

RecA protein promotes the regression of stalled replication forks *in vitro*

Mara E. Robu, Ross B. Inman, and Michael M. Cox*

Department of Biochemistry, University of Wisconsin, Madison, WI 53706-1544

Replication forks are halted by many types of DNA damage. At the site of a leading-strand DNA lesion, forks may stall and leave the lesion in a single-strand gap. Fork regression is the first step in several proposed pathways that permit repair without generating a double-strand break. Using model DNA substrates designed to mimic one of the known structures of a fork stalled at a leading-strand lesion, we show here that RecA protein of *Escherichia coli* will promote a fork regression reaction *in vitro*. The regression process exhibits an absolute requirement for ATP hydrolysis and is enhanced when dATP replaces ATP. The reaction is not affected by the inclusion of the RecO and R proteins. We present this reaction as one of several potential RecA protein roles in the repair of stalled and/or collapsed replication forks in bacteria.

Under normal growth conditions, bacterial replication forks are halted by DNA damage in virtually every cell and every cell generation (1–4). Replication forks can stall for a variety of reasons. They may encounter a DNA lesion or a single-stranded (ss)DNA nick, leading to collapse of the fork and formation of a DNA gap or double-strand break, respectively. They might encounter a region of secondary structure in DNA that constitutes a pause site. They might also collide with another protein or protein complex on DNA, such as an RNA polymerase itself stalled at a DNA lesion (5). Proteins bound on DNA can temporarily block a replication fork by inhibiting the progression of the replicative helicase.

Genetic and structural evidence suggests that a first step in the repair of a stalled fork is fork regression (Fig. 1A) (5–7). The products of fork regression have been observed by electron microscopy (EM) (6, 8). The resulting recombination intermediates can subsequently be processed to function as substrates for replication restart in an origin-independent manner (4). This process does not repair the lesion itself but instead establishes the necessary conditions for accurate repair by supplying an undamaged complementary strand opposite the lesion. The lesion can then be removed via a process like excision repair. Another key outcome of fork repair pathways is the nonmutagenic restoration of the replication fork.

In principle, a lesion on the lagging strand presents fewer problems to the replication machinery, because another Okazaki fragment can be initiated downstream (9). However, a lesion on the leading strand may block replication and lead to the formation of a ssDNA gap on this strand, whereas synthesis can become uncoupled and continue for some distance on the lagging strand under at least some circumstances (9, 10). Thus a fork with a leading-strand ssDNA gap at the branch is generated (Fig. 1A). Recombination is required to generate a DNA structure on which replication can resume and to allow for repair and/or bypass of the lesion. As part of recombinational DNA repair, *Escherichia coli* RecA protein is thought to promote DNA pairing, strand exchange, and branch migration beyond the lesion. These activities allow the lesion-containing strand to pair with an intact complementary strand so that repair can ensue (11). In some early models, a nuclease was proposed to generate

a free end in the lagging-strand template for the RecA-catalyzed strand invasion (12). However, the required nuclease has not been found. The alternative fork regression pathway (Fig. 1A) accomplishes the same purpose with no break in the parental strands (5–7, 11). A similar process of fork regression may occur spontaneously at stalled forks that do not have a ssDNA gap at the branch, facilitated by the superhelical stress generated during replication (13).

Fork regression leads to the formation of a Holliday intermediate (Fig. 1A). In principle, this intermediate could be processed in two ways. The tendency of stalled forks to generate double-strand breaks provides evidence that some intermediates are cleaved by the RuvABC system (7, 14). The resulting double-strand break would then be funneled into the RecBCD pathway for subsequent repair. The alternative path would involve a reverse branch migration to regenerate a fork-like structure, perhaps propelled by the RecG protein (5).

The regression of replication forks could be promoted by enzymes known to process branched DNAs, such as the RecG helicase (5) or the RuvAB system (14, 15). The RecA protein is also known to play an important role in replication fork repair. A role in the strand invasion step of double-strand break repair has extensive experimental support (16–20). Recent evidence suggests that a major pathway leading to the formation of cleavable Holliday intermediates at stalled forks requires the RecA protein (21), suggesting a role for RecA protein in fork regression. This would be a novel activity of RecA protein. At a stalled fork with a structure like that shown in Fig. 1A, RecA filaments would nucleate within the gap and extend away from the branch (22, 23). To promote fork regression, the RecA protein would thus act to move a DNA branch that was initially beyond one end of the filament. This study was initiated to investigate whether a replication fork with a gap at the branch could regress through the action of RecA protein. We report here that RecA protein does indeed promote fork regression *in vitro*.

Materials and Methods

Enzymes and Reagents. *E. coli* RecA (24), RecO (23), RecR (25), RecF (25), and single-strand DNA-binding protein (26) proteins were purified as described. The concentration of each protein was determined by absorbance at 280 nm by using their extinction coefficients: $\epsilon_{280} = 2.23 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ for RecA (27), $\epsilon_{280} = 5.6 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ for RecR (23), $\epsilon_{280} = 3.87 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ for RecF (28), and $\epsilon_{280} = 2.83 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ for SSB (29), with the exception of RecO protein, for which the concen-

This paper results from the National Academy of Sciences colloquium, "Links Between Recombination and Replication: Vital Roles of Recombination," held November 10–12, 2000, in Irvine, CA.

Abbreviations: ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; EM, electron microscopy; MM, model DNA molecule.

*To whom reprint requests should be addressed. E-mail: cox@biochem.wisc.edu.

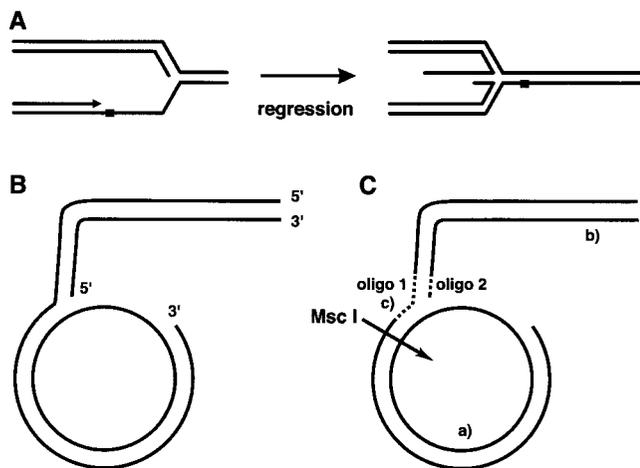


Fig. 1. Regression of a stalled fork, and experimental models. (A) In at least one replication system (9, 10), a lesion on the leading strand results in a stalled fork with a single-strand gap on the leading strand and with lagging-strand synthesis completed beyond the lesion, as shown. A postulated first step in the repair of such a fork is the regression reaction illustrated here, in which the original template strands are re-paired with each other and the newly synthesized strands are annealed as shown. (B) The structure of the model DNA molecules MM1 and MM2, designed to resemble a stalled replication fork. MM2 differs from MM1 only in that MM2 includes six mismatches at the branch to prevent spontaneous branch migration. The *MscI* site shown is the cleavage site used to generate the linearized version of MM2, called LMM2. (C) The components used to construct the model DNA substrates: (a) the gDNA molecule (7,261-bp circle with a 2,051-nt ssDNA gap); (b) the homologous dsDNA tail (7,261 bp), and (c) the linker (oligo1: 49-mer, oligo2: 30-mer), depicted with dashed lines.

tration was measured by Bradford assay by using BSA as a standard.

Restriction enzymes, λ DNA-*Bst*EII digest marker, and T4 DNA ligase were purchased from New England Biolabs, Tris buffer from Fisher Scientific, and ATP, dATP, phosphocreatine, and creatine phosphokinase (lot no. 78H 7824) from Sigma. ATP γ S was purchased from Boehringer Mannheim, SeaPlaque GTG low-melting agarose from FMC, Sephacryl 300 resin from Amersham Pharmacia, and hydroxylapatite DNA grade Bio-Gel HTP from Bio-Rad. The GeneCapsule kit was purchased from Geno Technology, St. Louis, MO, and oligonucleotides were synthesized by Operon Technologies, Alameda, CA.

DNA. Circular ssDNA from bacteriophage M13mp8.32 was prepared as described (30). Bacteriophage M13mp8.32 is derived from M13mp8, having a 32-bp insertion in the *Pst*I site. The insert was generated by annealing two synthetic oligonucleotides: 5'-GCACGTGGCACCGTGGCGCTCTTCGGTGCTGCA-3' (plus strand) and 5'-GCACCGAAGAGCGCACGGT-GCCACGTGCTGCA-3'. The insert contains a cleavage site for the restriction endonuclease *Sap*I. The double-stranded oligonucleotide was ligated into *Pst*I-cut M13mp8 dsDNA and cloned according to Maniatis *et al.* (31). Supercoiled M13mp8.32 DNA was purified from infected JM101 *E. coli* cells (32) with a QIAfilter Plasmid Mega Kit from Qiagen (Chatsworth, CA). ssDNA contamination was removed by using a hydroxylapatite DNA grade Bio-Gel HTP column. The concentrations of double-stranded (ds)DNA and ssDNA solutions were measured by absorbance at 260 nm, by using 50 and 36 μ g ml⁻¹, respectively, as conversion factors. The conversion factors used to calculate the concentration of gapped DNA and the branched model molecules were based on the fractions of ssDNA and dsDNA in the molecules. DNA concentrations are expressed in terms of total nucleotides, if not otherwise noted. For

DNA mixtures, the concentration of each component has been quantified by analyzing ethidium bromide-stained gels with IMAGEQUANT software (Molecular Dynamics).

Branched model DNA molecules [with model DNA molecule (MM) designations; Fig. 1B] were assembled from a gapped DNA GD2051 and homologous linear dsDNA (7.3 kb) ligated with a linker (Fig. 1C).

Gapped DNA with a precise gap length was prepared by a large-scale RecA reaction between circular ssDNA M13mp8.32 and a 5.2-kb linear dsDNA fragment, as previously described (33). The 5.2-kb linear dsDNA fragment was generated by complete digestion of supercoiled M13mp8.32 with *Eco*RI and *Bsr*GI enzymes. The fragment was gel purified by using 1% SeaPlaque GTG agarose gel and was extracted from gel pieces by using QIAquick Gel Extraction Kit from Qiagen. The product of the RecA reaction was GD2051, which has a 2,051-nt ssDNA gap. GD2051 was gel purified as described above.

Linear dsDNA (7.3 kb) was prepared by complete digestion of supercoiled M13mp8.32 DNA with *Sap*I and *Sma*I enzymes, which leave a nonpalindromic ssDNA overhang and a blunt end, respectively. The MM1 DNA was assembled by using a completely homologous linker prepared by annealing two synthetic oligonucleotides: oligo1 5'-CACCGAAGAGCGCACGGTGCACCGTGGCGCTGCGCTCTTCG-3' and oligo2 5'-CCTGCAGCACGTGGCACCGTGGCGCTCTTCG-3'. The MM2 DNA was assembled by using a linker that has six mismatches at the branch point (with respect to M13mp8.32) to prevent spontaneous branch migration. The linker was prepared by annealing two synthetic oligonucleotides: oligo1 M 5'-CACCGAAGAGCGCACGGTGGCGCGCGGGCGCCTCGACGGATCCCCGGG-3' and oligo2 M 5'-GGCGCCGCGCGCGCACCGTGGCGCTCTTCG-3'. The nucleotides in bold correspond to the mismatches. Either one of the linkers was ligated with the nonpalindromic ssDNA overhang of linear dsDNA (7.3 kb). The ligated products were purified from excess oligonucleotides by gel filtration on a Sephacryl 300 column. Thus two-tailed linear dsDNA were generated: lds (1–2) and lds[1 M-2 M]. Each has a 16-nt ssDNA overhang that is complementary to the 5' region of the gap in GD2051.

The MM1 and MM2 DNAs were assembled from GD2051 and lds [1–2] or lds[1 M-2 M], respectively (Fig. 1C). Annealing of the tailed linear dsDNA molecules to GD2051 and ligation with GD2051 were performed in the same reaction. Typically, GD2051 and either lds (1–2) or lds[1 M-2 M] were incubated with T4 DNA ligase in reaction buffer containing 50 mM Tris-HCl (pH 7.5), 10 mM DTT, 1 mM ATP, and 25 μ g/ml BSA at 25°C for 16 h. The mixtures were concentrated, and T4 DNA ligase was removed by using Microcon 100 centrifugal filter devices from Millipore. The reactions yielded MM1 or MM2, mixed with unreacted GD2051 and lds (1–2) or lds[1 M-2 M], respectively. The DNA mixtures were in some cases used directly in reactions with RecA protein. For many key reactions using MM2, the branched DNA was purified further on a 1% SeaPlaque GTG agarose gel. The MM2 DNA was extracted from gel pieces by electroelution using a GeneCapsule kit.

For some experiments, the MM2 DNA was linearized (before final purification) by complete digestion with *Msc*I enzyme and purified from *Msc*I by using Microcon 100 columns. This linear MM2 was mixed with linear GD2051 and 6.2-kb linear dsDNA generated by *Msc*I digestion of unreacted GD2051 and lds[1 M-2 M]. Some of these linear MM2 DNA mixtures were used directly in reactions with RecA protein, for some routine trials, and others were purified further. The mixtures were gel purified by using 1% SeaPlaque GTG agarose gel, and the linear MM2 DNA was extracted from gel pieces by electroelution by using a GeneCapsule kit. The purified linear MM2 preparations were typically about 90% pure, as determined from the fluorescence of bands in an ethidium bromide-stained gel.

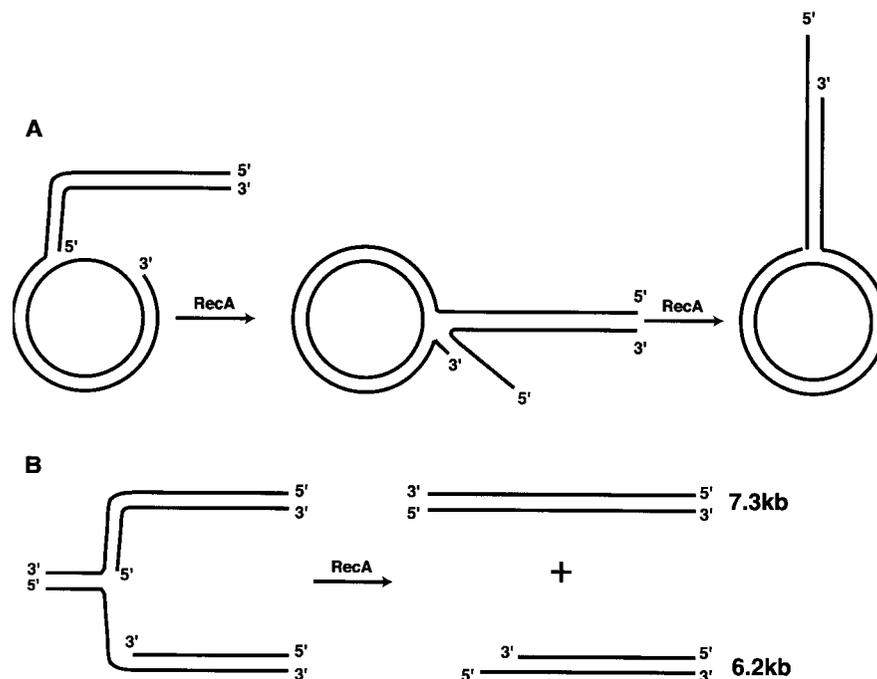


Fig. 2. Schematic of RecA-promoted reactions of model molecules. A schematic of RecA protein reaction with circular (A; MM1 or MM2) or linearized model molecules (B; LMM2) as predicted by the fork regression model.

RecA Protein Reactions. Reactions were carried out in 25 mM Tris acetate (80% cation)/10 mM Mg(OAc)/2.5% (wt/vol) glycerol/3 mM potassium glutamate/1 mM DTT in a final volume of 10 μ l. ATP, dATP, or ATP γ S at 3 mM was included, as indicated. An ATP or dATP regenerating system (10 units/ml creatine phosphokinase and 12 mM phosphocreatine) was also included. RecA and DNA concentrations are provided in the text and figure legends. Typically, RecA protein was incubated with DNA in reaction buffer and regenerating system for 10 min. Then, SSB protein (to a final ratio of 1 monomer per 10 single-stranded nucleotides) and the nucleotide cofactor were added as indicated. Reaction mixtures were incubated for different amounts of time, as indicated. When RecORF proteins were used, they were added as part of a preincubation mixture with the nucleotide cofactor and SSB. All incubations were carried out at 37°C. Reactions were stopped and deproteinized by adding SDS to 1%, EDTA to 12 mM, and proteinase K to 1 mg/ml (all final concentrations). Proteinase K was incubated for 30 min at 37°C before addition for predigestion. After addition, incubation was carried out for an additional 30 min at 37°C. Control reactions lacking one or more components were carried out as indicated in the text and legends. Reaction mixtures were electrophoresed on 1.4% agarose gels. The gels were run at 25–30 mV, at 4°C, typically for 16 h. DNA was stained with ethidium bromide and visualized with a charge-coupled device camera on a transilluminator.

EM. Samples for EM were obtained by spreading the reaction mixtures or DNA solutions as noted with a cytochrome spreading technique to visualize the DNA (25). Reaction mixtures were first deproteinized as described above and dialyzed against 20 mM NaCl and 5 mM EDTA for at least 4 h at 25°C on Millipore type VM (0.05 mm) filters. The samples were then diluted to a final concentration of 0.004 μ g/ μ l and spread, as described previously (34). Photography and measurements of the DNA molecules were performed as described (35).

Results

Experimental Design. The purpose of this study was to investigate the action of RecA protein on a stalled replication fork. We

therefore designed and constructed a DNA molecule that mimics one of the few stalled replication fork structures observed and characterized to date (Fig. 1A) (10). This branched model molecule was designated MM1 (Fig. 1B). It consists of a gapped DNA circle with a homologous dsDNA tail attached at the 5' region of the gap. RecA protein can bind the ssDNA region from the gap starting at the branching point in the 5' to 3' direction (36). The molecule was assembled from a gapped DNA circle and a homologous linear dsDNA joined by a synthetic linker (Fig. 1C).

To prevent or diminish spontaneous branch migration at the branch, we changed six base pairs in the linear dsDNA tail in positions 1, 2, 3, 6, 9, and 12 from the branch. These mismatches with the ssDNA in the gap were designed to prevent both end fraying of the linear dsDNA tail and spontaneous branch migration. The molecule constructed in this way was designated MM2. Except for the mismatches at the branch point, the overall structure of MM2 is the same as the structure of MM1 (Fig. 1A).

On a gapped DNA, RecA protein nucleates on the ssDNA region and extends the filament 5' to 3'. For the MM1 or MM2 branched molecules, filament extension within the gap will progress away from the branch, into the adjacent duplex DNA. Given enough protein, the filament can extend all the way around the circle, straddling the branch. Under some conditions, some binding might also occur on the linear dsDNA tail if enough protein were available. If a cellular replication fork stalled so as to leave a leading-strand gap as in Fig. 1A, the 5' to 3' direction of filament extension in the gap (away from the branch) would ensure that the RecA filament appeared only on the gap side of the branch. To better mimic this situation, we linearized MM2 with a restriction enzyme MscI that generates blunt ends (Fig. 1C). This molecule will be called LMM2. On LMM2, RecA filaments nucleated in the gap would be extended to the end of the DNA but would not be expected to straddle the branch itself.

The potential regression reactions and expected products of a RecA reaction with MM1 or MM2 and LMM2 are shown in Fig. 2. In the reactions, we used either unpurified branched model

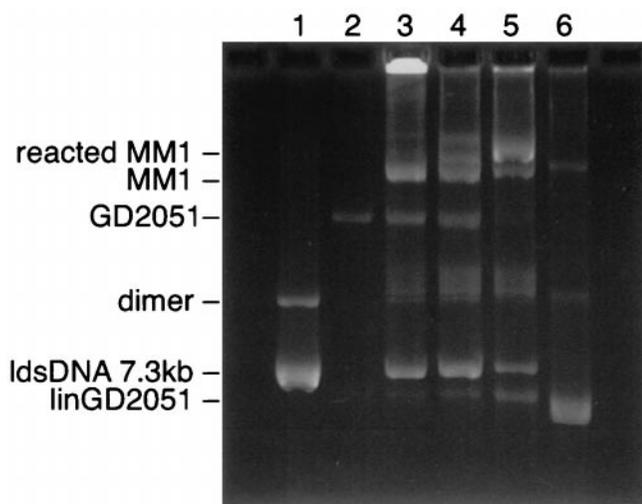


Fig. 3. RecA-promoted reactions of MM1. Gel electrophoresis of RecA protein-promoted reactions with unpurified MM1 DNA (branched MM1 molecules were >40% of the total DNA). Reaction mixtures were incubated with (lane 5) or without (lane 4) RecA protein. DNA concentration was 13.24 μ M total nucleotides, which corresponds to 1.15 μ M single-stranded nucleotides. RecA protein concentration was 3.8 μ M = $10\times$ ($1\times$ RecA protein represents the stoichiometric amount of RecA protein corresponding to the ssDNA in the mixture). Reaction mixtures were loaded on a 1.4% agarose gel, next to markers of MM mixture solution (lane 3), GD2051 (lane 2), linear dsDNA[1–2] (lane 1), and dsmp8.32 DNA (lane 6).

molecules (that also contained some unreacted substrates: GD2051 and one dsDNA ligated with the appropriate linker) or purified branched model molecules, as indicated. In cases where the unpurified DNA was used, the branched model molecules made up at least 40% of the total DNA molecules present. The background of unreacted DNA substrate molecules provided useful standards for recognition in experiments involving EM.

RecA Protein Reactions with MM1 DNA. Reaction mixtures were incubated either with (Fig. 3, lane 5) or without (Fig. 3, lane 4) RecA protein, in the presence of dATP and SSB protein for 2 h. The dATP is used for the standard reaction conditions to minimize end-dependent RecA filament disassembly (23). The reactions were deproteinized and divided in two for analysis by gel electrophoresis (Fig. 3) or by EM (Fig. 4).

In the absence of RecA protein, spontaneous branch migration occurs to a limited degree, as evidenced by the appearance of the minor upper bands in lane 4, Fig. 3. Counts of molecules observed in the electron microscope revealed that $\approx 60\%$ of the branched MM1 molecules had undergone no branch migration (Fig. 4A), and about 21% showed branch migration occurring over less than $1/3$ of the ssDNA gap, which is 2,051 nt long. Approximately 18% of the model molecules showed branch migration occurring over more than $1/3$ of the ssDNA gap and less than 2,051 nt (Fig. 4B). None of the molecules had undergone branch migration beyond the ssDNA gap.

In the presence of RecA protein, essentially all model molecules reacted to generate the final product after 2 h, represented by the intense upper band in lane 5, Fig. 3. This is the product expected for a complete RecA protein-mediated regression reaction of the model fork in the MM2 DNA, as confirmed by EM analysis. The second band from the top represents the nicked circles, one of the products of strand exchange between the excess of linear dsDNA and gapped circles still present in the reaction mixture. EM counts revealed that almost 100% of the model molecules had the structure of the final product (with the ssDNA region at the end of the tail) (Fig. 4C).

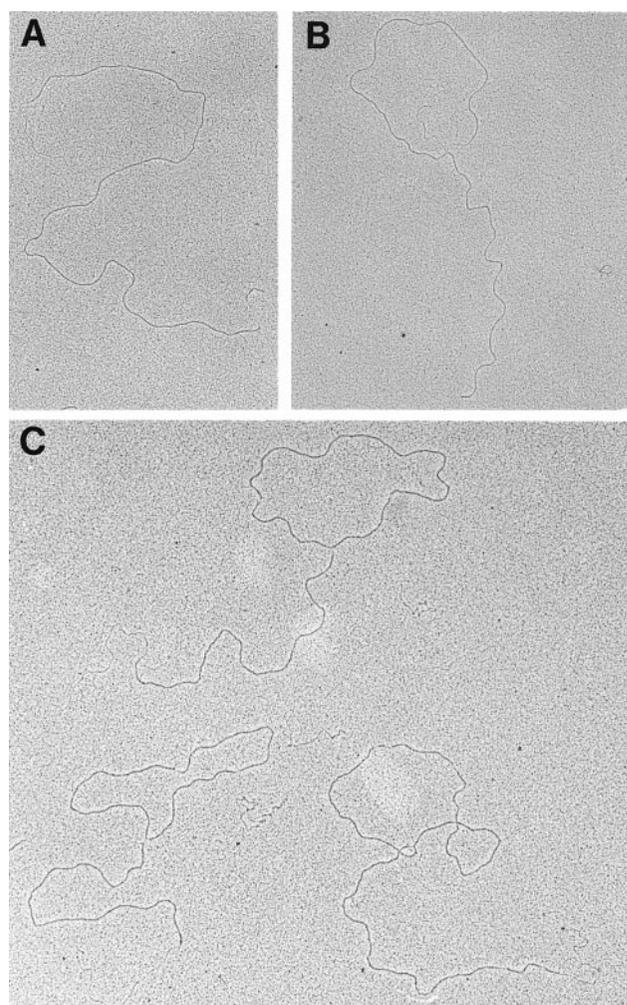


Fig. 4. RecA-promoted reactions of MM1 electron micrographs of the MM1 DNA substrate (A), an intermediate of branch migration (taken in this case from a tube that had undergone spontaneous branch migration only) (B), and final products of a regression reaction promoted by RecA protein (C). The intermediate has a single strand displaced at the branch, whereas the final products have the ssDNA regions at the end of the dsDNA tail.

RecA Protein Reactions with MM2 DNA. To diminish the background of spontaneous branch migration of the MM1 DNA, we introduced six mismatches at the branch of MM1 to generate MM2, as described above. MM2 did not exhibit spontaneous branch migration, as shown in Fig. 5A, lane 4. Control experiments quantified by EM confirmed that no spontaneous branch migration occurred when MM2 DNA was incubated in the absence of RecA protein. In the presence of RecA protein, after 2 h, a very pronounced fork regression reaction was observed. EM counts revealed that $\approx 81\%$ of the purified MM2 reacted to generate the final product. Another 5% was in an intermediate stage of branch migration, with the remainder unreacted. Four different experiments analyzed by gel electrophoresis with both unpurified and purified MM2 confirmed that the reaction is both very efficient and completely dependent on RecA protein.

Nucleotide Cofactor and Protein Requirements for RecA Protein-Promoted Fork Regression *in Vitro*. We wished to determine whether ATP hydrolysis was required for the *in vitro* fork regression reaction. We also wanted to explore the nucleotide cofactor requirements for the reaction. On ssDNA, dATP tends to prevent end-dependent filament disassembly (23), which would remove RecA from the filament end nearest the branch.

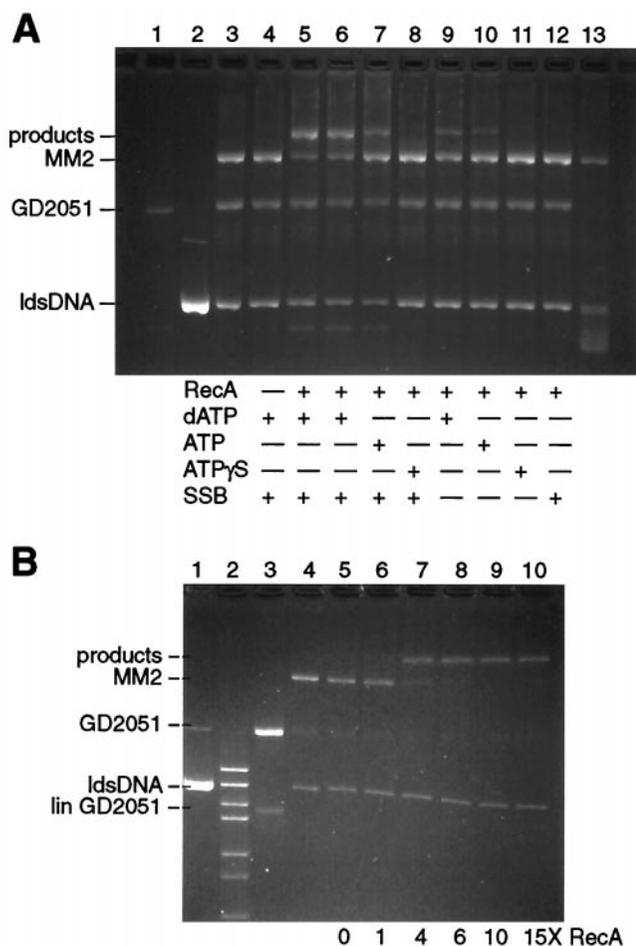


Fig. 5. RecA protein-promoted reactions with MM2. (A) Gel electrophoresis of reactions to test nucleotide cofactor requirements for RecA protein reaction with purified MM2 DNA. The reaction contained MM2 DNA at a concentration of 17.7 μ M total nucleotides, corresponding to 1.42 μ M single-stranded nucleotides (in the gaps). This was incubated with 7.05 μ M = 15 \times RecA protein (lane 5) or 4.7 μ M = 10 \times RecA protein (lanes 6–12) in the specified conditions. Lane 1, GD2051 marker; lane 2, ldsDNA [1 M-2 M] marker; lane 3, MM2 mixture; lanes 4–12, RecA reactions with MM2 DNA as indicated; lane 13, supercoiled M13mp8.32 DNA marker. (B) RecA protein titration in reactions with MM2 DNA. Purified MM2 DNA at a concentration of 10 μ M total nucleotides, corresponding to 0.57 μ M single-stranded nucleotides in the gaps, was incubated with increasing concentrations of RecA protein. 1 \times RecA represents 0.19 μ M RecA, or enough to bind all of the available single-stranded gap DNA. Lanes: 1, linear dsDNA [1 M-2 M] marker; 2, λ DNA-*Bst*II DNA marker (from top: 8,454, 7,242, 6,369, 5,686, 4,822, 4,324, 3,675 bp); 3, GD2051 marker; 4, purified MM2; 5–10, reactions with increasing concentrations of RecA protein as indicated.

Therefore, additional RecA reactions with MM2 were carried out in the presence of dATP, ATP, or ATP γ S, and in the presence or absence of SSB. These reactions contained unpurified MM2 (40% MM2 molecules) and were analyzed by gel electrophoresis. The RecA protein-mediated fork regression reaction is most efficient in the presence of dATP (Fig. 5A, lanes 5 and 6). ATP makes the reaction less efficient than that observed with dATP (Fig. 5A, lane 7), perhaps because RecA filaments are less stable in the presence of ATP (23). ATP γ S (an ATP analog that is not significantly hydrolyzed by RecA on this time scale) abolishes the reaction completely (Fig. 5A, lane 8). The preparation of ATP γ S used in this experiment was also used with consistent success in several dozen other RecA protein-promoted DNA-pairing reactions in the laboratory (data not shown), so that the lack of detectable regression in this exper-

iment was not because of an inactive or degraded ATP γ S reagent. Also in the absence of any nucleotide cofactor, the reaction is blocked (Fig. 5A, lane 12). Thus ATP hydrolysis is required for RecA-mediated fork regression.

Reactions in the absence of SSB protein were less efficient than those that contained it (Fig. 5A, lanes 9–11, as compared with lanes 6–8, respectively). A RecA titration revealed that an optimal reaction also required sufficient RecA protein to saturate the circular part of the MM2 DNA. A concentration of 1 \times RecA represents the amount of RecA protein required to saturate only the single-stranded gap in MM2 (and any other ssDNA in a particular reaction) at one RecA per three nucleotides. A concentration of \approx 3.5 \times represents enough RecA protein to coat the circular part of MM2. Approximately 9.5 \times RecA protein represents enough RecA protein to coat all of the DNA in the reaction mixture. No product is formed in the absence of RecA protein (Fig. 5B, lane 4). With increasing concentrations of RecA protein, more product is formed (upper band) as the MM2 substrate is consumed (second band from the top, Fig. 5B, lanes 5–10). A concentration of 4 \times RecA protein provided an essentially complete reaction with MM2.

To examine the kinetics of RecA-mediated *in vitro* fork regression, purified MM2 DNA was incubated with RecA protein at 4 \times concentration for up to 3 h. At different time points, aliquots were removed from the reaction mixture and were analyzed by gel electrophoresis (data not shown). The reaction approaches maximum efficiency at 40 min and is essentially complete after 60 min.

RecA Protein Reactions with LMM2 DNA. To better mimic the DNA structures that could be present *in vivo* and to better define the requirements for RecA protein, we carried out a series of similar experiments with the linearized version of MM2, LMM2. Purified LMM2 molecules were spread by using the cytochrome spreading technique and analyzed with EM. A representative molecule is shown in Fig. 6A. Purified LMM2 DNA was incubated with or without RecA protein for 150 min, and the reaction mixtures were analyzed by EM after cytochrome spreading. More than 30% of the LMM2 reacted to generate final products (linear dsDNA 7.3 kb and linear gDNA 6.2 kb with a 2-kb single-strand extension at one end; Fig. 2). This reaction was carried out under standard reaction conditions in the presence of dATP.

Purified LMM2 DNA was incubated with increasing amounts of RecA protein for 2 h, and reaction mixtures were analyzed by gel electrophoresis (Fig. 6). As in the previous set of experiments, a concentration of 1 \times RecA protein represents the amount of RecA protein needed to saturate the available DNA-binding sites in the single-stranded gaps of the DNA molecules at one RecA per three nucleotides. Under these conditions, the RecA may not all be bound to the gaps, with some bound in filaments extending beyond the gaps into adjacent duplex DNA and some gap DNA unbound. A concentration of \approx 2.7 \times represents enough RecA protein to coat the arm of the forked MM2 that contains the ssDNA gap. Approximately 6 \times RecA protein represents enough RecA protein to coat all DNA in the mixture. The reaction with LMM2 depends completely on RecA protein, as no product is formed in its absence (Fig. 6B, lane 3). With increasing concentrations of RecA protein, more products are formed [linear dsDNA 7.3 kb (second band from the top) and linear gDNA 6.2 kb with the 2-kb gap at one end (bottom band)] as the LMM2 substrate is consumed (Fig. 6B, lanes 4–8). The efficiency of the reaction starts to level off at 3 \times RecA protein.

The kinetics of the reaction are shown in Fig. 6C. Purified LMM2 DNA was incubated with RecA protein at 4 \times concentration for up to 3 h. At different times, aliquots were removed as indicated and were analyzed by gel electrophoresis. The reaction approached an endpoint of about 64% after \approx 150 min and increased little more out to the 3 h time point.

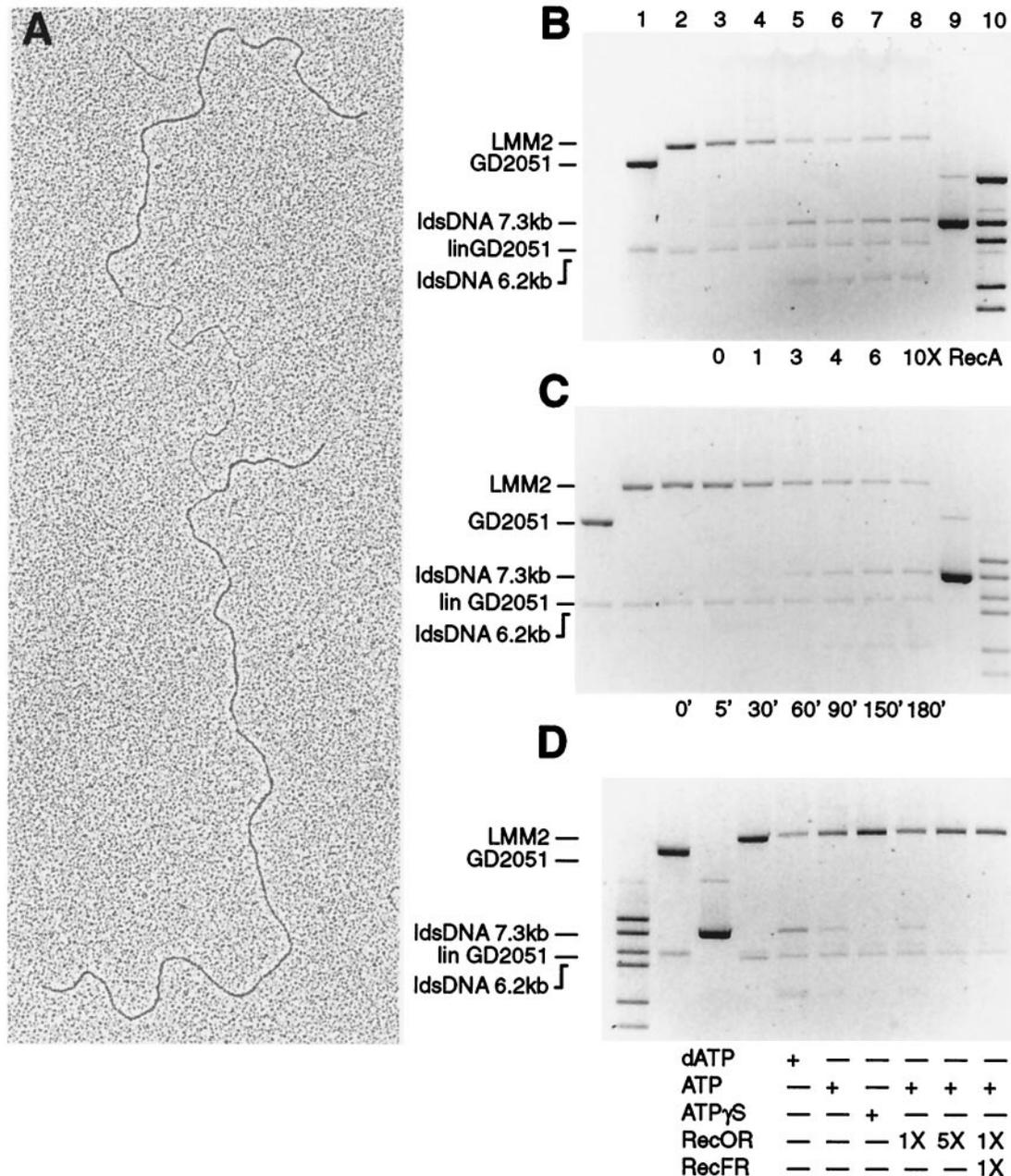


Fig. 6. RecA-promoted reactions of LMM2. (A) Electron micrograph of the LMM2 DNA substrate. (B) RecA protein titration in reactions with LMM2 DNA. Purified LMM2 at a concentration of $10.54 \mu\text{M}$ total nucleotides, corresponding to $0.93 \mu\text{M}$ single-stranded nucleotides in the gaps, was incubated with increasing concentrations of RecA protein. $1\times$ RecA protein represents $0.31 \mu\text{M}$ RecA. Lanes: 1, GD2051 marker; 2, purified LMM2 DNA; 3–8, reactions with increasing concentrations of RecA protein, as indicated; 9, linear dsDNA [1 M–2 M] marker; 10, lambda DNA–*Bst*II digest. (C) Time course of RecA protein reaction with LMM2. Purified LMM2 at a concentration of $10.6 \mu\text{M}$ total nucleotides, corresponding to $0.86 \mu\text{M}$ single-stranded nucleotides in the gaps, was incubated with $1.16 \mu\text{M} = 4\times$ RecA protein. Aliquots ($10 \mu\text{l}$) were removed from the reaction mixture at the indicated time points and analyzed by gel electrophoresis. Lanes (Left to Right): 1, GD2051 marker; 2, purified LMM2; 3–9, time points of the reaction as indicated; 10, linear dsDNA [1 M–2 M] marker; 11, lambda DNA–*Bst*II digest. (D) Nucleotide cofactor requirements and effect of RecORF proteins on RecA protein-promoted reaction of LMM2. Purified LMM2 at a concentration of $13.25 \mu\text{M}$ total nucleotides, corresponding to $1.1 \mu\text{M}$ single-stranded nucleotides in the gaps, was incubated with $1.08 \mu\text{M} = 3\times$ RecA under the specified conditions for 3 h. $1\times$ RecOR proteins represents RecO protein at 27.5 nM (1 RecO monomer/40 single-stranded nucleotides in the gaps), and RecR protein at 65 nM (1/17 nt). $1\times$ RecFR proteins represents RecF protein at a concentration of 27.5 nM (1 RecF monomer/40 single-stranded nucleotides), and RecR protein at 55 nM (1/20 nt). Thus in lane 10, the final concentration of RecR protein was 120 nM . Lanes (Left to Right): 1, lambda DNA–*Bst*II DNA markers (from top: 8,454, 7,242, 6,369, 5,686, 4,822, 4,324, 3,675 bp); 2, GD2051 marker; 3, ldsDNA [1 M–2 M] marker; 4, purified LMM2 DNA; 5–10, RecA reactions with LMM2, as indicated.

Reaction requirements were again explored and found to be similar to those observed for the MM2 substrate. Various RecA reactions with LMM2 were carried out with dATP, ATP, or ATP γ S (Fig. 6D). As with circular model molecules, the reaction with dATP is more efficient than the one with ATP (Fig. 6D, lanes 5 and 6). Also, ATP hydrolysis is required for the reaction, as ATP γ S

abolishes the RecA protein reaction (Fig. 6D, lane 7). At a low concentration, RecOR proteins do not exhibit a significant effect on the RecA reaction under our conditions (Fig. 6D, lane 8). When added at a higher concentration, RecOR proteins inhibited the reaction, probably by binding nonspecifically to DNA and preventing RecA-catalyzed fork strand exchange (Fig. 6D, lane 9). We have

tried using different concentrations of RecOR proteins, sometimes in combination with RecF protein, using MM2 instead of LMM2 and with different orders of addition. In three different experiments with both unpurified and purified LMM2, we did not observe any stimulation of the reaction by RecOR in the presence of ATP (data not shown). When RecF protein was added under conditions where RecOR did not have any effect on the RecA reaction, RecF protein inhibited the reaction completely (Fig. 6D, lane 10).

Discussion

RecA-Promoted Fork Regression *in Vitro*. Our primary conclusion is that RecA protein can promote fork regression in a model stalled replication fork *in vitro*. We designed a DNA molecule that mimics a structure observed for a replication fork stalled at a lesion in the leading strand (with a ssDNA gap on the leading strand). The existence of stalled forks with this structure has been postulated *in vivo* and has been demonstrated in the replication of DNA plasmids with a simian virus (SV)40 origin of replication with human cell extracts (10). The RecA-catalyzed fork regression we observed *in vitro* with these DNA substrates depends on ATP hydrolysis. RecO and RecR proteins do not stimulate this RecA-promoted reaction under our conditions. However, replacement of ATP with dATP has a stimulatory effect.

Spontaneous fork regression can occur only to a limited extent in our *in vitro* experiments with DNA MM1 (Fig. 3A). *In vivo*, it is possible that some of the numerous proteins present at the replication fork may further limit spontaneous fork regression. In many of our experiments, the spontaneous branch migration has been prevented by changing six base pairs in the arm of the model molecules (MM2). Fork regression with MM2 depends completely on RecA protein (Fig. 5A).

RecA protein reactions with MM2 resulted in an almost complete conversion to reaction products (Fig. 5). Most of this reaction is a four-strand exchange of a type that RecA protein actively drives in one direction when ATP is hydrolyzed (37). The reverse reaction would require dissociation of the RecA filament and its reassociation on the product molecule's tail, a process blocked in part by binding of the ssDNA segment by SSB. Reactions with the linearized LMM2 were somewhat less efficient, with just over 60% of the branches regressing in the best experiments (Fig. 6). With the linear molecules, there is potential for end-dependent RecA filament disassembly, which could reduce overall reaction efficiency. Indeed, we observed that dATP, which limits the disassembly process (23), provided a better reaction than ATP. This suggests that the end of the RecA filament must be at or near the DNA branch for a fork regression reaction to take place. For the circular MM2 DNAs, the RecA filament can always be near the branch by virtue of filament extension around the circle, ultimately straddling the branch. For the linear DNA substrates, the end of the RecA filament is less likely to be near the branch, and the filament is much less likely to straddle the branch.

For both MM2 and LMM2, the amount of RecA protein required to form a filament on the circle, or the arm with the ssDNA gap, respectively, is enough to promote maximum efficiency of the reaction (Figs. 5B and 6B). However, higher RecA protein amounts up to four times more than the minimum amount required for maximum efficiency did not inhibit the reaction under our conditions. When there is more RecA protein than that required to saturate the ssDNA in the gaps, the regression can clearly proceed beyond the gap to create Holliday junctions.

ATP hydrolysis is required for RecA-catalyzed fork regression (Figs. 5A and 6D). No reaction is seen at all with the model DNA substrates in the presence of ATP γ S, even though RecA can promote the formation of short regions of hybrid DNA under similar conditions in the classic three-strand exchange reaction (36, 38–40). RecA protein is a DNA-dependent ATPase. ATP hydrolysis is clearly coupled to the late phases of RecA protein-mediated DNA strand-exchange reactions *in vitro* (24, 33, 41, 42). ATP

hydrolysis enhances these DNA strand-exchange reactions in several ways, for example allowing the reaction to proceed past structural barriers and incorporate four DNA strands (11, 24, 36, 38, 39). The capacity of RecA to promote DNA strand exchange past barriers such as DNA lesions and heterologous inserts may be particularly important in the fork regression reaction. However, we note that there is no lesion or barrier in the model DNA substrates used in this study, other than the six mismatches present near the branch in MM2 and LMM2. ATP hydrolysis is not needed simply to bypass these mismatches in our model system, because MM1 (which has no mismatches) also exhibited no reaction with ATP γ S (although spontaneous branch migration could obscure a minimal reaction in this case; data not shown).

There are two major models for the mechanism by which the coupling between ATP hydrolysis and DNA strand exchange occurs. In one, ATP hydrolysis is coupled only to the dissociation of RecA filaments from the DNA (40, 43, 44). ATP hydrolysis can result in the disassembly of RecA filaments from one end (23). However, filament disassembly appears to be deleterious in the fork regression reaction we have observed in this study, and it has been noted elsewhere that much ATP hydrolysis takes place in the filament interior that does not result in RecA dissociation from DNA (11, 45). The alternative has been called the facilitated DNA rotation model. In this scenario, RecA protein is a motor protein, actively promoting DNA strand exchange and branch migration by the rotation of a DNA duplex around the homologous DNA molecule coated with the RecA filament (11), coupled to the hydrolysis of ATP. Several kinetic tests have provided evidence for this activity [(24, 33); T. A. Arenson and M.M.C., unpublished results], although the postulated motor has lacked an obvious *in vivo* function. We propose that this motor activity of RecA protein might be important for DNA pairing and branch migration to promote regression of a stalled replication fork. Regression of the fork requires a re-pairing of template DNA strands, and one must be wound around the other to recreate the helical duplex. A facilitated DNA rotation promoted by RecA protein, coupled to ATP hydrolysis, would provide the energy required to bring about fork regression in a cellular environment where the DNA had limited freedom of movement. An extensive regression reaction would also require the assistance of topoisomerases, but this would be true regardless of the mechanism by which regression occurred.

ATP hydrolysis would also be required to promote the four-strand exchange step in our *in vitro* reaction, leading to the formation of a Holliday intermediate. In the cell, this intermediate could then be processed by other enzymes such as the RuvABC system (15, 21, 46). ATP hydrolysis is clearly required for the four-strand DNA strand-exchange reactions promoted by RecA protein *in vitro* (24, 38). The facilitated DNA rotation model provides a mechanism to effect a four-strand exchange (11, 45). No attempt has been made to explain the requirement for ATP hydrolysis in four-strand exchange reactions within the context of the alternative RecA dissociation model.

If RecA protein is to promote fork regression, RecA filaments must be positioned in the DNA gap of a leading strand and maintained there. Some other factor or protein might be needed in the cell to prevent RecA filament disassembly and/or ensure RecA filament nucleation near the branch. Substitution of dATP for ATP seems unlikely to be a significant option *in vivo*. Complexes of the RecO and RecR proteins have been shown to stabilize the RecA filament in the presence of ATP by preventing a net end-dependent dissociation of RecA monomers (23). The RecF and RecR protein complexes attenuate *in vitro* extension of RecA filaments formed on gapped DNA molecules beyond the ssDNA gap (47). Modulation of RecA filament dynamics by the RecOR and RecFR complexes is thought to be important in the recombinational repair process by targeting and confining the RecA protein to the ssDNA gap that needs to be repaired

(47). However, the RecO and RecR proteins do not appear to be good candidates for such a regression-stimulating factor, because they had no effect on the fork regression reaction *in vitro*. Moreover, high concentrations of RecOR proteins inhibited the reaction, probably by binding nonspecifically to DNA and impeding RecA-catalyzed fork regression. When RecF protein was added with RecOR proteins, under conditions where RecOR did not have any effect on RecA-protein reaction, RecF inhibited the reaction (Fig. 6D). This is consistent with recent results of Michel *et al.*, who demonstrated that RecA protein, but not RecO or RecF, is required for certain *in vivo* reactions that

lead up to formation of an intermediate that can be processed by the RuvABC system (46). Although many questions remain about the protein transactions that might take place in replication fork repair, it is now clear that RecA protein can promote the potentially critical fork regression process in at least some situations.

We thank Maria Schnös for assistance with the EM experiments and David Dwyer for protein purification. This work was supported by Grants GM14711 (to R.B.I.) and GM52725 (to M.M.C.) from the National Institutes of Health.

1. Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J. & Marians, K. J. (2000) *Nature (London)* **404**, 37–41.
2. Kowalczykowski, S. C. (2000) *Trends Biochem. Sci.* **25**, 156–165.
3. Kuzminov, A. (1999) *Microbiol. Mol. Biol. Rev.* **63**, 751–813.
4. Marians, K. J. (2000) *Trends Biochem. Sci.* **25**, 185–189.
5. McGlynn, P. & Lloyd, R. G. (2000) *Cell* **101**, 35–45.
6. Higgins, N. P., Kato, K. & Strauss, B. (1976) *J. Mol. Biol.* **101**, 417–425.
7. Michel, B., Ehrlich, S. D. & Uzzest, M. (1997) *EMBO J.* **16**, 430–438.
8. Viguera, E., Hernandez, P., Krimer, D. B., Lurz, R. & Schwartzman, J. B. (2000) *Nucleic Acids Res.* **28**, 498–503.
9. Svoboda, D. L. & Vos, J. M. H. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 11975–11981.
10. Cordeiro-Stone, M., Makhov, A. M., Zaritskaya, L. S. & Griffith, J. D. (1999) *J. Mol. Biol.* **289**, 1207–1218.
11. Cox, M. M. (1999) *Prog. Nucleic Acids Res. Mol. Biol.* **63**, 310–366.
12. West, S. C., Cassuto, E. & Howard-Flanders, P. (1981) *Nature (London)* **294**, 659–662.
13. Postow, L., Ullsperger, C., Keller, R. W., Bustamante, C., Vologodskii, A. V. & Cozzarelli, N. R. (2001) *J. Biol. Chem.* **276**, 2790–2796.
14. Michel, B. (2000) *Trends Biochem. Sci.* **25**, 173–178.
15. Seigneur, M., Bidnenko, V., Ehrlich, S. D. & Michel, B. (1998) *Cell* **95**, 419–430.
16. Shibata, T., Das Gupta, C., Cunningham, R. P. & Radding, C. M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1638–1642.
17. Shibata, T., Das Gupta, C., Cunningham, R. P. & Radding, C. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2606–2610.
18. Kowalczykowski, S. C. (1994) *Experientia* **50**, 204–215.
19. Anderson, D. G. & Kowalczykowski, S. C. (1997) *Cell* **90**, 77–86.
20. Churchill, J. J., Anderson, D. G. & Kowalczykowski, S. C. (1999) *Genes Dev.* **13**, 901–911.
21. Seigneur, M., Ehrlich, S. D. & Michel, B. (2000) *Mol. Microbiol.* **38**, 565–574.
22. Register, J. C., III, & Griffith, J. (1985) *J. Biol. Chem.* **260**, 12308–12312.
23. Shan, Q., Bork, J. M., Webb, B. L., Inman, R. B. & Cox, M. M. (1997) *J. Mol. Biol.* **265**, 519–540.
24. Shan, Q., Cox, M. M. & Inman, R. B. (1996) *J. Biol. Chem.* **271**, 5712–5724.
25. Webb, B. L., Cox, M. M. & Inman, R. B. (1995) *J. Biol. Chem.* **270**, 31397–31404.
26. Lohman, T. M., Green, J. M. & Beyer, R. S. (1986) *Biochemistry* **25**, 21–25.
27. Craig, N. L. & Roberts, J. W. (1981) *J. Biol. Chem.* **256**, 8039–8044.
28. Webb, B. L., Cox, M. M. & Inman, R. B. (1999) *J. Biol. Chem.* **274**, 15367–15374.
29. Lohman, T. M. & Overman, L. B. (1985) *J. Biol. Chem.* **260**, 3594–3603.
30. Neuendorf, S. K. & Cox, M. M. (1986) *J. Biol. Chem.* **261**, 8276–8282.
31. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
32. Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276.
33. MacFarland, K. J., Shan, Q., Inman, R. B. & Cox, M. M. (1997) *J. Biol. Chem.* **272**, 17675–17685.
34. Inman, R. B. & Schnös, M. (1970) *J. Mol. Biol.* **49**, 93–98.
35. Littlewood, R. K. & Inman, R. B. (1982) *Nucleic Acids Res.* **10**, 1691–1706.
36. Jain, S. K., Cox, M. M. & Inman, R. B. (1994) *J. Biol. Chem.* **269**, 20653–20661.
37. Shan, Q. & Cox, M. M. (1998) *Mol. Cell* **1**, 309–317.
38. Kim, J. I., Cox, M. M. & Inman, R. B. (1992) *J. Biol. Chem.* **267**, 16444–16449.
39. Kim, J. I., Cox, M. M. & Inman, R. B. (1992) *J. Biol. Chem.* **267**, 16438–16443.
40. Menetski, J. P., Bear, D. G. & Kowalczykowski, S. C. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 21–25.
41. Bedale, W. A. & Cox, M. (1996) *J. Biol. Chem.* **271**, 5725–5732.
42. Nayak, S. & Bryant, F. R. (1999) *J. Biol. Chem.* **274**, 25979–25982.
43. Rehrauer, W. M. & Kowalczykowski, S. C. (1993) *J. Biol. Chem.* **268**, 1292–1297.
44. Kowalczykowski, S. C. & Krupp, R. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3478–3482.
45. Cox, M. M. (1994) *Trends Biochem. Sci.* **19**, 217–222.
46. Michel, B., Recchia, G. D., Penel-Colin, M., Ehrlich, S. D. & Sherratt, D. J. (2000) *Mol. Microbiol.* **37**, 180–191.
47. Webb, B. L., Cox, M. M. & Inman, R. B. (1997) *Cell* **91**, 347–356.