Polarity of heteroduplex formation promoted by *Escherichia coli* recA protein*

*(genetic recombination/homologous pairing/strand transfer/joint molecules)*

ROGER KAHN, RICHARD P. CUNNINGHAM, CHANCHAL DASGUPTA, AND CHARLES M. RADDING

The Departments of Human Genetics and Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Communicated by Aaron B. Lerner, April 24, 1981

**ABSTRACT** When recA protein pairs circular single strands with linear duplex DNA, the circular strand displaces its homolog from only one end of the duplex molecule and rapidly creates heteroduplex joints that are thousands of base pairs long [DasGupta, C., Shibata, T., Cunningham, R. P. & Radding, C. M. (1980) Cell 22, 437-446]. To examine this apparently polar reaction, we prepared chimeric duplex fragments of DNA that had M13 nucleotide sequences at one end and G4 sequences at the other. Circular single strands homologous to M13 DNA paired with a chimeric fragment when M13 sequences were located at the 3' end of the complementary strand but did not pair when the M13 sequences were located at the 5' end. Likewise circular single-stranded G4 DNA paired with chimeric fragments only when G4 sequences were located at the 3' end of the complementary strand. To confirm these observations, we prepared fd DNA labeled only at the 5' or 3' end of the plus strand, and we examined the susceptibility of these labeled ends to digestion by exonucleases when joint molecules were formed. Eighty percent of the 5' label in joint molecules became sensitive to exonuclease VII. Displacement of that 5' end by recA protein was concerted because it did not occur in the absence of single-stranded DNA or in the presence of heterologous single strands. By contrast, only a small fraction of the 3' label became sensitive to exonuclease VII or exonuclease I. These observations show that recA protein forms heteroduplex joints in a concerted and polarized way.

The product of the recA gene of *Escherichia coli* is a DNA-dependent ATPase that regulates a pathway of inducible repair by means of a specific protease activity (see ref. 1 for a review) and that also has a remarkably versatile ability to promote the homologous pairing of DNA molecules. With a few exceptions, recA protein can stably pair all manner of structural variants of DNA molecules: complementary single strands with each other, single strands with duplex DNA, and partially single-stranded DNA with duplex DNA (2-9). The rule that governs stable pairing appears to be that one molecule must be at least partially single-stranded and that a free end must exist somewhere in one of the two molecules (4, 8). We have experimentally distinguished two steps in the formation of stable joint molecules by recA protein: (i) synopsis, which appears to require single-stranded DNA but does not require a free end, and (ii) strand transfer, which requires a homologous free end and produces classical heteroduplex joints in which a strand from each parental molecule is helically intertwined with its complement (8, 10).

From an experimental point of view, one of the most interesting pairings promoted by recA protein is that of circular single strands with linear duplex DNA. This reaction is rapid and efficient; its substrates and products are well defined. By electron microscopy, we observed that the product of this reaction is a joint molecule in which the circular single strand has displaced its homolog, starting from one end of the duplex molecule (ref. 8 and this paper). In 15-30 min, some heteroduplex joints become thousands of nucleotides long (8). The rapid creation of these long heteroduplex joints, and the absence of molecules in which circular single strands invaded the duplex DNA from both ends (unpublished data) led us to infer that recA protein actively drives the formation of heteroduplex DNA in one direction. The experiments described in this paper were designed to examine this apparent polarity.

**METHODS**

**Enzymes.** *E. coli* exonuclease I was purified by John Williams (11) according to the procedure of Lehman and Nussbaum (12). *E. coli* exonuclease VII was a gift of John Chase (13); recA protein (3, 14) and 32P exonuclease (15) were purified as described. Bacterial alkaline phosphatase was a gift of Jan Clevelovsky. Polynucleotide kinase was purchased from Boehringer. *E. coli* polymerase I (large fragment) and exonuclease Sau 96 I were purchased from New England Biolabs. Endonuclease Hpa I was purchased from New England Biolabs and Bethesda Research Laboratories (Rockville, MD). Endonuclease Bam HI was purchased from Bethesda Research Laboratories.

**Preparation of DNA.** Single-stranded and replicative form DNA from phages fd, qX174, and G4 were prepared as described (7, 16). Superhelical DNA from M13 GoriI was prepared according to the procedures used for fd DNA (7).

**Duplex DNA 32P-Labeled at the 5' End of the Plus Strand.** We cut form I fd *[3H]DNA* at a single site by digesting it with *Hpa* I and labeled the 5' ends of this DNA by treating it with bacterial alkaline phosphatase following by polynucleotide kinase essentially as described by Maxam and Gilbert (17). To verify that the 32P label was present only at the ends, we cleaved an aliquot of this DNA with BamHI which cuts *Hpa* I-linearized fd DNA into three fragments: two end fragments and an internal fragment 3425 base pairs long. The internal fragment, separated by gel electrophoresis, contained less than 1% of the total 32P recovered in the three fragments.

The linear DNA with 32P label at both ends was cleaved into two fragments by *Sau* 96 I: a large fragment (5725 base pairs) labeled at the 5' end of the plus strand and a small one (683 base pairs) labeled at the 5' end of the minus strand. The fragments were separated by centrifugation through a 5-20% gradient of sucrose. According to the approximate specific activity of the *[γ-32P]ATP* used to label the ends, we estimated that 30% of 5' ends of plus strands were phosphorylated.

**Duplex DNA 32P-Labeled at the 3' End of the Plus Strand.** This DNA, which also consisted of the large fragment produced by cutting fd form I *[3H]DNA* with both *Sau* 96 I and *Hpa* I,
was labeled by a method to be described elsewhere. Other methods were as described (7, 8, 10, 14).

RESULTS

Time Course of Formation of Joint Molecules from Circular Single-Stranded DNA and Linear Duplex DNA by recA Protein. In a reaction mixture containing about twice as many circular single strands as linear duplex molecules, 80% of the duplex DNA appeared in joint molecules in 10 min (Figs. 1 and 2). When we heated deproteinized samples for 4 min at 41°C prior to assaying joint molecules, we found nearly as much product as in the unheated samples (Fig. 2). At 50°C there was a significant loss of joint molecules, but the fraction of joint molecules that survived heating at 50°C increased from 27% at 5 min to 91% at 30 min. This increase in thermal stability of the products of the reaction as a function of the duration of the reaction is consistent with a progressive increase in the average length of heteroduplex joints.

As in the formation of joint molecules by single-stranded fragments and duplex DNA (21), the kinetics of formation of the joint molecules described here resemble Michaelis–Menten kinetics (unpublished data), which suggests that the mechanism of forming joint molecules from single-stranded circles and linear duplex DNA is related to the mechanism of forming D loops (3, 21, 22).

Specific Pairing of Circular Single-Stranded DNA with the 3′ End of Its Complementary Strand. Circular single-stranded DNA isolated from phage particles consists of viral or plus strands that can pair only with the minus strand of the duplex DNA. Therefore, if the growth of heteroduplex joints is unidirectional, the single-stranded DNA should pair with only one end of a duplex fragment and should displace a specific end of the plus strand from the duplex DNA. To look for pairing with a specific end of the duplex DNA we used Hpa I to cleave M13GorI1 DNA into two fragments that had M13 DNA at one end and G4 DNA at the other (Fig. 3). Hpa I cuts the two strands at the same position, creating blunt ends without single-stranded extensions (27). The large fragment, which was 6783 base pairs long, had sequences from M13 at the 3′ end of the complementary strand, whereas the small fragment, which was 2040 base pairs long, had sequences from G4 at the 3′ end, a conclusion based on the published nucleotide sequences (24–26) and on the orientation of the inserted G4 fragment in M13GorI1 (23). The large difference in size of the two fragments made it easy to identify them by electron microscopy (Fig. 1A).

In one tube we mixed the preparation of M13GorI1 DNA containing both Hpa I fragments with circular single-stranded G4 DNA, and in another tube we mixed them with circular sin-
ingle-stranded fd DNA, which is 97% homologous to M13 DNA (26). Fd DNA paired only with the large fragments, whereas the G4 DNA paired only with the small fragments (Fig. 1; Table 1). The yield of joint molecules in this experiment was low, possibly because we did not repurify the fragments of duplex DNA after we prepared them by restriction cleavage of form I DNA. This experiment, however, was controlled internally: large fragments and small fragments were present together in the reaction mixtures, and circular single-stranded fd DNA made as many joint molecules with the large fragment as G4 DNA did with the small fragment.

The specific pairing of fd DNA with the large fragment and of G4 DNA with the small fragment indicates that the formation of heteroduplex joints is polarized and shows that the incoming strand pairs only with the 3' end of the complementary strand.

**Concerted Displacement of the 5' End of Plus Strands from Joint Molecules.** The above observations imply that the formation of a joint molecule results in specific displacement of the 5' end of the plus strand. To test this inference, we prepared two duplex substrates, one labeled with 32P at the 5' end of the plus strand and the other labeled at the 3' end of the plus strand. To prepare the 5'-labeled substrates, we cleaved form I fd DNA at a single site with Hpa I and phosphorylated the 5' ends by using [γ-32P]ATP. We removed the labeled 5' end of the minus strand by digestion with Sau 96I, which cut 683 base pairs from the appropriate end. The rest of the molecule, consisting of 5725 base pairs labeled at the 5' end of the plus strand, was resolated by sedimentation in sucrose.

The pairing of this substrate with circular single-stranded fd DNA produced material that sedimented faster than the linear DNA in a discrete peak with a long leading shoulder (Fig. 4A). The faster sedimenting material did not appear when heterologous DNA replaced single-stranded fd DNA (Fig. 4C), when recA protein was omitted (Fig. 4D), or when single-stranded DNA was omitted (data not shown). To look for displacement of the 5'-labeled end, we used E. coli exonuclease VII, which acts specifically on single-stranded DNA at both 5' and 3' ends (13). When the product of the complete reaction was treated with exonuclease VII for 10 min, most of the 32P label was cleaved from the material in the faster sedimenting peak and appeared near the top of the gradient (Fig. 4B). Some 32P was also removed from the slower sedimenting material. None of the 32P label was susceptible to exonuclease VII when joint molecules were not formed because heterologous single strands were used (Fig. 4C) or recA protein was omitted (Fig. 4D) or single-stranded DNA was omitted (data not shown). We conclude that the formation of joint molecules made the 5' end of the viral strand susceptible to exonuclease VII.

**Lack of Displacement of the 3' End of the Plus Strand.** A substrate labeled specifically at the 3' end of the plus strand was
prepared by a procedure similar to that used to make the 5' labeled substrate. We first cleaved fd form I DNA with Ssv 96 I, which makes a single staggered cut, leaving 5' extensions three nucleotides long. The 3' ends were then labeled by using E. coli polymerase I and [α-32P]dCTP plus unlabelled deoxynucleoside triphosphates. Hpa I removed the unwanted label at the 3' end of the minus strand by cutting 683 base pairs from the appropriate end. The large fragment was purified by sedimentation in sucrose.

Label at the 3' end of the plus strand was resistant to removal by exonuclease VII whether joint molecules were formed or not (Fig. 5 A–C). As a positive control (Fig. 5D), we partially digested the 3'-labeled substrate with λ exonuclease to make single-stranded 3' ends whose average length was 560 nucleotides; most of the 3' label in this material was removed by exonuclease VII.

The peak of more rapidly sedimenting material shown in Fig. 5D is attributable to joint molecules that recA protein makes from DNA with single-stranded tails, a reaction that we will describe elsewhere.

To confirm that recA protein does not displace the 3' end, we repeated the same experiment with exonuclease I, which degrades single-stranded DNA from its 3' end (12). We used enough exonuclease I to make the label at single-stranded 3' ends completely acid soluble (Table 2, reaction d). In the complete reaction containing the 3' labeled duplex substrate, homologous circular single-stranded DNA, recA protein, and exonuclease I, we observed that, after subtraction of the blank, 12% of the 3' ends became acid soluble. On the basis of the complete digestion of an authentic single-stranded 3' end in the

Table 2. Lack of displacement of the 3' end of the complementary strand

<table>
<thead>
<tr>
<th>Reaction</th>
<th>32P (a)/32P (b)</th>
<th>32P (a)/32P (b)</th>
<th>Acid-soluble, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Complete</td>
<td>75</td>
<td>78</td>
<td>2</td>
</tr>
<tr>
<td>b. Heterologous ssDNA</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>c. No ssDNA</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>d. 3'-Tailed DNA</td>
<td>—</td>
<td>—</td>
<td>25</td>
</tr>
</tbody>
</table>

This experiment was identical to that in Fig. 5 except that, instead of exonuclease VII, exonuclease I at 230 units/ml was added at 15 min, with no other reagents. We measured the fraction of joint molecules formed at 15 min (reaction a) and the amount of label made acid soluble by exonuclease I in the ensuing 6 min. No detergent or heat treatment was used in assay D. Heterologous DNA was circular single-stranded φX174 DNA. Tailed DNA was the substrate described in Fig. 5D; it had 3' single-stranded tails 560 nucleotides long. This reaction mixture contained recA protein and heterologous single-stranded circular DNA plus all the other reagents present in reactions a–c.

positive control (Table 2, reaction d), we conclude that the 3' end of the plus strand was not displaced in most joint molecules.

DISCUSSION

The pairing of circular single-stranded DNA with linear duplex DNA by recA protein is a rapid and efficient reaction that produces long heteroduplex joints (Figs. 1 and 2; ref. 8). The increase in stability of these joint molecules as a function of time is consistent with a progressive growth in the average size of heteroduplex joints. The formation of these heteroduplex joints is polar (Fig. 1; Table 1); there is a strong bias that favors the pairing of the circular plus strand with the 3' end of its complementary strand. In the sample of 41 joint molecules observed by electron microscopy (Table 1), no joints were found at the 5' end of the complementary strand, which indicates that the preference for pairing at the 3' end is greater than 40:1. By preparing specific terminally labeled substrates, we showed that only the 5' end of the plus strand is displaced from the duplex DNA in a concerted reaction that occurs only when homologous single-stranded DNA is present. The latter data support our previous inference on the concerted nature of strand transfer by recA protein (7, 8, 21) and agree with other observations (ref. 28; unpublished data).

recA protein forms heteroduplex joints at the 3' end of the complementary strand of linear duplex DNA; 15–30 min later, heteroduplex joints have grown to thousands of base pairs in length. How does the joint grow? recA protein might specifically recognize the 3' end of the complementary strand, initiate the formation of a joint, and either block its dissociation or rapidly promote reassociation. According to the formulation of Thompson et al. (29), in 30 min a random one-dimensional walk starting from such a reflecting barrier could result in joints that are hundreds to thousands of nucleotides long, whether the jump time per base pair is 170 μsec (29) or 12 μsec (30). Alternatively, recA protein might form the joint and actively enlarge it in one direction—for example, by binding or acting at the site of strand exchange in a polar way.

In order to make a stable joint molecule, recA protein must put molecules in homologous alignment, and it must find the free end of a DNA chain from which it can initiate a strand transfer. It might find the free end first by specifically binding to ends and then searching for homology, or it might find homol-
ogy first and then look for an end. Some observations favor the latter explanation: (i) recA protein can find homology in the absence of any end (8, 10), and (ii) recA protein can use any physical variant of an end to initiate a strand transfer—e.g., the end of a single strand, the end of a duplex molecule, or the ends present at nicks and gaps (7, 8). Because these observations regale the physical nature of an end to secondary importance, they also favor the view that the polar formation of heteroduplex joints reported here does not result from the binding of recA protein to a specific end but rather results from a directional action of recA protein analogous to that of DNA helicases (31).

While this paper was in preparation, Cox and Lehman (32) kindly sent us their manuscript which reports that, in a system containing both E. coli single-strand binding protein and recA protein, the extension of heteroduplex joints requires the continued action of recA protein.

In meiotic recombination there appear to be special sites at which the formation of heteroduplex DNA is initiated and from which the heteroduplex regions grow in a polar fashion (33–36). There are many nucleic acid enzymes whose action is polar, including polymerases, nuclease, and the ATP-dependent helicases (31, 37, 38). A whole set of such enzymes may determine the polarity of recombination in any particular pathway. Because recA protein is directly implicated in recombination (39, 40) and has recombination-like activities in vitro (2–10), the observation that it forms joint molecules in a polar fashion brings us a step closer to understanding the molecular basis of polarity in genetic recombination.

We thank John Chase for a gift of E. coli exonuclease VII, John Flory for expert guidance in our electron microscopic studies, and Anna Wu and John Flory for helpful comments on our manuscript. This research was sponsored by grants from the American Cancer Society and the National Cancer Institute.