Annu. Rev. Biochem. 2002. 71:71–100
DOI: 10.1146/annurev.biochem.71.083101.133940
Copyright © 2002 by Annual Reviews. All rights reserved

THE BACTERIAL RECA PROTEIN AND THE RECOMBINATIONAL DNA REPAIR OF STALLED REPLICATION FORKS

Shelley L. Lusetti and Michael M. Cox
Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Drive, Madison, Wisconsin 53706-1544; e-mail: slusetti@biochem.wisc.edu; cox@biochem.wisc.edu

Key Words recombination, motor, DNA damage, double-strand break, SOS

■ **Abstract** The primary function of bacterial recombination systems is the nonmutagenic repair of stalled or collapsed replication forks. The RecA protein plays a central role in these repair pathways, and its biochemistry must be considered in this context. RecA protein promotes DNA strand exchange, a reaction that contributes to fork regression and DNA end invasion steps. RecA protein activities, especially formation and disassembly of its filaments, affect many additional steps. So far, *Escherichia coli* RecA appears to be unique among its nearly ubiquitous family of homologous proteins in that it possesses a motorlike activity that can couple the branch movement in DNA strand exchange to ATP hydrolysis. RecA is also a multifunctional protein, serving in different biochemical roles for recombinational processes, SOS induction, and mutagenic lesion bypass. New biochemical and structural information highlights both the similarities and distinctions between RecA and its homologs. Increasingly, those differences can be rationalized in terms of biological function.

CONTENTS

INTRODUCTION: THE REPAIR OF STALLED REPLICATION FORKS

Bacterial replication forks are halted at the sites of DNA damage or other barriers. Halting of the replication fork occurs often even under normal aerobic growth conditions, perhaps during every replication cycle (1–7). The nonmutagenic repair of stalled replication forks involves an elaborate collaboration between enzymes from every mode of DNA metabolism. Recombination enzymes remodel the DNA structures at the broken fork to regenerate a viable fork structure. Repair DNA polymerases fill in gaps as needed. Specialized complexes to restart replication (independent of the origin) reestablish normal DNA replication. The proposed pathways involved are truly myriad (Figure 1), determined to a degree by the type of damage or barrier encountered. Work to date has served to classify these pathways collectively as a major new DNA repair process, and has advanced an appreciation for the extraordinary degree of integration that exists among all facets of bacterial DNA metabolism (4–7). Recent reviews have detailed what is currently known about the genetics and biochemistry of nonmutagenic recombinational repair of stalled replication forks (1–8). This modern synthesis of bacterial DNA metabolism has both a rich and interesting history (9) and a long list of questions to enliven future research (10). The repair of stalled replication forks is probably a significant function of recombination systems in eukaryotes as well where recombination has critical roles in a number of other processes (11, 12).

It is now possible to define two major classes of fork repair pathways. If a replication fork encounters a DNA strand break, repair is shunted into a set of pathways that can be quite distinct from those enlisted when the collision involves DNA lesions or other nonreplicable barriers [such as a stalled RNA

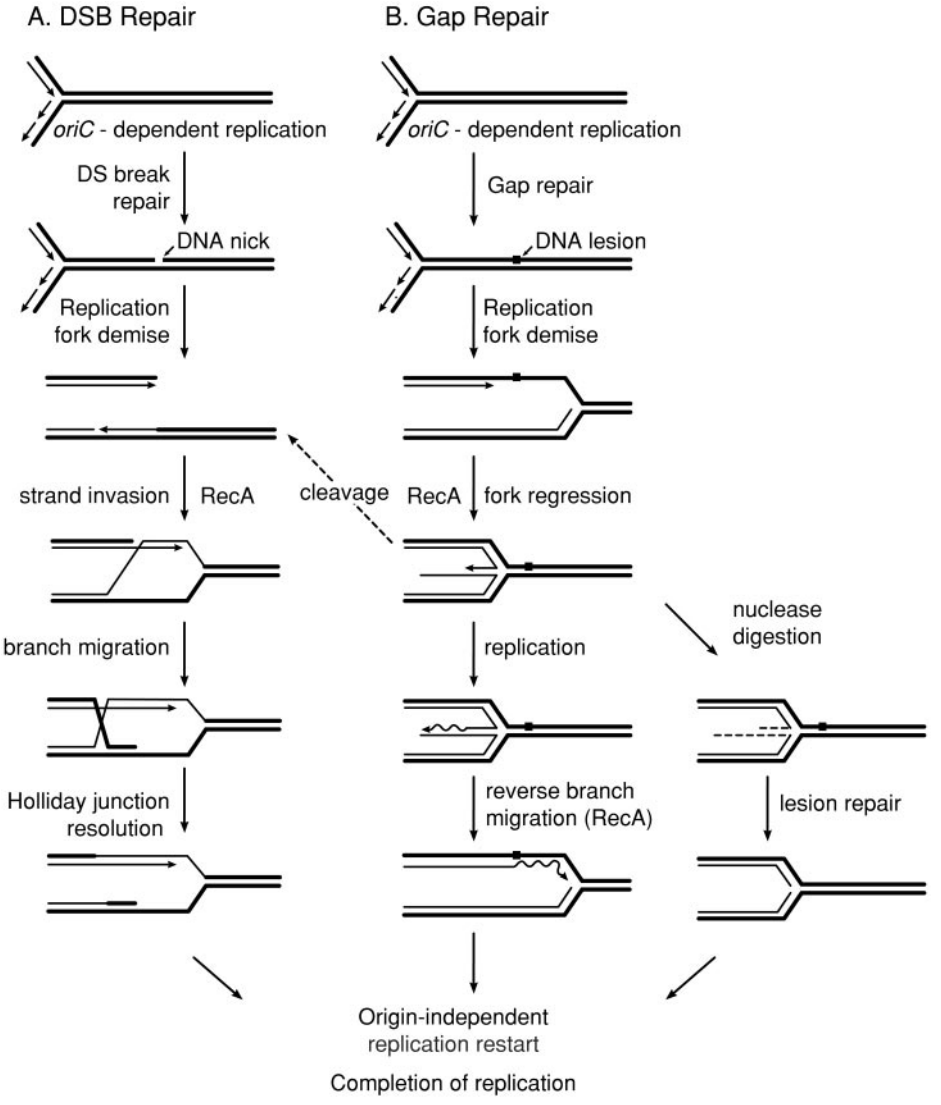

Figure 1 A sampling of proposed pathways for the recombinational DNA repair of replication forks. The likely path varies according to the type of lesion encountered and its location (leading or lagging strand). One major division in the pathways concerns the encounter of (*A*) a strand break [which creates a double-strand break (DSB)] versus (*B*) an encounter with a blocking lesion or other barrier. The pathways shown may be oversimplified deficient in the case of blocking lesions, with many additional paths proposed in the references cited in the text.

polymerase (13)]. A collision with a strand break creates a double-strand break (DSB) separating one branch of the fork from the parental DNA (Figure 1*A*). In the DSB repair pathways, the broken end is processed to generate a 3′ single-strand extension, and this end must then be used in a homology-dependent strand invasion step to reconnect the branch with the rest of the chromosome. In contrast, a collision with a lesion can leave single-strand gaps in the fork structure (Figure 1*B*). Gap repair may often involve a facilitated migration of the branched fork structure. Movement of the fork backward, so as to reanneal the original template strands and displace the newly synthesized strands, is called fork regression. The displaced strands may themselves anneal, generating a Holliday structure with a short arm (14), a structure recently dubbed the "chicken foot" (15). Fork regression may occur spontaneously in some situations (15), and is also promoted enzymatically by the RecG, RuvAB, and RecA proteins in *Escherichia coli*. The chicken foot can be processed, by a second branch migration in the opposite direction (logically called fork progression) or by digestion of the short arm of the chicken foot, to restore a forklike structure (16). Alternatively, the four-way junction may be cleaved, shunting the repair into the DSB repair pathways (7). The scheme in Figure 1*B* illustrates only one of the potential structures that might be present in a fork stalled at a barrier, and many variations of these repair pathways may exist to accommodate other structural situations (10). In general, the major pathways for replication fork repair make use of some subset of the bacterial recombination functions, as well as the origin-independent replication restart system featuring the restart primosome (17).

Most of the available repair paths appear to make use of the bacterial RecA protein at one step or another (Figure 1). The bacterial RecA protein is the first-known member of a still-expanding list of the family of proteins sharing the capacity to catalyze DNA strand exchange reactions. The family is a rich source of enlightening structural and functional comparisons. The family includes the UvsX protein of bacteriophage T4 (18, 19), the archaeal RadA protein (20), and the eukaryotic Dmc1 (21) and Rad51 (22, 23) proteins, all characterized in vitro. Dmc1 protein forms octameric rings that so far appear to be its functional form (24). All of the other proteins form a filament on DNA with a strikingly similar appearance in the electron microscope (20, 22, 25, 26). This nucleoprotein filament then aligns the bound single strand with a homologous duplex, and promotes a strand exchange between them. The most common model systems used for DNA strand exchange in vitro are shown in Figure 2.

Although considered both structural and functional homologs, the RecA, Rad51, and other proteins exhibit structural and functional differences that have been highlighted in recent work. The primary sequence conservation is seen in the core domain of the RecA protein (Figure 3). The bacterial RecA proteins include a carboxyl-terminal domain not found in the others, while the RadA, Rad51, and Dmc1 proteins all have an amino-terminal domain not shared by RecA protein (27, 28). ATP is hydrolyzed by the UvsX and RecA proteins some one to two orders of magnitude faster than by the other proteins. All the proteins promote the basic process of DNA strand exchange, without a requirement for

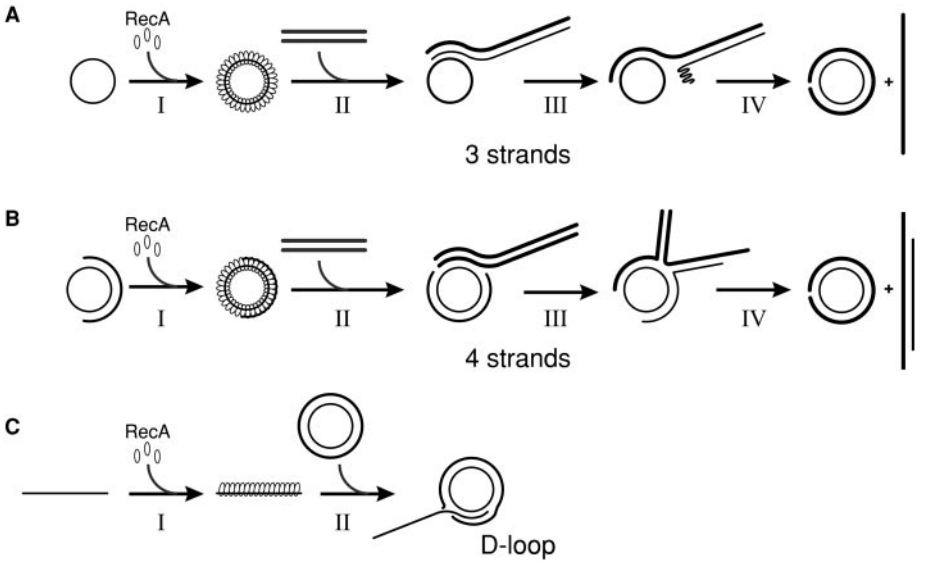

Figure 2 The major models of DNA strand exchange used to explore RecA protein function in vitro. DNA substrates for the three- and four-strand exchange reactions (*A* and *B*) are generally derived from bacteriophage DNAs. Roman numerals refer to distinct phases in the reactions that are detailed in the text. The D-loop formation reaction (*C*) probably mimics RecA function in the repair path for double-strand breaks shown in Figure 1A.

ATP hydrolysis (29–35). The bacterial RecA protein is unique so far in augmenting this process with an ATP-driven motor function, as detailed below. Bacterial RecA proteins such as that of *E. coli* are also multifunctional. In addition to the direct role in recombination manifested in its DNA strand exchange activity, RecA protein has two other distinct activities. (*a*) The RecA protein has a regulatory role in the induction of the SOS response, where RecA filaments on the DNA promote the autocatalytic proteolysis of the LexA repressor, the UmuD protein, and certain other proteins (36). (*b*) The RecA protein is required for the mutagenic bypass of DNA lesions during the SOS response, a reaction catalyzed by DNA polymerase V (37, 38).

If the primary function of bacterial recombination is the repair of stalled replication forks, then the biochemistry of bacterial recombination proteins is best considered in this light. This review focuses on the bacterial RecA protein, incorporating information about RecA homologs where that information provides insight into RecA function. Activities are presented in the context of replication fork repair, emphasizing the DNA strand invasion and strand exchange functions, with a detailed listing of potential RecA effects on repair pathways for replication forks provided at the end. The roles of the RecA protein

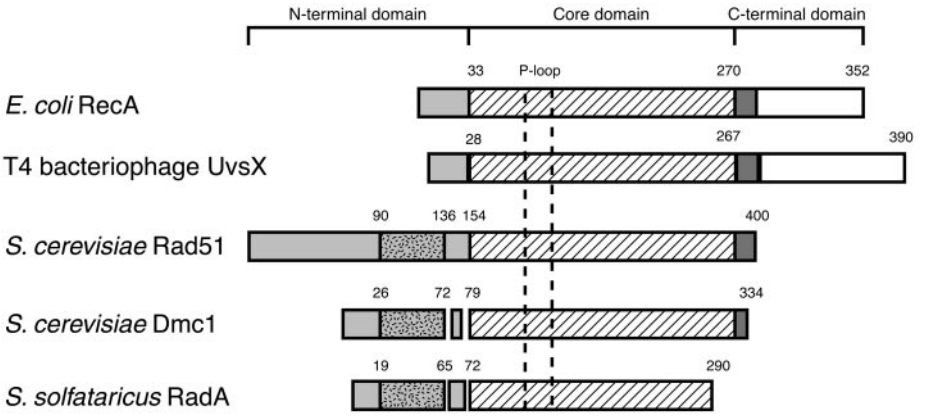

Figure 3 Comparison of structural domains in the *E. coli* RecA protein with the bacteriophage T4 UvsX, yeast (*Saccharomyces cerevisiae*) Rad51 and Dmc1, and the archaeal (*Sulfolobus solfataricus*) RadA proteins. Core domain conservation is depicted with diagonal lines. N-terminal domain conservation between Rad51, Dmc1, and RadA is shown as a spotted region. Regions with no sequence homology include the light gray shaded regions of the N-terminal domains and all regions C-terminal to the core domain, including the C-terminal domains of the *E. coli* RecA and T4 UvsX proteins.

in SOS induction and in SOS mutagenic lesion bypass have recently been reviewed (36, 38), and are not considered further here.

RecA PROTEIN ACTIVITIES

Fundamentals

The RecA protein of *E. coli* is DNA-dependent ATPase and an ATP-dependent DNA binding protein. It comprises 352 amino acid residues and has a calculated molecular weight of 37,842 after the initiating methionine residue is removed. The measured isoelectric point of the protein is 5.6 (27).

DNA Binding

Binding to DNA is a fundamental aspect of the role of RecA protein in DNA metabolism, and is a multistep process that results in the establishment of a nucleoprotein filament. The first step consists of the nucleation of a RecA monomer onto the DNA. This is the slow step of the binding process (39–41). Nucleation is considerably faster on single-stranded DNA (ssDNA) than on double-stranded DNA (dsDNA), and the most effective nucleation site on a chromosome would be an ssDNA gap. The binding of RecA to dsDNA is slower

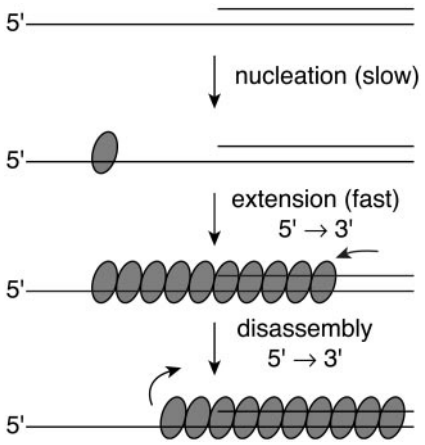

Figure 4 DNA binding pathway for the RecA protein. DNA binding includes distinct nucleation and filament extension steps, with nucleation occurring most readily on single-stranded DNA. Both the extension of the filament during filament assembly and the dissociation of the filament are end dependent and proceed 5′ to 3′ along a single strand.

at physiological pH and above than at pH values below 6.5, and the lag in binding at the higher pH values is due to slow nucleation (39–41). Nucleation to dsDNA can be facilitated by distortions in the DNA structure, especially DNA unwinding (40).

Once nucleation has occurred, monomers are added in an extension phase to form the nucleoprotein filament. Monomers polymerize rapidly and cooperatively in the 5′ to 3′ direction (Figure 4) (42–44). The cooperativity of DNA binding complicates the measurement of binding affinity by classical biochemical techniques.

In general, the binding of RecA to DNA is nonspecific with respect to sequence, but it apparently exhibits an enhanced binding affinity for poly(dT) polymers greater than 50 nucleotides in length (45–47) and for some GT-rich sequences that are similar to the recombination hot spot Chi (48, 49). Furthermore, nucleation is hindered by stacked purine sequences and secondary structure (50).

RecA Nucleoprotein Filaments

In the presence of ATP, dATP, or an adenine nucleotide analog with a triphosphate-like structure (ATPγS or ADP·AlF_4^-), RecA monomers polymerize on ssDNA to form an extended nucleoprotein filament. Electron microscopy analysis has shown this "active" form is a right-handed helical filament with a pitch of 95 Å, an axial rise per nucleotide of 5.1 Å, a diameter of 100 Å, six RecA monomers per turn, and three nucleotides per RecA monomer (51–54). When RecA is bound to dsDNA, the same extended filament is observed and the DNA is underwound relative to the B-form helix (55, 56). In contrast, the RecA filament formed on DNA in the absence of cofactor or with ADP is the "collapsed" nucleoprotein filament. This compact, "inactive" form exhibits a helical pitch of 64 Å and an axial rise of 2.1 Å per nucleotide (57, 58).

RecA protein filaments both assemble and disassemble in a filament end–dependent manner, with disassembly also occurring 5′ to 3′ on ssDNA (59) and at the end opposite to that of assembly (44, 60). The rate of disassembly is approximately 70 min^{-1} on ssDNA at 37°C (61). ATP hydrolysis occurs uniformly throughout the filament (62), with no enhancement at a filament end. Some exchange of RecA monomers has been detected within filament interiors under some conditions (63), although this might be attributed to DNA ends present at filament discontinuities.

DNA Strand Exchange Reactions

The homologous recombination of DNA in vivo is studied in vitro by a reaction that can be broken down conceptually into distinct steps (Figure 2). There is an ordered binding of ssDNA first, followed by the dsDNA substrate, focusing the reaction on single-strand gaps and ends where repair is likely to be needed. In phase I, RecA forms a nucleoprotein filament stoichiometrically on ssDNA, as described above. In the four-strand exchange reaction, the protein nucleates on the single-stranded region of a gapped duplex molecule but can readily extend into the duplex portion of the molecule. Phase II is characterized by the alignment of homology. The nucleoprotein filament recruits a linear dsDNA molecule and the homology between the two DNAs is aligned. In the four-strand reaction, this pairing occurs within the single-stranded region of the gap. Several hundred base pairs of hybrid DNA are formed in phase III as the aligned DNA strands are rapidly switched. Branched DNA intermediates are migrated by the displacement of the like strand in the duplex molecule (Figure 2). Only the binding of ATP, and not its hydrolysis, is required for phases I, II, and III of the three-strand reactions (30, 33–35). The four-strand reactions do not occur without ATP hydrolysis (31). The segment of hybrid DNA formed in phase III is extended in phase IV of the reaction. This RecA-facilitated migration of branched intermediates (extension) proceeds in the 5′ to 3′ direction (with respect to the initiating strand) until products are formed and requires ATP hydrolysis (64).

DNA Pairing

The homologous alignment that occurs in phase II deserves some elaboration, as it represents one of the central processes of recombination. The mechanism by which RecA mediates homologous pairing remains a focus of investigation, and new complexities continue to be revealed that have not yet been fully explained. Analysis of the kinetics of DNA pairing with short oligonucleotides reveals at least two steps, a rapid second-order alignment of the two homologous DNA molecules, followed by a slow first-order process (65–67). The initial alignment may involve base flipping by some bases in the duplex (especially A and T bases) to align the DNAs via Watson-Crick pairing, as revealed for the human Rad51

protein (68). In this mechanism, the slow first-order process may be a completion of the strand switch to include all remaining base pairs over an extended region.

The fundamental process of DNA pairing and strand transfer occurs within the filament groove, an arrangement originally proposed by Paul Howard-Flanders (69). The DNA pairing promoted by the *E. coli* RecA features an ordered binding of DNA substrates, in which a single strand is bound first and the homologous duplex is drawn in second. An active site constructed to promote this reaction should in principle be able to promote the same process in the opposite direction. An inverse reaction has been reported for the *E. coli* RecA protein, in which the duplex DNA is bound first and the single strand is drawn in (70). The reaction has been demonstrated only for short duplex DNA substrates (63 base pairs) in which RecA protein filaments are stabilized by the presence of long circular single strands contiguous with the duplex.

In principle, the duplex could approach the ssDNA within the filament via either its major or minor groove. Within the major-groove-first path, a novel triplex DNA intermediate, formed prior to a strand switch or base flipping, has been proposed. This triplex has been called R-form DNA (71, 72). In the alternative minor-groove-first path, homologous alignment would involve only the standard Watson-Crick base pairing provided by base flipping. As the duplex bound, it would be extended and underwound such that its bases would be free to rotate and "sample" the bound ssDNA for complementarity. Both the RecA and Rad51 proteins promote DNA strand exchange readily with DNA substrates heavily substituted with base analogs that disrupt many of the interactions proposed to stabilize the R-form triplex as it has been presented (73–75). Overall, the weight of the evidence currently supports a minor-groove-first pathway for DNA pairing that excludes an R-form triplex intermediate (68, 76–81), although arguments for a major-groove-first path have also appeared (82). The RecA filament appears to stabilize the products of DNA strand exchange, using binding energy to promote the strand switch (83).

The "search for homology" involves the transient underwinding of the duplex DNA substrate, an effect that can be observed with heterologous DNA substrates (84). In addition, the binding of a heterologous duplex DNA to a RecA-ssDNA complex activates the duplex for a strand exchange reaction with a homologous single strand introduced external to the filament (85). This strand exchange in *trans* is a much weaker reaction than the conventional strand exchange, and has been demonstrated only for short oligonucleotides. At a minimum, interesting structural changes are conferred on a duplex DNA when it binds to a RecA nucleoprotein filament, and the alterations are likely to facilitate strand exchange reactions in general.

ATPase Activity

The RecA protein hydrolyzes ATP to ADP and inorganic phosphate. As with DNA binding, ATP binding and hydrolysis is a cooperative process. Except at very high salt concentrations (85a) ATP hydrolysis is DNA-dependent. RecA

monomers in the filament hydrolyze ATP with a rate of catalysis (k_{cat}) of about 30 min^{-1} on ssDNA. This agrees with the measured ATPase rate during phase I (see above) of a three-strand exchange reaction. The rate of ATP hydrolysis drops to about 16–20 min^{-1} when the duplex molecule is added to the reaction (phase II) or when a filament is formed on dsDNA. The role of ATP hydrolysis in RecA-mediated strand exchange is explored below.

RecA PROTEIN STRUCTURE

Sequence Alignments

The bacterial RecA protein is a highly conserved polypeptide chain. Primary sequence alignments of the RecA proteins from 67 bacterial species have been published in two independent reports (27, 86). Based on these alignments, a consensus sequence for all bacterial RecA proteins was generated (Figure 5). With the *E. coli* RecA protein as a reference, the percentage of identical amino acid residues in bacterial homologs ranges from 49% for *Mycoplasma pulmonis* to 100% for *Shigella flexneri*. Additionally, many nonidentical aligned residues retain chemical similarity (Figure 5). In the following discussion, all amino acid residue numbers refer to the *E. coli* protein sequence for simplicity.

Structural Information from X-Ray Crystallography

There are four reported X-ray crystal structures of bacterial RecA proteins, two from *E. coli* (87, 88) and, more recently, two from *Mycobacterium tuberculosis* (89). Even though *M. tuberculosis* RecA is only 62% identical to *E. coli* RecA, the structures are nearly identical with a root mean squared deviation (RMSD) of only 0.6 to 1.1 Å, depending on the extent of alignment. All of these structures lack DNA, but one *E. coli* structure includes an ADP cofactor and one *M. tuberculosis* structure includes a nonhydrolyzable ATP analog. The RecA structures have revealed a central core domain and two smaller domains at the amino (N) and carboxyl (C) termini (Figure 6).

The core domain of the RecA protein is structurally homologous to several proteins, to which it bears very little to no sequence similarity. The structural units of hexameric helicases are RecA-like domains (90). Furthermore, 120 α-carbon atoms of the core domain of RecA can be spatially aligned with the mitochondrial F_1-ATPase and the cobalamin nucleotide loop assembly protein CobU, with an RMSD of less than 2 Å (91, 92).

The crystal structures provide information about both the RecA monomer and the filaments it forms. Polymerization of RecA monomers in the absence of DNA has been observed in vitro and in the crystal structure. In the crystal, monomers are packed so as to form a right-handed helical filament with six monomers/turn (Figure 6). Thus, the monomer structure (Figure 6) is a

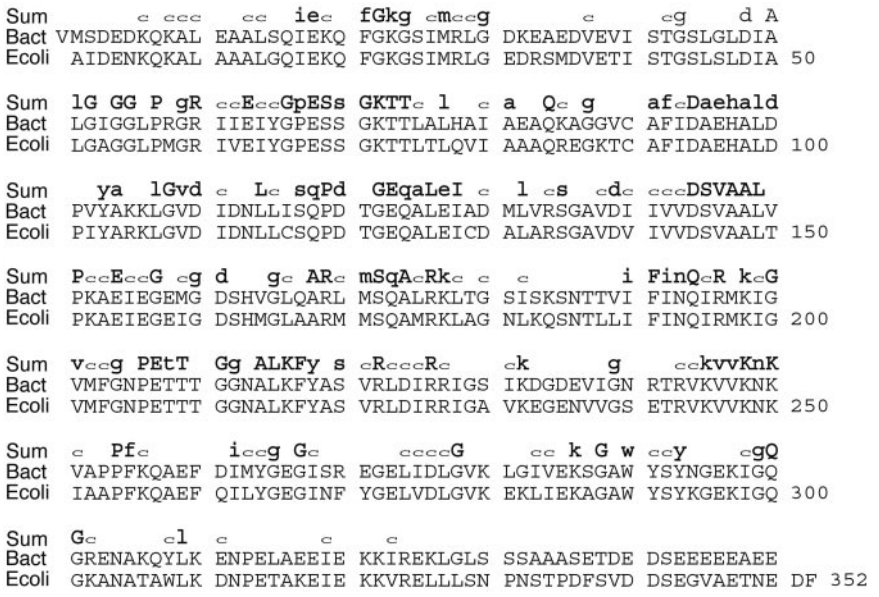

Figure 5 Bacterial RecA primary sequence alignment. The bacterial RecA consensus sequence (Bact) based on 67 sequence alignments was generated by determining the residue most conserved in that position. Above this consensus sequence is a summary (Sum) of the most conserved residues. A capital letter above a residue in the consensus sequence means it is invariant across all bacterial RecA proteins in the alignment. A lowercase letter refers to those residues whose identity is conserved more than 90% of the time. A small c denotes a residue that is chemically conserved in more than 90% of the aligned sequences, based on the following criteria: aromatic (FWY), hydrophobic (ILMV), hydrophilic (ST), small aliphatic (AG), amides (NQ), acidic (DE), and basic (HKR). The consensus sequence is listed directly above the RecA proteins from *E. coli* (Ecoli). Although the average length of all the sequences aligned is 352 amino acid residues (the range was 318 residues in *Bacteroides fragilis* to 388 residues in *Streptococcus pneumoniae*), the consensus sequence is, on average, one residue longer on the N terminus and two residues shorter on the C terminus relative to the *E. coli* RecA protein. The numbers refer to the residues of the *E. coli* protein from the N terminus to the C terminus with the initiating methionine residue omitted.

representative unit of the polymer and not the monomeric form of the protein. The core domain of the RecA protein (residues 34–269) is the part of the protein most highly conserved among bacterial species. This conservation also extends to eukaryotic homologs. In the crystal structures, this domain consists of a mixed, eight-stranded, twisted β-sheet flanked by four α-helices.

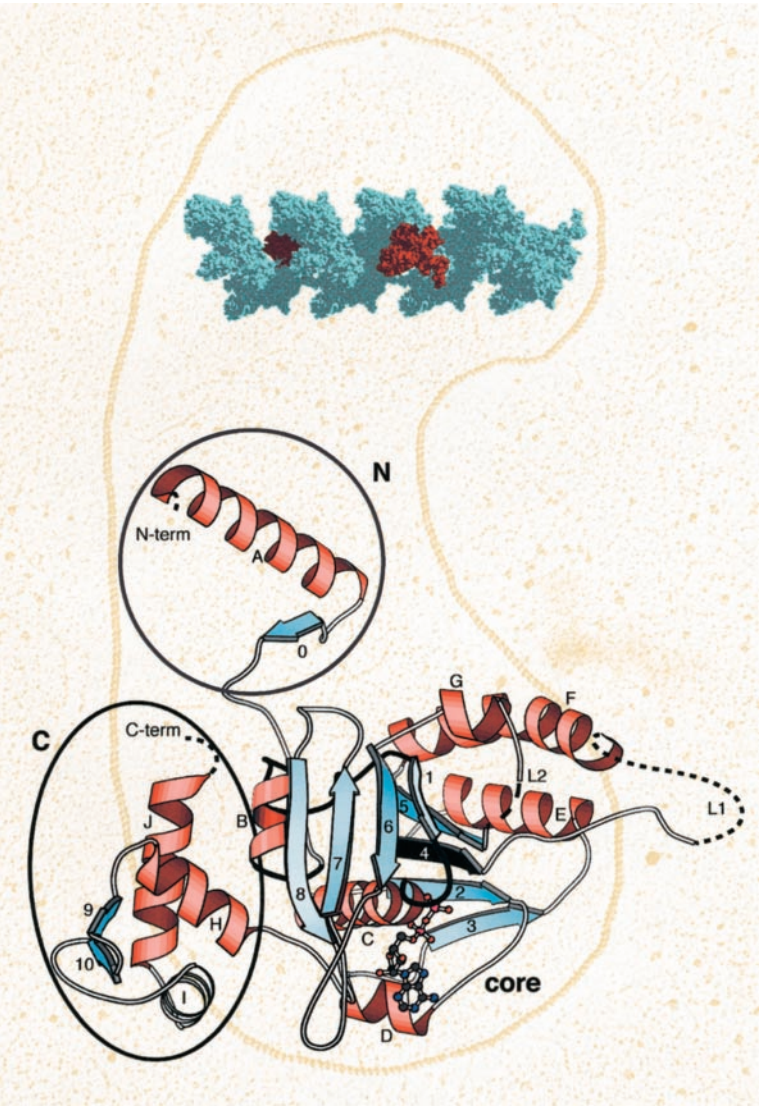

Figure 6 The structure of RecA protein and its filaments. A RecA monomer is shown in ribbon form at the bottom, with β-sheets numbered and α-helices lettered according to the established conventions used in the text. The N- and C-terminal domains are circled, and a bound ADP is seen at the bottom of the structure within the core domain. A RecA filament with 24 monomers is shown in the inset at the top, with one monomer shown in a darker color. In the background is an electron microscope (EM) image showing a RecA protein filament formed on a circular duplex DNA molecule in the presence of ATPγS. The helical groove in the filament is manifested by striations in the EM image.

The Monomer-Monomer Interface

Steitz and coworkers determined those residues whose solvent-accessible surface area decreased significantly upon polymerization (87). In the filament visualized by X-ray crystallography, β-strand 3, the loop C-terminal to β-strand 2, and α-helix D of the core domain form one surface that packs against α-helix A and β-strand 0 (the N-terminal domain) of an adjacent monomer. The molecular details of this monomer/monomer interface are generally supported by mutagenesis studies. Residues 95, 97, and 99 have solvent-accessible surface areas that decrease by more than 15 $Å^2$ upon polymer formation (87), and the RecA His97Ala mutant is deficient in polymerization (93). The region of residues 117–128 is also highly conserved; many of these residues contribute to the monomer/monomer interface in the crystal structure, and the Ser117Phe mutation affects RecA polymerization (94). The region of α-helix G and amino terminal side of β-strand 6 (residues 214–227) is another well-conserved segment. The crystal structures implicate this region in monomer/monomer contacts from the same face as the N terminus itself. Additionally, Knight and coworkers have shown that Lys216, Phe217, and Arg222 are critical residues in stabilizing the RecA polymer (95–97).

ATP Binding Site

The core domain region of residues 47–74 is quite well conserved with 14 invariant residues. It contains the nucleotide binding loop between β-strand 1 and α-helix C. The *E. coli* sequence GPESSGKT matches the consensus sequence of amino acids (G/A)XXXXGK(T/S) for the Walker A box (98) (also referred to as the P-loop) found in a number of nucleoside triphosphate (NTP)-binding proteins. The RecA Lys72Arg mutant protein is ATPase deficient while retaining nucleotide binding (33, 34).

Other parts of the ATP binding site are also highly conserved, including residues 140–150. This region contains six invariant residues and includes another nucleotide binding motif, the Walker B box (98). This motif is found at β-strand 4 in the RecA structure (residues 140–144). The Walker B box is characterized by four hydrophobic amino acids followed by an acidic residue (usually aspartate). Interestingly, a non-prolyl *cis*-peptide bond was found between Asp144 and Ser145 in the structure. This uncommon peptide bond configuration was also found in both subunits of the F_1-ATPase as well as the CobU protein. This is likely a conserved motif that is a feature of the RecA-like fold of the core domain, and it will be interesting to see if this is conserved in more structural homologs of the RecA protein as they are discovered.

Nucleotide specificity and additional ATP binding interactions are contributed by the amino acid residues at β-strand 2 and the loop C-terminal to that strand (residues 90–100), all of which are greater than 90% conserved among bacterial RecA proteins. Ala95 is almost completely conserved and is responsible for stabilizing the non-prolyl *cis*-peptide bond described above through a main chain

hydrogen bond. Glu96 is thought to activate a water molecule for in-line attack of the γ-phosphate of ATP, and mutation of this residue to an aspartate renders the protein ATPase deficient (99). Asp100 forms a hydrogen bond with the N^6 amino group of the adenosine base (88). This specificity can be changed from ATP to ITP, which has an N^6 carbonyl group, by replacing the aspartate with an asparagine (100).

DNA Binding

In the absence of a RecA-DNA co-crystal, much of the discussion about the DNA binding sites of RecA protein has focused on two loops within the highly conserved regions of residues 151–176 and 190–227. These loops, disordered in the crystal structure, are commonly referred to as L1 (residues 157–164) and L2 (residues 195–209). Both of these regions are moderately well conserved among bacterial RecA proteins, although the conservation does not extend to the archaeal and eukaryotic homologs. Although the loops are disordered in the crystal structure, it seems likely that they may become ordered upon DNA binding. The implication of a DNA binding function for these loops stems in part from the location of the peptide endpoints deep within the filament groove (87) that was determined by Egelman and coworkers, using electron microscopy, to be the primary DNA binding site (101). It is worthwhile to note that residues that make up L1 in MtRecA are ordered in that structure and, in fact, L1 is oriented into the groove (89). A number of mutagenesis and DNA cross-linking studies have further implicated these regions in DNA binding (for a review see 27). Detailed mutagenesis of loops L1 and L2 have been carried out by the Knight (102) and Camerini-Otero (103) groups, respectively. DNA cross-linking studies have supported a role for these loops in DNA binding (104, 105). Conversely, residues outside of these regions, such as Tyr103 (106), Lys183 (106, 107), and in the region of residues 233–243 (107), have also been shown to cross-link to DNA. Many details of the RecA-DNA interaction remain to be elucidated.

The Carboxyl Terminal Domain

The C-terminal domain of the RecA protein (residues 270–352) exhibits the least amount of sequence conservation. This domain folds into α-helices H, I, J and β-strands 9 and 10 and is positioned distal to the filament axis in the polymer structure (Figure 6). This solvent-exposed domain appears as lobes on the exterior of the filament in the electron microscope. Recently, Egelman and coworkers have observed C-terminal domain movement relative to the core domain that may be responsible for (or diagnostic of) the active or inactive state of the RecA filament (26). The last 25 residues of the RecA protein are disordered in the crystal structure. This region includes a high concentration of negatively charged residues. In fact, while there is little sequence identity in the C terminus of bacterial RecA proteins, most bacterial RecA sequences display a concentration of negatively charged residues. Other ssDNA binding proteins

such as SSB of *E. coli* (108) and the gene 32 protein of phage T4 (109) also have highly negatively charged C-terminal regions. Upon C-terminal truncation to remove these negative charges, RecA, SSB, and the gene 32 protein show increased dsDNA affinity relative to the intact protein (108–111; SL Lusetti & MM Cox, unpublished data). It has been proposed that the negatively charged C terminus of RecA regulates the direct binding of RecA to dsDNA (in the primary binding site) by electrostatically repelling the phosphate backbone of the DNA (110, 111).

During strand exchange, RecA protein generally forms a filament on an ssDNA. The RecA-ssDNA nucleoprotein filament must then bind a homologous dsDNA molecule. The proposed binding site for this second DNA molecule is in the RecA filament cleft between the C-terminal domain at β-strand 9 and α-helix I and the core domain at β-strand 7 (113–115).

THE MOLECULAR FUNCTION OF RecA PROTEIN–MEDIATED ATP HYDROLYSIS

There are many differences between RecA protein and homologs like the Rad51 and RadA proteins. Few are as dramatic as the utilization of ATP by the RecA protein. Comparisons in this review focus on the Rad51 protein, since in most cases the key experiments have not yet been attempted for the Dmc1 and RadA proteins. As already noted, RecA protein catalyzes a much faster hydrolysis of ATP and dATP. ATP hydrolysis has two clear functions. First, it is coupled to the 5′ to 3′ end-dependent disassembly process (44, 59, 61). ATP hydrolysis by the monomer at the 5′ proximal end of the filament may result in dissociation, with the overall likelihood determined by pH and other solution parameters (44). Second, ATP hydrolysis is directly coupled to DNA strand exchange.

Although DNA pairing and some strand exchange can occur in its absence, ATP hydrolysis confers several new properties to the RecA-promoted reaction. The movement of DNA branches during the late stages of strand exchange is rendered unidirectional (34, 64). RecA can promote extensive (multi-kilobase-pair) DNA strand exchanges much more efficiently (34, 64). The strand exchange can bypass heterologous insertions or other barriers that may extend for a few hundred base pairs (32, 34). Finally, the reaction can accommodate four DNA strands (31, 34, 116).

Extensive DNA strand exchange is promoted by the yeast and human Rad51 proteins in the absence of high levels of ATP hydrolysis (75, 117). It is not yet clear if ATP hydrolysis is coupled to dissociation of Rad51 protein from DNA. However, the Rad51 protein cannot promote DNA strand exchange through significant heterologous insertions (118, 119) or promote four-strand exchange reactions (P Sung, personal communication), whether or not ATP is hydrolyzed. The bias in the direction of DNA strand exchange seen in some studies is largely accounted for by a requirement for short single-strand extensions at the end

where strand exchange is initiated (118, 120), as opposed to an active process in the RecA filament driving the reaction in one direction. ATP hydrolysis appears to play no role in the polarity of a Rad51-mediated DNA strand exchange reaction. Thus, Rad51 appears to lack much of the functionality that requires ATP hydrolysis in the RecA system.

The coupling of ATP hydrolysis to DNA strand exchange by RecA protein, as manifested in the activity enhancements noted above, suggests the presence of a motor function in RecA protein that Rad51 protein appears to lack. The proposal that RecA protein possesses a motor function has evolved over the last 15 years (3, 34, 121–124). The current iteration may help to explain RecA function at a stalled replication fork.

THE RecA MOTOR

To explain the coupling of ATP hydrolysis to RecA protein–mediated DNA strand exchange, a model must explain (*a*) the requirement for ATP hydrolysis to bypass heterologous insertions (32, 34), (*b*) the requirement for ATP hydrolysis in four-strand exchange reactions (31, 34), (*c*) the apparent limitation for facile coordination of only three DNA strands within the RecA filament (125–130), and (*d*) the known disposition of RecA protein filaments during DNA strand exchange reactions in vitro (131, 132). Some ideas for the use of ATP hydrolysis by RecA (69, 133, 134) can be eliminated based on one or more of these criteria (most were proposed prior to the experimentation that established the criteria). Note that all of these models have elements unrelated to ATP hydrolysis that still contribute prominently to current thinking.

Two models for the coupling of ATP hydrolysis to DNA strand exchange have retained a capacity to explain most current data. One involves a simple redistribution of RecA monomers within a filament coupled to ATP hydrolysis (30, 33, 35). This proposal obviates the need for a motor by explaining most experimental observations in terms of a simple DNA association/dissociation carried out by RecA monomers. The other involves a motor function in which the ATP hydrolysis is coupled to a rotation of the two DNA substrates about one another (3, 121, 123).

The RecA Redistribution Model

This model posits that the only limitation to extensive DNA strand exchange comes in the form of filament discontinuities. When ATP is not hydrolyzed (e.g., when a non-hydrolyzed analog such as ATPγS replaces ATP), these discontinuities halt DNA strand exchange. ATP hydrolysis allows the dissociation and redistribution of RecA monomers to fill in these discontinuities. An exchange of RecA monomers between free and bound forms occurs during DNA strand exchange (135), as predicted by this model. The model also provides a mecha-

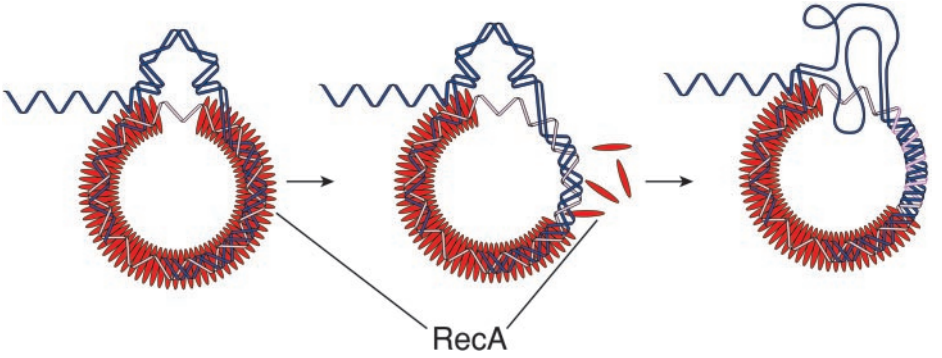

Figure 7 A model for the bypass of a heterologous insertion via RecA protein redistribution. Duplex DNA is paired with the RecA-ssDNA complex on both sides of the insertion, and strand exchange proceeds up to the insertion on one side. Dissociation of RecA protein then releases underwound DNA. With no way to eliminate the DNA torsional stress via rotation of the DNA at an end (the DNAs are circular and thus tethered), the torsional stress can be transmitted to the insertion and used to unwind it. Strand exchange can then proceed again on the opposite side of the unwound insertion. Note that if the substrate single strand is linear rather than circular, the torsional stress resulting from RecA dissociation can in principle be dissipated by DNA rotation at the ends.

nism for strand exchange to bypass heterologous insertions in the duplex DNA substrate (Figure 7) (136). In brief, strand exchange would proceed in an ATPase-independent way up to the heterology. The homologous duplex DNA would be taken into the filament on both sides of the heterology. Since RecA-bound DNA is highly unwound, ATPase-dependent dissociation of RecA monomers would release underwound DNA. If the superhelicity of this DNA was not dissipated by free rotation of the DNA at an end (and it would be so constrained in the standard circular DNA substrates), the resulting torsional stress could be transmitted into an unwinding of the duplex DNA within the heterologous insertion. This strand separation would allow strand exchange to continue on the other side of the insertion. In effect, the bypass of a heterologous insertion requires only that the RecA protein bind so as to underwind the DNA and dissociate in a reaction coupled to ATP hydrolysis.

There are now four experimental tests that conflict to some degree with the RecA redistribution model. First, in the absence of ATP hydrolysis, the addition of excess RecA protein (to fill in any filament discontinuities) is not enough to remove the obstacles to extensive DNA strand exchange with long DNA substrates (34). This test is limited in that the DNA exposed within filament discontinuities may not always occur in multiples of the three-nucleotide binding site needed by RecA. Second, the exchange of RecA protein into and out of filaments during strand exchange is not enough in itself to complete strand exchange and bypass heterologous insertions. In mixed filaments that include

small amounts of the RecA mutant K72R, which binds but does not hydrolyze ATP, DNA strand exchange can be blocked, although some RecA monomer exchange still occurs (135). Third, the mechanism of Figure 7 predicts that heterology bypass should not occur if both of the DNA strand exchange substrates are linear. Free rotation at the ends would dissipate the superhelical stress in the DNA released by RecA and thus prevent the needed unwinding of the insert DNA. However, heterologous insertions are bypassed on linear DNA substrates as well as in a reaction using circular ssDNA (124). Finally, mutant RecA proteins with short C-terminal deletions exhibit nearly normal ssDNA binding, DNA extension, ATPase, DNA dissociation, and DNA strand exchange activities (110, 111, 137; S. L. Lusetti & M. M. Cox, unpublished data). These mutants exhibit somewhat enhanced binding to dsDNA (110, 111, 137; S. L. Lusetti & M. M. Cox, unpublished data). These proteins can thus bind to and extend DNA, and dissociate in a reaction coupled to ATP hydrolysis. However, this is clearly not enough to bypass a heterologous insertion, as these proteins have greatly reduced or absent bypass activities. The deletion mutants also cannot promote a four-strand exchange reaction (S. L. Lusetti & M. M. Cox, unpublished data).

Another problem is the lack of a proposal to explain the requirement for ATP hydrolysis in four-strand exchange reactions within the RecA redistribution model.

The Facilitated DNA Rotation Model

This model (a RecA motor) posits instead that the barrier to extensive RecA-mediated DNA strand exchange in vitro is the occurrence of unproductive DNA pairing or other interactions between the RecA-ssDNA complex and the duplex DNA substrate. When DNA strand exchange is initiated with a RecA-class recombinase, extension of the initial segment of hybrid DNA requires a continued spooling of the duplex DNA into the filament. A secondary DNA pairing event, or a nonspecific interaction of the duplex with the filament, can halt the exchange (Figure 8*A*) (34, 75, 117). In the case of RecA protein, these loops of DNA outside the filament can be acted on and rotated about the outside of the filament, in a reaction coupled to ATP hydrolysis. This model postulates a set of DNA binding sites on the outside and extending longitudinally along the surface of the filament. With six RecA monomers per turn in the helical filament, one postulates six of these exterior binding sites per filament, with DNA passed unidirectionally from one to another as ATP is hydrolyzed (Figure 8*B* and Figure 8*C*). This unidirectional rotation would generate a unidirectional DNA strand exchange.

Like all molecular motors, the facilitated DNA rotation model postulates a precise relationship between ATP hydrolysis and the motion it affects, in this case the branch movement during DNA strand exchange. If DNA is rotated in six steps around the outside of the filament, then each RecA monomer should hydrolyze one ATP during each 360-degree rotation of the DNA. Since the DNA within a RecA filament is underwound to approximately 18 base pairs (bp) per turn, that same rotation should move the branch by 18 base pairs. Thus, the

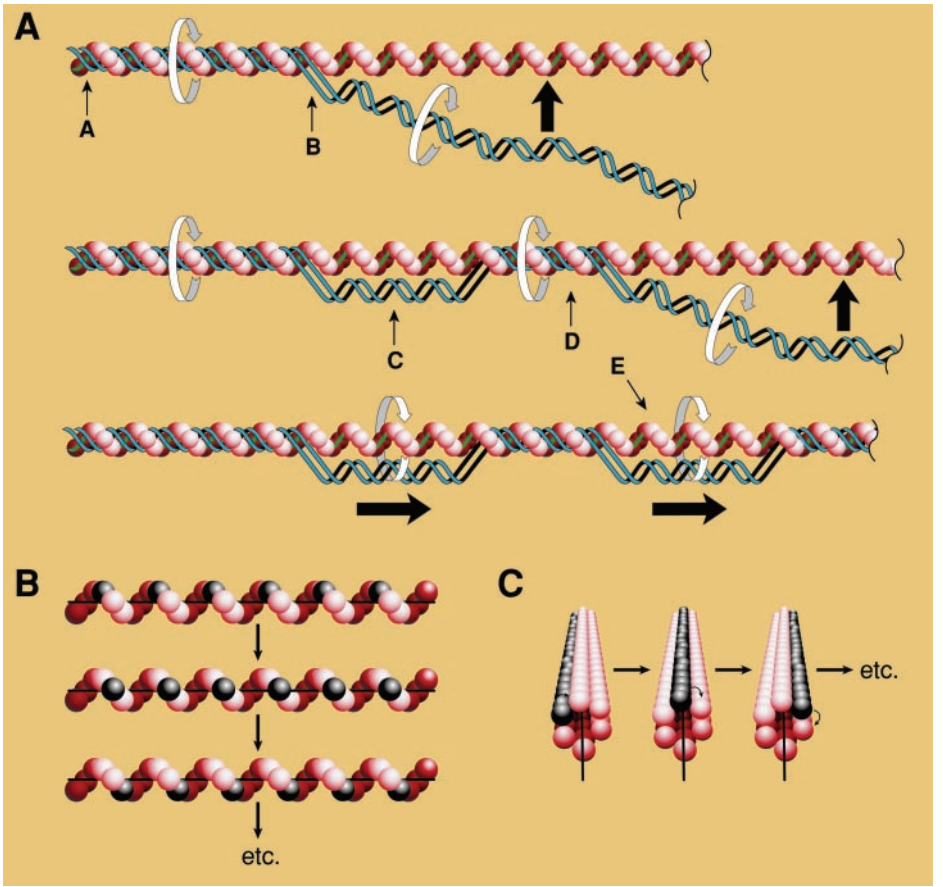

Figure 8 The facilitated DNA rotation model for RecA protein–mediated DNA strand exchange. (*A*) In this model, extensive DNA strand exchange in the absence of ATP hydrolysis is blocked by secondary DNA pairing events that leave segments of the DNA on the exterior of the filament. When ATP is hydrolyzed, these loops are rotated about the filament axis and about the bound DNA within the filament. The rotation results in unidirectional movement of the DNA branch. (*B* and *C*) The model posits a degree of organization of ATP hydrolytic events, such that the actions of every sixth monomer are coordinated as shown (i.e., the darkly colored monomers would hydrolyze ATP at the same time). The filament structure may have 6.2 RecA monomers per turn rather than the 6 shown, and in this case the "stripe" of ATP hydrolytic events would spiral somewhat about the filament axis rather than proceed longitudinally as seen here.

turnover number for ATP hydrolysis during strand exchange (in units of min^{-1}) should be related to the rate of branch movement (in units of bp min^{-1}), by a factor of 18 bp. Both parameters were examined carefully over the widest

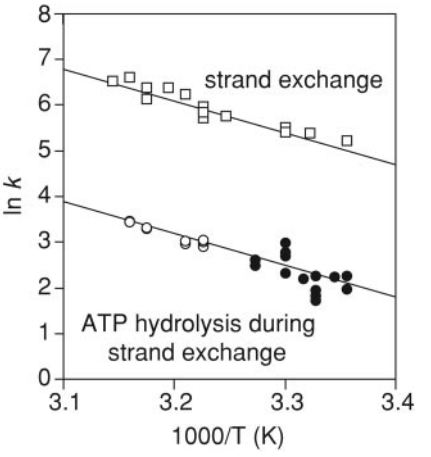

Figure 9 The relationship of ATP hydrolysis to the movement of the DNA branch during DNA strand exchange. [Data is an Arrhenius plot relating reaction rate constants (*k*) to temperature (in degrees kelvin) from Bedale et al. (138).]

temperature range possible, to generate the Arrhenius plot shown in Figure 9 (138). The lines described by each set of data are parallel within experimental error. More important, the lines drawn in the figure are separated by a factor of exactly 18 bp, substantiating a key prediction of the model. Work with the S69G mutant RecA protein, which affects the rate of nucleoside triphosphate (NTP) hydrolysis, has also demonstrated that increases or decreases in the rates of ATP hydrolysis can result in similar changes in the rate of DNA strand exchange. This provides more evidence for a direct link between ATP hydrolysis and branch movement in this system (139).

The model makes some additional predictions. The rotation of one DNA about the other should, in some circumstances, generate torsional stress that could be used to unwind DNA. The predicted generation of torsional stress can be seen in an indirect helicase activity demonstrated for RecA (124) and in the bypass of double-strand breaks during a four-strand exchange reaction (116) (Figure 10). The generation of torsional stress is also necessary for the bypass of heterologous insertions in the duplex DNA substrate in a three-strand exchange (140), although in this case such stress would be predicted by either the RecA redistribution or the facilitated DNA rotation models.

The DNA binding sites on the exterior of the filament, as predicted by the model, have not been identified. However, the carboxyl-terminal lobe of RecA protein is positioned so that it could be part of such external binding sites (26). As already mentioned, short deletions in the C-terminal domain of RecA largely eliminate what we have defined as motor functions (the capacity to bypass heterologous insertions or to carry out four-strand exchanges) while affecting few other RecA activities. The absence of a comparable C-terminal domain in Rad51 is consistent with the lack of heterology bypass or four-strand exchange activities in Rad51. This work suggests that the C-terminal domain plays a significant role in the coupling of ATP hydrolysis to DNA strand exchange.

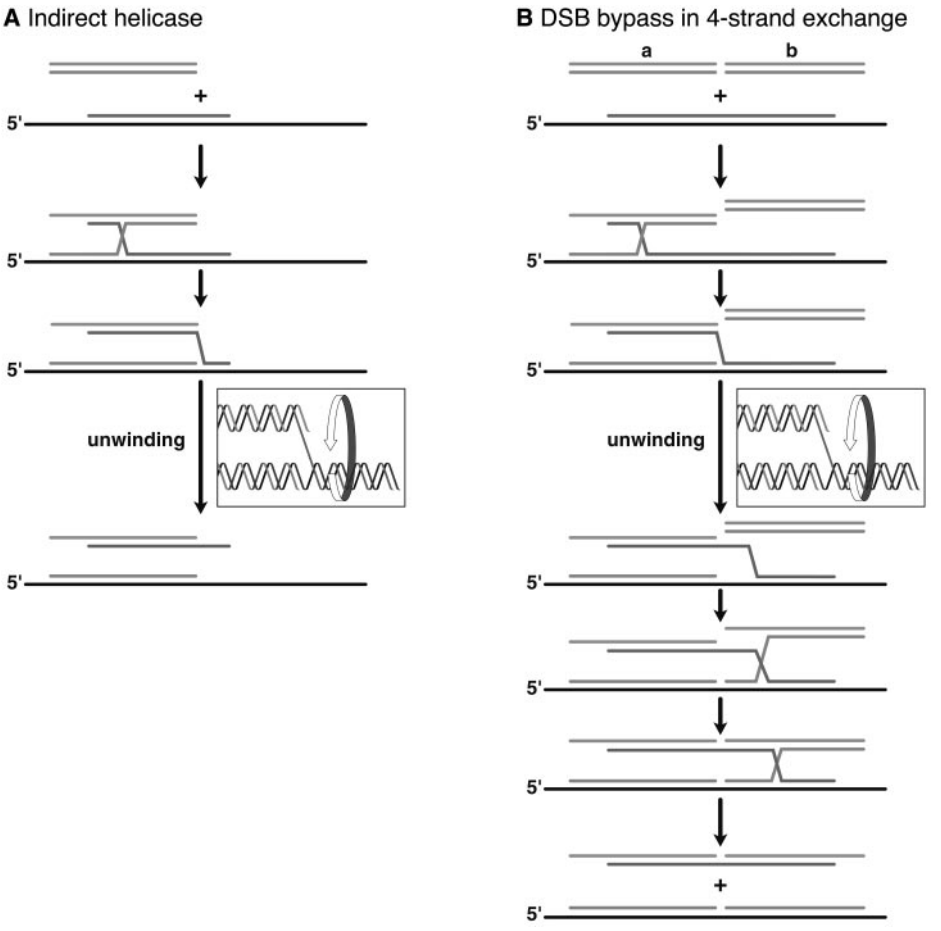

Figure 10 An indirect helicase function of RecA protein. In a four-strand exchange reaction such as that illustrated (with the exchanged strand in the gapped duplex longer than the linear duplex substrate), complete exchange is observed and requires a helicase-like activity of RecA protein. This type of strand separation is needed to bypass strand breaks in a four-strand exchange as shown.

THE MOLECULAR FUNCTION OF RecA PROTEIN AT STALLED REPLICATION FORKS

On the basis of the activities of RecA protein characterized to date, one can propose at least six different ways that RecA protein could affect the repair of stalled replication forks. Some of these are firmly based in experimentation in vitro and to some extent in vivo. Others are hypothetical and a few are just new ideas. There are many key reaction types among the pathways of Figure 1, and

RecA can contribute to or affect, at least in principle, a surprisingly diverse array of them.

3′ End Invasion

In the DSB repair pathways, the RecA protein promotes 3′ end invasion. The free dsDNA end is processed by the RecBCD enzyme to generate a 3′ single-strand extension. The RecBCD enzyme loads the RecA protein onto the single-stranded DNA (141, 142). This RecA nucleoprotein filament then promotes the strand invasion at a site in the dsDNA that is homologous with the bound single strand. This is perhaps the most established activity of RecA protein at replication forks, as well as the classical RecA function during conjugation and a variety of other recombinational processes. In the context of fork repair, strand invasion would occur at the site of a double-strand break generated by a fork encounter with a strand break as shown in Figure 1*A*. The RecA protein and all of its homologs promote the formation of D-loops in vitro (Figure 2*C*), a reaction that probably accurately mimics the 3′ end invasion process. There is no apparent need for a motor function in this process, and RecA mutant proteins that bind but do not hydrolyze ATP can adequately promote this reaction. The key advantage of a 3′ end invasion is that the 3′ end can subsequently be used as a replication primer.

5′ End Invasion

The RecA-mediated formation of D-loops in vitro exhibits a strong 3′ end bias (143, 144). This reflects the RecA filament assembly and disassembly process rather than a distinct end preference, since 5′ ends can also be used if RecA protein can be stabilized there (145, 146). In particular, the 3′ end bias is generally eliminated in the presence of the RecO and RecR proteins (147). Rosenberg and Hastings have proposed a split-end model for recombination (148, 149) that might be applied to double-strand break repair at a damaged fork. A 5′ end invasion could come in other forms as well. Since RecBCD requires nearly flush ends with less than 25 bp nucleotides of single-stranded DNA (150), any broken end with a 5′ extension longer than this would require a different repair path. Such ends might be generated by the action of exonuclease III or by the occurrence of a strand break within a single-strand gap. A RecAOR filament could provide the alternative to RecBCD needed for the repair of such ends.

Fork Regression of a Stalled Fork with a Leading Strand Gap

When a blocking DNA lesion is encountered by the replication fork on the leading strand, the lesion can be left in a long (~1 kilobase) single-strand gap (151), as illustrated in Figure 1*B*. An often-postulated pathway for repair of stalled forks of this kind is via fork regression. The original template strands are reannealed, and the newly synthesized strands are displaced, forcing the fork

backward. At least three plausible pathways have been identified for fork regression. These include a spontaneous, topology-driven process (151), catalysis by a branch helicase such as the RecG (13, 152) or RuvAB proteins (7), and catalysis by the RecA protein. Given a long gap in the leading strand, RecA protein promotes in vitro the regression of DNA substrate models for stalled forks (153). The reaction appears to be entirely dependent on ATP hydrolysis (153) and represents a potential application of the postulated RecA motor function. In a stalled fork, rotation of one branch about the other, in the manner prescribed by the polarity of RecA filament assembly in a leading strand gap, leads to reconstitution of the double helix of the original template and to fork regression. In vivo, such a fork regression reaction presumes the involvement of topoisomerases downstream to eliminate the DNA tangles that would result.

Fork Regression with a Lagging Strand Gap

RecA appears to be required for fork regression in vivo in cells deficient in DnaB function (154), lending further support to the idea that RecA protein is involved in at least some fork regression reactions. However, the involvement of the DnaB helicase, which travels along the lagging strand, suggests a role on lagging rather than leading strand gaps. The RecA protein cannot promote regression of a fork with a lagging strand gap by the kind of reversal of branch migration illustrated in Figure 1*B*. This is because the polarity of the RecA filament assembly on the lagging strand template, and any subsequent DNA strand exchange, would be in the same direction as normal fork progression (not regression). This does not exclude RecA from a role in fork regression from a lagging strand gap; it simply changes the nature of the reaction that must occur. What would have to happen is illustrated in Figure 11. RecA filaments would form in the lagging strand gap, and assembly would proceed up to (and perhaps beyond) the branch at the fork itself. The filament would then promote strand invasion of the single strand in the gap into the leading strand duplex. There would be no free DNA ends in such a reaction, and topology would prevent a net strand exchange without the assistance of either a nuclease or a topoisomerase. The hypothetical pathway outlined in Figure 1 represents what might occur by the combined action of RecA and a topoisomerase. Radding and colleagues have demonstrated that a reaction such as that depicted can be promoted with RecA protein and the *E. coli* topoisomerase I in vitro (155). In effect, the DNA strand exchange would create a bubble behind the fork, and the bubble (including the stranded fork itself) would move forward (in the direction of normal fork movement) until all newly synthesized DNA had been displaced. At that point, the newly synthesized DNA would become the short branch of a “chicken foot” intermediate, while the separated template strands in the bubble would reanneal. The end result would be the same as the regression process already described, but the pathway bringing it about would be quite distinct. Given the topological barriers involved, the motor activity of RecA postulated above could, in principle, be needed to force this reaction in the needed direction.

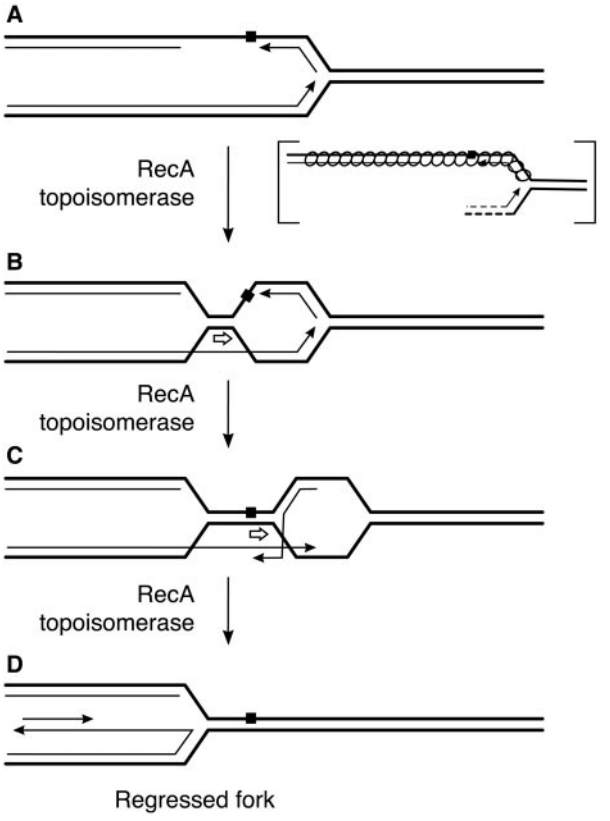

Figure 11 A proposal for RecA protein–mediated fork regression at a lagging strand gap. RecA protein binds in the gap and initiates DNA pairing. Since no suitable DNA ends are present, a net strand exchange requires the coordinated activity of a DNA topoisomerase. The polarity of the lagging strand gap dictates that RecA filament assembly, and the subsequent DNA strand exchange, progress toward the branch point of the fork. The movement of the resulting exchange "bubble" would gradually strip the newly synthesized DNA strands from the template. Bubble movement would entail the indirect helicase activity of RecA (Figure 10). When separation is complete, the newly synthesized strands would become the short branch within a new "chicken foot," and the bubble (now consisting of separated but complementary single strands) would collapse to re-form template duplex DNA.

Topology-Driven Progression of a Regressed Replication Fork

If regression of a fork with a leading strand gap is carried out by RecA protein, there is a potentially interesting side effect. The DNA within the RecA filament is extensively underwound. As the chicken foot structure is formed and proceeds to the

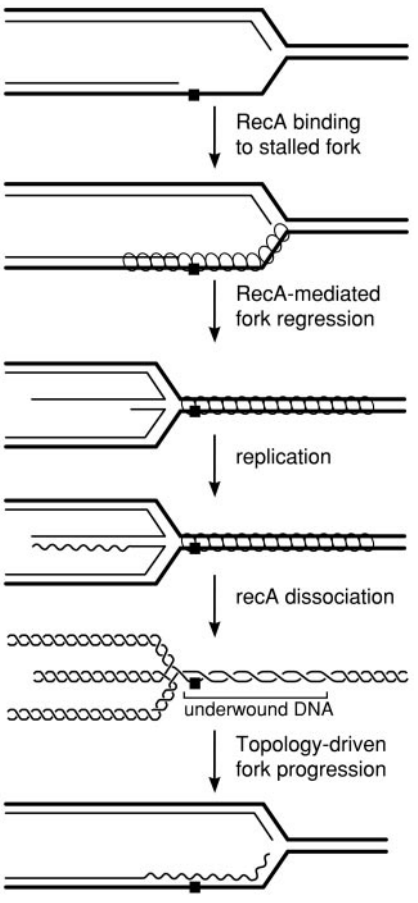

Figure 12 The potential for topology-driven fork progression. If the RecA protein promotes fork regression from a leading or lagging strand gap, the end product would be a chicken foot structure in which the re-paired template strands would feature bound RecA protein (as shown here for regression of a leading strand gap). When this RecA dissociates, it would leave underwound DNA ahead of the new branch point. This underwinding could provide the torsional stress needed to propel fork progression so as to restore the fork structure (ideally after a polymerase has filled in the gap on the chicken foot, or short branch).

end of the RecA filament, and the short branch is filled in by a polymerase, one branch of the new four-way junction (the one with the reunited template strands) remains underwound within the RecA filament. If this DNA is released by RecA dissociation, the underwound DNA would remain constrained by the four-way junction and its contiguous linkage with the larger chromosome. The only nonenzymatic path to eliminating the trapped underwinding would be rotation of the end of the short arm. If a topoisomerase were again available to eliminate tangles resulting from rotation of the two fork branches about each other, relaxation of the DNA in front of the fork would in principle drive a fork progression that would reverse the regression process and restore the fork structure (Figure 12). In effect, if the positive supercoiling generated by replication ahead of the fork can promote spontaneous fork regression (15), then negative supercoiling ahead of the fork generated by RecA protein binding may promote spontaneous fork progression. The extent of the fork progression would depend on the length of the RecA filament

promoting the regression process. This idea has not been tested thoroughly, but RecA-mediated underwinding can be trapped in DNA via RecA-promoted fork regression of model DNAs and the formation of chicken foot structures in vitro (ME Robu, RB Inman, & MM Cox, unpublished results).

RecA Filament Blockage of Replication Initiation

The 3′ end invasion reaction promoted by RecA protein introduces a 3′ end that can be used as a replication primer. However, the 3′ end is buried within a RecA filament. To make the end accessible to DNA polymerases, the RecA filament must dissociate. This circumstance has been demonstrated in vitro (K Marians, personal communication). It illustrates the importance of obtaining a better understanding of RecA filament assembly and disassembly, and the enzymes that modulate these processes. It also helps illustrate how little is known about the replication/recombination interface in recombinational DNA repair, where recombination enzymes must hand off to replication enzymes and vice versa.

ACKNOWLEDGMENT

Work in our own laboratory is supported by grants GM32335 and GM52725 from the National Institutes of Health (to MMC).

The *Annual Review of Biochemistry* is online at http://biochem.annualreviews.org

LITERATURE CITED

1. Cox MM. 1997. *Proc. Natl. Acad. Sci. USA* 94:11764–66
2. Cox MM. 1998. *Genes Cells* 3:65–78
3. Cox MM. 1999. *Prog. Nucleic Acids Res. Mol. Biol.* 63:310–66
4. Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ, Marians KJ. 2000. *Nature* 404:37–41
5. Kuzminov A. 1999. *Microbiol. Mol. Biol. Rev.* 63:751–813
6. Kuzminov A. 2001. *Proc. Natl. Acad. Sci. USA* 98:8461–68
7. Michel B, Flores MJ, Viguera E, Grompone G, Seigneur M, Bidnenko V. 2001. *Proc. Natl. Acad. Sci. USA* 98:8181–88
8. Kowalczykowski SC. 2000. *Trends Biochem. Sci.* 25:156–65
9. Cox MM. 2001. *Proc. Natl. Acad. Sci. USA* 98:8173–80
10. Cox MM. 2001. *Annu. Rev. Genet.* 35:53–82
11. Haber JE. 1999. *Trends Biochem. Sci.* 24:271–75
12. Paques F, Haber JE. 1999. *Microbiol. Mol. Biol. Rev.* 63:349–404
13. McGlynn P, Lloyd RG. 2000. *Cell* 101: 35–45
14. Higgins NP, Kato K, Strauss B. 1976. *J. Mol. Biol.* 101:417–25
15. Postow L, Ullsperger C, Keller RW, Bustamante C, Vologodskii AV, Cozzarelli NR. 2001. *J. Biol. Chem.* 267:2790–96
16. Michel B. 2000. *Trends Biochem. Sci.* 25:173–78
17. Sandler SJ, Marians KJ. 2000. *J. Bacteriol.* 182:9–13
18. Kodadek T, Wong ML, Alberts BM. 1988. *J. Biol. Chem.* 263:9427–36

19. Beernik HTH, Morrical SW. 1999. *Trends Biochem. Sci.* 24:385–89
20. Seitz EM, Brockman JP, Sandler SJ, Clark AJ, Kowalczykowski SC. 1998. *Genes Dev.* 12:1248–53
21. Gupta RC, Golub E, Bi B, Radding CM. 2001. *Proc. Natl. Acad. Sci. USA* 98:8433–39
22. Ogawa T, Yu X, Shinohara A, Egelman EH. 1993. *Science* 259:1896–99
23. Sung P. 1994. *Science* 265:1241–43
24. Passy SI, Yu X, Li ZF, Radding CM, Masson JY, et al. 1999. *Proc. Natl. Acad. Sci. USA* 96:10684–88
25. Griffith J, Formosa T. 1985. *J. Biol. Chem.* 260:4484–91
26. Yu X, Jacobs SA, West SC, Ogawa T, Egelman EH. 2001. *Proc. Natl. Acad. Sci. USA* 98:8419–24
27. Roca AI, Cox MM. 1997. *Prog. Nucleic Acids Res. Mol. Biol.* 56:129–223
28. Brendel V, Brocchieri L, Sandler SJ, Clark AJ, Karlin S. 1997. *J. Mol. Evol.* 44:528–41
29. Sung P, Stratton SA. 1996. *J. Biol. Chem.* 271:27983–86
30. Menetski JP, Bear DG, Kowalczykowski SC. 1990. *Proc. Natl. Acad. Sci. USA* 87:21–25
31. Kim JI, Cox MM, Inman RB. 1992. *J. Biol. Chem.* 267:16444–49
32. Kim JI, Cox MM, Inman RB. 1992. *J. Biol. Chem.* 267:16438–43
33. Rehrauer WM, Kowalczykowski SC. 1993. *J. Biol. Chem.* 268:1292–97
34. Shan Q, Cox MM, Inman RB. 1996. *J. Biol. Chem.* 271:5712–24
35. Kowalczykowski SC, Krupp RA. 1995. *Proc. Natl. Acad. Sci. USA* 92:3478–82
36. Walker GC, Smith BT, Sutton MD. 2000. In *Bacterial Stress Responses*, ed. G Storz, R HenggeAronis, pp. 131–44. Washington, DC: Am. Soc. Microbiol.
37. Pham P, Bertram JG, O'Donnell M, Woodgate R, Goodman MF. 2001. *Nature* 409:366–70
38. Goodman MF. 2002. *Annu. Rev. Biochem.* 71:17–50
39. Kowalczykowski SC, Clow J, Krupp RA. 1987. *Proc. Natl. Acad. Sci. USA* 84:3127–31
40. Pugh BF, Cox MM. 1988. *J. Mol. Biol.* 203:479–93
41. Pugh BF, Cox MM. 1987. *J. Biol. Chem.* 262:1326–36
42. Kowalczykowski SC, Eggleston AK. 1994. *Annu. Rev. Biochem.* 63:991–1043
43. Register JC III, Griffith J. 1985. *J. Biol. Chem.* 260:12308–12
44. Shan Q, Bork JM, Webb BL, Inman RB, Cox MM. 1997. *J. Mol. Biol.* 265: 519–40
45. Amaratunga M, Benight AS. 1988. *Biochem. Biophys. Res. Commun.* 157: 127–33
46. Cazenave C, Chabbert M, Toulme JJ, Helene C. 1984. *Biochim. Biophys. Acta* 781:7–13
47. McEntee K, Weinstock GM, Lehman IR. 1981. *J. Biol. Chem.* 256:8835–44
48. Tracy RB, Chedin F, Kowalczykowski SC. 1997. *Cell* 90:205–6
49. Tracy RB, Kowalczykowski SC. 1996. *Genes Dev.* 10:1890–903
50. Bar-Ziv R, Libchaber A. 2001. *Proc. Natl. Acad. Sci. USA* 98:9068–73
51. Di Capua E, Engel A, Stasiak A, Koller T. 1982. *J. Mol. Biol.* 157:87–103
52. Egelman EH, Stasiak A. 1986. *J. Mol. Biol.* 191:677–97
53. Nordén B, Elvingson C, Kubista M, Sjoberg B, Ryberg H, et al. 1992. *J. Mol. Biol.* 226:1175–91
54. Yu X, Egelman EH. 1992. *J. Mol. Biol.* 227:334–46
55. Stasiak A, Di Capua E. 1982. *Nature* 299:185–86
56. Pugh BF, Schutte BC, Cox MM. 1989. *J. Mol. Biol.* 205:487–92
57. Stasiak A, Egelman EH. 1988. In *Genetic Recombination*, ed. R Kucherlapati, GR Smith, pp. 265–307. Washington, DC: Am. Society Microbiol.
58. Heuser J, Griffith J. 1989. *J. Mol. Biol.* 210:473–84

59. Bork JM, Cox MM, Inman RB. 2001. *J. Biol. Chem.* 276:45740–43
60. Lindsley JE, Cox MM. 1990. *J. Biol. Chem.* 265:9043–54
61. Arenson TA, Tsodikov OV, Cox MM. 1999. *J. Mol. Biol.* 288:391–401
62. Brenner SL, Mitchell RS, Morrical SW, Neuendorf SK, Schutte BC, Cox MM. 1987. *J. Biol. Chem.* 262:4011–16
63. Shan Q, Cox MM. 1996. *J. Mol. Biol.* 257:756–74
64. Jain SK, Cox MM, Inman RB. 1994. *J. Biol. Chem.* 269:20653–61
65. Yancey-Wrona JE, Camerini-Otero RD. 1995. *Curr. Biol.* 5:1149–58
66. Bazemore LR, Takahashi M, Radding CM. 1997. *J. Biol. Chem.* 272:14672–82
67. Bazemore LR, Folta-Stogniew E, Takahashi M, Radding CM. 1997. *Proc. Natl. Acad. Sci. USA* 94:11863–68
68. Gupta RC, Folta-Stogniew E, O'Malley S, Takahashi M, Radding CM. 1999. *Mol. Cell* 4:705–14
69. Howard-Flanders P, West SC, Stasiak A. 1984. *Nature* 309:215–19
70. Zaitsev EN, Kowalczykowski SC. 2000. *Genes Dev.* 14:740–49
71. Zhurkin VB, Raghunathan G, Ulyanov NB, Camerini-Otero RD, Jernigan RL. 1994. *J. Mol. Biol.* 239:181–200
72. Kim MG, Zhurkin VB, Jernigan RL, Camerini-Otero RD. 1995. *J. Mol. Biol.* 247:874–89
73. Jain SK, Inman RB, Cox MM. 1992. *J. Biol. Chem.* 267:4215–22
74. Rice KP, Chaput JC, Cox MM, Switzer CY. 2000. *Biochemistry* 39:10177–88
75. Rice KP, Eggler AL, Sung P, Cox MM. 2001. *J. Biol. Chem.* 276:38570–81
76. Kumar KA, Muniyappa K. 1992. *J. Biol. Chem.* 267:24824–32
77. Baliga R, Singleton JW, Dervan PB. 1995. *Proc. Natl. Acad. Sci. USA* 92:10393–97
78. Podyminogin MA, Meyer RB, Gamper HB. 1995. *Biochemistry* 34:13098–108
79. Podyminogin MA, Meyer RB, Gamper HB. 1996. *Biochemistry* 35:7267–74
80. Zhou X, Adzuma K. 1997. *Biochemistry* 36:4650–61
81. Frank-Kamenetskii MD, Mirkin SM. 1995. *Annu. Rev. Biochem.* 64:65–95
82. Malkov VA, Panyutin IG, Neumann RD, Zhurkin VB, Camerini-Otero RD. 2000. *J. Mol. Biol.* 299:629–40
83. Adzuma K. 1992. *Genes Dev.* 6:1679–94
84. Rould E, Muniyappa K, Radding CM. 1992. *J. Mol. Biol.* 226:127–39
85. Mazin AV, Kowalczykowski SC. 1999. *Genes Dev.* 13:2005–16
85a. Pugh BF, Cox MM. 1988. *J. Biol. Chem.* 163:76–83
86. Karlin S, Brocchieri L. 1996. *J. Bacteriol.* 178:1881–94
87. Story RM, Weber IT, Steitz TA. 1992. *Nature* 355:318–25
88. Story RM, Steitz TA. 1992. *Nature* 355: 374–76
89. Datta S, Prabu MM, Vaze MB, Ganesh N, Chandra NR, et al. 2000. *Nucleic Acids Res.* 28:4964–73
90. Bird LE, Subramanya HS, Wigley DB. 1998. *Curr. Opin. Struct. Biol.* 8:14–18
91. Story RM, Bishop DK, Kleckner N, Steitz TA. 1993. *Science* 259:1892–96
92. Thompson TB, Thomas MG, Escalante-Semerena JC, Rayment IR. 1998. *Biochemistry* 37:686–95
93. Nguyen TT, Muench KA, Bryant FR. 1993. *J. Biol. Chem.* 268:3107–13
94. Dutreix M, Moreau PL, Bailone A, Galibert F, Battista JR, et al. 1989. *J. Bacteriol.* 171:2415–23
95. Skiba MC, Knight KL. 1994. *J. Biol. Chem.* 269:3823–28
96. Logan KM, Skiba MC, Eldin S, Knight KL. 1997. *J. Mol. Biol.* 266:306–16
97. Skiba MC, Logan KM, Knight KL. 1999. *Biochemistry* 38:11933–41
98. Walker JE, Saraste M, Runswick MJ, Gay NJ. 1982. *EMBO J.* 1:945–51
99. Campbell MJ, Davis RW. 1999. *J. Mol. Biol.* 286:437–45
100. Stole E, Bryant FR. 1996. *J. Biol. Chem.* 271:18326–28

101. Egelman EH, Yu X. 1989. *Science* 245: 404–7
102. Nastri HG, Knight KL. 1994. *J. Biol. Chem.* 269:26311–22
103. Hortnagel K, Voloshin ON, Kinal HH, Ma N, Schaffer-Judge C, Camerini-Otero RD. 1999. *J. Mol. Biol*. 286:1097–106
104. Malkov VA, Camerini-Otero RD. 1995. *J. Biol. Chem.* 270:30230–33
105. Wang Y, Adzuma K. 1996. *Biochemistry* 35:3563–71
106. Morimatsu K, Horii T. 1995. *Eur. J. Biochem.* 228:772–78
107. Rehrauer WM, Kowalczykowski SC. 1996. *J. Biol. Chem.* 271:11996–2002
108. Williams KR, Spicer EK, LoPresti MB, Guggenheimer RA, Chase JA. 1983. *J. Biol. Chem.* 258:3346–55
109. Lonberg N, Kowalczykowski S, Paul L, von Hippel P. 1981. *J. Mol. Biol.* 145: 123–38
110. Tateishi S, Horii T, Ogawa T, Ogawa H. 1992. *J. Mol. Biol.* 223:115–29
111. Benedict RC, Kowalczykowski SC. 1988. *J. Biol. Chem.* 263:15513–20
112. Deleted in proof
113. Morimatsu K, Horii T. 1995. *Adv. Biophys.* 31:23–48
114. Kurumizaka H, Aihara H, Ikawa S, Kashima T, Bazemore LR, et al. 1996. *J. Biol. Chem.* 271:33515–24
115. Aihara H, Ito Y, Kurumizaka H, Terada T, Yokoyama S, Shibata T. 1997. *J. Mol. Biol.* 274:213–21
116. Shan Q, Cox MM. 1998. *Mol. Cell* 1:309–17
117. Sigurdsson S, Trujillo K, Song BW, Stratton S, Sung P. 2001. *J. Biol. Chem.* 276:8798–806
118. Namsaraev EA, Berg P. 2000. *J. Biol. Chem.* 275:3970–76
119. Sung P, Robberson DL. 1995. *Cell* 82:453–61
120. Namsaraev EA, Berg P. 1998. *Proc. Natl. Acad. Sci. USA* 95:10477–81
121. Cox MM, Lehman IR. 1987. *Annu. Rev. Biochem.* 56:229–62
122. Cox MM. 1990. *Binding Two DNA Molecules at Once: The RecA Protein*, pp. 171–96. Boca Raton, FL: CRC Press
123. Cox MM. 1994. *Trends Biochem. Sci.* 19:217–22
124. MacFarland KJ, Shan Q, Inman RB, Cox MM. 1997. *J. Biol. Chem.* 272: 17675–85
125. Müller B, Koller T, Stasiak A. 1990. *J. Mol. Biol.* 212:97–112
126. Takahashi M, Kubista M, Nordén B. 1991. *Biochimie* 73:219–26
127. Takahashi M, Nordén B. 1994. *Adv. Biophys.* 30:1–35
128. Kubista M, Simonson T, Sjöback R, Widlund H, Johansson A. 1996. In *Biological Structure and Function: Proceedings of the Ninth Conversation, The State University of New York*, ed. RH Sarma, MH Sarma, pp. 49–59. New York: Adenine
129. Kurumizaka H, Rao BJ, Ogawa T, Radding CM, Shibata T. 1994. *Nucleic Acids Res.* 22:3387–91
130. Kurumizaka H, Shibata T. 1996. *J. Biol. Chem.* 119:216–23
131. Pugh BF, Cox MM. 1987. *J. Biol. Chem.* 262:1337–43
132. Ullsperger CJ, Cox MM. 1995. *Biochemistry* 34:10859–66
133. Konforti BB, Davis RW. 1992. *J. Mol. Biol.* 227:38–53
134. Morel P, Stasiak A, Ehrlich SD, Cassuto E. 1994. *J. Biol. Chem.* 269:19830–35
135. Shan Q, Cox MM. 1997. *J. Biol. Chem.* 272:11063–73
136. Bianco PR, Tracy RB, Kowalczykowski SC. 1998. *Front. Biosci.* 3:560–603. *(on-line journal)* http://www.bioscience.org/1998/v3/d/bianco/list.htm
137. Horii T, Ozawa N, Ogawa T, Ogawa H. 1992. *J. Mol. Biol.* 223:105–14
138. Bedale WA, Cox M. 1996. *J. Biol. Chem.* 271:5725–32
139. Nayak S, Bryant FR. 1999. *J. Biol. Chem.* 274:25979–82
140. Jwang B, Radding CM. 1992. *Proc. Natl. Acad. Sci. USA* 89:7596–600

141. Anderson DG, Kowalczykowski SC. 1997. *Cell* 90:77–86
142. Churchill JJ, Anderson DG, Kowalczykowski SC. 1999. *Genes Dev.* 13:901–11
143. Konforti BB, Davis RW. 1987. *Proc. Natl. Acad. Sci. USA* 84:690–94
144. Konforti BB, Davis RW. 1990. *J. Biol. Chem.* 265:6916–20
145. Dutreix M, Rao BJ, Radding CM. 1991. *J. Mol. Biol.* 219:645–54
146. McIlwraith MJ, West SC. 2001. *J. Mol. Biol.* 305:23–31
147. Bork JM, Cox MM, Inman RB. 2001. *EMBO J.* 20:7313–22
148. Razavy H, Szigety SK, Rosenberg SM. 1996. *Genetics* 142:333–39
149. Rosenberg SM, Hastings PJ. 1991. *Biochimie* 73:385–97
150. Taylor AF, Smith GR. 1985. *J. Mol. Biol.* 185:431–43
151. Cordeiro-Stone M, Makhov AM, Zaritskaya LS, Griffith JD. 1999. *J. Mol. Biol.* 289:1207–18
152. McGlynn P, Lloyd RG. 2001. *Proc. Natl. Acad. Sci. USA* 98:8227–34
153. Robu ME, Inman RB, Cox MM. 2001. *Proc. Natl. Acad. Sci. USA* 98: 8211–18
154. Seigneur M, Ehrlich SD, Michel B. 2000. *Mol. Microbiol.* 38:565–74
155. Cunningham RP, Wu AM, Shibata T, Das Gupta C, Radding CM. 1981. *Cell* 24:213–23