No Sliding during Homology Search by RecA Protein*

Kenji Adzuma‡

From The Rockefeller University, New York, New York 10021

The RecA protein of Escherichia coli is a prototype of the RecA/Rad51 family of proteins that exist in virtually all the organisms. In a process called DNA synapsis, RecA first polymerizes onto a single-stranded DNA (ssDNA) molecule; the resulting RecA-ssDNA complex then searches for and binds to a double-stranded DNA (dsDNA) molecule containing the almost identical, or “homologous,” sequence. The RecA-ssDNA complex thus can be envisioned as a sequence-specific binding entity. How does the complex search for its target buried within nonspecific sequences? One possible mechanism is the sliding mechanism, in which the complex first binds to a dsDNA molecule nonspecifically and then linearly diffuses, or slides, along the dsDNA. To understand the mechanism of homology search by RecA, this sliding model was tested. A plasmid containing four homologous targets in tandem was constructed and used as the dsDNA substrate in the synapsis reaction. If the sliding is the predominant search mode, the two outermost targets should act as more efficient targets than the inner targets. No such positional preference was observed, indicating that a long range sliding of the RecA-ssDNA complex does not occur. These and other available data can be adequately explained by a simple three-dimensional random collision mechanism.

An early phase of homologous recombination involves a reaction by which two DNA molecules with almost identical, or “homologous,” sequences are aligned and the strands are exchanged between them. This reaction is called the “DNA strand exchange reaction.” The RecA protein of Escherichia coli is a catalyst of this reaction.

RecA is a member of the RecA/Rad51 family of proteins, which are ubiquitous among organisms, including humans (1, 2). These proteins are generally thought to catalyze recombination. A different perspective also exists in which these proteins are regarded as the enzymes primarily designed to repair damaged DNA by recombination (3, 4). In this regard, it is interesting to note that a human RecA homologue appears to be interacting with proteins involved in oncogenesis, including the BRCA1 and BRCA2 proteins (5, 6), whose mutations are implicated in the onset of breast cancer. It is not yet clear what exactly those eukaryotic RecA homologues (Rad51 proteins) do in vivo. However, at least in vitro, they can catalyze the DNA strand exchange reaction in a manner reminiscent of the E. coli RecA (7–9).

The strand exchange takes place between a ssDNA1 and the homologous dsDNA. The reaction can be carried out in vitro using only defined components: a homologous pair of dsDNA and ssDNA molecules, RecA, Mg2+, and ATP, as essential components (for reviews, see (4, 10)). The reaction produces a pair of DNA products: a hybrid-duplex molecule, which comprises the original ssDNA and the complementary strand in the original dsDNA, and a displaced strand from the original dsDNA. The molecular mechanism of how this reaction works, however, is not well understood, and many important questions remain unanswered.

The strand exchange reaction initiates with polymerization of RecA molecules onto the ssDNA substrate in the presence of ATP. The resulting filamentous RecA-ssDNA complex is called the presynaptic complex (PSC) and acts as a central intermediate (11, 12). RecA has at least two DNA binding sites (13–15). The primary site is occupied by the ssDNA in the PSC. Using the unoccupied secondary site, the PSC searches for and eventually binds to a homologous sequence in the dsDNA substrate very stably. This binding, or “synapsis” of the dsDNA and ssDNA molecules, gives rise to another nucleoprotein filament called the stable synaptic complex (SSC), which involves all three DNA strands. Finally, in the presence of ATP hydrolysis, the synapsis step is followed by the release of a displaced strand from the SSC. Because at least one role of ATP hydrolysis is to release the strand exchange product(s) (16, 17), the SSC accumulates in the presence of practically nonhydrolyzable ATPγS. In this paper, the term “DNA synapsis” is operational and refers to the reaction to form the SSC by using ATPγS.

As a first approximation, the DNA synapsis reaction can be viewed as a binding reaction between a large “sequence-specific” binding entity (i.e. PSC) and its target (homologous sequence). Just like any other sequence-specific binding proteins, the PSC must find its target among vast arrays of nonspecific sequences and must do so within a time scale that biology of cells demands. How does the PSC do this? The initial contact between the complex and a dsDNA molecule must necessarily be nonhomologous. After this nonproductive contact, the PSC might fall off from the dsDNA molecule and start over another round of random search. Alternatively, the PSC might sample the next potential target without leaving the dsDNA molecule to which it is already bound. In the latter scenario, the initial nonspecific PSC-dsDNA complex is not an abortive complex destined to fall apart but instead acts as an essential intermediate in homology search.

Berg et al. (18) have rigorously discussed the two-step binding mechanism, in which nonspecifically bound complexes act as intermediates (see also Ref. 19). Only those aspects directly

* This work was supported by National Institutes of Health Grant R01 GM052862. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence may be addressed: The Rockefeller University, Box 256, 1230 York Ave., New York, NY 10021. Tel.: 212-327-7471; Fax: 212-327-7183; E-mail: adzuma@rockvax.rockefeller.edu.

1 The abbreviations used are: ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; PSC, presynaptic complex; SSC, stable synaptic complex; bp, base pair(s); ATPγS, adenosine 5′-O-(3-thio)triphosphate; KOAc, potassium acetate; BSA, bovine serum albumin; DTT, dithiothreitol.
relevant to this article are summarized here. Compared with the one-step search mechanism, which relies only on the three-dimensional random collision, the two-step mechanism could make the search more efficient by accelerating the diffusion of reactants. One such mechanism is called the “sliding” mechanism (20). In this model, a protein first binds to a DNA molecule nonspecifically and then diffuses linearly, or “slides,” along the contour of the DNA molecule. It is crucial to recognize that the sliding is a one-dimensional random walk, not a unidirectional movement at the expense of external energy. Thus, at any point along the DNA, probabilities for a sliding protein to move to the left and right are in principle equal and not dependent on the prior movement. Sliding could accelerate diffusion of a protein, because it circumvents the activation energy involved in dissociation-reassociation with DNA. Also, the sliding in effect expands the size of a target to a distance that a protein “scans” during sliding. Note that diffusion such as sliding progresses in proportion to not the time but the square root of the time (e.g. it takes 4 times more time to diffuse twice as far). Therefore, the kinetic gain by sliding tends to diminish as the diffusing distance becomes longer.

For at least some sequence-specific binding proteins, several lines of evidence support the involvement of sliding during target location. These examples include, but are not limited to, the Lac repressor (21–23), EcoRI endonuclease (24, 25), and E. coli RNA polymerase (26, 27). Might the sliding also be involved in the homology search by RecA? Attempts to address this problem in the mid-1980s produced mixed results (28–30); however, no further systematic attempts have been made to date. This article describes the results of simple functional assays that should signal the existence of sliding if this is a predominant mechanism for homology search.

**EXPERIMENTAL PROCEDURES**

DNA—The plasmid pKN106 was constructed as follows. A 106-bp PstI–Smal fragment was cut out from the pUC19 DNA at positions 306–412. The purified fragment was joined together by T4 DNA ligase in the presence of PstI and Smal to produce only the head-to-tail joined products (PstI and Smal both generate blunt ends). The joined fragment of trimer size was purified through polyacrylamide gel electrophoresis and then inserted back into the original Smal site of pUC19. The resulting plasmid thus carries four copies of the PstI–Smal 106-bp sequence in tandem repeats. The sequence of Oligo-1 corresponds to positions 373–412 (GACGTT...CGGTAC) of pUC19 in this orientation; Oligo-2 is complementary to Oligo-1. The E. coli genomic DNA was purified from strain DH5α by a conventional lysis method, followed by deproteinization and CsCl-ethidium bromide ultracentrifugation. Unless otherwise noted, DNA concentrations refer to those of the DNA molecules, which are based on their nucleotide concentrations calculated from the absorbance at 260 nm using nominal extinction coefficients of 1.0 and 0.94 at 260 and 280 nm, respectively. DNA concentrations were calculated from the absorbance at 260 nm using the extinction coefficients of 1.0 and 0.94 at 260 and 280 nm, respectively.

**Proteins**—RecA was purified from an overproducing strain JC12772 (31), according to a published protocol (32) with minor modifications. The final fraction was dialyzed against the RecA diluent containing 20 mM Tris acetate (pH 7.5), 0.1 mM EDTA, 1 mM DTT, and 10% (v/v) glycerol. The concentration of RecA was determined from the absorbance at 278 nm in the RecA diluent using the extinction coefficient of 2.16 × 104 M−1 cm−1 (33). The EcoRI nuclease was purchased from New England Biolabs. Because the commercial EcoRI is in the buffer already (31), according to a published protocol (32) with minor modifications.

**Formation of the Synaptic Complex and Methyase Protection Assay**—To make the PSC, 75 nM of Oligo-1 or Oligo-2 was preincubated with 2.0 μM RecA in 10 μl of the synapsis buffer (40 mM Tris acetate, pH 7.5, 25 mM potassium acetate, 5 mM magnesium acetate, 1 mM EDTA, 250 μM ATPγS, and 100 μg/ml BSA) at 37 °C for 30 min. The PSC mixture was then added with 10 μl of the target DNA mixture containing 17 μg pKN106 DNA and 1.0 mM (in nucleotide) HindIII-digested E. coli genomic DNA in the same buffer. Thus, the final concentrations of reactants are 1.0 μM RecA, 38 mM (1.5 mM in nucleotide) ssDNA, 8.5 mM (5 μM in nucleotide) pKN106 DNA, and 500 μM (in nucleotide) nonhomologous E. coli DNA. After incubation at 37 °C for 2–3 h, 1 μl each of 2.0 mM S-adenosylmethionine and EcoRI methylase (40 units) were added to 20 μl of the synapsis mixture, and the incubation was continued for another 20 min. The reaction was stopped by extraction with phenol. The DNA was precipitated with ethanol twice, resuspended, and digested with EcoRI nuclease to completion. The nonhomologous E. coli DNA was included because it reduces the background caused by nonspecific binding of RecA to the pKN106 DNA. Without adding the nonhomologous DNA, intensities of the background bands (L1 and R4; see "Results" and Fig. 4) would become a few times higher than the other bands.

**Image Analyses**—Electrophoretic patterns of radioactive DNA fragments were digitized by using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Since this instrument creates a full-scale 16-bit image, the grayness of each image was scaled down to a more manageable 8-bit image, which was then exported to "PhotoShop" (Adobe) to correct any tilt of the image. Except for this modification, none of the digital filtering was performed. An inherent problem in accurately quantifying the bands on an agarose gel is their tendency to produce significant amounts of background and main peaks, resulting in uneven background throughout the lanes and overlapping peaks. Therefore, instead of encircling a band and taking the sum of all the pixel values within, the following method of quantification was used. The image was first imported into NIH Image software (written by Wayne Rasband at NIH). Then each lane was scanned, and the pixel densities were recorded. At this point, it was made certain that no pixel had reached the saturation or complete black. "PeakFit" (Jandel Scientific Software) was used to separate and quantify each peak in the scan. Since little is known about a theoretical shape of electrophoretic peaks of DNA fragments, the peaks were fitted into a general Exponentially Modified Gaussian function, which can deal with asymmetric peaks with minimal number of parameters. For further simplicity, the degree of asymmetry was fixed for all of the theoretical peaks in a single scan, so that the fundamental shapes of the peaks remained the same. The widths of the theoretical peaks were allowed to vary; this is appropriate, because faster migrating bands are broader. See Fig. 4b for an example of this type of quantification.

**Random Walk Simulation**—The simulation was done using a homemade program. The program first places one molecule of the PSC at a random position on the pKN106 DNA and then moves it either to the left or to the right with a 50:50 probability by a prespecified step length ("step length" here means the number of base pairs the PSC moves in a single movement). In addition, with a prespecified probability (p), the PSC also has an option of falling off from the DNA instead of moving. This fall off probability is related to the average sliding distance by the equation,

$$p = \frac{v}{d} \approx \frac{1}{2b} \cdot \left(1 - \frac{d}{s}ight)$$

where s is the average sliding distance in bp and d is the step length in bp. Note that the square of s/d corresponds to the number of steps it takes for the PSC to diffuse s bp on average. The simulation iterates indefinitely until the PSC falls off from the DNA or reaches one of the
homologous targets. In the latter case, the program assigns a certain probability by which the PSC may overlook the homology to continue random walk. Once the PSC thus overlooks a homology, however, it no longer “remembers” the fact that it just missed the homology. Eventually, the PSC settles down on one of the following fates: fall off (either from the end or the middle of DNA) or synapsis at one of the four targets. The program records this final outcome and then starts over another cycle of the virtual walk.

RESULTS

Experimental Idea—To test the idea of sliding during homology search, a plasmid (pKN106) was constructed that carries four 106-bp sequences in tandem array (Fig. 1a). The plasmid DNA was linearized by a restriction enzyme digestion and then used as the dsDNA substrate. The ssDNA substrate is a 40-mer synthetic oligonucleotide, called Oligo-1 or Oligo-2 (see “Experimental Procedures”). The two oligonucleotides are complementary to each other and homologous to the same 40-bp sequence located within each tandem array. The positions of these homologous regions are indicated by filled rectangles. Each homologous sequence contains a single EcoRI site (E) at the center as indicated. The numbers indicate the distances in bp between two adjacent restriction sites. b, experimental idea. Shown here is an example in which the PSC (stippled ball) binds to the 1,775-bp flanking sequence and starts sliding; for details, see text. The names of the homologous targets, Target-1 to Target-4, are indicated by T1–T4, respectively.

Test of System Using a Known Example—The test plasmid (pKN106) contains a single EcoRI site in the middle of each homologous target. This site serves two purposes: detecting the SSC (see below) and testing the system itself. Although evidence is still indirect, the EcoRI nuclease is believed to be able to slide over a significant distance (24, 25). If so, in the pKN106 DNA, one could expect the EcoRI cleavage to be more efficient at the two outermost targets (Target-1 and -4) than at two inner targets (Target-2 and -3).

The pKN106 DNA was linearized by ScaI, labeled with 32P at both ends, and digested by EcoRI. The digestion was carried out in the presence of 5–105 mM NaCl. The higher the salt concentration, the shorter the sliding distance should be (24, 25). The rate that sliding might accelerate is the “on” rate of EcoRI with the outermost EcoRI sites. However, the experiment does not measure this rate directly; it instead measures the rate of cleavage at each target site. It is therefore essential to set things up so that the possible difference in the on rates would result in different cleavage rates. To achieve this, the following steps were taken. 1) The level of EcoRI was titrated so as to produce a “single hit” cleavage (i.e. about a half of the input DNA molecules remained intact after the digestion). Otherwise, cleavage at multiple sites would complicate the measurement of independent cleavage rate at each site. 2) A stoichiometric excess of target sites over EcoRI molecules was used, ranging from ~55:1 (sites versus EcoRI dimer) at 5 mM NaCl to ~440:1 at 105 mM NaCl. This is to establish a steady state.

The meaning of these conditions is explained below. At steady state, cleavage rate at a target site i, $v_i$, should be proportional to the concentration of uncleaved EcoRI-DNA complex at site i, [ES]i, which is related to the K_m of EcoRI with respect to this site.

$$v_i = \frac{k_{cat} \times [ES] \times [E][S]}{K_m + [S]} \quad (Eq. 2)$$

where $k_{cat}$ is the first-order catalytic rate constant, and [E] and [S], are concentrations of free EcoRI and free site i, respectively. Because the four EcoRI sites exist at the same concentrations to begin with, [S], should also be about the same for each i, as long as each site is mostly unoccupied. A stoichiometric excess of target sites over EcoRI molecules ensures that each site is far from being saturated, as long as the cleavage is stopped early (i.e. condition of “single hit cleavage”). Under these conditions, the ratio of rates becomes equal to the ratio of association rate constants if the other constants (dissociation rate constant and catalytic rate constant) are independent on i.

Fig. 2 shows a list of the EcoRI fragments containing the 32P

![Figure 1](image1.png)  
**Fig. 1. Experimental strategy.** a, the dsDNA substrate. The plasmid pKN106, carrying four 106-bp tandem repeats (indicated by leftward arrows), was linearized by digestion with ScaI (S) and then used as the dsDNA substrate. The ssDNA substrate is a 40-mer synthetic oligonucleotide, called Oligo-1 or Oligo-2 (see “Experimental Procedures”). The two oligonucleotides are complementary to each other and homologous to the same 40-bp sequence located within each tandem array. The positions of these homologous regions are indicated by filled rectangles. Each homologous sequence contains a single EcoRI site (E) at the center as indicated. The numbers indicate the distances in bp between two adjacent restriction sites. b, experimental idea. Shown here is an example in which the PSC (stippled ball) binds to the 1,775-bp flanking sequence and starts sliding; for details, see text. The names of the homologous targets, Target-1 to Target-4, are indicated by T1–T4, respectively.

![Figure 2](image2.png)  
**Fig. 2. List of possible DNA fragments generated by single EcoRI cleavage.** The pKN106 DNA is linearized by ScaI (S) and labeled with 32P (asterisk) at both ends. Single cleavage of this DNA by EcoRI produces four pairs of DNA fragments; the fragments L1 and R1 correspond to the cleavage at Target-1 (T1), L2 and R2 correspond to the cleavage at Target-2 (T2), and so on. The lengths of the fragments are also indicated.
and if this region was close to any homologous target, these experiments should be able to signal the existence of such region(s). 3) Occupancy of a given target by the SSC, which is what experiments measure (see below), must be proportional to the “on” rate of the PSC with this target. Otherwise, one will not detect the sliding even if it occurs. To satisfy this condition, the experiments were set up to produce a low (usually <10%) fractional occupancy, so that there would be virtually no dsDNA molecule with two or more SSC molecules. Under these conditions, the SSC distribution among four potential targets should correspond to the relative encounter frequency between the PSC and each target, because the SSC formation is practically irreversible. This encounter frequency should be proportional to the association rate constant (also see “Discussion” for further details of the points argued here). 4) The PSC must not frequently pass through a homologous sequence. This point will be discussed later.

**Four Potential Targets Are Used Equally in the DNA Synapsis**—To determine the occupancy of each target by the SSC, a methylase protection assay (35) was carried out as described under “Experimental Procedures.” In brief, after the synapsis reaction, the sample containing the SSC was treated with the EcoRI methylase. Since each homologous target contains an EcoRI site, formation of the SSC would prevent the site within the complex from methylation, whereas the other three sites on the same DNA molecule would be methylated. After destroying the SSC, the methylase-treated DNA was digested with the EcoRI nuclease. Because the EcoRI nuclease does not cut a methylated site, cleavage would be limited to the unmethylated site, which in turn corresponds to the original site of synapsis.

After the final cleavage, the DNA samples were fractionated by agarose gel electrophoresis, and the pattern was visualized by autoradiography. Without methylase treatment (Fig. 4a, lane 1), EcoRI should cleave all the EcoRI sites, producing only L1 and R4 on the autoradiogram, which is exactly what is seen in this lane. Without RecA (lane 2), all of the EcoRI sites should become fully methylated; indeed, there was no cleavage by EcoRI. The EcoRI methylase (36, 37), as well as EcoRI endonuclease (24, 25), appears to be able to slide along DNA. However, the two controls just described (lanes 1 and 2) show that these enzymes nonetheless can act upon all of the target sites under the conditions used. Without any ssDNA substrate (lane 3), there should be no SSC formed, and all of the EcoRI sites should become methylated. Indeed, the majority of the dsDNA molecules were resistant to the EcoRI cleavage; however, a small fraction of the dsDNA produced L1 and R4. A possible explanation for these bands is to assume the formation of a “patch” of nonspecifically bound RecA molecules covering all of the EcoRI sites. In this case, all of these sites would be cleaved later by EcoRI, producing only L1 and R4 in the autoradiogram. Because of its highly “cooperative” binding, RecA in the presence of excess DNA tends to cluster in some DNA molecules while leaving other DNA molecules empty. Consistent with this interpretation, intensities of L1 and R4 were greatly reduced by the addition of a nonhomologous oligonucleotide (lane 4). Here, presumably, most RecA molecules were absorbed by the ssDNA, to which RecA is known to bind much more rapidly than to dsDNA (38).

**Lanes 5 and 6 of Fig. 4a** show the results with homologous ssDNA substrates Oligo-1 and Oligo-2. To quantify the band intensities, the same pattern was exposed to a PhosphorImager screen; the resulting intensity scan of lane 5 is shown in Fig. 4b as an example (open circles). A curve fit software was used to separate these peaks, and the outcome of separation is included in Fig. 4b (thin solid lines). The histogram of area integration of these peaks is shown in Fig. 5a.
There was no difference between the histogram obtained with Oligo-1 (Fig. 5a, filled bars) and that with Oligo-2 (which is complementary to Oligo-1; open bars), indicating that polarities of PSC relative to the dsDNA substrate do not affect the results. Recall that the sliding model predicted a preferential production of L4 and R1 relative to L2 and R2 and R3, respectively. Clearly, there was no such positional preference. The bands L1 and R4 were significantly more intense than other bands. This, however, is most likely because of the formation of a RecA patch covering multiple EcoRI sites. Such a cluster could form around the site of synapsis if free RecA molecules polymerized onto the ends of SSC.

The salt condition used in the analyses described above and in the rest of this article was 25 mM KOAc plus 5 mM magnesium acetate. The results were indistinguishable at 100 mM KOAc (data not shown). Fig. 4c compares the result at 25 mM KOAc with that in the absence of any KOAc. The band intensities were quantified, and the resulting histograms are shown in Fig. 5b. Thus, even at 0 mM KOAc (Fig. 5b, open bars), there was no preferential production of L4 or R1. The condition without KOAc was not used extensively in this study, because this condition tends to produce a much higher background (L1 and R4) presumably because of increased nonspecific binding of RecA to dsDNA.

The kinetics of SSC formation appears to be very similar for all four targets (data not shown). The formation of SSC levels off usually after about 1.5–2 h of incubation under the conditions used. The samples for the quantification summarized in Fig. 5 were taken at the 2–3-h time point. The samples taken before the level off (at the 1-h time point) also produced qualitatively indistinguishable patterns; just the intensities were fainter.

No Detrimental Effect by Removing a Flanking Segment—If the sliding is the major search mechanism, the PSC must approach a homologous target through a flanking DNA segment. Therefore, in parallel to the experiments described above, additional experiments were carried out in which one of the DNA segments flanking the target sequences was removed. The sliding model predicts that the removal would poison the target efficiency of the homologous sequence to which the removed segment used to be connected. As shown below, however, the removal had no effect on the target efficiency. In one set of experiments, the ScaI-linearized pKN106 DNA carrying the 32P labels at both ends was cleaved further by a restriction enzyme EcoRI and then used as the dsDNA substrate. As shown in Fig. 6a, the EcoRI cuts the DNA at a position 161 bp away from the center (EcoRI site) of Target-1, eliminating most of the flanking segment originally located on the left of Target-1. The experimental conditions were as described above, and the results are shown in Fig. 6, b and c. Note that band T1, which indicates the synapsis at the target (Target-1) that is closest to the removed flanking segment, is as intense as bands T2 and T3, each of which corresponds to the synapsis at each inner target. Compare this cleavage pattern (Fig. 6c) with the pattern obtained earlier with the ScaI-cleaved substrate...
The main conclusion of this study is that, during a search of homology, sliding of the PSC along a dsDNA molecule does not occur to any significant extent. Because of the way experiments were set up, “significant extent” here means a few hundred bp or longer. This conclusion is based on the finding that four homologous targets placed in tandem in the test substrate (pKN106 DNA) acted as equally efficient targets in the synapsis reaction. In addition, removal of either of the two flanking sequences did not reduce the target efficiency of the homologous targets placed in tandem in the test substrate (pKN106 DNA) as the ssDNA substrate. The pKN106 DNA was cleaved by ScaI (S), labeled with 32P at the ends (asterisks), and cleaved further by Ehel (Eh). The distance between the Ehel site and the EcoRI (E) site closest to the end is 161 bp as indicated. The smaller ScaI–Ehel fragment that does not contain a homologous target was also present in the same reaction mixture. b, the samples were analyzed by electrophoresis on an 0.8% HGT-agarose gel, followed by autoradiography. Each lane corresponds to the same reaction condition as each lane in Fig. 4a, c, summary of the analyses obtained with Oligo-1 (filled bars), Oligo-2 (open bars), and a nonhomologous 40-mer oligonucleotide (stippled bars) as the ssDNA substrates, respectively. Error bars show S.D. values.

DISCUSSION

The main conclusion of this study is that, during a search of homology, sliding of the PSC along a dsDNA molecule does not occur to any significant extent. Because of the way experiments were set up, “significant extent” here means a few hundred bp or longer. This conclusion is based on the finding that four homologous targets placed in tandem in the test substrate (pKN106 DNA) acted as equally efficient targets in the synapsis reaction. In addition, removal of either of the two flanking sequences did not reduce the target efficiency of the homologous sequence that used to be connected to the removed flanking segment.

The Sliding of the EcoRI Nuclease—The lack of sliding of the PSC is further supported by the finding that the same experimental system can signal the existence of sliding of the EcoRI nuclease, an enzyme that has been shown to slide along DNA (24, 25). At a low salt condition, EcoRI cleaved the pKN106 DNA preferentially at the two outermost targets, and this preferential cleavage disappeared at a high salt condition where the sliding distance is expected to become negligible.

The sliding distance of EcoRI has been reported to be at least several hundred bp under a catalytic condition (25) or even ~1,300 bp under a noncatalytic condition (24). Considering these distances, the level of preferential cleavage observed in this study (~40%) might appear too small. A possible interpretation of this level of preference might be that the average sliding distance is just ~150 bp (~106 bp × 1.4). However, Terry et al. (25) have shown that the EcoRI can cut the DNA “processively”; i.e. after cutting one site, the nuclease may proceed to cut a nearby site located in cis without leaving the DNA. The idea of processive cleavage can indeed explain why the partial digestion of the pKN106 DNA still predominantly produced the bands of complete cleavage (L1 and R4; see Fig. 3). A possible sliding distance mentioned above (~150 bp) is based on the intensities of two particular bands, R1 and L4, relative to R2 and -3 and L2 and -3, respectively. If the cleavage by EcoRI was very processive, however, it must have greatly reduced the relative intensities of R1 and L4, because the production of these bands strictly requires the DNA to be cut only once (see Fig. 2). Therefore, the ~150-bp estimate most probably represents just the lower end of possible sliding distance of EcoRI.

Nonideal Behavior of the DNA Synapsis Reaction—To analyze a complex reaction such as DNA synapsis, it is important to consider what is not being measured: the side reactions that are most likely occurring simultaneously with the “real” reaction. In the synapsis reaction, these side reactions may include any combination of the following: self-aggregation of RecA (39, 40), binding of free RecA to dsDNA (38), side-by-side bundle formation of RecA-DNA filaments (41), and the synapsis at pseudohomologous sites. These side reactions are probably irreversible in the presence of ATP and result in the inhibition of the synapsis reaction, by effectively taking away the reactants (RecA and DNA).

The existence of side reactions at least partly explains why the extent of synapsis is usually much lower than what is expected to be at equilibrium; i.e. because the “off” rate of the SSC is negligible, the extent of synapsis in principle should reach near 100%. In reality, it reaches usually only up to ~50% with respect to the dsDNA substrate, even when there are twice as many ssDNA molecules as dsDNA substrates (for example, see Ref. 42). Furthermore, the extent of synapsis appears to be controlled kinetically, i.e. determined primarily by how quickly the “real” reaction occurs relative to “unwanted” reactions. Under these conditions, the reaction outcome becomes very sensitive to initial conditions, since even a slight
change of condition would affect not just the equilibrium but also the partition of reactants among competing reactions. Consequently, a condition that decreases the rate of synopsis tends to decrease its extent as well rather than simply result in a longer time to reach the same extent, because more reactants participate in side reactions in this case. For the same reason, “stimulation” of the DNA synopsis can be achieved by counteracting the inhibitory effects of “unwanted” reactions, although the biological significance of such stimulation should be interpreted carefully.

Rationale for Experimental Design—A traditional method for testing the sliding model is to embed a target in nonspecific DNA molecules of various lengths and then measure the rates of binding with these DNA molecules. If the sliding is operative, binding to the DNA molecules whose lengths are significantly shorter than the sliding distance should be slower than the binding to longer DNA molecules (20). This strategy can be straightforward so long as one is measuring the rate of a simple binding reaction. However, the situation becomes complicated if the rate being measured is a composite of the initial association rate (on rate) and the rates of various first-order steps (such as a chemical step) that follow it. In this case, one must set up the experiment so that the measured rate becomes proportional to the on rate; otherwise, the results could become misleading. For example, when a later step was limiting the measured rate, effects of sliding would be masked. A problem, however, is that it is not easy to determine where the rate-limiting step is in a complex reaction. Moreover, in the synopsis reaction, interference from side reactions has made it difficult for us to obtain the kinetic data whose quality warrants vigorous postmeasurement analyses needed for the determination of rate-limiting step(s).

The experimental strategy used in this study was a compromise. The strategy was originally inspired by related studies on the E. coli RNA polymerase (27) and the Lac repressor (21) and aims at deducing a qualitative conclusion as to the existence of sliding. The principle of the experiment is straightforward; first let four potential targets arrayed in tandem compete for a single PSC molecule, and then record which target won and by what degree. Since the experiment focuses on relative efficiencies of the targets present in the same test tube, interference from side reactions should not be a crucial problem.

What is crucial, however, is to ensure that any potential change of the on rates would show up as the corresponding difference in target occupancy, which is what experiments measure. To achieve this correlation, the reaction was set up so that the fractional occupancy was far below saturation when the reaction reached a plateau. The SSC detected under these conditions should correspond to essentially the first and only reaction reached a plateau. The SSC detected under these conditions should correspond to essentially the first and only reaction reached a plateau.

Might the PSC Pass through a Homologous Target?—In the foregoing arguments, a possibility that the PSC might somehow pass through a homologous target has been set aside. It was assumed that the PSC would become the SSC at the first target it encountered. If this assumption is valid, then the even distribution of the SSC seen in Fig. 5 speaks against the sliding model. However, if the PSC can “overlook” a homology, the same data might still be consistent with the sliding model. In this case, the PSC might pass through a homologous target(s) to become equilibrated among all of the targets.

To explore this issue more quantitatively, a computer simulation was carried out to ask the following question. Let us assume that the PSC actually slides and that the even distribution of the SSC is a result of the PSC having passed through the targets. How frequently does the PSC need to overlook a homology in order for the SSC to become more or less equally distributed among the targets? It turned out that the PSC must overlook a homology most of the time it encounters a homologous target. An example is shown in Fig. 8. Here the sliding distance was set to be 400 bp. Under this condition, the probability for overlooking a homology needs to be as high as ~99% to produce a more or less even distribution of the SSC (Fig. 8, filled bars). The probability of 90% is clearly not enough to produce an even distribution (stippled bars). These results may seem strange, but they are a consequence of the fact that a sliding particle tends to go over the same point back and forth many times before it wanders away for good. Because of this redundancy, overlooking a homology, for example just 50% of the time, would not significantly affect the efficiency of target capture.

In brief, if the sliding were to be involved in homology search at all, the test of homology during sliding would have to be very “sporadic,” such as only once during −100 or more sliding steps. Otherwise, there would be little chance for the PSC to pass through the target it meets first. Since the PSC must recognize a homology in the first place, a very high level of overlooking a homology poses a paradox and thus seems counterintuitive. Nevertheless, the question of how frequently the PSC might overlook a homology requires further study to answer definitively. In the case of the EcoRI nuclease, however, the enzyme appears to be acting on the site it encounters first during sliding (34).

Is Homology Search Facilitated?—Is it necessary to consider a facilitated diffusion mechanism other than sliding, such as “intersegmental transfer” (for a review, see Ref. 18), in the DNA synopsis reaction? Alternatively, can a simple random collision mechanism explain the available data at least for the time being? To answer this question, one must look at the rate of synopsis and compare this to the theoretical upper limit of association rate based entirely on a three-dimensional collision. If the actual rate exceeds this theoretical limit, the existence of some kind of facilitated mechanism can be inferred.

Using a restriction enzyme protection assay, Yancey-Wrona

---

2 K. Adzuma, unpublished observation. Also see Refs. 43–45 for similar observations.
and Camerini-Otero (46) measured the association rate constant of the synopsis reaction between a 27-mer oligonucleotide coated with RecA and the homologous dsDNA molecule. With a long (~2,700-bp) plasmid DNA as the dsDNA target, the authors found the rate constant to be on the order of $10^4$ to $10^6$ M$^{-1}$ s$^{-1}$. The rate constants of these magnitudes are too small to be considered diffusion-limited. One can estimate the upper limit of a bimolecular association rate constant using the so-called Smoluchowski’s equation (for an example of such a calculation, see Ref. 47). Under the condition used by Yancey-Wrona and Camerini-Otero (46), our crude estimate for the upper limit would be on the order of $10^7$ M$^{-1}$ s$^{-1}$ (parameters used are $5 \times 10^{-7}$ cm$^2$ s$^{-1}$ for diffusion constant of the 27-mer PSC; 80 Å for interaction radius; we also assumed that only 1% of the total collision would be productive). This estimate, although very crude, is still 3–4 orders of magnitude larger than the value reported by Yancey-Wrona and Camerini-Otero (46). Thus, at least when the length of homology is short (tens of bp), the speed of DNA synopsis appears to be much slower than the diffusion limit of the system involving a rapid equilibrium such as, slow binding reaction.

In general, the simplest model consistent with a slow second-order kinetics is to assume some kind of two-step mechanism involving a rapid equilibrium such as,

![Scheme I](image)

or

![Scheme II](image)

In either scheme, the rate of the complex (C) formation will be as follows,

$$
\frac{d[C]}{dt} = \frac{k}{K_a}[A][B] 
$$

(Eq. 3)

where [A] and [B] are free concentrations of A and B. In terms of total concentrations, the equation will become a complicated one, unless the concentration of one reactant is much greater than that of the other. When the $K_a$ is very large compared with total concentrations of A and B, however, the concentration of AB (Scheme I) or $A^*$ (Scheme II) becomes negligible, and free concentrations of A and B can approximate their total concentrations. Only under these conditions, the overall rate will exhibit a pseudo-second-order kinetics with the $k/K_a$ as the apparent association rate constant, which can be smaller than the real association rate constant. Thermal annealing of two short ssDNA molecules is an example of this kind of slow second-order reaction (50).

Other kinetics data published a long time ago (28–30), in which the length of homology was very large (ranging from 975 to 8,600 bp), were also reanalyzed. Those reactions appear to proceed rapidly, with the apparent association rate constant on the order of greater than $10^5$ to $10^8$ M$^{-1}$ s$^{-1}$ (note that these numbers are probably underestimate, because the reactions were already saturated with at least one reactant). Therefore, when dsDNA and dsDNA substrates share more than 1 kilobase of homology, it is possible that the rate of homology search might exceed the diffusion limit. A very long RecA-ssDNA filament can interact with multiple dsDNA segments simultaneously, and this might explain the rate enhancement. The enhancement of this type would ultimately be limited by segmental diffusion of DNA (18).

Acknowledgments—I am grateful to Drs. K. Mizuuchi (National Institutes of Health), P. Modrich (Duke University), and J. F. Feng (Rockefeller University) for comments on the manuscript and to Dr. Qing Sheng (Credit Suisse First Boston) for stimulating discussion about random walk.

**REFERENCES**

1. Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T. (1993) *Nat. Genet.* 4, 239–243
2. Yoshimura, Y., Morita, T., Yamamoto, A., and Matsumoto, A. (1993) *Nucleic Acids Res.* 21, 1665
3. Cox, M. M. (1993) *BioEssays* 15, 617–623
4. Roca, A., and Cox, M. M. (1997) *Prog. Nucleic Acids Res. Mol. Biol.* 56, 129–223
5. Scully, R., Chen, J., Plag, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T., and Livengood, D. M. (1997) *Cell* 88, 265–275
6. Sharan, S. K., Morimoto, M., Albrecht, U., Lim, D. S., Regel, E., Dinh, C., Sand, A., Eichele, G., Hasty, P., and Bradley, A. (1997) *Nature* 386, 804–810
7. Baumann, P., Benson, F. E., and West, S. C. (1996) *Cell* 87, 757–766
8. Gupta, R. C., Bazemore, L. R., Golub, E. I., and Radding, C. M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 463–468
9. Sunagawa, S., and Inoue, T. (1998) *Science* 281, 1243–1248
10. Wrona and Camerini-Otero (46), our crude estimate for the kinetic diffusion of DNA (18).

**REFERENCES**

1. Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T. (1993) *Nat. Genet.* 4, 239–243
2. Yoshimura, Y., Morita, T., Yamamoto, A., and Matsumoto, A. (1993) *Nucleic Acids Res.* 21, 1665
3. Cox, M. M. (1993) *BioEssays* 15, 617–623
4. Roca, A., and Cox, M. M. (1997) *Prog. Nucleic Acids Res. Mol. Biol.* 56, 129–223
5. Scully, R., Chen, J., Plag, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T., and Livengood, D. M. (1997) *Cell* 88, 265–275
6. Sharan, S. K., Morimoto, M., Albrecht, U., Lim, D. S., Regel, E., Dinh, C., Sand, A., Eichele, G., Hasty, P., and Bradley, A. (1997) *Nature* 386, 804–810
7. Baumann, P., Benson, F. E., and West, S. C. (1996) *Cell* 87, 757–766
8. Gupta, R. C., Bazemore, L. R., Golub, E. I., and Radding, C. M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 463–468
9. Sunagawa, S., and Inoue, T. (1998) *Science* 281, 1243–1248
10. Wrona and Camerini-Otero (46), our crude estimate for the kinetic diffusion of DNA (18).

**REFERENCES**

1. Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T. (1993) *Nat. Genet.* 4, 239–243
2. Yoshimura, Y., Morita, T., Yamamoto, A., and Matsumoto, A. (1993) *Nucleic Acids Res.* 21, 1665
3. Cox, M. M. (1993) *BioEssays* 15, 617–623
4. Roca, A., and Cox, M. M. (1997) *Prog. Nucleic Acids Res. Mol. Biol.* 56, 129–223
5. Scully, R., Chen, J., Plag, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T., and Livengood, D. M. (1997) *Cell* 88, 265–275
6. Sharan, S. K., Morimoto, M., Albrecht, U., Lim, D. S., Regel, E., Dinh, C., Sand, A., Eichele, G., Hasty, P., and Bradley, A. (1997) *Nature* 386, 804–810
7. Baumann, P., Benson, F. E., and West, S. C. (1996) *Cell* 87, 757–766
8. Gupta, R. C., Bazemore, L. R., Golub, E. I., and Radding, C. M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 463–468
9. Sunagawa, S., and Inoue, T. (1998) *Science* 281, 1243–1248
10. Wrona and Camerini-Otero (46), our crude estimate for the kinetic diffusion of DNA (18).
959–972
40. Morrical, S. W., and Cox, M. M. (1985) Biochemistry 24, 760–767
41. Egelman, E. H., and Stasiak, A. (1988) J. Mol. Biol. 200, 329–349
42. Adzuma, K. (1992) Genes Dev. 6, 1679–1694
43. Baliga, R., Singleton, J. W., and Dervan, P. B. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 10393–10397
44. Golub, E. I., Ward, D. C., and Radding, C. M. (1992) Nucleic Acids Res. 20, 3121–3125
45. Podyminogin, M. A., Meyer, R. B., and Gamper, H. B. (1995) Biochemistry 34, 13098–13108
46. Yancey-Wrona, J. E., and Camerini-Otero, R. D. (1995) Curr. Biol. 5, 1149–1158
47. von Hippel, P. H., and Berg, O. G. (1989) J. Biol. Chem. 264, 675–678
48. Bazemore, L. R., Takahashi, M., and Radding, C. M. (1997) J. Biol. Chem. 272, 14672–14682
49. Lapham, J., Rife, J. P., Moore, P. B., and Crothers, D. M. (1997) J. Biomol. NMR 10, 255–262
50. Cantor, C. R., and Schimmel, P. R. (1980) Biophysical Chemistry Part III: The Behavior of Biological Macromolecules, pp. 1183–1264, W. H. Freeman and Co., New York