THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 52, Issue of December 26, pp. 52710–52723, 2003
Printed in U.S.A.

A DNA Pairing-enhanced Conformation of Bacterial RecA Proteins*

Received for publication, August 4, 2003, and in revised form, October 3, 2003
Published, JBC Papers in Press, October 6, 2003, DOI 10.1074/jbc.M308563200

Nami Haruta‡, Xiong Yu§, Shixin Yang§, Edward H. Egelman§, and Michael M. Cox‡¶

From the ‡Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706 and the §Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia 22908-0733

The RecA proteins of *Escherichia coli* (Ec) and *Deinococcus radiodurans* (Dr) both promote a DNA strand exchange reaction involving two duplex DNAs. The four-strand exchange reaction promoted by the DrRecA protein is similar to that promoted by EcRecA, except that key parts of the reaction are inhibited by Ec single-stranded DNA-binding protein (SSB). In the absence of SSB, the initiation of strand exchange is greatly enhanced by dsDNA-ssDNA junctions at the ends of DNA gaps. This same trend is seen with the EcRecA protein. The results lead to an expansion of published hypotheses for the pathway for RecA-mediated DNA pairing, in which the slow first order step (observed in several studies) involves a structural transition to a state we designate P. The P state is identical to the state found when RecA is bound to double-stranded (ds) DNA. The structural state present when the RecA protein is bound to single-stranded (ss) DNA is designated A. The DNA pairing model in turn facilitates an articulation of three additional conclusions arising from the present work. 1) When a segment of a RecA filament bound to ssDNA is forced into the P state (as RecA bound to the ssDNA immediately adjacent to dsDNA-ssDNA junction), the segment becomes "pairing enhanced." 2) The unusual DNA pairing properties of the *D. radiodurans* RecA protein can be explained by postulating this protein has a more stringent requirement to initiate DNA strand exchange from the P state. 3) RecA filaments bound to dsDNA (P state) have directly observable structural changes relative to RecA filaments bound to ssDNA (A state), involving the C-terminal domain.

The RecA protein of *Escherichia coli* (EcRecA)[1] is the prototype of a class of proteins playing a central role in the recombinational DNA repair in all organisms. The 352-amino acid polypeptide (M_r 37,842) functions as part of a helical nucleoprotein filament (1, 2). RecA protein promotes a DNA strand exchange reaction that reflects its major function in cellular recombination (Fig. 1*A*). The RecA filament forms on the circular single-stranded DNA (ssDNA). A linear duplex is then aligned with the bound ssDNA, and strand exchange ensues. This order of substrate binding, ssDNA first and dsDNA second, is characteristic of the DNA strand exchange reactions promoted by all RecA family proteins, with one exception treated below. Strand exchange can initiate on either end of the linear duplex, but the reaction is propagated uniquely 5′ to 3′ relative to the single-stranded DNA substrate (3–5). RecA protein is a DNA-dependent ATPase, with a k_{cat} under most conditions of 20–30 min^{-1}. DNA pairing is initiated, and considerable strand exchange can occur without ATP hydrolysis (*e.g.* with the use of ATPγS or RecA mutants that bind but do not hydrolyze ATP) (6–12). However, when ATP (or dATP) is hydrolyzed, the reaction becomes unidirectional and can bypass substantial structural barriers such as heterologous sequence insertions in one of the DNA substrates (6, 8, 12). The present work focuses on the initial DNA pairing that occurs during DNA strand exchange, as promoted by RecA and related proteins. A range of issues are addressed, and these require an introduction to five separate but related topics.

First, the single-stranded DNA-binding protein (EcSSB) has a major effect on all RecA reactions, playing a role in at least two steps in DNA strand exchange. SSB can both inhibit and enhance the RecA filament formation. The assembly of a filament on ssDNA occurs with distinct nucleation and extension steps, with the extension proceeding 5′ to 3′. In the absence of SSB, RecA protein nucleates well onto ssDNA but does not extend into regions containing secondary structure. The effect of SSB depends on when it is added. If bound to the DNA prior to addition of RecA, the SSB inhibits the nucleation step (13, 14). If SSB is added after RecA nucleation has occurred, the SSB facilitates the extension phase of the process, melting the regions of secondary structure before being displaced by RecA (13, 14). In addition, SSB plays an important role during DNA strand exchange, binding to the displaced strand of the duplex substrate as the reaction proceeds (15, 16). The stimulatory effects of SSB can be partially compensated for by forming RecA filaments on ssDNA under conditions in which little secondary structure exists, and then adding the Mg^{2+} needed for efficient reaction (a Mg^{2+}-shift protocol) (17, 18).

Second, The RecA protein from *Deinococcus radiodurans* (DrRecA) promotes DNA strand exchange a little differently than a typical RecA recombinase. *D. radiodurans* is part of a small family of bacterial species that are the most radiation-resistant organisms known (19–21). Although this organism efficiently repairs many types of DNA damage (21), resistance to ionizing radiation is especially remarkable. *D. radiodurans* is able to survive exposures to γ radiation of 1.7 megarads without lethality or induced mutation (22), and can survive hundreds of irradiation-induced DNA double-stranded breaks per haploid genome (22, 23). The mechanisms underlying the very efficient DNA repair of this organism remain poorly understood (24). A functional DrRecA protein is required for expression of the radiation-resistant phenotypes (25).

* This work was supported by National Institutes of Health Grants GM32335 (to M. M. C.) and GM35269 (to E. H. E.). 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 should be addressed: Dept. of Biochemistry, University of Wisconsin, 433 Babcock Dr., Madison, WI 53706-1544. Tel.: 608-262-1181; Fax: 608-265-2603; E-mail: cox@biochem.wisc.edu.

[1] The abbreviations used are: EcRecA, RecA protein of *E. coli*; ss, single-stranded; ds, double-stranded; SSB, single-stranded DNA-binding protein; ATPγS, adenosine 5′-*O*-(thiotriphosphate); LDS, linear Duplex substrate; css DNA, circular ssDNA; AMP-PNP, adenosine 5′-(β,γ-imino)triphosphate; GD, gapped duplex; nt, nucleotides; DrRecA, RecA protein of *D. radiodurans*.

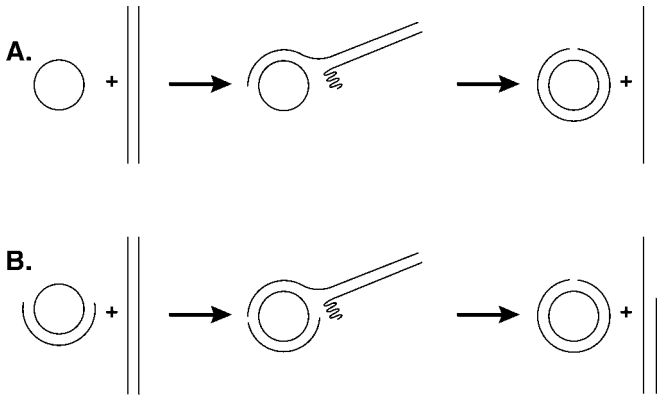

FIG. 1. **DNA strand exchange reaction promoted by bacterial RecA proteins.** The DNA substrates are derived from bacteriophage DNAs. *A*, in the 3-strand reaction, a RecA filament is formed on the circular ssDNA. The linear duplex DNA is then aligned with the bound single strand, and exchange is initiated to form a branched intermediate. Strand exchange is propagated unidirectionally, 5′ to 3′ relative to the ssDNA in the circle, until the nicked circular product is generated, with release of the displaced single strand. For the DrRecA protein, the inverse pathway, in which the filament first forms on the linear duplex DNA, predominates (27). *B*, in the 4-strand reaction, the filament forms on the gapped circular DNA. DNA strand exchange is initiated in the gap as a 3-strand reaction. The remainder of the reaction proceeds as in *A*, except that progression through the 4-stranded portion of the reaction exhibits an absolute requirement for ATP hydrolysis (7, 12, 38).

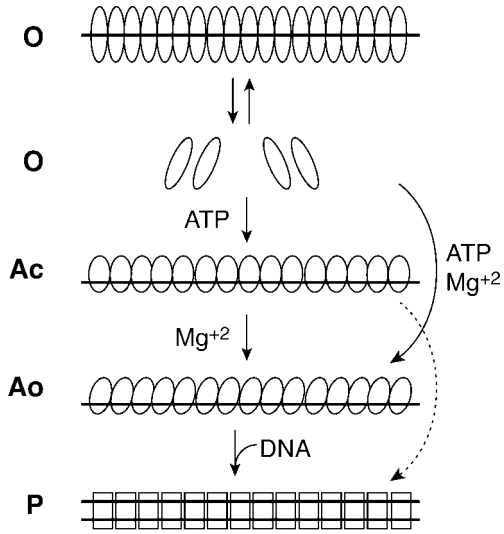

FIG. 2. **Structural states of RecA protein.** Based primarily on functional distinctions, and in part on structural distinctions evident in electron micrographs, there are at least four significant structural states of the RecA protein. These are designated O, Ac, Ao, and P. The O state is the one existing in the absence of nucleotide cofactor and with bound ADP. This state will bind to ssDNA. In the presence of ATP, RecA will bind to ssDNA and acquire an activated (A) structural state. Although the O and A states both bind to ssDNA, they are not interconvertible while bound to ssDNA (41, 43). There are at least two forms of the A state. Ac is present at low free Mg^{2+} concentrations and is "closed" to most DNA pairing processes (40). The Ao state is open to DNA pairing with a wide range of DNA substrates and is obtained *in vitro* in the presence of at least 6–8 mM free Mg^{2+} (40). The P state is obtained when a second complementary DNA strand is introduced to the filament, either by DNA strand exchange or by direct binding to duplex DNA. The P state is characterized by a higher degree of cooperativity and lower rates of ATP hydrolysis than the A states. The A and P states all hydrolyze ATP and probably cycle through multiple conformational states as they do so (54). There may also be multiple forms of the O state, depending on the presence of bound ADP and/or ssDNA.

The DrRecA protein (362 amino acids, M_r 38,013) has an amino acid sequence that is 57% identical (72% similar) to that of the EcRecA protein. In most respects, the DrRecA protein is similar to other bacterial RecA proteins. It forms helical filaments on DNA, hydrolyzes ATP and dATP, and promotes DNA strand exchange (26). It does bind to dsDNA much faster than the EcRecA protein. However, the most substantial difference documented to date concerns the pathway for DNA strand exchange, where the order of substrate binding changes in reactions promoted by the DrRecA protein. In the reaction of Fig. 1*A*, the filament is formed first on the duplex DNA, and the single-stranded DNA is then brought in for exchange (27).

Third, the initial alignment and pairing of the two DNA molecules by the EcRecA protein and the eukaryotic Rad51 protein has been the focus of numerous studies, reviewed in some detail elsewhere (2, 28). There are several distinct kinetic steps in the pairing process, as explored by the Camerini-Otero and Radding groups (29–31). A theme of the kinetic studies is the existence of a fast second order step followed by a slower first order step in the pairing and exchange pathway (29–31), with both steps being reversible in the presence of ATP (30, 31). Also important is another contribution by Radding and coworkers (32) that, when combined with earlier work, provides a good picture of the likely pairing process. The effort focuses on the human Rad51 protein, but we assume that the DNA pairing process is similar in Rad51 and RecA. In brief, the duplex DNA probably approaches the bound single strand via the minor groove in the duplex. The homology search likely involves base flipping, with A:T base pairs playing a particularly prominent role. DNAs with high GC content block or reduce DNA strand exchange involving short oligonucleotides (32).

Fourth, there is an important contrast between the 3- and 4-strand exchange processes diagrammed in Fig. 1. The *E. coli* RecA protein will promote a 4-strand exchange reaction (Fig. 1*B*). One of the duplex substrates must have a single-strand gap. The gap serves as both a nucleation site for RecA filament formation and as an initiation site for DNA strand exchange. The second linear duplex must overlap the gap by at least 50–100 bp for an efficient strand exchange to occur (33, 34). These reactions are also unidirectional, proceeding 5′ to 3′ relative to the ssDNA in the gap (35–38). The 4-stranded portion of the reaction also requires ATP hydrolysis. When ATP is not hydrolyzed, reactions are initiated in the gap but do not proceed beyond the gap region (7, 12). Even though the exchange is largely between two duplex DNAs, SSB plays a significant role in the 4-strand reaction. Some early reports indicated that SSB did not stimulate the 4-strand exchange reactions (35, 39). However, subsequent studies found that SSB greatly enhanced most 4-strand exchange processes (34), and SSB has generally been included in these reactions (7, 12, 34, 37, 38). The apparent requirement for SSB in many but not all 4-strand exchange reactions has never been explained.

Fifth, there is now evidence for at least four different functional states of RecA protein, occurring at different reaction stages. These have recently been designated O, Ac, Ao, and P (1, 40) (Fig. 2).

The O state is largely inactive and is found in the absence of nucleotide cofactor or in the presence of ADP (41). This form can bind to DNA, creating a helical filament with a pitch of 76 Å (41). There may be multiple forms of this state, depending on binding to ADP and/or ssDNA (42, 43).

Addition of ATP, ATPγS, or dATP results in a conformation change to an active form that is manifested by extended filaments on DNA with a pitch of 95 Å (41–49). This form is given the general designation A, and this is the form that hydrolyzes ATP with a k_{cat} of about 30 min^{-1} when bound to ssDNA. The A and O states are not directly interconvertible, such that RecA bound to ssDNA in the O state must dissociate and re-bind in the A state (41–43). There are two forms of the "A" state. A closed form, found at relatively low Mg^{2+} concentrations, has a

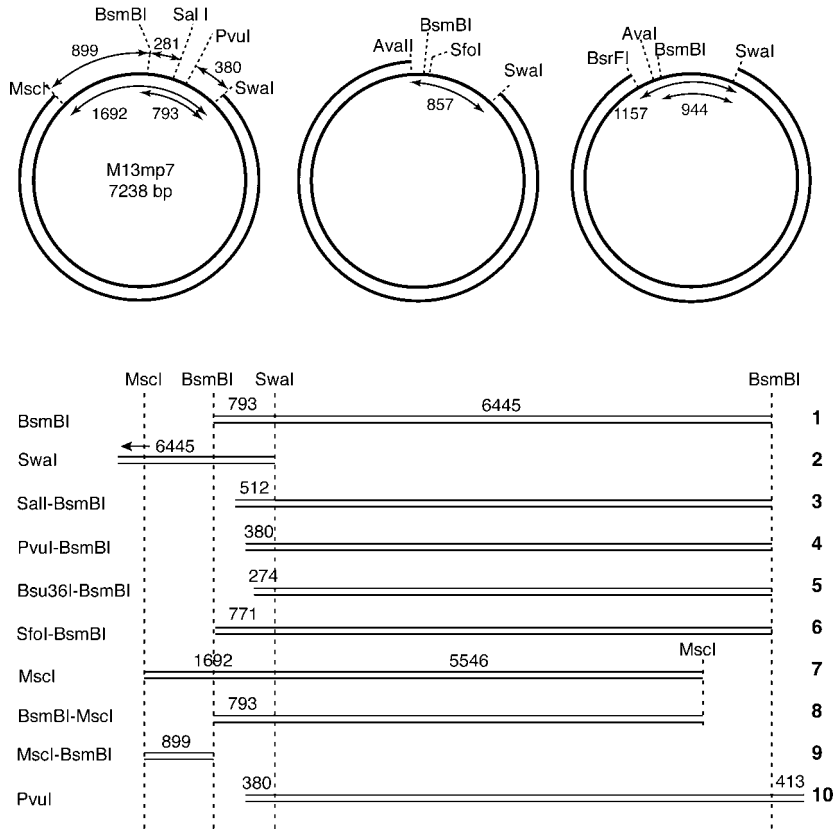

FIG. 3. **DNA substrates used in this study.** A range of gapped DNAs and linear duplex DNAs were constructed for this work, and all are illustrated here. All are based on the bacteriophage M13mp7. The restriction sites listed are those that represent either gap boundaries or linear duplex ends in various constructs. *Numbers 1–10 to the right* of the linear duplexes are used in subsequent figure legends to better identify the substrates employed in individual experiments. Lengths of various segments, in bp or nucleotides, are given *below the arrows* denoting the segment in the circular gapped duplex illustrations, or *above the lines* in the linear duplex illustrations.

restricted capacity for the pairing of the bound ssDNA to homologous duplexes and is thus designated Ac (for closed form of A) (40). In this state, the DNA pairing is enhanced if the linear duplex has a single-strand extension at the end (40). An open form of A (Ao) is found at higher Mg^{2+} concentrations (6–8 mM above the ATP concentration) and is much more permissive in DNA pairing reactions (40). The Ac to Ao transition involves the C terminus of RecA protein (40). Besides the DNA pairing modulation, many properties of Ac and Ao appear to be similar, such as the capacity to hydrolyze ATP. However, there is additional evidence that Mg^{2+} concentrations affect RecA function and structural state. The contour lengths of RecA-ssDNA nucleoprotein filaments formed in 1 mM magnesium are 116–120% relative to duplex DNA in the presence of 1 mM ATPγS (50) and 137% relative to duplex DNA in the presence of 1.3 mM ATP (17), which translates to less filament extension than the 150% extension observed when the filaments are formed with 10 mM magnesium ion (41, 44, 51). Elevated Mg^{2+} levels also result in the formation of RecA filament bundles observable by electron microscopy under some conditions (44, 45, 52). Finally, with elevated Mg^{2+} levels, RecA protein better resists displacement by SSB (13). Both states, Ac and Ao, hydrolyze ATP at similar rates. Image reconstructions from electron micrographs have shown that RecA, particularly the C-terminal domain, undergoes significant conformation changes in response to ATP hydrolysis (53, 54). We anticipate that both A states exist in a family of conformations, with the exact conformation present in a certain protomer dependent on the stage of the ATP hydrolytic cycle.

When a second strand of DNA is present, either by direct binding to dsDNA or as a result of DNA strand exchange, another structural change occurs to a state we designate P (1). The evidence for this is functional and indirect, but nonetheless extensive. The P structural form is characterized by 30% lower rates of ATP hydrolysis (55, 56), higher rates of exchange of RecA monomers into and out of the filament (57, 58), and a higher degree of cooperativity (57, 59) than the A conformations.

A distinction between the A and P states can also be seen in the RecA-stimulated autocatalytic cleavage of the LexA repressor (60). On ssDNA (the A state), RecA protein stimulates LexA cleavage, whereas stimulation is much reduced when RecA is bound to dsDNA (60). Most studies of RecA protein-mediated stimulation of LexA cleavage have been carried out in the presence of relatively low Mg^{2+} concentrations (61, 62), so that the Ac form of RecA may be the most active in this reaction.

The present study began as an effort to explore further the DNA pairing properties of the DrRecA protein, specifically focused on 4-strand exchange reactions in an attempt to explain the unusual DNA pairing pathway of the protein. The study evolved into an investigation of fundamental DNA pairing mechanisms, although the results do suggest an explanation for the unique DNA pairing properties of the DrRecA protein. The work provides some data consistent with the observations of Radding and co-workers (32) that A:T base pairs are important for the initiation of DNA pairing and extends the observation to RecA protein. In addition, evidence is provided that the P conformation of the RecA protein does not simply occur as a result of DNA strand exchange but is instead a conformation that greatly facilitates the process of DNA strand exchange.

The primary observation around which these conclusions are built is that a 4-strand exchange reaction is greatly enhanced if the end of the linear DNA coincides with the very end of the gap in the gapped DNA substrate. On a gapped DNA, the RecA filament is in effect a mixture of segments in the A conformation (in the gap) and in the P conformation (on the duplex DNA). We hypothesize that the P conformation, which is the more cooperative of the two, extends for a short distance into the gap, leading to a large enhancement of DNA pairing near the dsDNA-ssDNA junction at the end of the gap. Finally, the structural alterations that relate the A and P forms of RecA

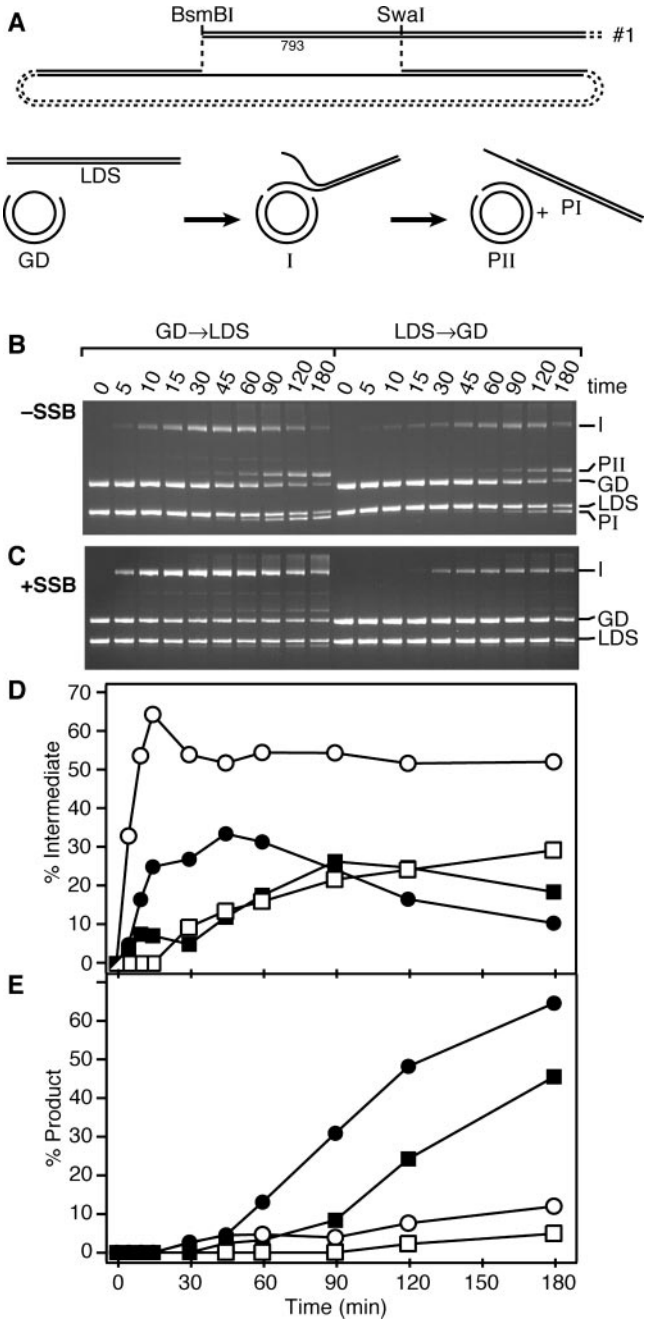

FIG. 4. **The RecA protein of *D. radiodurans* promotes a 4-strand exchange reaction.** Reactions were carried out as described under "Experimental Procedures" and contained 1.8 μM DrRecA protein, 0.3 μM EcSSB protein (*C* only), 10 μM gapped DNA, with a 793-nucleotide gap as defined in *A*, and 10.6 μM linear DNA (*substrate 1* in Fig. 3). The two DNA substrate concentrations are designed to be equivalent in terms of total molecules. The DrRecA concentration is designed to be stoichiometric with the available DNA-binding sites (3 nucleotides or bp per site) on one but not both DNA substrates. The DrRecA protein was preincubated with either the gapped DNA (GD → LDS reactions) or linear duplex (LDS → GD reactions) for 10 min at 37 °C. ATP (2 mM) and SSB (*B* only) were added, and incubation was continued for another 40 min before addition of the second DNA substrate. Aliquots were taken at the indicated times, deproteinized with SDS and proteinase as described under "Experimental Procedures," and subjected to electrophoresis on a 1% agarose gel. DNA species labeled in this and subsequent figures are identified in the reaction schematic in *A*. The GD → LDS reaction order produces the better reactions. Note the high concentration of reaction intermediates that are only slowly converted to products in the GD → LDS reaction carried out in the presence of SSB. Reactions are quantified in *D* and *E*. The % intermediates are defined as [I]/([I] + [PII] + [GD]). The % products are defined as [PII]/([I] + [PII] + [GD]). As a rule, the PI and LDS bands were not included in the calculations because these bands were too close for useful quantitation in some experiments. The GD → LDS reactions are indicated with *circles*, and the LDS → GD reactions are denoted with *squares*. *Closed* and *open symbols* indicate reactions carried out without and with the addition of SSB, respectively.

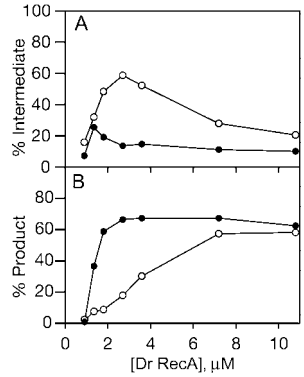

FIG. 5. **Effects of DrRecA protein concentration on the inhibition of 4-strand exchange reactions by EcSSB protein.** Reactions were carried out as described under "Experimental Procedures" and contained the indicated concentration of DrRecA in the absence (*closed circles*) or presence of 0.3 μM EcSSB protein (*open circles*). The DNA substrates and their concentrations are as in Fig. 4. The 1.8 μM DrRecA is sufficient to fully cover one but not both DNA substrates. Gapped DNA was incubated with RecA for 10 min at 37 °C followed by addition of 2 mM ATP and SSB (*open circles*). After 40 min of preincubation, the linear dsDNA was added to initiate the reaction. The reactions were carried out for 120 min. *A* is the yield of intermediate ([I]/([I] + [PII] + [GD])), and *B* is the yield of product PII ([PII]/([I] + [PII] + [GD]) plotted against the DrRecA concentration. The reaction without SSB exhibited a slight generation of network products retained in the well (data not shown), but these are not estimated here.

protein (on ssDNA *versus* on dsDNA, respectively) are explored via the use of image reconstruction from electron micrographs.

EXPERIMENTAL PROCEDURES

Enzymes and Reagents—The EcRecA and DrRecA proteins were purified by polyethyleneimine precipitation followed by a DEAE-Sepharose column (Amersham Biosciences) and a hydroxyapatite column (Bio-Rad) as described (51). Protein concentrations were determined by absorbance at 280 nm using the extinction coefficients $\epsilon_{280} = 2.23 \times 10^4$ M^{-1} cm^{-1} for EcRecA (63), 1.41×10^4 M^{-1} cm^{-1} for DrRecA (26), and 2.38×10^4 M^{-1} cm^{-1} for *E. coli* single-stranded DNA-binding protein (SSB) (64). Restriction enzymes were purchased from New England Biolabs. EcSSB and all other reagents were obtained from Sigma unless otherwise described. Glycerol, Tris buffer, and CsCl were purchased from Fisher.

Preparation of Gapped DNA and Linear Duplex (LDS) DNA Substrates—Duplex supercoiled DNA and circular ssDNA (css DNA) substrates were derived from bacteriophage M13mp7 (7238 bp). Supercoiled DNA and css DNA were purified by the CsCl banding method with the following modifications (65, 66). Supercoiled DNA was purified using a QIAfilter Plasmid Mega Kit (Qiagen) and then isolated once in 1.59 g/ml CsCl by centrifugation for 20 h at 55,000 rpm in a Beckman Vti 65.2 rotor at 20 °C. For css DNA purification, phage particles were purified by precipitation with polyethylene glycol (Acros Organics) followed by twice spinning in a Beckman SW 41 Ti rotor at 28,000 rpm for 18 h at 25 °C.

Gapped duplex DNA substrates were prepared by large scale RecA protein-mediated three-strand exchange reactions between circular ssDNA M13mp7 and linear dsDNA as described (12). The linear dsDNA fragment was prepared by complete digestion of supercoiled M13mp7 with appropriate restriction enzymes followed by purification by size exclusion chromatography (Sephacryl S-500, Amersham Biosciences) or by electrophoresis on 1.0% agarose gel. The gapped DNA product was isolated by electrophoresis in a 1.0% agarose (Genemate) gel and electroeluted in the dialysis bag (Pierce) followed by concentration through Centricon 30 (Millipore).

The linear dsDNA substrates were prepared from M13mp7 supercoiled dsDNA, which were cut with appropriate restriction endonucleases using conditions recommended by the supplier. The LDS DNA substrates that were not full length were separated from other frag-

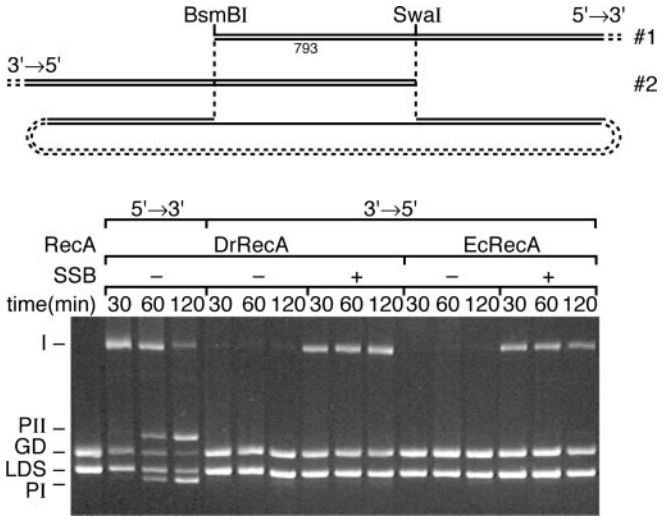

FIG. 6. **The 4-strand exchange reaction promoted by the DrRecA protein proceeds uniquely 5′ to 3′ relative to the single-stranded DNA in the gap.** Reactions were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA or EcRecA protein (as indicated), 0.6 μM EcSSB protein (where indicated), 10 μM gapped DNA, with a 793-nucleotide gap as defined in the *top illustration*, and 10.6 μM linear DNA (*substrate 1* in Fig. 3 for the 5′ to 3′ reactions and *substrate 2* for the 3′ to 5′ reactions). Aliquots were taken at the times indicated and processed as described under "Experimental Procedures" and the legend to Fig. 4 except for the use of a 0.8% agarose gel. The labels for the various DNA bands are identified in Fig. 4. Note the complete absence of reaction products in the 3′ to 5′ reactions.

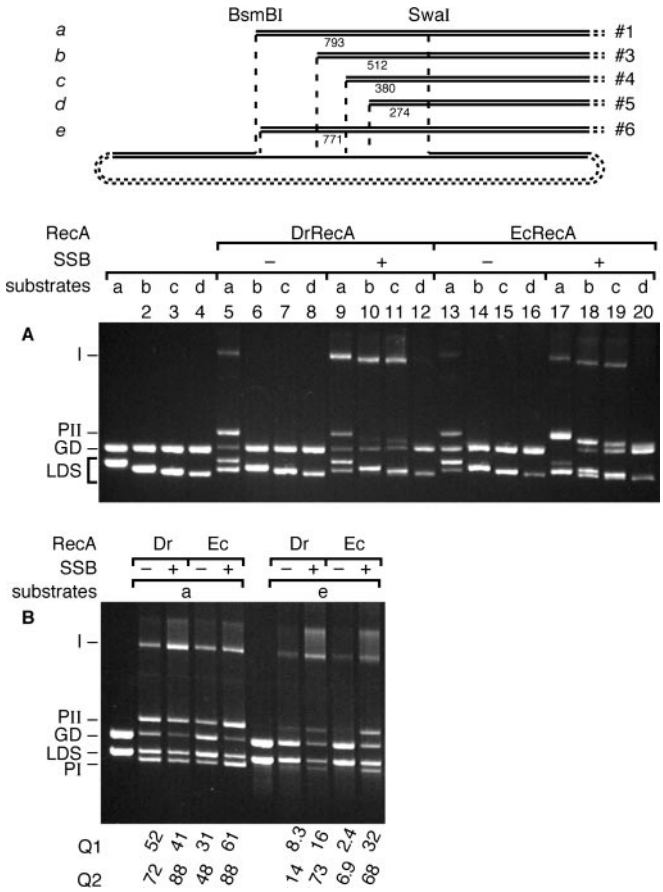

FIG. 7. **The 4-strand exchange reactions of bacterial RecA proteins are greatly enhanced if the linear duplex end coincides with the dsDNA-ssDNA junction in the gapped DNA substrate.** Reactions were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA or EcRecA protein (as indicated), 0.6 μM EcSSB protein (where indicated), 10 μM gapped DNA, with a 793-nucleotide gap as defined in the *top illustration*, and 9.7–10.6 μM linear DNA. The linear DNA concentrations were adjusted according to the length of the duplexes to provide a 1:1 molar match with the gapped DNA in terms of total molecules. The linear DNAs in reactions (*a–e*) are *substrates 1* and *3–6* in Fig. 3, respectively. Following preincubation (see legend to Fig. 5), reactions were carried out for 120 min. All other reaction details are as in the legend to Fig. 5. *A* and *B* are the same except for the DNA substrates used. The reactions in *B* were quantitated, with results presented at the bottom of the gel as % final products (*Q1*, [PII]/([I] + [PII] + [GD])) or % total paired DNAs (*Q2*, [I] + [PII]/([I] + [PII] + [GD])).

ments by size exclusion chromatography or gel purification as descried above and tested for nuclease contamination.

The concentrations of ssDNA and dsDNA solutions were determined by absorbance at 260 nm, using 36 and 50 $\mu g\ ml^{-1}\ A_{260}^{-1}$, respectively, as conversion factors. The concentrations of DNA and proteins reported below are the final concentrations after addition of all components, and DNA concentrations are in terms of total nucleotides.

DNA Strand Exchange Reaction—Unless otherwise stated, both three- and four-strand exchange reactions were carried out at 37 °C in the solutions containing 20 mM Tris-HCl (80% H^+, pH 7.5), 10 mM magnesium chloride (Mallincraft), 2 mM dithiothreitol (Research Organics), 5% (w/v) glycerol, and an ATP-regeneration system (8 mM phosphocreatine and 8 units/ml phosphocreatine kinase (Roche Applied Science)) based on the protocol described previously (38) with a few modifications. Typically, 5 μM RecA and 10 μM gapped DNA were preincubated in the reaction buffer and regeneration system for 10 min before addition of 2 mM ATP. Where indicated, SSB was also added at this time point. The reaction was started by addition of equivalent amounts in molecules of LDS DNA and gapped DNA (10.6 μM for LDS DNA cut by BsmBI) after 20 min of incubation with ATP. The reaction was incubated for 2 h, unless otherwise specified. Reaction aliquots (9.5 μl) were deproteinized by addition of 1.2 μl of 10% SDS, 0.3 μl of 0.5 M EDTA, and 0.6 μl of 20 mg/ml proteinase K and incubated for 30 min at 37 °C. Aliquots mixed with 2.5 μl of 6× loading buffer (15% Ficoll, 0.25% bromphenol blue, 0.25% xylene cyanol FF) were loaded on a 0.8 or 1% agarose gel and electrophoresed at 25–35 V for 16 h at room temperature. To visualize the DNA bands, the gels were stained with ethidium bromide and exposed to ultraviolet light. Gel images were captured with a digital CCD camera utilizing GelExpert software (Nucleotech). When indicated, the intensity of DNA bands was quantitated with the software package TotalLab version 1.10 from Phoretix.

Electron Microscopy and Image Reconstructions—The RecA-dsDNA-AMP-PNP reconstruction was described previously (54). The RecA-ssDNA-AMP-PNP reconstruction was generated using identical conditions but with ssDNA instead of dsDNA. A total of 14,604 filament segments were selected from the images of the RecA-ssDNA filaments, and these were sorted by pitch. A subset of 2,525 segments having a pitch of ~90 Å was then used for the reconstruction shown. Multiple cycles of the Iterative Helical Real Space reconstruction method (67) converged to a symmetry of 6.14 subunits per 89-Å pitch turn for this subset.

RESULTS

We first characterize briefly the 4-strand exchange reaction promoted by the DrRecA protein. The study then introduced two new factors affecting the efficiency of 4-strand reactions promoted by both the EcRecA and DrRecA proteins. This study required the construction of a wide variety of circular gapped duplex DNA (GD) and linear double-strand DNA substrates, all of which are detailed in Fig. 3. Substrate and product DNA bands on the gels are labeled as described in the scheme presented in Fig. 4. Simplified schematics of the DNA substrates used in particular experiments are provided in each figure. Where SSB is employed, it is in all cases the SSB from *E. coli* (EcSSB). To directly explore the structural basis for the different properties of RecA filaments bound to ssDNA *versus* dsDNA, we have made use of image reconstruction from electron micrographs.

The DrRecA Protein Promotes a 4-Strand Exchange Reaction—In the 3-strand exchange reaction promoted by the DrRecA protein, the duplex DNA is bound first and the ssDNA second in the predominant reaction pathway (27). A 4-strand

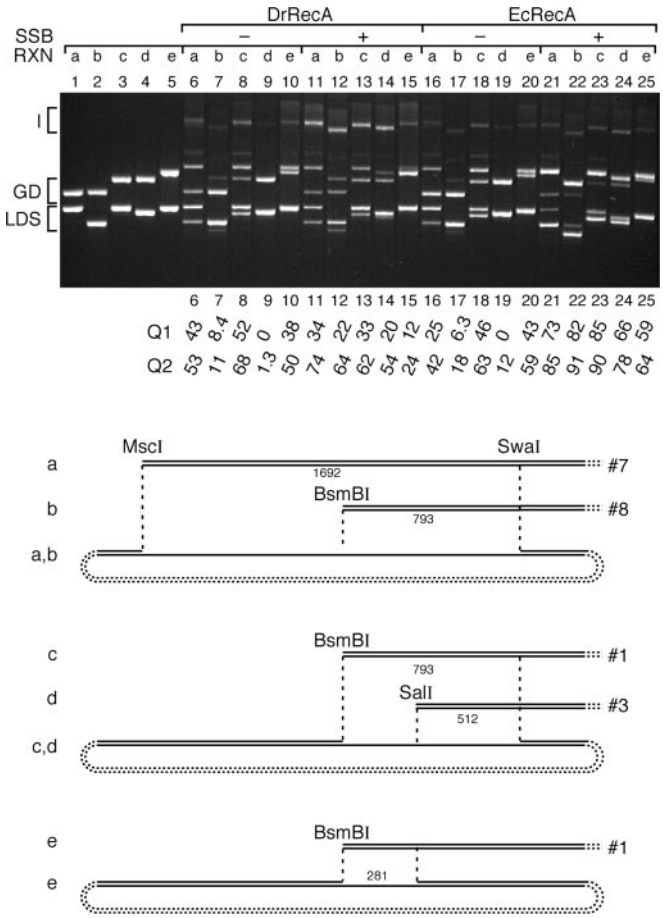

FIG. 8. **The coincidence of the end of the linear duplex with the dsDNA-ssDNA junction in the gapped DNA substrate, rather than extent of gap overlap, is the key factor in overall reaction enhancement.** Reactions were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA or EcRecA protein (as indicated), 0.6 μM EcSSB protein (where indicated), 10 μM gapped DNA, with the various gaps defined in the bottom illustration, and 10.1–11.2 μM linear DNA (adjusted in each reaction to provide a 1:1 molar match with the gapped DNA in terms of total molecules). The linear DNAs in reactions (*a–e*) are *substrates 7*, *8*, *1*, *3*, and *1* in Fig. 3, respectively. Following preincubation (see legend to Fig. 5), reactions were carried out for 120 min. All other reaction details are as in the legend to Fig. 5. A 1% agarose gel was used in this experiment. The reactions were quantitated, with results presented at the bottom of the gel as % final products (*Q1*, [PII]/([I] + [PII] + [GD])) or % total paired DNAs (*Q2*, [I] + [PII]/([I] + [PII] + [GD])).

exchange reaction was carried out, using the DNA substrates illustrated (Fig. 4*A*). Two different orders of addition were tried. In one, the DrRecA protein was preincubated with the gapped duplex substrate and ATP for 40 min prior to addition of the linear duplex. In the other, the DrRecA protein was instead preincubated with the linear duplex. The results are shown in Fig. 4*B*. The DrRecA protein promotes a substantial 4-strand exchange reaction with both protocols. However, the preincubation with the GD leads to a faster reaction and somewhat greater final product yield over this time course. Based on the observations from the 3-strand reaction, where the DrRecA protein binds first to the duplex DNA (27), this result was somewhat unexpected.

Also surprising was the effect of EcSSB. In the 3-strand reaction, addition of EcSSB has a very large enhancing effect on DrRecA-promoted reactions (26). In contrast, the 4-strand reaction was suppressed by EcSSB (Fig. 4*C*). The effects are perhaps most instructive when the DrRecA protein is preincubated with the gapped duplex. Here much of the substrate DNA is converted into intermediates, but fewer of the intermediates are resolved to products. The time courses for the DrRecA protein are quantitated in Fig. 4, *D* and *E*. The fraction of the DNA substrates converted to pairing intermediates is actually increased slightly with SSB, but the yield of final products is reduced by 6-fold after 180 min of reaction. This suggests that the SSB is enhancing the initiation of the reaction (the 3-strand segment) but interfering with the later reaction stages in which two duplexes are undergoing exchange.

The effect of EcSSB is further explored in Fig. 5. The concentration of DrRecA protein in the reactions of Fig. 4 is just enough to saturate the gapped DNA or the linear duplex DNA substrate but not both. In the Fig. 5 experiment, the concentration of DrRecA protein is varied to see if the effect of EcSSB can be moderated. At relatively low concentrations of DrRecA protein, the EcSSB protein dramatically inhibits the formation of full reaction products, with greatly elevated levels of reaction intermediates observed as in Fig. 4. Increasing the DrRecA concentration moderated the effects of EcSSB, but the generation of final products was not comparable with that seen in the absence of EcSSB until the DrRecA protein was present in about 2-fold excess relative to that required to bind all of the DNA in the reaction mixture. In subsequent experiments, the DrRecA concentration is increased somewhat relative to that in Fig. 4, to generate a more readily visible yield of products in the presence of EcSSB.

Also in subsequent DNA strand exchange assays, selected results are quantified in two ways. First, the yield of all paired DNAs, intermediates, and products is reported as a percentage of the total DNA in a given lane. Second, the percentage of the total DNA present as completed products of DNA strand exchange (PII) is reported.

DrRecA Protein-promoted 4-Strand Exchange Reactions Proceed 5′ to 3′ Relative to the ssDNA in the Gap of the GD—The reaction of Fig. 4 was compared with a similar one in which the linear dsDNA was cut with SwaI and thus overlapped the gap from the opposite side. The 4-strand exchange with this new linear dsDNA, after initiation in the gap, would have to proceed 3′ to 5′ relative to the ssDNA in the gap, whereas the reaction of Fig. 4 is designed to go 5′ to 3′. Results are provided in Fig. 6. The 5′ to 3′ reaction proceeds efficiently, as before. The 3′ to 5′ reaction generates intermediates (at least when SSB is provided) but no detectable products. Without SSB to stabilize the intermediates resulting from the 3-strand portion of the reaction, essentially no reaction is observed in this "backwards" reaction. The full 4-strand exchange reaction is thus unidirectional, proceeding 5′ to 3′. This is the same direction as the EcRecA-promoted reaction, and results with this protein and the SwaI-cleaved linear dsDNA are provided in Fig. 6 for comparison. For both proteins, the initiation of a 3-strand exchange can occur on either end of the linear duplex, but the reaction is propagated through the 4-stranded segment unidirectionally. The results also demonstrated that a 4-strand exchange reaction must be preceded by a 3-strand exchange reaction, for the DrRecA-promoted process as it is for the EcRecA protein-promoted reaction.

The dsDNA-ssDNA Junction at the End of the Gap Has a Large Stimulatory Effect on 4-Strand Exchange Reactions—We next carried out a reaction to determine how much overlap was required between the linear duplex DNA and the gap in the GD substrate. In the reactions of Figs. 4 and 5, the overlap for the BsmBI-cleaved linear DNA is 793 bp. We will refer to this as the full overlap substrate in experiments where, as here, the end of the linear duplex exactly coincides with the dsDNA-ssDNA junction at the end of the gap. A series of reactions were carried out with linear dsDNAs with different end points within the gap (the opposite end, where strand exchange is

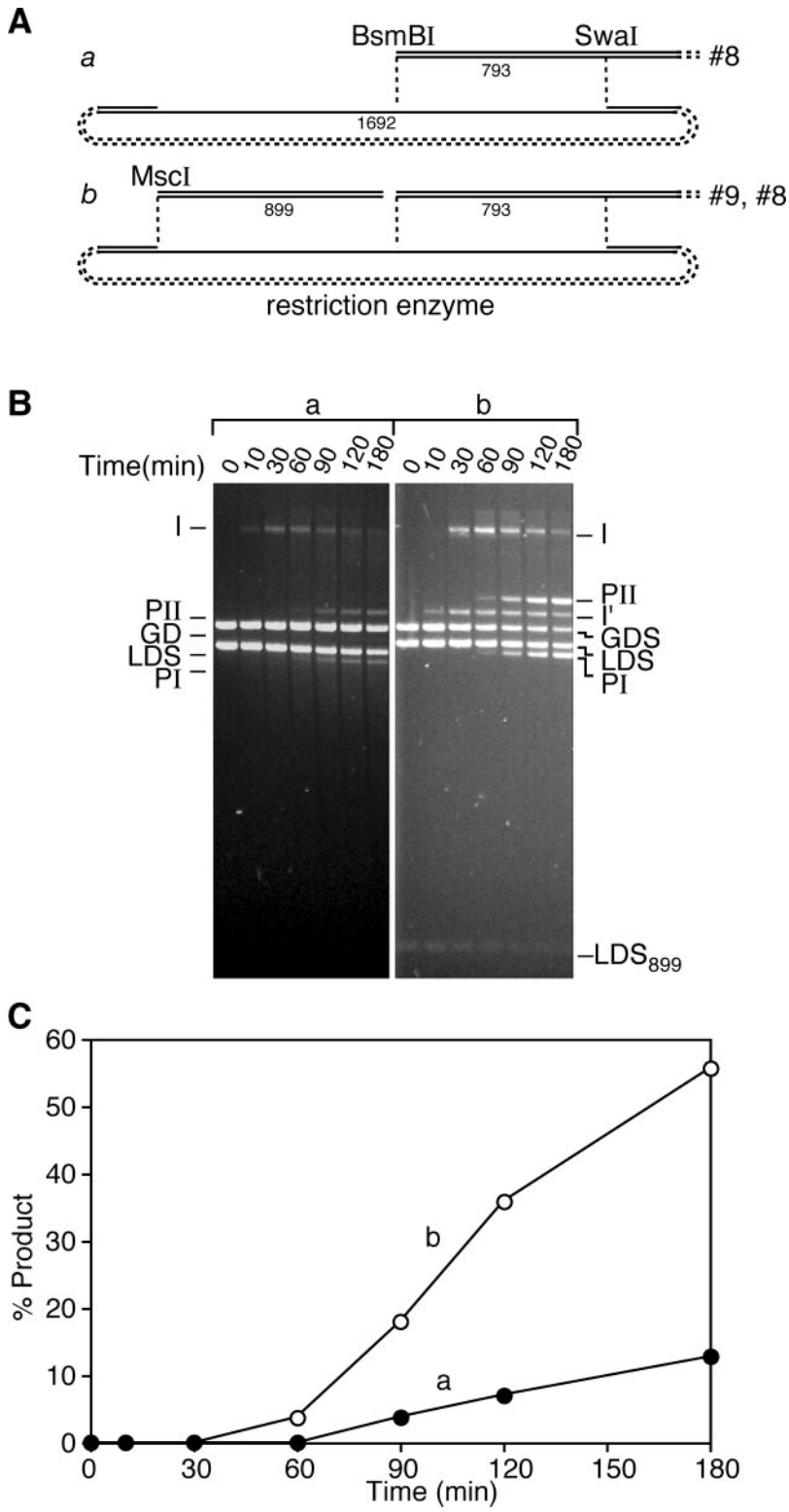

FIG. 9. **A 4-strand exchange initiating in the middle of a gap can be rescued.** Reactions were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA protein, 10 μM gapped DNA, with the gap defined in *A*, 10.1 μM linear DNA (*substrate 8*, Fig. 3). In *reaction* (*RXN*) *b* only (the *right reaction* in *B*), 1.1 μM linear duplex 9 (Fig. 3) was included in the reaction. The linear duplexes were adjusted in each reaction to provide a 1:1 molar match with the gapped DNA in terms of total molecules. Following preincubation (see legend to Fig. 5), reactions were carried out for 120 min. All other reaction details are as in the legend to Fig. 5. A 0.8% agarose gel was used in this experiment. In *reaction b*, the 4-strand exchange with *LDS 8* is preceded by a 3-strand exchange with *LDS 9*. Both exchange reactions are stimulated by the coincidence of the LDS end with a dsDNA-ssDNA junction at the end of the gap, which is not available in *reaction a*. In *reaction b*, *I′* is the transition product during a 3-strand exchange with *LDS 9* followed by 4-strand exchange with *LDS 8*. *Reaction b* is stimulated by the coincidence of the LDS end with the new dsDNA-ssDNA junction (51% PII formation at 180 min), which is not available in *reaction a* (14% at 180 min). The yield of final reaction products ([PI] + [PII]/total DNA) for the reactions in *B* is plotted in *C*.

completed, was at the BsmBI site in all cases) (Fig. 7*A*). All reactions were carried out for 120 min. The original DNA substrate again gave a good reaction in the absence of SSB. However, when the linear DNAs were shortened, the reaction was essentially abolished. A linear duplex with a 512-bp overlap of the gap gave no detectable reaction. The addition of SSB to the DrRecA reaction increased the yield of paired DNAs, permitting intermediates to be generated in two of the other three reactions, and allowing limited product formation in the reactions where the duplex partially overlapped the gap. However, the SSB again suppressed the generation of completed products in reaction a (the full overlap) in favor of an increase in reaction intermediates, as in Figs. 4 and 5.

Somewhat surprisingly, the pattern observed in the absence of SSB was similar for the EcRecA protein. Without SSB, the full overlap DNA substrate was the only linear duplex to generate a good yield of reaction products, albeit with a somewhat lower efficiency than seen in the comparable DrRecA reaction. In contrast to the results obtained with DrRecA protein, the EcSSB *enhanced* all of the reactions promoted by EcRecA pro-

TABLE I

A:T content at the initiation site of DNA strand exchange

All linear duplexes are derived from M13mp7 DNA. The end structure generated by the indicated restriction enzymes is given. For each restriction site, there are two rows of information provided. The % A:T is given as a function of individual 10-bp segments (top lines) as well as cumulative A:T content from the initiating end (bottom lines), assuming that strand exchange is progressing 5′ to 3′ relative to the single strand bound in the filament. The duplex DNA substrates generated by cleavage with the top three enzymes give good reactions, and those generated by cleavage with the bottom two enzymes give poor reactions. The major differences in A:T content are concentrated in the first 20 bp from the cleaved end. The structure of the ends (blunt or overhang) appears not to affect these reactions.

M13mp7		A:T content in 10-bp segments from the cleaved end				
Cleaved end		1-10	11-20 0-20	21-30 0-30	31-40 0-40	41-50 0-50
BsmBI	5′ (4)	80	40 60	40 53.3	30 47.5	20 42
MscI	Blunt	50	80 65	40 56.7	50 55	50 54
BtsI	3′ (2)	40	70 55	60 56.6	50 55	40 52
SfoI	Blunt	40	40 40	10 30	40 32.5	70 40
PvuI	3′ (2)	20	50 35	50 40	30 37.5	40 38

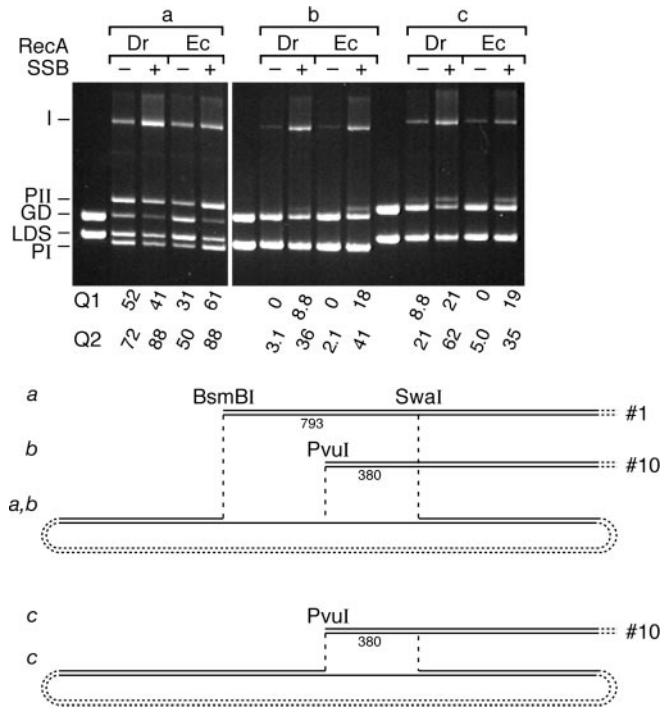

FIG. 10. **The A:T content of the DNA at the initiation site also affects the efficiency of initiation of DNA strand exchange.** Reactions were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA or EcRecA protein (as indicated), 0.6 μM EcSSB protein (where indicated), 10 μM gapped DNA, with the two different gaps defined in the bottom illustration, and 10.3–10.6 μM linear DNA (adjusted in each reaction to provide a 1:1 molar match with the gapped DNA in terms of total molecules). The linear DNAs in reactions *a–c* are *substrates 1*, *4*, and *10* in Fig. 3, respectively. Following preincubation (see legend to Fig. 5), reactions were carried out for 120 min. All other details are as in the legend to Fig. 5. A 0.8% agarose gel was used in this experiment. The reactions were quantitated, with results presented at the *bottom* of the gel as % final products (*Q1*, [PII]/([I] + [PII] + [GD])) or % total paired DNAs (*Q2*, [I] + [PII]/([I] + [PII] + [GD])).

tein greatly, but the best product yield was again seen in the full overlap reaction.

The great enhancement seen in the full overlap reaction relative to the others, especially in the absence of SSB, could simply mean that overlaps approaching 800 bp have some advantage in these reactions, or it could mean that there is a previously undetected effect of the dsDNA-ssDNA junction itself. The reactions of Fig. 7*B* provide some insight. Here the full overlap substrate again reacts well (with reaction yields quantified in the figure). However, when this substrate is shortened by just 22 bp, the reaction is reduced nearly 6-fold in the absence of SSB. The addition of SSB again allows some reaction to occur, but it is still nearly 3-fold less than the reaction observed for the full overlap substrate. We doubted that a DNA substrate with 793 bp of overlap should have much advantage over one with a 771-bp overlap, and this suggested that the dsDNA-ssDNA junction itself was playing a stimulatory role in the reaction.

This idea is given more substance in Fig. 8. Reactions are shown with three different GD substrates, with gaps of 1692, 793, and 281 nucleotides. Reactions a, c, and e are full overlap reactions with each of these gapped DNAs, respectively. Strong reactions are seen in each case, even in the absence of SSB, for both RecA proteins. Thus, the actual length of the overlapping DNA (over this range) does not seem to play a role in the yield of products. In the absence of SSB, both RecA proteins promote a consistently weak reaction (less than 10% conversion to final products) when the linear duplex is sized so that the strand exchange must initiate in the middle of a gap rather than at the dsDNA-ssDNA junction (Fig. 8, *reactions b* and *d*). The DrRecA protein-promoted reaction is reduced more than 6-fold when the linear duplex overlaps the 1692-bp gap by 793 bp (*reaction b*), compared with the reaction observed when this same linear duplex is paired with the GD with the 793-nucleotide gap (*reaction c*). If the linear duplex is shortened so that it overlaps the 793-nucleotide gap by 512 bp, the reaction is all but abolished (*reaction d*). However, the same linear duplex (this time with a 281-bp overlap) gives a good reaction when paired with the GD where the end corresponds to a dsDNA-ssDNA junction (*reaction e*). Thus a coincidence of the linear duplex end of one DNA substrate with the end of the gap of the other DNA substrate is important. In the reactions examined here, the actual length of the overlap is not a factor. In the absence of SSB, a similar pattern can be seen in the reactions promoted by EcRecA protein, so the dsDNA-ssDNA junction is having a priming effect on these reactions as well. SSB again has a somewhat inhibitory effect on the full overlap DrRecA reactions (reaction intermediates increase, but full products decrease), while allowing some reaction to occur where a full overlap does not exist. As before, the same SSB has a stimulatory effect on all of the EcRecA reactions.

A final experiment to establish the importance of the dsDNA-ssDNA junction at the end of the gap is shown in Fig. 9. When the GD with a 1692-bp gap is paired with a linear duplex that overlaps the gap by 793 bp, a poor reaction ensues with the DrRecA protein (no SSB). Only 14% of the substrate is converted to products in the 3-h reaction, and 84% of the DNA remains in the form of substrates. If a second short duplex DNA that spans the region between the dsDNA-ssDNA junction and the end where the linear duplex is initiating, the generation of final products in the 4-strand exchange reaction is enhanced by more than 4-fold. In effect, the 3-strand exchange establishes a new dsDNA-ssDNA junction, rescuing the reaction of the linear duplex with the 793-bp overlap.

The Initiation of 4-Strand Exchange Reactions Is Enhanced Both by dsDNA-ssDNA Junctions and the A:T Content of the Initiation Region—We wanted to determine whether other fac-

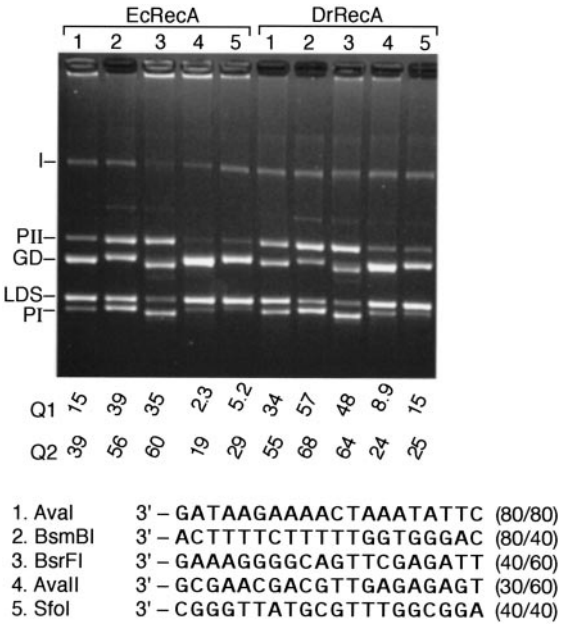

FIG. 11. **Determining the effect of A:T content on the efficiency of reaction initiation at the ends of the dsDNA-ssDNA junction at the end of the gap.** Each reaction was carried out with a different gapped DNA, which has one end of the ssDNA gap region at the indicated restriction enzyme position and the other at an SwaI site. All the duplexes completely overlap the gap, and thus one end corresponds in terms of sequence to the end of the gap. Gap sizes for the GD DNA substrates are BsrFI (1157-nt ssDNA gap), AvaI (944 nt), AvaII (857 nt), BsmBI (793 nt), and SfoI (771 nt). LDS SfoI has a blunt end. Substrates are aligned according to the order of A:T content in the first 20 nucleotides that would be paired after the initiation of strand exchange at the 5′ end of the gap region. The *numbers* in *parentheses* (*X/Y*) after the listing of the dsDNA-ssDNA junction proximal DNA sequence in the gap represent the A:T content in the first 10 and 11–20 bp, respectively. The reactions were quantitated, with results presented at the *bottom* of the gel as % final products (*Q1*, [PII]/([I] + [PII] + [GD + [N])) or % total paired DNAs (*Q2*, [I] + [PII]/([I] + [PII] + [GD] + [N])), where N denotes DNA in networks that are in the wells and will not resolve.

tors could affect the initiation of 4-strand exchange reactions. Based on the work of Radding and co-workers (32), the A:T content of the DNA in the initiation region could plausibly affect the reaction. Most of the reactions described so far feature an A:T content of greater than 50% in the first 20 bp at the initiating end of the linear duplex (Table I). However, the reaction initiated 22 bp from the dsDNA-ssDNA junction in Fig. 7*B* could be suppressed as a result of low A:T content at the initiation site (only 40% A:T in the first 20 bp) rather than its lack of coincidence with the dsDNA-ssDNA junction. We therefore constructed a linear duplex with an end at the PvuI site of M13, which delivers an initiation site with only 20% A:T content in the first 10 bp and 35% in the first 20 bp. As can be seen in Fig. 10, this has a large effect on the results. The reaction initiated at the dsDNA-ssDNA junction defined by the BsmBI restriction site is strong, and DrRecA converts over half of the substrate to complete products in the absence of SSB. The reactions using the duplex ending at the PvuI site are reduced substantially. This is true whether the reaction is initiated in the middle of a gap (*reaction b*) or at a dsDNA-ssDNA junction at the end of the gap (*reaction c*), although about 4-fold more reaction intermediates are observed in the latter case. SSB generally has a stimulating effect on the reactions with duplexes ending at the PvuI site, allowing about 20% of the substrates to be converted to completed products in most reactions, and increasing the yield of reaction intermediates by 5–16-fold.

The effects of A:T content are further explored in Fig. 11. A series of substrate combinations are used. In each case, the duplex overlaps the gap completely, and all of the gaps are sized within a relatively narrow range (771–1157 nucleotides in length). The initiating ends of the linear duplexes all have blunt ends or 5′ extensions, so that end of the duplex portion of the linear DNA corresponds to the dsDNA-ssDNA junction (because the 3′-ending strand is transferred to initiate exchange, a DNA with a 3′ extension would initiate with several nucleotides of ssDNA). The sequences near the 3′ end of the transferred strand are given in Table I. The duplexes with 50% or more of the first 20 bp as A:T all gave easily detectable reactions in the absence of SSB (15–57% conversion to final products). The reaction was diminished to low and in some cases nearly undetectable levels when the A:T content of the first 20 bp was reduced to below 50%. The trend is distinct but not perfect. The DNAs ending in AvaI have an A:T content of 80% near the end, but produce final products about half as efficiently as the DNAs ending in BsmBI or BsrFI sites, even though the latter have somewhat lower A:T contents.

DNA Secondary Structure in the Gap Affects the Requirement for SSB in RecA Protein-mediated 4-Strand Exchange Reactions—Another factor that might affect the efficiency of initiation in these reactions is the presence of secondary structure in the ssDNA in the gaps. The presence of a dsDNA-ssDNA junction could limit the secondary structure in the ssDNA near that point, whereas initiation in the middle of the gap could be impeded by secondary structure. We therefore left out the SSB altogether, and we compared reactions at dsDNA-ssDNA junctions with reactions initiated by the magnesium-shift protocol pioneered by Radding and co-workers (17, 18). In brief, the RecA filaments are allowed to form on the gapped DNA at 2 mM Mg^{2+}, where secondary structure is not an obstacle, and then the reaction is initiated by adding the linear duplex DNA and adjusting the Mg^{2+} to 10 mM at the same time. In Fig. 12, it can be seen that the magnesium shift protocol does stimulate the 4-strand exchange reaction when it is initiated in the middle of the gap (*reaction b*). The effect is quite modest for the DrRecA protein (about 3-fold), whereas the coincidence with a dsDNA-ssDNA junction improves the overall pairing reaction by 12-fold. A greater (3-fold) effect of the magnesium-shift protocol is seen in the EcRecA reactions, although the dsDNA-ssDNA junction improved the overall yield of products initiated at the BsmBI site (*reaction c versus reaction b*) by another 2-fold.

One final strand exchange experiment was carried out to control for all of these factors and reevaluate the effects of the dsDNA-ssDNA junctions (Fig. 13). We used the BsmBI-ended linear duplex, which has a high A:T content at the initiating end and generates a demonstrably strong reaction when its end corresponds to a dsDNA-ssDNA junction at the end of the gap as seen in many of the preceding figures. This was paired with two different gapped DNAs, one that had a dsDNA-ssDNA junction at BsmBI, so that the linear duplex end corresponded to the dsDNA-ssDNA junction at the end of the gap, and another with the dsDNA-ssDNA junction further back at an AvaII site, so that initiation of strand exchange would occur 64 nucleotides away from the dsDNA-ssDNA junction at the end of the gap (but still at the BsmBI site). In the absence of SSB, coincidence with the dsDNA-ssDNA junction had a dramatic effect in all cases, providing a 7–9-fold enhancement for the DrRecA protein reactions and 5–6-fold for the EcRecA reactions. The magnesium shift protocol, designed to eliminate secondary structure in the ssDNA in the gap, has no measurable effect in these reactions, for either the DrRecA or the EcRecA proteins, *i.e.* the same stimulatory effect of ssDNA-dsDNA junctions was seen whether the magnesium-shift or the

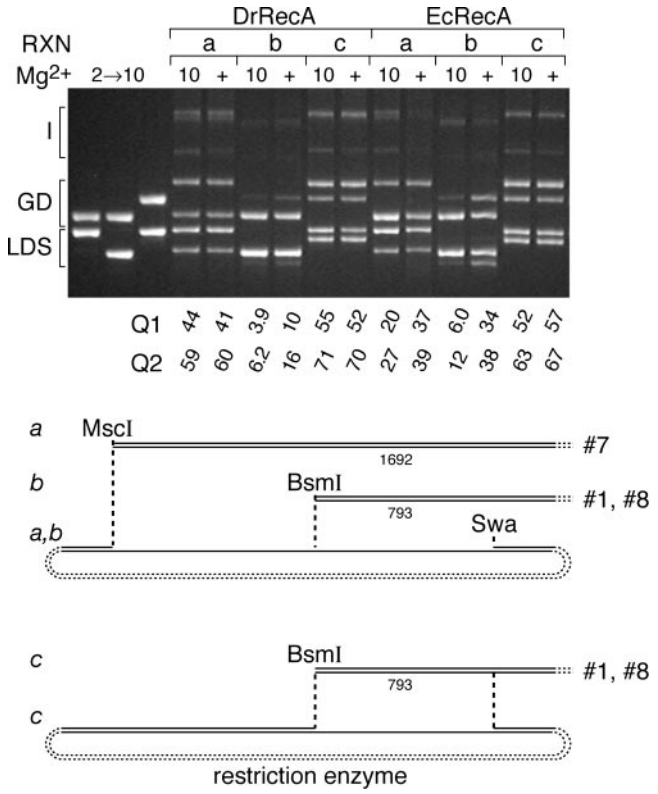

FIG. 12. **The dsDNA-ssDNA junction at the end of the gap has a greater stimulating effect on DNA strand exchange than elimination of DNA secondary structure within the gap.** Reactions (*RXN*) were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA or EcRecA protein (as indicated), 0.6 μM EcSSB protein (where indicated), 10 μM gapped DNA, with the two different gaps defined in the *bottom illustration*, and 10.1–11.2 μM linear DNA (adjusted in each reaction to provide a 1:1 molar match with the gapped DNA in terms of total molecules). The linear DNAs in reactions *a–c* are *substrates 7*, *8*, and *1* in Fig. 3, respectively. Reactions either contained 10 mM Mg^{2+} throughout or were subjected to a magnesium-shift protocol (reactions labeled +), as indicated. All preincubations and the experiment were carried out at 37 °C. The 10 mM Mg^{2+} reactions followed the preincubation protocol of Fig. 5. The magnesium shift protocol included only 2 mM Mg^{2+} in the 10-min preincubation with the RecA protein and the gapped DNA. The Mg^{2+} concentration was not changed after addition of the 2 mM ATP and during the subsequent 20-min incubation. The Mg^{2+} concentration was increased to 10 mM when the linear duplex DNA was added to initiate DNA strand exchange. All reactions were carried out for 120 min. All other reaction details are as in the legend to Fig. 5. A 1% agarose gel was used in this experiment. The reactions were quantitated, with results presented at the *bottom* of the gel as % final products (*Q1*, [PII]/([I] + [PII] + [GD])) or % total paired DNAs (*Q2*, [I] + [PII]/([I] + [PII] + [GD])).

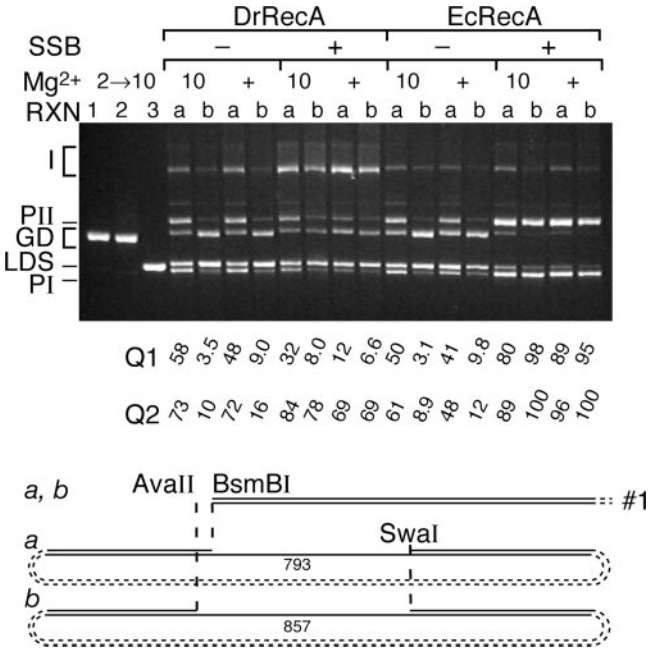

FIG. 13. **dsDNA-ssDNA junctions greatly enhance the initiation of DNA strand exchange even when initiation is constrained to a high A:T site.** Reactions (*RXN*) were carried out as described under "Experimental Procedures" and contained 5 μM DrRecA or EcRecA protein (as indicated), 0.6 μM EcSSB protein (where indicated), 10 μM gapped DNA, with the two different gaps defined in the *bottom illustration*, and 10.6 μM linear DNA. The linear DNA is *substrate 1* in Fig. 3 for all reactions. Reactions either contained 10 mM Mg^{2+} throughout or were subjected to a magnesium shift protocol (reactions labeled +), as indicated. Preincubation protocols are described in the legend to Fig. 10. All reactions were carried out for 120 min. All other reaction details are as in the legend to Fig. 5. A 1% agarose gel was used in this experiment. The reactions were quantitated, with results presented at the *bottom* of the gel as % final products (*Q1*, [PII]/([I] + [PII] + [GD])) or % total paired DNAs (*Q2*, [I] + [PII]/([I] + [PII] + [GD])).

normal protocol was used. Because the magnesium shift protocol does enhance some of these reactions (see Fig. 12), the lack of an effect here may mean that DNA secondary structure is limited near the dsDNA-ssDNA junction or in this sequence region in general. However, the presence of a dsDNA-ssDNA junction at the initiation site clearly has a major effect.

When SSB was present, the patterns were again consistent with what was seen in the earlier figures. In the DrRecA reactions, SSB stimulated the reactions initiated in the middle of the gap, although complete reactions in the full overlap case (reaction a) were again impeded somewhat (final product yield drops by nearly 2-fold, although overall yield of paired DNAs, including intermediates, improves by about 15%). The EcRecA reactions were all stimulated substantially as shown above.

Electron Microscopy and Image Reconstruction—Our explanation of the effects of dsDNA-ssDNA junctions on these DNA strand exchange reactions (see "Discussion") is that the RecA filaments on a gapped DNA are in two different structural states, depending on whether a particular segment is bound to dsDNA or to the ssDNA in the gap. The extension of the more cooperative dsDNA (P) conformation into the gap leads to the observed enhancement of strand exchange initiation near the dsDNA-ssDNA junction at the end of the gap. We were interested to know whether RecA filaments formed on ssDNA should be in a structural state detectably different from the corresponding filaments formed on dsDNA. To test this, we have reconstructed RecA filaments formed on ssDNA and compared them with filaments formed on dsDNA (54). The two reconstructions are similar (Fig. 14), except that differences appear at the highest radius in the reconstruction, within the C-terminal domain. This is surprising, as the C-terminal domain is relatively distant from the DNA bound at a small radius near the helical axis of the filament (68). However, this result is consistent with many other observations showing a coupling between this C-terminal domain and both the nucleotide- and DNA-binding sites located within the RecA "core" domain (40, 54, 69–71). The main difference is that the C-terminal lobes are significantly smaller in the RecA-dsDNA filaments than they are in the RecA-ssDNA filaments.

The simplest explanation for this difference in size is that the C-terminal domain is more disordered or that this domain exists in a number of discrete conformations, when the filament is formed on dsDNA. However, when the filament is formed on ssDNA, this C-terminal domain is more stable. When many subunits are averaged together (in this case, the average contains 2,525 segments, with each segment containing ~16 sub-

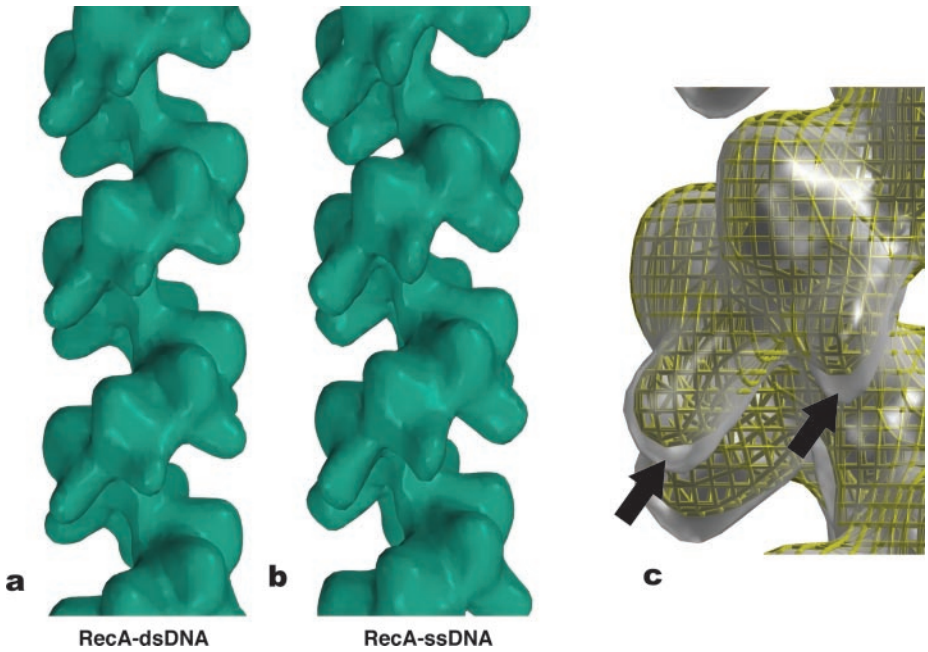

FIG. 14. **Three-dimensional reconstructions of RecA filaments.** Filaments were formed on dsDNA (*a*) and ssDNA (*b*) in the presence of the ATP analog AMP-PNP. When the two volumes are superimposed (*c*), it can be seen that in the C-terminal lobe of RecA (*c, arrows*) the RecA-ssDNA reconstruction (*gray transparent surface*) is significantly larger than the RecA-dsDNA reconstruction (*yellow mesh*). However, for the nucleotide-binding core that forms the continuous helical backbone or "yoke" of the filament, the two maps superimpose perfectly. The smaller C-terminal lobe is interpreted as being due to greater disorder in this region.

units, so ~40,000 subunits have been averaged together), this disorder will therefore result in less than the full density being apparent. Evidence in support of this explanation is that the C-terminal domain of RecA has already been observed by EM to be quite dynamic in conformation. Under some conditions (such as when using the ATPase-deficient RecA E96D mutant protein and ATP), there is a significant increase in the amount of the C-terminal domain visualized after averaging (54). Furthermore, in filaments formed by the archaeal RadA (72) and the eukaryotic RAD51 (53) proteins (both RecA homologs), part or all of the N-terminal domains are not visualized under some conditions by EM after averaging.

DISCUSSION

Our primary observation is that a dsDNA-ssDNA junction has a major enhancing effect on the initiation of DNA strand exchange in the adjacent gap. The A:T content of the DNA at the initiation site also may play a significant role in the efficiency of the reaction, although this is not as clearly evident as the effect of the dsDNA-ssDNA junction. The coincidence of a dsDNA-ssDNA junction with the end of the linear duplex can have a very large effect on the overall efficiency of a 4-strand exchange reaction. A strong 4-strand exchange reaction is promoted by the DrRecA protein without SSB, and the EcRecA protein also promotes its strongest reactions with the full overlap substrates. Somewhat ironically (because this study was originally set up to investigate 4-strand exchange reactions), all of the effects we observe are directed mainly at the initiation of the reaction in the gap, *i.e.* the 3-stranded portion of the reaction.

When the A:T content is controlled, the juxtaposition of a dsDNA-ssDNA junction still has a dramatic stimulatory effect on the strand exchange reactions (Fig. 13). This is true even when the effects of DNA secondary structure are minimized with a magnesium-shift protocol (Fig. 12). Given the extension of DNA within an activated RecA filament, which should eliminate normal base stacking (73), it is difficult to propose an explanation for this phenomenon based on DNA structure. Instead, we explain it within the context of a proposed model for RecA-mediated DNA pairing (Fig. 15), which is itself based on known properties of RecA protein. RecA protein filaments bound to ssDNA exhibit extensive functional differences with respect to RecA bound to dsDNA, suggesting that they are in

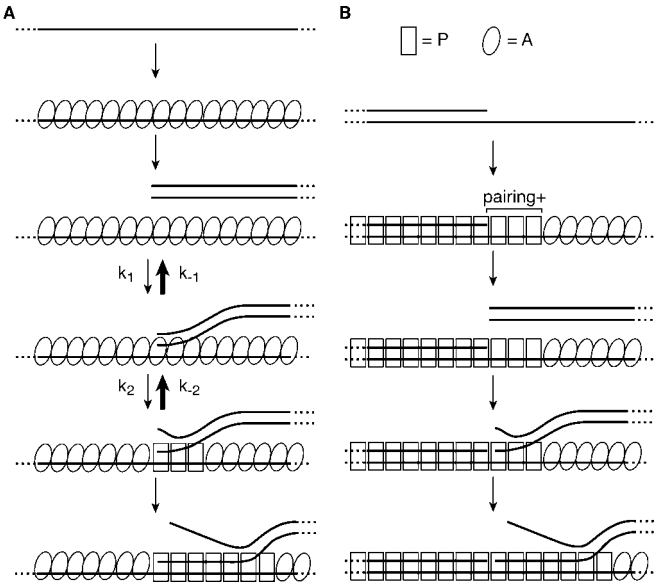

FIG. 15. **The DNA pairing process promoted by bacterial RecA proteins, and the effect of dsDNA-ssDNA junction at the end of the gaps.** *A*, RecA protein bound to ssDNA is in the A state, when ATP is present. In the experiments carried out in the present study, the Mg^{2+} concentration places the EcRecA protein in the Ao state. Initiation of DNA strand exchange occurs in two reversible steps, as described in the text, labeled k_1 and k_2. We propose that the slower first order step, described by k_2, involves an A to P structural transition. The overall reaction is reversible, by virtue of the re-pairing of the substrate duplex and perhaps the relative instability of the P state. *B*, the presence of a dsDNA-ssDNA junction greatly enhances the initiation of DNA strand exchange in the adjacent DNA gap. The gapped duplexes are bound with RecA protein, with the A and P states predominating in the single-stranded and duplex regions, respectively. We propose that the more cooperative P state extends for a short distance into the single-strand gap, creating a pairing-enhanced site near the dsDNA-ssDNA junction. Because the P state in these monomers is stabilized by proximity to dsDNA-bound monomers in the adjacent duplex DNA, DNA pairing is much less reversible at this site. DNA strand exchange can be propagated from this site to encompass the entire DNA molecule without the aid of SSB.

quite different structural states. As indicated in the Introduction, we designate these states as A and P, respectively. This difference in structural state is examined more directly with

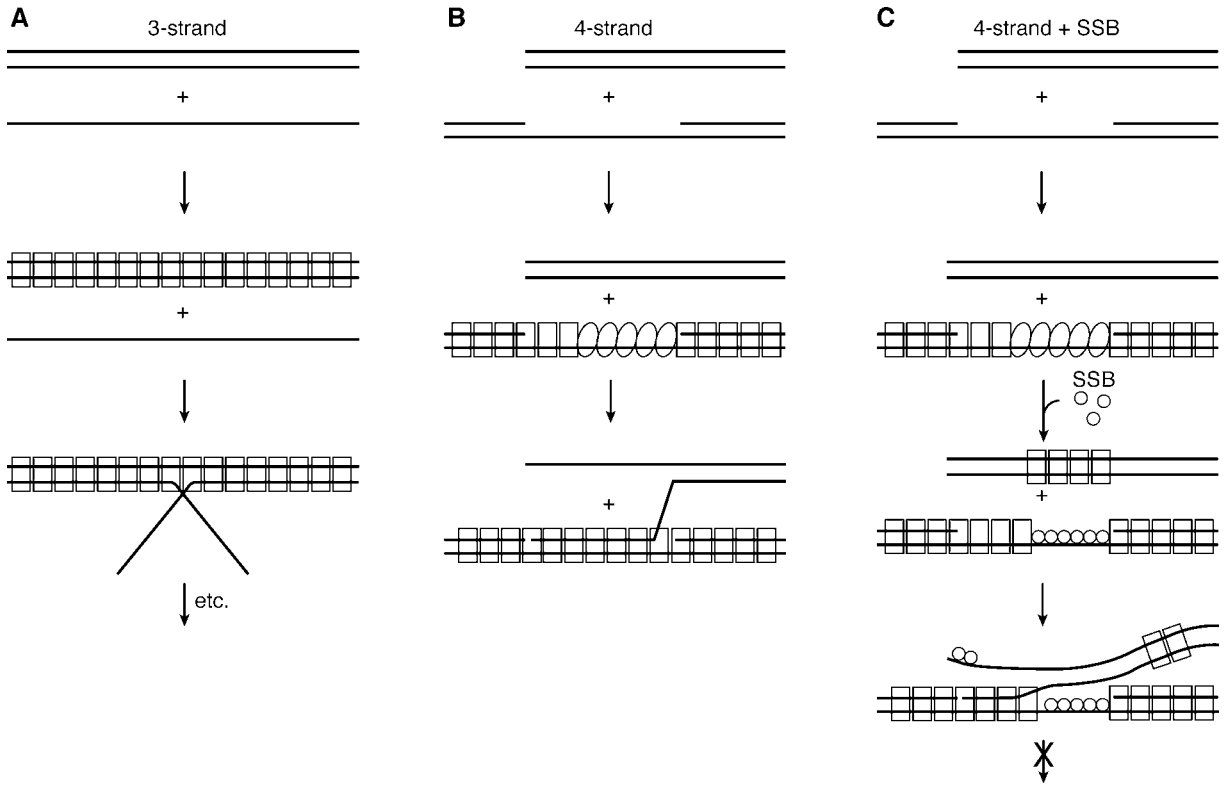

FIG. 16. **DNA pairing and strand exchange mediated by the DrRecA protein.** *A*, the predominant pathway for DrRecA protein-mediated 3-strand exchange reactions involves the formation of DrRecA filaments on the dsDNA substrate, followed by incorporation of the single-stranded substrate to initiate exchange. We propose that the DrRecA bound to dsDNA is in a state equivalent to the P state defined for EcRecA protein, and that there is a more stringent requirement for the DrRecA protein to initiate exchange from the P state. *B*, in a 4-strand reaction, the DrRecA protein bound to the duplex portions of the gapped DNA is in the P state, as is the DrRecA in the ssDNA gap near the dsDNA-ssDNA junction. Thus, DNA strand exchange can be initiated from the gapped DNA, with no need to migrate to the linear duplex. *C*, the presence of the EcSSB protein facilitates some DrRecA protein migration to the linear duplex DNA. EcSSB stimulates the initiation of the 3-strand phase of the reaction in the gap. However, the presence of DrRecA bound to both DNA substrates impedes the 4-strand phase of the reaction. Some products are generated, either from substrates that are not thus impeded or on substrates where redistribution of the DrRecA protein permits it.

the aid of image reconstruction from electron micrographs in Fig. 14. The model, detailed below, provides a context within which to articulate three mechanistic conclusions arising from the data in this and previous studies. 1) The P state is a pairing-enhanced state. If RecA protein can be stabilized in this state while on ssDNA, the initiation of strand exchange will be stimulated at that site (2). The unusual DNA pairing properties of the DrRecA protein can be explained by a more stringent requirement to initiate strand exchange from the P state (3). The A and P states are structurally distinct, with the differences centered at the C terminus of the protein.

The Model—In a normal 3-strand exchange reaction, initiation occurs as in Fig. 15*A*. The kinetic schemes of Camerini-Otero and co-workers (29) and Radding and co-workers (30, 31) both envision two key steps, with a slow first order step following a faster second order step. We propose that the transition from the A to P structural states represents a key component of the slower first order process. In Fig. 15*A,* we have labeled these steps k_1 and k_2, according to the Radding scheme. Both are reversible, and if the P state is less stable than the A state, many pairing initiation events will be eliminated as the substrate duplex DNA is re-paired. A successful initiation would be propagated down the filament as shown, as A state monomers were converted into the more cooperative P state. We note that Camerini-Otero and co-workers (29) speculated that a conformational change might be part of the slow first order process.

Conclusions—1) On a gapped DNA, the situation would be quite different. Here the RecA filament is a hybrid of segments in the A and P conformations, with state A in the single-strand gap and state P on the duplex portions. To explain our results, we propose that the highly cooperative P conformation extends a very short distance into the gap, providing an enhanced DNA pairing site immediately adjacent to the dsDNA-ssDNA junction. From that point, as strand exchange proceeds, the remainder of the RecA monomers in the gap is converted from the A to the P form (Fig. 15*B*). Reversal of the initiation steps is minimized because the RecA monomers near the dsDNA-ssDNA junction are stabilized in the P state by virtue of their proximity to the dsDNA-bound RecA protein beyond the dsDNA-ssDNA junction. This leads directly to our first conclusion above. The P state, in effect, is the state that binds to the products of strand exchange most tightly, and thus the barrier to pairing is reduced when RecA bound to ssDNA is in this state, *i.e.* it is pairing-enhanced.

2) If the DrRecA protein *must* be in the P state in order to initiate DNA strand exchange, then we can begin to explain the unusual DNA strand exchange reactions promoted by this protein. The relevant extension of our model is laid out in Fig. 16. In a pure 3-strand exchange (Fig. 16*A*), the only way to achieve the P state is to bind to dsDNA, and thus the DrRecA initiates 3-strand exchanges by binding to dsDNA first. On a gapped DNA (Fig. 16*B*), some of the DrRecA protein near the dsDNA-ssDNA junction will be in the P state, so that initiation can occur efficiently in this location with no need for the protein to migrate to the linear dsDNA first. The EcSSB protein stimulates the DrRecA-mediated 3-strand exchange reactions, in part by displacing the DrRecA from the ssDNA (27). In the 4-strand exchange, the EcSSB stimulates strand exchange initiation both at or away from a dsDNA-ssDNA junction. How-

ever, completion of the reactions initiated at the dsDNA-ssDNA is partially suppressed, and it is this effect that provides contrast with the reaction promoted by EcRecA. Because SSB should displace some DrRecA protein from the gap of the gapped DNA, we anticipate that most of the inhibition of product formation is due to the generation of discontinuities in the DrRecA filament promoting the reaction. This view is supported by the results of Fig. 5, where increased DrRecA protein concentrations overcome the effects of EcSSB. The high levels of DrRecA required to do this are enough to more than saturate both of the DNA substrates in the reaction. However, it is unlikely that two filaments participate in strand exchange. An excellent reaction is seen in the absence of EcSSB when there is only enough DrRecA protein to coat one of the DNA substrates (Figs. 4 and 5). Thus, our second conclusion is that the DrRecA protein has a more stringent requirement for initiating DNA strand exchange from the P state.

Other factors, of course, can affect the initiation of strand exchange, and these reinforce the idea that there is a considerable thermodynamic barrier to initiation. A possible additional factor is the A:T content of the DNA at the site where the strand exchange is initiated. A study by Radding and co-workers (32) pinpointed A:T content as a major factor in the overall DNA pairing process. This work employed the human Rad51 protein, but the present study suggests that this constraint may be universal. Another factor is secondary structure in the ssDNA. This is removed by SSB during the presynaptic facilitation of contiguous RecA filaments on ssDNA (13, 14). However, in the 4-strand reactions investigated, a strong yield of final products was seen even when the 3-strand phase of the reaction had to proceed through single-strand gaps nearly 1700 nucleotides in length. Thus, the P conformation seems able to propagate a strand exchange, one initiated, through whatever secondary structure was present in these molecules. It is possible that the usual role of SSB is most important at the point of initiation of DNA strand exchange, where a sequestering of the displaced strand would help to stabilize a segment of RecA filament in the P state.

3) The observable structural distinctions between RecA filaments bound to ssDNA (A state) or to dsDNA (P state) are not large and appear to affect only the C terminus of the protein (Fig. 14). Although some structural changes might be missed at this resolution, a large change involving the C terminus seems readily rationalized. The C terminus of RecA has a large role in modulating the initiation of DNA strand exchange (40, 51, 71). An enhancement of strand exchange initiation, as we propose for the P state, should logically be accompanied by a structural change in the C terminus.

We do not know how far into the gap the stimulatory effect of a dsDNA-ssDNA junction extends. Where the A:T content at the initiation site is controlled (Fig. 13), we know that it is less than 64 bp. Because the data in Fig. 11 suggest that there is some correlation between the extent of enhancement at a dsDNA-ssDNA junction and the A:T content in the first 20 bp, the effect of the dsDNA-ssDNA junctions may not need to extend more than 20 bp into the gap to bring about a significant enhancement of the overall reaction.

The results presented here help explain the variability seen in the DNA strand exchange reactions reported by many different groups, with a range of different DNA substrates. As noted in the Introduction, there were discrepancies in the reported requirements for SSB in 4-strand exchange reactions promoted by EcRecA protein that can now be rationalized. The early reports of good reactions without SSB employed DNA substrates of the "full overlap" type (35, 36, 39), whereas the later work requiring SSB did not. The A:T content at the end of the linear duplex DNA is also revealed as a factor that may influence the efficiency of an *in vitro* DNA strand exchange reaction. Whether all details of the model presented above prove correct or not, the thermodynamics of DNA strand exchange clearly make the reaction sensitive to a range of RecA structure and DNA sequence factors. The path to an efficient DNA pairing reaction, for *in vitro* study or for purposes of biotechnology application, clearly depends upon attention to the structural status of RecA and the structure and sequence of the DNAs employed.

Acknowledgment—We thank Shelley Lusetti for reading and commenting on the manuscript.

REFERENCES

1. Cox, M. M. (2003) *Annu. Rev. Microbiol.* **57,** 551–577
2. Lusetti, S. L., and Cox, M. M. (2002) *Annu. Rev. Biochem.* **71,** 71–100
3. Cox, M. M., and Lehman, I. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6018–6022
4. Kahn, R., Cunningham, R. P., Das Gupta, C., and Radding, C. M. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 4786–4790
5. West, S. C., Cassuto, E., and Howard-Flanders, P. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6149–6153
6. Jain, S. K., Cox, M. M., and Inman, R. B. (1994) *J. Biol. Chem.* **269,** 20653–20661
7. Kim, J. I., Cox, M. M., and Inman, R. B. (1992) *J. Biol. Chem.* **267,** 16444–16449
8. Kim, J. I., Cox, M. M., and Inman, R. B. (1992) *J. Biol. Chem.* **267,** 16438–16443
9. Kowalczykowski, S. C., and Krupp, R. A. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 3478–3482
10. Menetski, J. P., Bear, D. G., and Kowalczykowski, S. C. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 21–25
11. Rehrauer, W. M., and Kowalczykowski, S. C. (1993) *J. Biol. Chem.* **268,** 1292–1297
12. Shan, Q., Cox, M. M., and Inman, R. B. (1996) *J. Biol. Chem.* **271,** 5712–5724
13. Kowalczykowski, S. C., and Krupp, R. A. (1987) *J. Mol. Biol.* **193,** 97–113
14. Kowalczykowski, S. C. (1991) *Biochimie* (*Paris*) **73,** 289–304
15. Lavery, P. E., and Kowalczykowski, S. C. (1992) *J. Biol. Chem.* **267,** 9315–9320
16. Ullsperger, C. J., and Cox, M. M. (1995) *Biochemistry* **34,** 10859–10866
17. Flory, J., Tsang, S. S., and Muniyappa, K. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 7026–7030
18. Muniyappa, K., Shaner, S. L., Tsang, S. S., and Radding, C. M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 2757–2761
19. Battista, J. R., Earl, A. M., and Park, M. J. (1999) *Trends Microbiol.* **7,** 362–365
20. Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., and da Costa, M. S. (1997) *Int. J. Syst. Bacteriol.* **47,** 939–947
21. Minton, K. W. (1994) *Mol. Microbiol.* **13,** 9–15
22. Daly, M. J., Ouyang, L., Fuchs, P., and Minton, K. W. (1994) *J. Bacteriol.* **176,** 3508–3517
23. Lin, J. Y., Qi, R., Aston, C., Jing, J. P., Anantharaman, T. S., Mishra, B., White, O., Daly, M. J., Minton, K. W., Venter, J. C., and Schwartz, D. C. (1999) *Science* **285,** 1558–1562
24. Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., and Daly, M. J. (2001) *Microbiol. Mol. Biol. Rev.* **65,** 44–79
25. Minton, K. W. (1996) *Mutat. Res.* **363,** 1–7
26. Kim, J. I., Sharma, A. K., Abbott, S. N., Wood, E. A., Dwyer, D. W., Jambura, A., Minton, K. W., Inman, R. B., Daly, M. J., and Cox, M. M. (2002) *J. Bacteriol.* **184,** 1649–1660
27. Kim, J.-I., and Cox, M. M. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7917–7921
28. Cox, M. M. (1999) *Prog. Nucleic Acids Res. Mol. Biol.* **63,** 310–366
29. Yancey-Wrona, J. E., and Camerini-Otero, R. D. (1995) *Curr. Biol.* **5,** 1149–1158
30. Bazemore, L. R., Takahashi, M., and Radding, C. M. (1997) *J. Biol. Chem.* **272,** 14672–14682
31. Bazemore, L. R., Foltastogniew, E., Takahashi, M., and Radding, C. M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 11863–11868
32. Gupta, R. C., Folta-Stogniew, E., O'Malley, S., Takahashi, M., and Radding, C. M. (1999) *Mol. Cell* **4,** 705–714
33. Conley, E. C., and West, S. C. (1990) *J. Biol. Chem.* **265,** 10156–10163
34. Lindsley, J. E., and Cox, M. M. (1990) *J. Biol. Chem.* **265,** 10164–10171
35. West, S. C., Cassuto, E., and Howard, F. P. (1982) *Mol. Gen. Genet.* **187,** 209–217
36. West, S. C., and Howard-Flanders, P. (1984) *Cell* **37,** 683–691
37. Chow, S. A., Chiu, S. K., and Wong, B. C. (1992) *J. Mol. Biol.* **223,** 79–93
38. Shan, Q., and Cox, M. M. (1998) *Mol. Cell* **1,** 309–317
39. West, S. C., Cassuto, E., and Howard-Flanders, P. (1982) *Mol. Gen. Genet.* **186,** 333–338
40. Lusetti, S. L., Shaw, J. J., and Cox, M. M. (2003) *J. Biol. Chem.* **278,** 16381–16388
41. Yu, X., and Egelman, E. H. (1992) *J. Mol. Biol.* **227,** 334–346
42. Lee, J. W., and Cox, M. M. (1990) *Biochemistry* **29,** 7666–7676
43. Lee, J. W., and Cox, M. M. (1990) *Biochemistry* **29,** 7677–7683
44. Egelman, E. H., and Stasiak, A. (1986) *J. Mol. Biol.* **191,** 677–697
45. Egelman, E. H., and Stasiak, A. (1988) *J. Mol. Biol.* **200,** 329–349
46. Stasiak, A., Egelman, E. H., and Howard-Flanders, P. (1988) *J. Mol. Biol.* **202,** 659–662
47. Stole, E., and Bryant, F. R. (1994) *J. Biol. Chem.* **269,** 7919–7925

48. Stole, E., and Bryant, F. R. (1995) *J. Biol. Chem.* **270,** 20322–20328
49. Stole, E., and Bryant, F. R. (1997) *Biochemistry* **36,** 3483–3490
50. Honigberg, S. M., Gonda, D. K., Flory, J., and Radding, C. M. (1985) *J. Biol. Chem.* **260,** 11845–11851
51. Lusetti, S. L., Wood, E. A., Fleming, C. D., Modica, M. J., Korth, J., Abbott, L., Dwyer, D. W., Roca, A. I., Inman, R. B., and Cox, M. M. (2003) *J. Biol. Chem.* **278,** 16372–16380
52. Yu, X., and Egelman, E. H. (1992) *J. Mol. Biol.* **225,** 193–216
53. Yu, X., Jacobs, S. A., West, S. C., Ogawa, T., and Egelman, E. H. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8419–8424
54. VanLoock, M. S., Yu, X., Yang, S., Lai, A. L., Low, C., Campbell, M. J., and Egelman, E. H. (2003) *Structure* **11,** 1–20
55. Bedale, W. A., and Cox, M. (1996) *J. Biol. Chem.* **271,** 5725–5732
56. Schutte, B. C., and Cox, M. M. (1987) *Biochemistry* **26,** 5616–5625
57. Shan, Q., and Cox, M. M. (1996) *J. Mol. Biol.* **257,** 756–774
58. Shan, Q., and Cox, M. M. (1997) *J. Biol. Chem.* **272,** 11063–11073
59. Arenson, T. A., Tsodikov, O. V., and Cox, M. M. (1999) *J. Mol. Biol.* **288,** 391–401
60. Takahashi, M., and Schnarr, M. (1989) *Eur. J. Biochem.* **183,** 617–622
61. Little, J. W., Edmiston, S. H., Pacelli, L. Z., and Mount, D. W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 3225–3229
62. Little, J. W. (1991) *Biochimie* (*Paris*) **73,** 411–422
63. Craig, N. L., and Roberts, J. W. (1981) *J. Biol. Chem.* **256,** 8039–8044
64. Lohman, T. M., and Overman, L. B. (1985) *J. Biol. Chem.* **260,** 3594–3603
65. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78
66. Neuendorf, S. K., and Cox, M. M. (1986) *J. Biol. Chem.* **261,** 8276–8282
67. Egelman, E. H. (2000) *Ultramicroscopy* **85,** 225–234
68. Egelman, E. H., and Yu, X. (1989) *Science* **245,** 404–407
69. Benedict, R. C., and Kowalczykowski, S. C. (1988) *J. Biol. Chem.* **263,** 15513–15520
70. Tateishi, S., Horii, T., Ogawa, T., and Ogawa, H. (1992) *J. Mol. Biol.* **223,** 115–129
71. Eggler, A. L., Lusetti, S. L., and Cox, M. M. (2003) *J. Biol. Chem.* **278,** 16389–16396
72. Yang, S. X., Yu, X., Seitz, E. M., Kowalczykowski, S. C., and Egelman, E. H. (2001) *J. Mol. Biol.* **314,** 1077–1085
73. Nishinaka, T., Ito, Y., Yokoyama, S., and Shibata, T. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6623–6628