked remotely by the mismatches at positions about 9 bp before the match–mismatch joint (MMJ). The distance coincides with the large step sizes observed in the strand-exchange process. Our results suggest a dynamic conformation in which a segment of ∼9-bp donor DNA is added to the front of the synapsis at a time, where it is checked for homology via base flipping, followed by a hypothetical transitional segment that connects the homologychecking frontier and the post-strand-exchanged duplex.

## Results

**Strand Exchange Terminates ∼9 bp before the MMJ.** We first used magnetic tweezers (MT) to trace the expansion of post-strand-exchanged synapsis. In our measurements, a DNA hairpin was used to mimic a half of the donor DNA (Fig. 1*A*, *Left*) and was connected to two DNA handles (Fig. 1 *A*, *Right*). The DNA hairpin was stable when the force exerted by the MT was lower than 10 pN and only naked I strands were added (*SI Appendix*, Fig. S1*A*). When strand exchange occurred between the DNA hairpin and the RecA-coated I strand, ∼3 pN was enough to pull the O strand

### Significance

Homologous recombination (HR) is a major pathway for repair of double-stranded DNA break. The RecA proteins form a helical nucleofilament on single-stranded DNA (ssDNA) with a periodicity of ∼18 bases, which searches a donor DNA that is homologous to the ssDNA. The key point in HR is how the filament senses the sequence of the donor DNA. We designed a series of mismatch-containing donor DNAs to study HR and found that the strand exchange was blocked remotely by the mismatches. Our data suggest that the strand exchange progresses iteratively with the sequences being checked successively in ∼9-bp steps in the frontier of the extending synapsis, followed by a transitional segment that leads the post-strand-exchanged duplex.

Author contributions: C.X. and M.L. designed research; X.H., Y.L., S.W., M.S., J.L., J.M., D.M., Q.J., S.H., and C.X. performed research; X.H., Y.L., S.W., C.X., and M.L. analyzed data; and S.W., C.X., and M.L. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

[1]X.H., Y.L., and S.W. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: xch@iphy.ac.cn or mingli@iphy.ac.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1920265117/-/DCSupplemental.



Downloaded at National Science Library on August 24, 2020

BIOCHEMISTRY

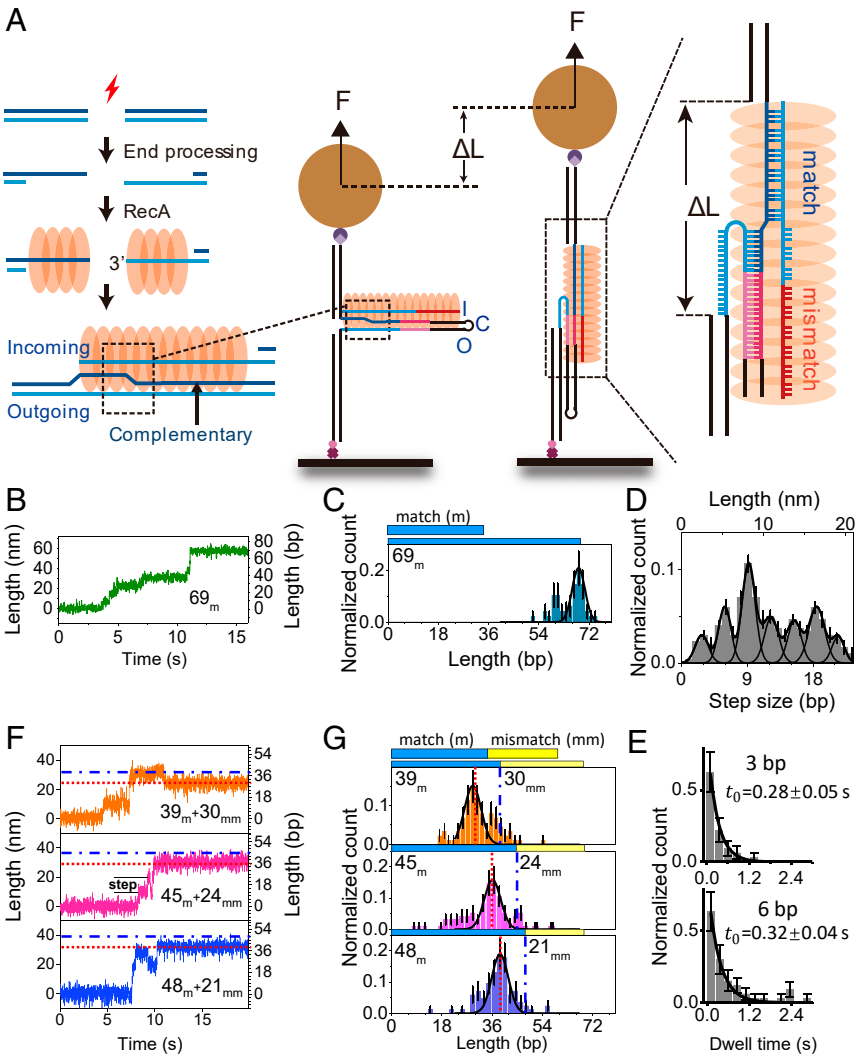


**Fig. 1.** MT assay of the strand exchange. (*A*) A 120-bp DNA hairpin connected to two handles (*Right*) to mimic a half of the donor DNA (*Left*). (*B*) A typical trace for strand exchange with a fully homologous 69-nt I strand. (*C*) Distribution of strand-exchanged lengths with a major peak at 68.2 ± 1.3 bp (SEM; $N_{curve}$ = 48). (*D*) Distribution of the strand-exchange step sizes. The narrow peaks (black lines) were from the fitting with a multi-Gaussian function, yielding peaks at 2.8 ± 0.3, 5.7 ± 0.3, 8.9 ± 0.4, 12.3 ± 0.7, 15.2 ± 0.5, 18.5 ± 0.6, and 21.4 ± 1.2 bp (SEM, $N_{step}$ = 308). The peak at ∼9 bp is the highest, with a fraction around 29%. (*E*) Dwell-time distributions of the 3- and 6-bp steps with a characteristic time $t_0$ = 0.28 ± 0.05 s (SEM, $N_{step}$ = 33) for the 3-bp steps and $t_0$ = 0.32 ± 0.04 s (SEM, $N_{step}$ = 48) for the 6-bp steps. (*F*) Typical traces for strand exchange of the donor DNA with partially nonhomologous I strands. (*G*) Distributions of the final lengths of the post-strand-exchanged segment displaying the main peak at 29.0 ± 1.5 bp (SEM, $N_{step}$ = 87) for orange, 36.0 ± 1.5 bp (SEM, $N_{step}$ = 95) for magenta and 38.8 ± 1.5 bp (SEM, $N_{step}$ = 84) for violet. The blue dash-dotted lines represent the position of MMJ. The red dotted lines represent the final position of the blocked HR.

out of the filament (*SI Appendix*, Fig. S2). The force was consistent with the measured strength of interaction between the O strand and site II (the filament's secondary DNA-binding site) (16). Unless otherwise indicated, we used ATPγS to monitor the synaptic events without turnover of RecA monomers during HR (12). The data in the following were acquired at ∼7 pN to gain a proper signal-to-noise ratio (SNR). We designed four types of I strands, which were all 69 nt long. One of the I strands was fully homologous to the O strand. The rest each consisted of two parts. The first part had a length of 39, 45, or 48 nt, respectively, and was homologous to the O strand. The second part was completely nonhomologous to the O strand, so that the HR process could not run through it. The MMJ was at the 39th, the 45th, and the 48th base, respectively. The strand exchange can start from the middle or from the left end of the DNA hairpin (17). It, however, cannot start from the right end of the DNA hairpin because the I strand was not homologous to the O strand there. We found that when the I strand joined the synapsis from the middle, nothing could be observed until the strand exchange progressed from the middle to the left end, resulting in a very large jump (tens of nanometers) of the end-to-end distance of the hairpin DNA (*SI Appendix*, Fig. S3). When the strand exchange proceeded from the left to the right end (Fig. 1*B* and *SI Appendix*, Fig. S4), we observed moderate steps (∼10 nm). The histogram of the sizes of such steps are shown in Fig. 1*D*, which confirms our previous observation that the strand exchange most probably progresses in ∼9-bp steps (14). The step size is widely distributed with the smallest one equal to 3 bp (Fig. 1*D* and *SI Appendix*, Fig. S4*B*). According to the Allan deviation plot (*SI Appendix*, Fig. S1*B*), the position error of our MT was 1.7 nm at an acquisition time of 30 ms (three data points). More than 90% of the 3-bp steps were identified with SNRs ≥ 1.5. Similar percentages of the 6- and 9-bp steps were identified with SNR ≥ 2 (please see details in *SI Appendix*).

Downloaded at National Science Library on August 24, 2020

Here, we focus on the termination position of the strand exchange. We expected that the position would be identical to the MMJ in the donor DNA. However, it turned out that, except for the case with the fully matched I strand (Fig. 1*C*), the most probable ending positions were about 9 bp before the MMJ (Fig. 1 *F* and *G*). We believe that the homology checking should have taken place before the strand exchange terminated. Otherwise, the presynaptic filament would not know whether the sequences beyond the MMJ are homologous or not. For the same reason, the sequences in the range from the MMJ−9 to the MMJ, which were homologous to the ones in the I strand, should also have been checked. Why were the strands not exchanged in this range?

**Nucleases Digest dsDNA That Is Not Incorporated into the Filament.** The termination position of HR is so surprising that it should be further studied. A scanning-force-microscopy analysis reported that the nonhomologous part of the donor DNA was specifically distorted and sharply bent out of RAD51-coated filaments (18). This promoted us to use endonucleases to analyze the length of the donor DNA not yet incorporated into the filament (Fig. 2). The I strand used in the assay was 48 nt long and was tethered to the surface by a short, biotinylated DNA duplex. The first 27 nt of the 45-nt-long C strand were homologous to the I strand, but the rest were completely not. The MMJ was, hence, at the 27th base. In the DNase I assay, Cy3 was labeled on the C strand at the 26th base that is just one base before the MMJ (Fig. 2*A*). If all of the homologous bases of the donor DNA were incorporated into the filament after strand exchange, the Cy3-labeling site would be buried in the groove of the filament, which is about 4 nm deep (3), and would therefore be protected from digestion. We, however, found that DNase I can digest up to 58% of the Cy3-labeled site (Fig. 2*C*), while only 11% was removed in a control with fully homologous sequences, indicating that the 26th base of the C strand was not protected by the synaptic filament. When Cy3 was shifted to the 24th base (Fig. 2*B*), only 37% of the Cy3-labeled bases were digested, indicating that the 24th base was more protected than the 26th base because the former was more deeply buried in the groove of the filament. In the endonuclease BmgB I assay, Cy3 was labeled on the right end of the donor dsDNA, while the recognition site was positioned at the 27th to 32nd bases covering the MMJ (Fig. 2*D*). The Cy3 counts were reduced by 62% after digestion (Fig. 2*F*), indicating that the MMJ was exposed to the surface of the filament. Altogether, the assays verified that the strand exchange did not reach the MMJ, and the rest of the donor DNA remained outside the filament.

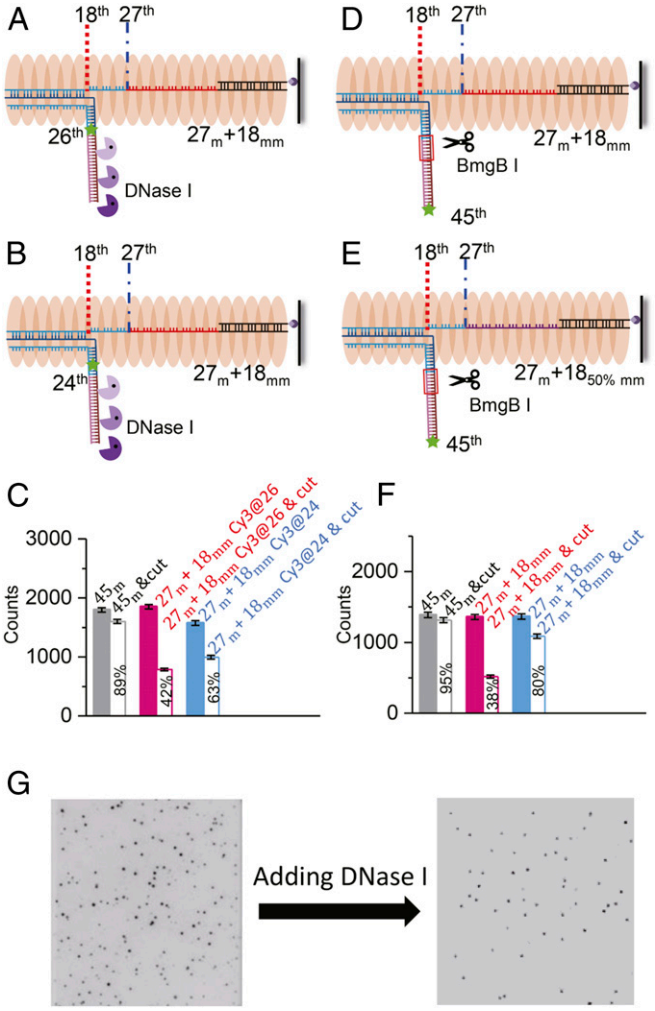


**Fig. 2.** Nuclease protection assays to locate the bending site. (*A*–*C*) DNase I was used to digest the donor DNA with Cy3 at the 26th base (*A*) and the 24th base (*B*). (*D*–*F*) A recognition site for the endonuclease BmgB I was located at the 27th to 32nd bases with Cy3 at the right end of the donor DNA. The MMJ was at the 27th base. Digestion fractions of 89%, 42%, and 63% relative to gray, red, and blue columns, respectively, were obtained by DNase I in *C* and 95%, 38%, and 80% by BmgB I in *F*. Data in *C* and *F* were from 30 images in three independent assays. (*G*) Images of Cy3 molecules before and after the digestion.

**The Donor DNA Bends out of the Filament at the Site ∼9 bp before the MMJ.** One needs to know what happens before the strand exchange occurs in order to understand why the strand exchange does not occur in the range from MMJ−9 to MMJ. The MT assay can only provide information about the separation of the O strand from the post-strand-exchanged synapsis. We therefore designed a Förster resonance energy transfer (FRET) assay to analyze the change in configuration of the donor DNA in the frontier of the expanding synapsis. As displayed in Fig. 3*A*, the I strand used here was the same as that in the nuclease protection assay. The first 27 nt of the 45-nt long C strand were homologous to the I strand, but the rest were completely not. The C strand was labeled with Cy5 at the 38th base from the left 5′ end, and the I strand was labeled with Cy3 at the 29th base from the left 3′ end. After a few intermediate states, most of the FRET signals converged to a main value around 0.18 when HR terminated (Fig. 3*B*). The value was too low to be explained by a configuration that the whole donor dsDNA was stretched and aligned parallel to the filament, in which case Cy3 and Cy5 were 9 bp apart inside the filament (see *SI Appendix* for distance calibration of FRET and *SI Appendix*, Fig. S5). In order to determine the configuration of the donor DNA when HR terminated, we designed two control experiments with Cy3 labeled at the 18th and the 24th base from the 3′ end of the I strand, respectively (*SI Appendix*, Fig. S6), for which the final FRET values became 0.17 and 0.26, respectively. Altogether, our results were consistent with a conformation that the first ∼18 nt, rather than 27 nt, of the C strand paired with the I strand, whereas the rest of the donor DNA was bent out of the filament with an average angle of about 89° between dsDNA and the filament axial (Fig. 3 *A*, *Bottom* and *SI Appendix*, Fig. S6*E*; see *SI Appendix* text for detailed calculations). The positions at which the bending occurred were again at about 9 bp before the MMJ. The result was further confirmed by the observation that the bending position shifted from the 18th to the 21st base when the MMJ was moved from the 27th to the 30th base (Fig. 3 *D* and *E*). The control experiment in Fig. 3 *F* and *G* showed that the dsDNA could be completely incorporated into the filament when the donor DNA was fully homologous to the I strand.

Kinetics studies have shown that the initial stage of sequence recognition does not depend on adenosine triphosphate (ATP) hydrolysis (19–21). We carried out a control experiment using ATP as the cofactor (*SI Appendix*, Fig. S7). Three stages, namely,



Downloaded at National Science Library on August 24, 2020

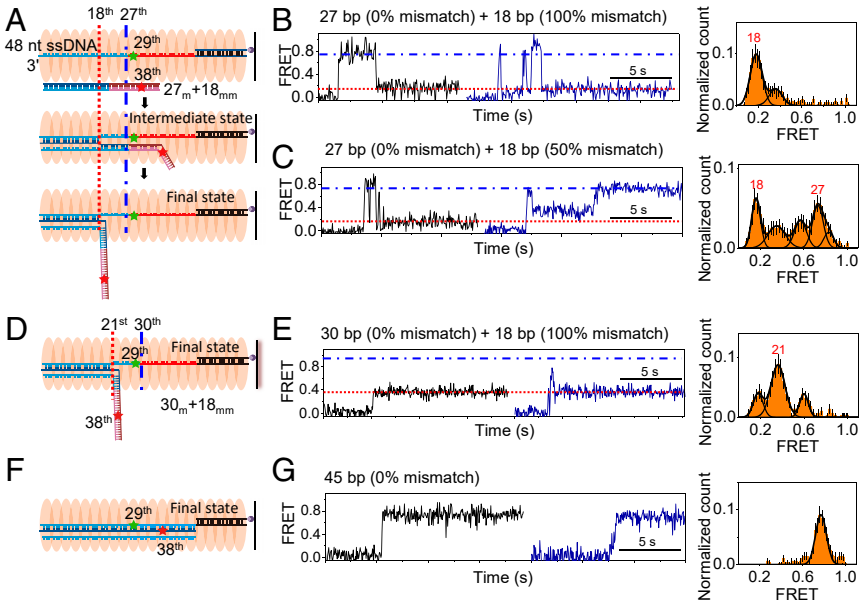


**Fig. 3.** FRET analysis of strand exchange. (*A*) Cartoons of strand exchange which ends up with a final configuration in which the donor DNA bends out of the filament. The blue dash-dotted lines represent the position of MMJ (the 27th base). The red dotted lines are 9 bp before the MMJ. *B*, *Left* displays FRET traces when the second part (18 bp) of the donor DNA was 100% mismatched to the I strand. *B*, *Right* shows the corresponding statistical distribution of the final FRET values, two peaks at 0.18 ± 0.01 and 0.34 ± 0.04 (SEM, $N_{curve}$ = 338), respectively. (*C*) The same as in *B*, except that the second part was 50% mismatched to the I strand, four peaks at 0.16 ± 0.01, 0.35 ± 0.02, 0.57 ± 0.02, and 0.75 ± 0.02 (SEM, $N_{curve}$ = 356), respectively. (*D* and *E*) The same as in *A* and *B* except that the MMJ was at the 30th base, three peaks at 0.18 ± 0.01, 0.36 ± 0.01, and 0.60 ± 0.01 (SEM, $N_{curve}$ = 315), respectively. The number labeled at each peak represents the bending position of B-DNA. (*F* and *G*) Data for the strand exchange between fully matched sequences. The peak at 0.78 ± 0.04 (SEM, $N_{curve}$ = 279) indicates that the donor DNA is completely incorporated to the filament.

*i*, *ii*, and *iii*, were observed, as indicated by different colors in *SI Appendix*, Fig. S7. It turned out that the results with ATP and with ATPγS were consistent in the first two stages. In the presence of ATP, stage *ii* was quickly succeeded by stage *iii*, in which the FRET values became irregular. The irregularity might arise from the partial dissociation of RecA upon ATP hydrolysis. We believe that the results we observed in the present work represent the true physical dynamics of HR before ATP hydrolysis.

**Strand-Exchange Termination May Shift to MMJ at a Lower Degree of Mismatch.** We used MT to analyze how the termination site of the strand exchange was affected by the degree of mismatch of the second part of the donor DNA. To this end, we constructed three types of I strands, each consisting of a 45-nt fully homologous segment and a 24-nt partially homologous segment (Fig. 4 *A* and *B*). The MMJ was at the 45th base. We observed that the termination sites for the DNA with 75% mismatch were mainly at the 36th base, while it was mainly at the 45th base for the DNA with 33% mismatch. The results indicate that the strand exchange can reach the MMJ when the second part has a higher degree of homology. That is to say, the termination site may shift by 9 bp to the MMJ when the degree of mismatch of the second part of the donor DNA is lowered. Interestingly, two main termination positions were observed, which were site 36 and site 45, respectively, when the second part of the donor DNA was 50% mismatched (Fig. 4 *B*, *Middle*). This is in accordance with the FRET result in Fig. 3*C* that the FRET values ended up at two main values, i.e., ∼0.16 and ∼0.75, when the second part of the C strand was 50% mismatched to the I strand. The two results can explain why the Cy3 counts were much less reduced in Fig. 2*E* (with 50% mismatch) than in Fig. 2*D* (with 100% mismatch) after the nuclease digestion.

**Strand-Exchange Termination Shifts to MMJ under High Forces.** The bending positions in the FRET assay were almost the same as that in the MT assay. Because no force was exerted on the DNA in the FRET assay, the agreement between the FRET and the MT assays implicated that the 7-pN force in the MT assay does not change the termination of the strand exchange. However, when the force was increased from 7 to 10 pN, the termination position of the strand exchange for an I strand composed of 45-nt fully matched segments and 24-nt completely mismatched segments had two values, one at MMJ−9 and the other at MMJ (Fig. 4 *C* and *D*). It implicated that the force may regulate the strand exchange when it is large enough. When the force was further increased, the DNA hairpin itself became unstable, such that it became difficult to observe the HR process.

**The Homology Checking Proceeds in ∼9-bp Steps.** We further examined by FRET how the bases of the donor DNA were aligned and elongated in the frontier of the expanding synapsis by positioning Cy3 at the right end of the I strand (the 47th base) and Cy5 at the 38th, 41st, or 45th base of the C strand (*SI Appendix*, Fig. S8), respectively. The MMJ was again at the 27th base, but with the second part 50% mismatched. Such a design allowed us to see as many as possible transient events beyond the MMJ to gain good statistics. Because the homology checking is dynamic, we were not able to determine the spatial position of the Cy5 molecule as we did in *SI Appendix*, Fig. S6. However, four preferred positions were recognized, as shown in the histograms in *SI Appendix*, Fig. S8. The results indicate that the homology checking proceeds mainly in ∼9-bp steps.

## Discussion

We designed a series of mismatches containing donor DNAs to study HR at the single-molecule level. Our study yielded a surprising result that, even though 3 bp is still the basic unit, both the homology checking and the strand exchange may proceed in multiple steps at a time, resulting in ∼9-bp large steps on average. More interestingly, the strand exchange is blocked at a position that is far before it arrives at the MMJ. The results

Downloaded at National Science Library on August 24, 2020

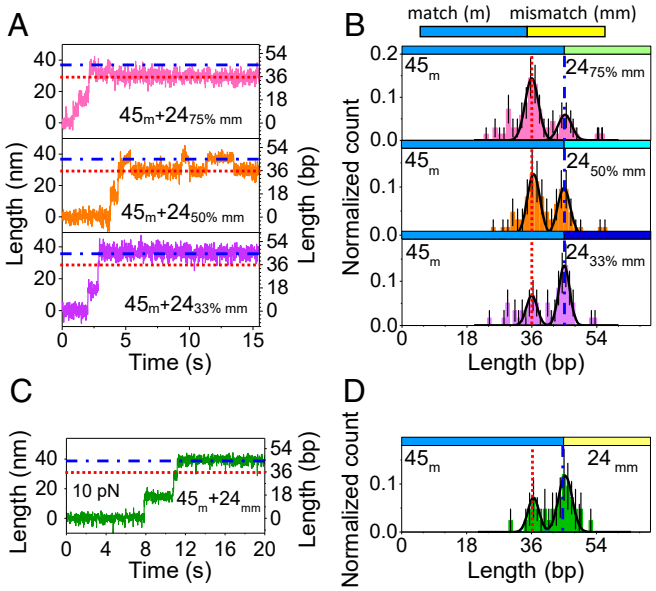


**Fig. 4.** Homology and force dependence of the strand exchange in the MT assay. (*A* and *B*) Three typical MT traces and the corresponding histograms of the final lengths of the post-strand-exchanged segments for three differently mismatched I strands. The degrees of mismatches of the second parts were 75%, 50%, and 33%, respectively. In 75% mismatch, two major peaks at 36.0 ± 1.9 bp (normalized amplitude = 0.14 ± 0.02) and 45.2 ± 1.8 bp (normalized amplitude = 0.05 ± 0.02), respectively. In 50% mismatch, two major peaks at 36.6 ± 1.8 bp (normalized amplitude = 0.12 ± 0.03) and 45.0 ± 2.5 bp (normalized amplitude = 0.10 ± 0.03), respectively. In 33% mismatch, there were two major peaks at 36.0 ± 1.5 bp (normalized amplitude = 0.07 ± 0.01) and 45.0 ± 1.5 bp (normalized amplitude = 0.13 ± 0.04), respectively. $N_{curve}$ = 69, pink; $N_{curve}$ = 63, orange; $N_{curve}$ = 59, purple. All of the errors listed are SEM. (*C* and *D*) A typical MT trace and the corresponding final length distribution for an I strand whose second part is 100% mismatched. The force was 10 pN. Two peaks at 36.5 ± 1.8 and 45.3 ± 2.4 bp (SEM, $N_{curve}$ = 42). The blue dash-dotted lines represent the position of MMJ. The red dotted lines are 9 bp before the MMJ.

suggest that the homology checking and the strand exchange are separated in space, with the strand exchange lagging behind the homology checking by ∼9 bp.

The X-ray structure analyses for HR revealed that the helical radius of the I strand (yellow line in Fig. 5) is ∼1 nm in both the presynaptic and the post-strand-exchanged filaments of RecA (3). The same study also showed that the helical radius of the C strand is ∼1.2 nm in a post-strand-exchanged filament (red line in Fig. 5). Neither X-ray structures nor cryo-electron microscopy structures were available for the helical radius of the O strand (gray line in Fig. 5). Yet, molecular-dynamics simulations and structural analysis by Peacock-Villada et al. (22) suggested a value of about 2.5 nm. Assuming that the bases of the C strand should pair with both the O strand and the I strand during the homology testing, the same group suggested that the homology-testing C strand (green line in Fig. 5) has a larger helical radius and rotates with a large angle with respect to the post-strand-exchanged C strand (red line in Fig. 5). Therefore, the post-strand-exchanged C strand (red) and the homology-testing C strand (green) cannot be connected directly (Fig. 5 *A* and *B*). We, hence, propose that there must exist a transitional region (blue segment in Fig. 5 *C* and *D*) in which the helical radius of the C strand decreases gradually from the homology-testing region (green segment) to the post-strand-exchanged region (red segment).

Combining our results with those of previous studies (3, 5, 23–25), we propose that, in the forefront of the synapsis (Fig. 6), a segment of donor DNA (white duplex) is elongated and paralleled with the I strand (yellow) through interaction between the O strand (gray) and site *II* of the RecA filament. The I strand is bound firmly to site *I* (the filament's primary DNA-binding site) (3, 26, 27). Based on molecular-dynamics simulations and energy analysis, Peacock-Villada et al. (22) suggested a conformation for the C strand in the green region in Fig. 6*A*, in which the bases of the C strand are able to flip from paring with the O strand to pairing with the I strand, facilitating homology checking. However, elongation of the donor DNA results in a nonlinear increase in elastic energy (13, 28–30). As a result, the total binding energy has a minimum when ∼9-bp donor DNA is bound to the filament. After that, adding more triplets to that given conformation becomes free-energetically unfavorable. However, the elastic energy density would be reduced if the C strand transfers from the homology testing state (green) to the hypothetic intermediate state (blue) because the C strand in the blue region has a smaller helical radius and is, therefore, less stretched than in the green region. The energy can be further reduced if the C strand in the blue state transfers to the post-strand-exchanged red state, in which the C strand has an even smaller radius. Although being designed to investigate the initiation of HR, we believe that the prediction of the theoretical calculations by Prentiss et al. (13, 29) might also be employed to explain the progressing of strand exchange. The model involved two major checkpoints at two lengths, ∼9 contiguous homologous base pairs for firmly binding (the first major checkpoint), and ∼18 contiguous homologous base pairs for strand exchange (the second major checkpoint). That is to say, if all of the bound 18 bp are contiguous and homologous, they would transit to the post-strand exchange state. Following the same idea, we may propose a mechanism to interpret our data. At the moment illustrated in Fig. 6*A*, ∼9-bp donor DNA can be added at the forefront of the progressing HR (Fig. 6*B*) because the green segment in Fig. 6*A* is homologous to the I strand, so that the first major checkpoint is passed. Consequently, the previously added two segments can make transitions, from green to blue and from blue to red, respectively, because the second major checkpoint is passed (Fig. 6 *A* and *B*). A similar process occurs from Fig. 6 *B* and *C*. In this way, the strand exchange progresses like a traveling wave to the end of the donor DNA (Fig. 6*D*) or until it is blocked by a mismatched segment (black segments in Fig. 6 *E*–*H*). Even in the latter situation, the homology checking must proceed beyond the MMJ (Fig. 6*G*). Otherwise, it would never know the degree of homology beyond the MMJ. The transition from Fig. 6 *F* and *G*

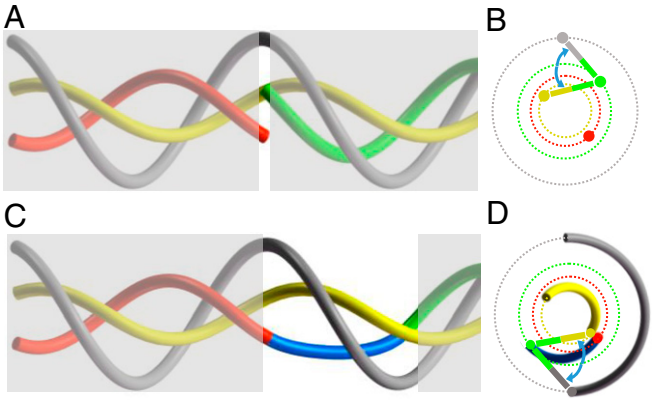


**Fig. 5.** Relative positions of the O (gray), I (yellow), and C (red + blue + green) strands in HR. (*A*) Disconnection between the post-strand-exchanged C strand (red segment) and the homology-testing C strand (green segment). (*B*) Two-dimensional projection of the central part in *A* that is not covered by the translucent rectangles. (*C*) Soft connection between the post-strand-exchanged C strand (red segment) and the homology-testing C strand (green segment). (*D*) Two-dimensional projection of the central part in *C* that is not covered by the translucent rectangles.



Downloaded at National Science Library on August 24, 2020

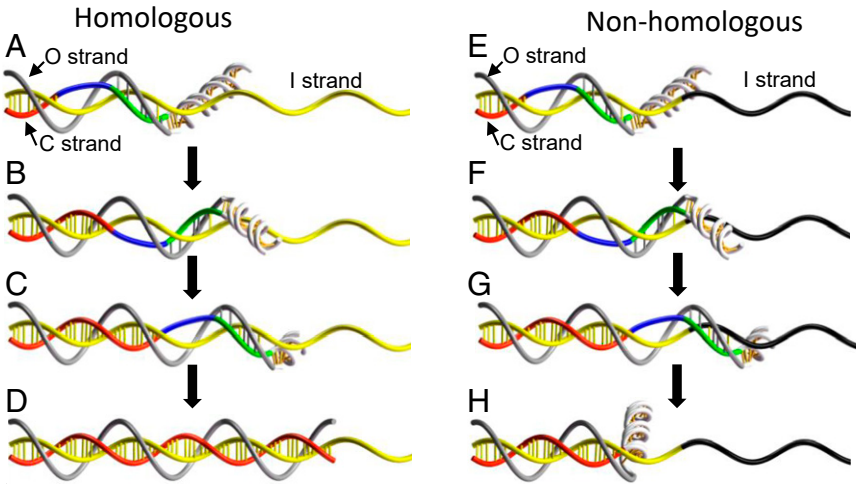


**Fig. 6.** Stepwise progression in HR. (*A–D*) HR with fully homologous DNA. The C strand is divided into three parts: post-strand-exchanged (red), homology testing (green), and transitional connection (blue). (*E–H*) HR blocked by the nonhomologous segment. The homology checking may proceed beyond the MMJ. The homologous I strand is colored yellow, and the nonhomologous I strand is colored black. Proteins are removed for clarity.

is allowed because the first major checkpoint is passed. However, the blue segment in Fig. 6*G* may not be allowed to transit to the red state (the strand-exchanged state) because the second major checkpoint is not passed in this range. Therefore, the ∼18-bp checking segment at the forefront of the progressing HR in Fig. 6*G* is unstable and will slowly unbind and bend out of the filament at a site which is ∼9 bp before the MMJ (Fig. 6*H*). The same argument may apply to the situation with a partially homologous barrier (please see details in *SI Appendix*, Fig. S9).

In summary, our data support a traveling-wave-like model for the extending of strand exchange during HR in which the donor DNA is incorporated successively into the ssDNA–RecA filament to check homology in ∼9-bp steps in the frontier, followed by a hypothetical transitional segment and then the post-strand-exchanged duplex.

## Materials and Methods

Experimental procedures of the MT and fluorescent assays and the DNA constructions and sequences can be found in *SI Appendix*. *SI Appendix* also includes calibrations of force or distance in MT and FRET assays and other data analyses and calculations, such as the statistics of step sizes and dwell times, resolution of steps, and estimation of the bending angle and bending site of the donor DNA.

**Data Availability.** All data are available in the main text or *SI Appendix*.

**ACKNOWLEDGMENTS.** This work was supported by National Natural Science Foundation of China Grants 11834018, 11574381, 91753104, and 11574382; Chinese Academy of Sciences (CAS) Key Research Program of Frontier Sciences Grant QYZDJ-SSW-SYS014; Youth Innovation Promotion Association of CAS Grant 2017015; and the K. C. Wong Education Foundation.

1. K. K. Khanna, S. P. Jackson, DNA double-strand breaks: Signaling, repair and the cancer connection. *Nat. Genet.* **27**, 247–254 (2001).
2. G. A. Cromie, J. C. Connelly, D. R. F. Leach, Recombination at double-strand breaks and DNA ends: Conserved mechanisms from phage to humans. *Mol. Cell* **8**, 1163–1174 (2001).
3. Z. Chen, H. Yang, N. P. Pavletich, Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489–4 (2008).
4. C. M. Radding, Recombination activities of *E. coli* recA protein. *Cell* **25**, 3–4 (1981).
5. A. V. Mazin, S. C. Kowalczykowski, The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. *EMBO J.* **17**, 1161–1168 (1998).
6. R. Fulconis *et al*., Twisting and untwisting a single DNA molecule covered by RecA protein. *Biophys. J.* **87**, 2552–2563 (2004).
7. L. R. Bazemore, M. Takahashi, C. M. Radding, Kinetic analysis of pairing and strand exchange catalyzed by RecA. Detection by fluorescence energy transfer. *J. Biol. Chem.* **272**, 14672–14682 (1997).
8. L. R. Bazemore, E. Folta-Stogniew, M. Takahashi, C. M. Radding, RecA tests homology at both pairing and strand exchange. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11863–11868 (1997).
9. J. Y. Lee *et al*., DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases. *Science* **349**, 977–981 (2015).
10. Z. Qi *et al*., DNA sequence alignment by microhomology sampling during homologous recombination. *Cell* **160**, 856–869 (2015).
11. J. Y. Lee *et al*., Sequence imperfections and base triplet recognition by the Rad51/RecA family of recombinases. *J. Biol. Chem.* **292**, 11125–11135 (2017).
12. K. Ragunathan, C. Joo, T. Ha, Real-time observation of strand exchange reaction with high spatiotemporal resolution. *Structure* **19**, 1064–1073 (2011).
13. J. Kates-Harbeck, A. Tilloy, M. Prentiss, Simplified biased random walk model for RecA-protein-mediated homology recognition offers rapid and accurate self-assembly of long linear arrays of binding sites. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **88**, 012702 (2013).
14. Y. W. Zhang *et al*., Iterative homology checking and non-uniform stepping during RecA-mediated strand exchange. *Biochem. Biophys. Res. Commun.* **478**, 1153–1157 (2016).
15. E. C. Greene, DNA sequence alignment during homologous recombination. *J. Biol. Chem.* **291**, 11572–11580 (2016).
16. I. De Vlaminck *et al*., Mechanism of homology recognition in DNA recombination from dual-molecule experiments. *Mol. Cell* **46**, 616–624 (2012).
17. J. Lee *et al*., Single-molecule four-color FRET. *Angew. Chem. Int. Ed. Engl.* **49**, 9922–9925 (2010).
18. D. Ristic, R. Kanaar, C. Wyman, Visualizing RAD51-mediated joint molecules: Implications for recombination mechanism and the effect of sequence heterology. *Nucleic Acids Res.* **39**, 155–167 (2011).
19. B. Boyer, C. Danilowicz, M. Prentiss, C. Prévost, Weaving DNA strands: Structural insight on ATP hydrolysis in RecA-induced homologous recombination. *Nucleic Acids Res.* **47**, 7798–7808 (2019).
20. S. C. Kowalczykowski, R. A. Krupp, DNA-strand exchange promoted by RecA protein in the absence of ATP: Implications for the mechanism of energy transduction in protein-promoted nucleic acid transactions. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3478–3482 (1995).
21. O. H. Gumbs, S. L. Shaner, 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 (1998).
22. A. Peacock-Villada *et al*., Complementary strand relocation may play vital roles in RecA-based homology recognition. *Nucleic Acids Res.* **40**, 10441–10451 (2012).
23. P. Howard-Flanders, S. C. West, A. Stasiak, Role of RecA protein spiral filaments in genetic recombination. *Nature* **309**, 215–219 (1984).
24. A. V. Mazin, S. C. Kowalczykowski, The specificity of the secondary DNA binding site of RecA protein defines its role in DNA strand exchange. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10673–10678 (1996).
25. D. Yang, B. Boyer, C. Prévost, C. Danilowicz, M. Prentiss, Integrating multi-scale data on homologous recombination into a new recognition mechanism based on simulations of the RecA-ssDNA/dsDNA structure. *Nucleic Acids Res.* **43**, 10251–10263 (2015).
26. D. A. McGrew, K. L. Knight, Molecular design and functional organization of the RecA protein. *Crit. Rev. Biochem. Mol. Biol.* **38**, 385–432 (2003).
27. B. Boyer, D. R. Yang, C. Danilowicz, M. Prentiss, C. Prevost, Structural insights into the mechanism of homologous recombination. *J. Biomol. Struct. Dyn.* **33**, 72 (2015).
28. L. Jiang, M. Prentiss, RecA-mediated sequence homology recognition as an example of how searching speed in self-assembly systems can be optimized by balancing entropic and enthalpic barriers. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **90**, 022704 (2014).
29. J. Vlassakis *et al*., Tension on dsDNA bound to ssDNA-RecA filaments may play an important role in driving efficient and accurate homology recognition and strand exchange. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **87**, 032702 (2013).
30. C. Danilowicz *et al*., RecA homology search is promoted by mechanical stress along the scanned duplex DNA. *Nucleic Acids Res.* **40**, 1717–1727 (2012).



Downloaded at National Science Library on August 24, 2020