Proc. Natl. Acad. Sci. USA
Vol. 78, No. 10, pp. 6018–6022, October 1981
Biochemistry

Directionality and polarity in recA protein-promoted branch migration

(*recA*/recombination/D loops/strand exchange/restriction analysis)

MICHAEL M. COX AND I. R. LEHMAN

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

Contributed by I. R. Lehman, June 15, 1981

ABSTRACT **The recA protein of *Escherichia coli* promotes the complete exchange of strands between full-length linear duplex and single-stranded circular φX174 DNA molecules. Analysis of the reaction by electron microscopy confirms that D loops containing short heteroduplex regions are rapidly formed at the ends of the linear duplex, followed by a relatively slow branch migration that converts the D loops to nicked circular duplexes (RFII) and displaced linear single strands. Heteroduplex extension and displacement of the linear single strand are concerted. Heterologous sequences within the linear duplex halt branch migration and lead to the accumulation of D loops. Although D loops can be formed at either end of the linear duplex, recA protein-promoted branch migration proceeds uniquely in the 3′ → 5′ direction relative to the (−) strand of the linear duplex.**

Analysis of meiotic gene conversion in *Ustilago maydis* has led to the conclusion that DNA structures containing heteroduplex regions, with strands contributed from each of two DNA molecules, are intermediates in genetic recombination (1, 2). The heteroduplex regions can be several thousand base pairs long (3). Consequently, models for recombination invariably include a step in which strands from different DNA molecules are paired, followed by a reaction in which the heteroduplex thus formed is extended (4, 5). The latter is termed "branch migration" and involves the movement of a crossover junction, shortening or extending the heteroduplex with a reciprocal loss or gain of parental base pairs.

In vitro, the recA protein promotes the pairing of single-stranded (ss) DNA with single strands containing homologous sequences (renaturation) (6) or with homologous sequences within duplex molecules (strand assimilation or D-loop formation) (7, 8). The latter is analogous to the pairing reaction believed to occur early in recombination.

We recently found that recA protein promotes the complete exchange of strands between φX174 (φX) linear duplex DNA and circular (+) ssDNA molecules with the formation of nicked circular duplexes (RFII) and linear (+) single strands. We further found that strand exchange is a two-phase reaction, consisting of the rapid formation of D loops, followed by relatively slow branch migration to generate RFII (9). recA protein-promoted branch migration has a continuous need for ATP hydrolysis, requires stoichiometric amounts of recA protein, and is greatly stimulated by *Escherichia coli* ssDNA binding protein (SSB).

The experiments reported here were designed both to confirm our earlier findings and to extend our investigation of recA protein-promoted branch migration by several different methods. Our principal new conclusion is that, in contrast to D-loop formation, recA protein-promoted branch migration exhibits a unique polarity.

MATERIALS AND METHODS

Materials. recA protein was purified to homogeneity as described (10). Concentrations of recA protein were determined by using $E^{1\%}_{280} = 5.16$ based upon its amino acid sequence (11, 12). SSB, purified from an overproducing strain (13) by a published procedure (14), was the generous gift of J. Kaguni of this department. Nuclease S1 and yeast hexokinase were purchased from Sigma. *Bam*HI was donated by P. Southern of this department. Other restriction endonucleases were purchased from New England BioLabs.

Preparation of linear duplex and circular φX ssDNA has been described (9). Circular M13 ssDNA and the RFI form of M13G*ori*1 [M13 DNA into which the G4 replication origin has been inserted (15)] were the generous gifts of J. Kobori and J. Kaguni, respectively, of this department. Full-length linear duplex M13G*ori*1 DNA was produced by *Bam*HI cleavage of M13G*ori*1 RFI followed by heating at 70°C for 10 min to inactivate the endonuclease. All DNA concentrations are reported as total nucleotide.

EN^{3}HANCE was purchased from New England Nuclear. DE 81 paper was from Whatman.

Methods. The standard reaction mixture for strand exchange was 25 mM Tris·HCl, 80% cation, final pH 7.2/5% (vol/vol) glycerol/10 mM $MgCl_2$/1 mM dithiothreitol/1 mM ATP/5.6 μM linear duplex φX (or M13G*ori*1) DNA/3.3 μM circular φX174 or M13 ssDNA (+ strands)/2.0 μM recA protein/0.15 μM SSB. Where indicated, adenosine 5′-[γ-thio]triphosphate (ATP[γ-S]) was used at a concentration of 200 μM. All reactions were performed at 37°C and were initiated by the addition of ATP or ATP[γ-S] after preincubation of the other components. Nitrocellulose filter binding and nuclease S1 assays for D-loop formation and heteroduplex formation, respectively, have been described (9).

Electron Microscopic Analysis of Branch Migration. The concentration of recA protein in the standard reaction mixture was decreased by 50% to decrease the background of denatured protein. This change decreased the apparent rate and extent of the reaction by about 10% (9). Aliquots (15 μl) of the reaction mixture were removed at the indicated times and added to 10 μl of 0.5 M Tris·HCl, pH 8.1/50 mM EDTA containing cytochrome *c* at 5 mg/ml; 35 μl of H_2O and 40 μl of formamide were added, and 50 μl of this mixture was spread immediately onto a layer of 10% formamide in 10 mM Tris·HCl (pH 8.1) at room

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: RFI, supercoiled circular double-stranded DNA; RFII, circular double-stranded DNA containing a nick in one strand; ss, single-stranded; SSB, *E. coli* ssDNA binding protein; φX, bacteriophage φX174; ATP[γ-S], adenosine 5′-[γ-thio]triphosphate.

temperature. The DNA was transferred to grids coated with 3.5% parlodion, stained with uranyl acetate, and subjected to rotary shadowing with Pt/20% Pd at an angle of 8°. DNA contour lengths were measured with a Hewlett–Packard 9810A calculator equipped with a digitizer at magnifications >60,000.

Restriction Endonuclease Analysis of Branch Migration. This assay measures the appearance of radioactivity in fragments obtained after restriction endonuclease cleavage of the products of the reaction of unlabeled linear duplex with ^{3}H-labeled ϕX ssDNA in the presence of recA protein and ATP. The appearance of label in a restriction fragment indicates that heteroduplex formation has occurred involving sequences within the fragment. By using the appropriate restriction map (see Fig. 4), the progress of branch migration can be followed.

In these experiments the concentrations of recA protein, SSB, and the ϕX DNA substrates were doubled. This change increased the rate of the reaction by approximately 20%, although it had no effect on the fraction of substrate molecules that were converted to RFII. The circular ϕX ssDNA was labeled with ^{3}H (120,000 cpm/μg). Aliquots (200 μl) were removed from the reaction mixture at the times indicated. Glucose (final concentration, 2.0 mM) and an amount of hexokinase, determined empirically, that converted the ATP to ADP with a $t_{1/2}$ of 10 sec, were added immediately to halt branch migration (9), along with 8 units of restriction endonuclease (*Hpa* II, *Hae* III, or *Hha* I). The mixtures were incubated at 37°C for 30 min; then 2.5 μl of 10% (wt/vol) $NaDodSO_4$ was added, followed by 60 μl of 250 mM sodium acetate, 40% anion, pH ≈4.4/5 mM zinc acetate. Nuclease S1 [2.0 Sigma units (1 unit is defined as the amount needed to digest 1 μg of ssDNA in 1 min at 37°C)], 3 μg of ϕX ssDNA, and H_2O were added to a final volume of 300 μl. The mixture was incubated for 10 min at 37°C. tRNA (25 μg) and 800 μl of absolute ethanol were added and the samples were left for 14 hr at −20°C. The resulting precipitates were collected by centrifugation and the supernatants were removed.

The pellets were resuspended in 20 μl of 1% $NaDodSO_4$/50% (vol/vol) glycerol/50 mM EDTA/0.04% bromphenol blue, and the samples were electrophoresed in an 8% polyacrylamide gel. The buffer used in constructing and running the gel contained 0.1% $NaDodSO_4$, 40 mM Tris base, 20 mM acetic acid, and 2 mM EDTA. After electrophoresis, the gel was soaked for 1 hr at 37°C in EN^{3}HANCE, with gentle shaking, and then in 2% trichloroacetic acid for 1 hr. The gel was then dried onto DE 81 paper, and the labeled bands were visualized by exposure of a sheet of Kodak XAR-5 x-ray film (preflashed) at −80°C for 4 days with the aid of a DuPont Cronex ZD intensifier.

RESULTS

Electron Microscopic Analysis of recA Protein-Promoted Strand Exchange. In the presence of SSB, ATP, and Mg^{2+}, the recA protein efficiently converted circular (+) ss and linear duplex ϕX DNA to RFII and linear (+) single strands (Fig. 1; Table 1).

These results provide strong confirmation for several features of the strand-exchange reaction reported previously (9). First, a large number of D loops appear before RFII forms can be detected. This finding establishes that the reaction does indeed consist of two phases: the rapid formation of D loops followed

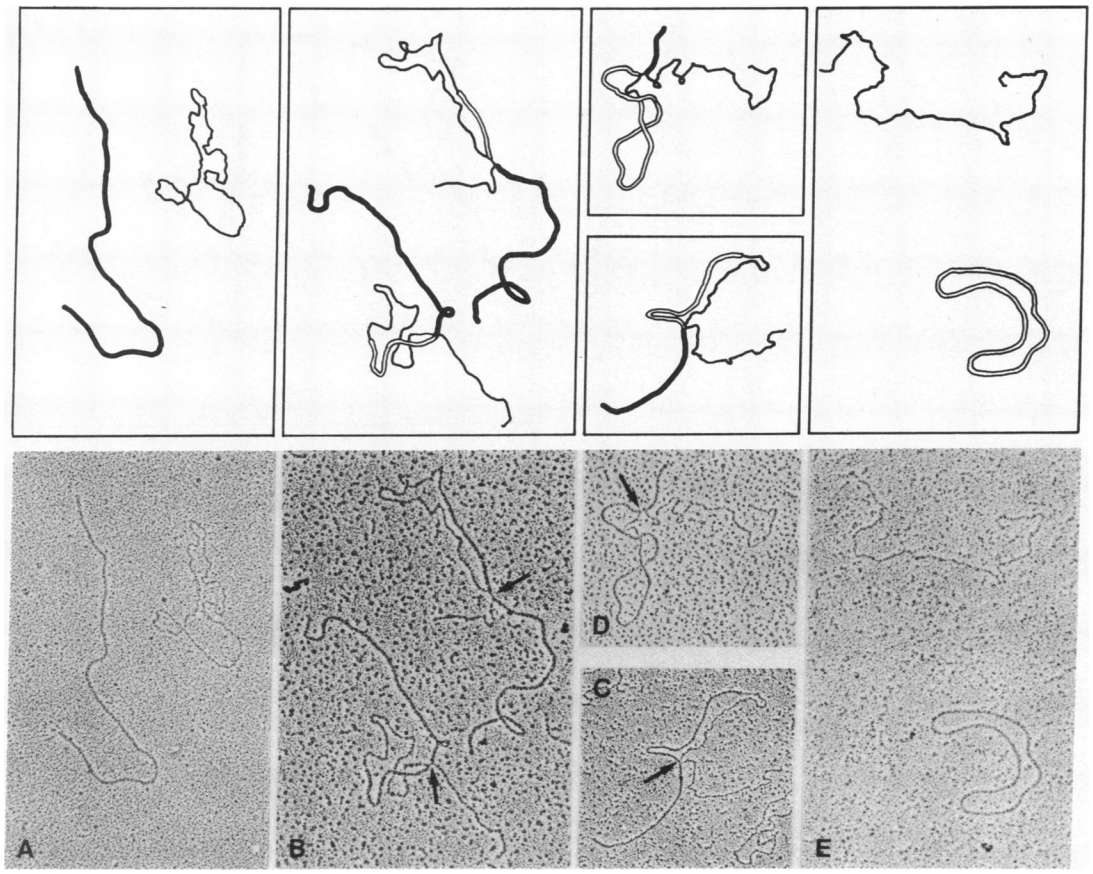


FIG. 1. Electron microscopy of recA protein-promoted branch migration. In the interpretive drawings, heavy lines indicate duplex DNA, thin lines indicate ssDNA, and parallel lines indicate heteroduplex. (*A*) Substrates: linear duplex DNA and circular ϕX ssDNA. (*B–D*) D loops in various stages of reaction. Arrows indicate positions of branch points. (*E*) Products: linear ssDNA and RFII.

Table 1. Time course of recA protein-promoted strand exchange as observed by electron microscopy

Reaction conditions		Time, min	Distribution, fraction of total as:			Molecules counted, no.
DNA	Mixture		Linear	D loop*	RFII	
φX	Complete	0	1.0	—	—	100
		10	0.66	0.33	0.01	244
		20	0.58	0.38	0.04	198
		40	0.47	0.19	0.34	255
		60	0.47	0.09	0.44	338
φX	Lacking ATP	30	1.0	—	—	100
φX	Lacking ATP; with ATP[γ-S]	45	0.87	0.13	0	463
M13G*ori*1	Complete	30	0.36	0.64	0	257
		60	0.32	0.68	0	276

* Includes double D loops.

by relatively slow branch migration. Second, in >99% of the D loops examined, one end of the linear duplex was clearly involved in the reaction, consistent with the conclusion that D-loop formation originates at an end. In a separate survey, three molecules were observed in which D-loop formation appeared to originate in the interior of the duplex. Because a small fraction (<5%) of the linear duplexes contained nicks, it is likely that the D loops in these molecules were generated at the nicks (16). Finally, it is apparent from Fig. 1 that incorporation of the circular ssDNA into heteroduplex is concerted with displacement of the linear ssDNA.

As expected, no reaction occurred in the absence of ATP. When ATP[γ-S] was substituted for ATP, D loops were again formed; however, the frequency was lower than that observed with ATP, and the D loops were not converted to RFII, confirming earlier results (9, 17). The heteroduplex regions in 11 of these D loops (chosen at random) were measured and found to be evenly distributed between 220 ± 50 and 1970 ± 150 base pairs. Because spontaneous branch migration (18) may occur after D loops have been formed, the longer heteroduplex regions observed in the presence of ATP[γ-S] are not necessarily a result of the action of recA protein. As with ATP, the reaction originated at the ends of the linear duplex.

Reactions were also performed with circular ssDNA from bacteriophage M13 and linear duplexes derived from M13G*ori*1. When cleaved by *Bam*HI, the RFI form of M13G*ori*1 is converted to a linear duplex that contains [from the 3′ end of the (−) strand] 3351 base pairs of M13 sequence followed by 2216 base pairs of the heterologous (G4) sequence and finally the remaining 3056 base pairs of M13 sequence (15). The heterologous sequences effectively blocked heteroduplex extension (Table 1). However, D loops were generated at a high frequency. The heteroduplex regions of 21 of the D loops observed at 60 min were measured. The shortest was 950 base pairs long and none exceeded 3300 ± 150 base pairs, consistent with the known boundary of the heterologous sequences in the M13G*ori*1 DNA.

The results presented in Table 1 are subject to several limitations. Eleven percent of the φX ssDNA molecules were linear, having been broken randomly during isolation. In each

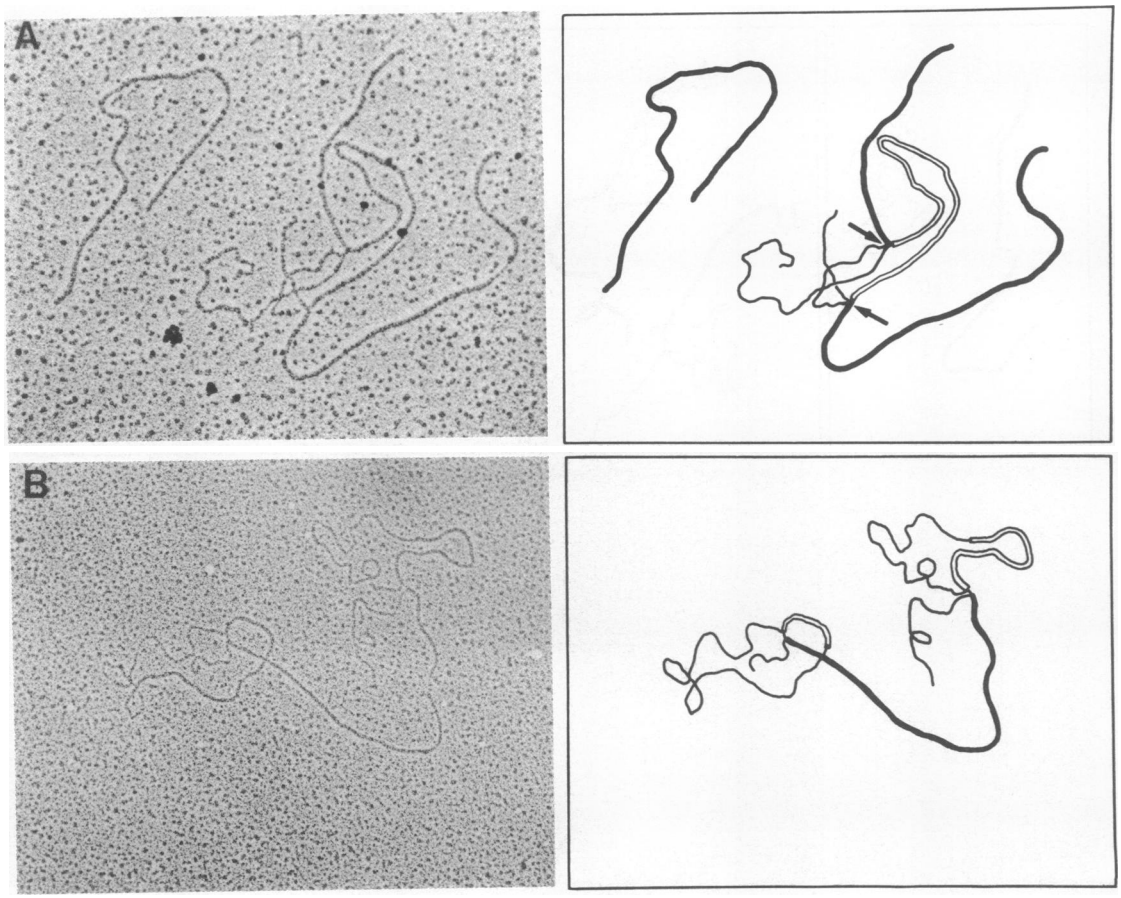


FIG. 2. Double D loops. Designations of duplex and single strands in the interpretive drawings are as in Fig. 1. (*A*) Single circular φX ssDNA molecule reacting with two linear duplexes. The unreacted linear duplex serves as an internal standard. (*B*) Two circular M13 ssDNA molecules reacting with a single M13G*ori*1 linear duplex.

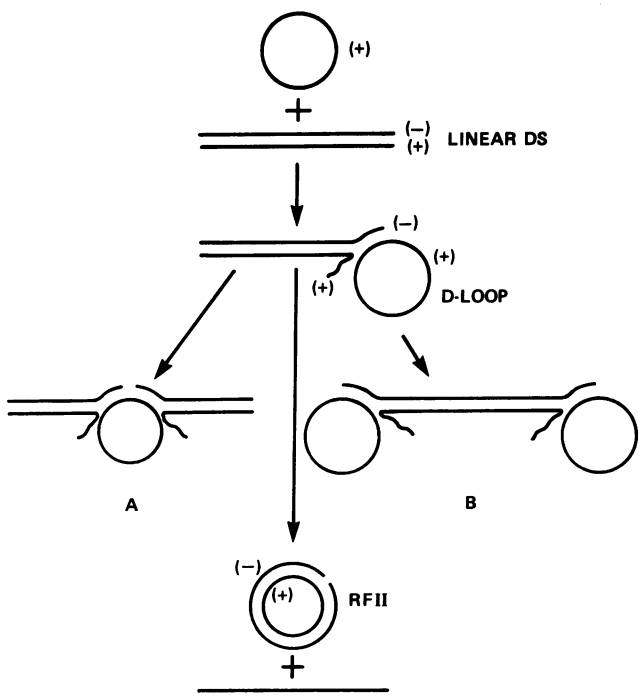


FIG. 3. Summary of structures generated during recA protein-promoted strand exchange. Structures A and B correspond to the structures in Fig. 2 *A* and *B*, respectively.

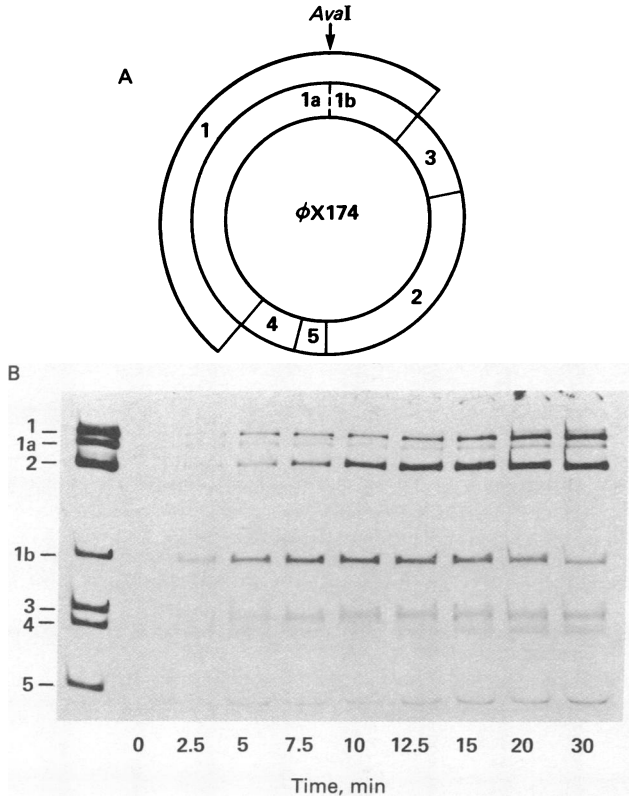


FIG. 4. Restriction endonuclease analysis of polarity of recA protein-promoted branch migration. (*A*) *Hpa* II restriction map of φX with the *Ava* I site indicated. (*B*) Fluorogram showing the appearance, with time, of radioactivity in *Hpa* II restriction fragments. Film was exposed for 4 days. The zero time point was obtained after 20-min incubation of the reaction mixture lacking ATP.

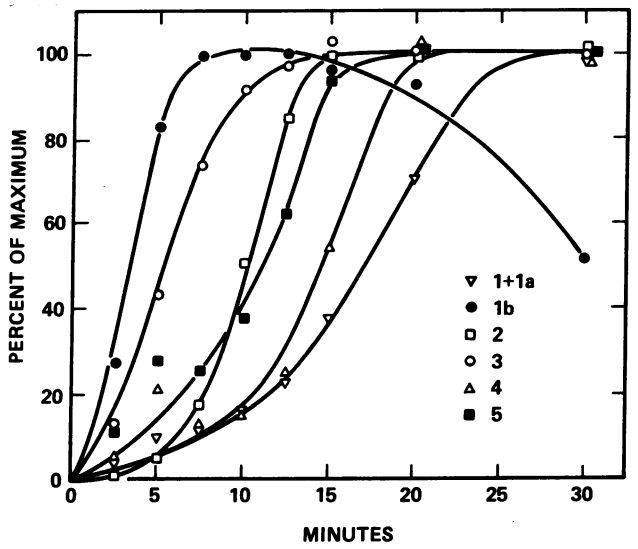


FIG. 5. Rate of appearance of radioactivity in restriction fragments. The exposed x-ray film was subject to spectrophotometric scanning using a Helena Laboratories "Quick Scan Jr." equipped with an integrator. Maximum labeling is taken to be that amount present at 30 min for all fragments except 1b. Since label in 1b begins to decrease after 20 min, maximum labeling is taken to be that present at 12.5 min. The larger fragments will contain proportionately higher amounts of label. To ensure that scanning data reflected relative amounts of label at different time points as accurately as possible, data for fragments 1, 1a, and 2 were obtained from a 24-hr exposure, while data for all other fragments were obtained from the 4-day exposure shown in Fig. 4*B*.

DNA spread, 10–20% of the duplex DNA was involved in complex aggregates, either as a result of nonspecific interaction with broken single strands or as an artifact of spreading. These aggregates were ignored in tabulating the data. Spontaneous branch migration, which may occur as recA protein is denatured during spreading, also limits electron microscopy as a quantitative tool. This process should result in the loss of some of the shorter D loops; hence, especially at the earlier time points, the values are probably underestimates.

At 10 min, about 70% of the duplex DNA was involved in D loops, as measured by the nitrocellulose filter binding assay (not shown). The discrepancy between this value and that shown in Table 1 (33%) probably reflects the loss of shorter D loops during spreading as well as the exclusion of complex aggregates, both of which would be scored in the filter binding but not the electron microscopic assay.

Lack of Polarity of D-Loop Formation. In both of the reactions described above, a significant proportion of the D loops were "double"—i.e., a single-stranded circle had reacted with two linear duplexes or vice versa. Examples of these structures are shown in Fig. 2. At 10 min, these structures represented approximately 2% of the total D loops. In a separate survey of 26 of these structures, 23 were structure A (Fig. 3). Four of these were photographed with an unreacted linear duplex as an internal standard. Measurement of these molecules showed that the duplex DNA was twice the length of the unreacted monomer. In the M13G*ori*1 reaction, 5% of the D loops present at 30 min were in the form of double D loops; this increased to 8% at 60 min. In a survey of 20 of these molecules at both time points, most proved to be structure B (60% at 30 min; 80% at 60 min). Because all of the circular single strands are (+), the double D loops demonstrate directly that D-loop formation can occur at both the 3′ and 5′ ends of the (−) strand in the duplex—i.e., it has no absolute polarity. A preference for D-loop

formation at one of the two ends of the linear duplex, however, would not be detected by these methods.

Polarity of Branch Migration. Analysis of branch migration by *Hpa* II cleavage (Figs. 4 and 5) demonstrated that, in contrast to D-loop formation, branch migration does have a unique polarity. The first *Hpa* II fragment to be labeled was 1b, located at the 3′ end of the (−) strand of the linear duplex generated by *Ava* I cleavage. The last to be labeled was fragment 1a, located on the opposite end of the linear duplex. As complete RFIIs were generated, fragment 1 became prominent with an accompanying decrease in label in 1b. Much of the increase in label in 1a resulted from nuclease S1 cleavage of fragment 1 at the nick at the *Ava* I site. Fragments 1 + 1a acquired 50% of their maximal radioactivity at least 11 min after this point was reached in fragment 1b. A similar result was obtained when the *Hae* III or *Hha* I was used in place of *Hpa* II (not shown).

This result clearly demonstrates that, although D loops can be generated at either end of the linear duplex, recA protein-promoted branch migration has a unique polarity. The polarity is 3′ → 5′ relative to the (−) strand in the linear duplex, which serves as the acceptor strand for heteroduplex formation with the circular (+) ϕX ssDNA.

There was a background of radioactivity in all the restriction fragments, even at early time points. This background could result from interaction between substrate molecules damaged during isolation (nicked duplexes or linearized single strands) which are present to some extent in all DNA preparations. Alternatively, it could result from low levels of spontaneous branch migration from D loops generated at the "wrong" end of the linear duplex. We also cannot eliminate the possibility that the polarity of recA protein-promoted branch migration is not absolute.

DISCUSSION

The demonstration that recA protein can promote branch migration leads directly to a consideration of what its role in this process might be. In model systems, branch migration occurs spontaneously, via a random walk, at up to 6000 steps per sec (18). In our system, D loops are formed at an end of the linear duplex. If branch migration were to proceed via a random walk, migration for the few hundred base pairs required to eliminate the D loop would be many times more probable than the migration for many thousands of base pairs in the other direction that is required to form an RFII molecule. Thus, the efficiency of branch migration would not be increased if the recA protein were simply to catalyze a random walk migration. The conversion of as much as 50% of the substrate to RFII therefore clearly implies that at least one contribution of recA protein is to confer directionality on the reaction. The findings reported here further demonstrate that the reaction has a unique polarity. Although D loops can be formed at either end of the linear duplex, subsequent branch migration can occur in only one direction (3′ → 5′) relative to the (−) strand and, as a consequence, as many as half the D loops formed are dissipated. This finding represents a second distinction between D-loop formation and branch migration. We demonstrated previously that D-loop formation, but not branch migration, can proceed in the presence of ATP[γ-S] (9).

We have argued that mechanisms of branch migration in which recA protein moves processively along the DNA, or acts distributively at a branch point, are not consistent with available data (9). The alternative we find most attractive is one in which the recA protein initially involved in D-loop formation acts as a core onto which other recA monomers can bind to form a growing filament. In fact, recA protein forms highly structured filaments in the presence of ATP or ATP[γ-S] (19). Filament growth would then drive branch migration. Possibly, the polarity of branch migration is a consequence of an asymmetry of the two ends of the filament coupled to the polarity of one or more of the DNA strands involved in the reaction.

Note Added in Proof. After this paper was submitted, Charles Radding and Stephen West kindly communicated to us their independent findings that D loops are found predominantly at the 3′ end of the (−) strand of a linear duplex in systems similar to the one described here.

This work was supported by grants from the National Institutes of Health (GM-06196) and the National Science Foundation (PCM 74-00856). M.M.C. is a Postdoctoral Fellow of the American Cancer Society.

1. Holliday, R. (1964) *Genet. Res.* **5**, 282–304.
2. Whitehouse, H. L. K. (1973) *Towards an Understanding of the Mechanism of Heredity*, 3rd Ed. (St. Martin's, New York).
3. Fox, M. S., Dudney, C. S. & Sodergren, E. J. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 999–1007.
4. Radding, C. M. (1978) *Annu. Rev. Biochem.* **47**, 847–880.
5. Stahl, F. W. (1979) *Genetic Recombination* (Freeman, San Francisco).
6. Weinstock, G. M., McEntee, K. & Lehman, I. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 126–130.
7. Shibata, T., Das Gupta, C., Cunningham, R. P. & Radding, C. M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1638–1642.
8. McEntee, K., Weinstock, G. M. & Lehman, I. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2615–2619.
9. Cox, M. M. & Lehman, I. R. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3433–3437.
10. Cox, M. M., McEntee, K. & Lehman, I. R. (1981) *J. Biol. Chem.* **256**, 4676–4678.
11. Horii, T., Ogawa, T. & Ogawa, H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 313–317.
12. Sancar, A., Stackhelek, C., Konigsberg, W. & Rupp, W. D. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2611–2615.
13. Chase, J. W., Whittier, R. F., Auerbach, J., Sancar, A. & Rupp, W. D. (1980) *Nucleic Acids Res.* **8**, 3215–3227.
14. Weiner, J. H., Bertsch, L. L. & Kornberg, A. (1975) *J. Biol. Chem.* **250**, 1972–1980.
15. Kaguni, J. & Ray, D. S. (1979) *J. Mol. Biol.* **135**, 863–878.
16. Das Gupta, C., Shibata, T., Cunningham, R. P. & Radding, C. M. (1980) *Cell* **22**, 437–446.
17. McEntee, K., Weinstock, G. M. & Lehman, I. R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 857–861.
18. Warner, R. C. & Tessman, I. T. (1978) in *The Single-Stranded DNA Phages*, eds. Denhardt, D. T., Dressler, D. & Ray, D. S. (Cold Spring Harbor Laboratories, Cold Spring Harbor, NY), pp. 417–432.
19. McEntee, K., Weinstock, G. M. & Lehman, I. R. (1981) *J. Biol. Chem.*, in press.