Regulation of Bacterial RecA Protein Function

ABSTRACT The RecA protein is a recombinase functioning in recombinational DNA repair in bacteria. RecA is regulated at many levels. The expression of the recA gene is regulated within the SOS response. The activity of the RecA protein itself is autoregulated by its own C-terminus. RecA is also regulated by the action of other proteins. To date, these include the RecF, RecO, RecR, DinI, RecX, RdgC, PsiB, and UvrD proteins. The SSB protein also indirectly affects RecA function by competing for ssDNA binding sites. The RecO and RecR, and possibly the RecF proteins, all facilitate RecA loading onto SSB-coated ssDNA. The RecX protein blocks RecA filament extension, and may have other effects on RecA activity. The DinI protein stabilizes RecA filaments. The RdgC protein binds to dsDNA and blocks RecA access to dsDNA. The PsiB protein, encoded by F plasmids, is uncharacterized, but may inhibit RecA in some manner. The UvrD helicase removes RecA filaments from RecA. All of these proteins function in a network that determines where and how RecA functions. Additional regulatory proteins may remain to be discovered. The elaborate regulatory pattern is likely to be reprised for RecA homologues in archaeans and eukaryotes.

KEYWORDS RecA, recombination, repair, DNA, replication fork, RecF, RecO, RecR, DinI, RecX

INTRODUCTION

Recombination plays a critical role in DNA repair and genome maintenance. However, it is equally critical to regulate where, when, and how recombination takes place. When unregulated, recombination can lead to genome instability and carcinogenesis. Meiotic recombination in eukaryotes is under tight regulation, ensuring the proper spacing and complete chromosomal coverage of the recombination events that are needed for proper chromosome segregation (Cohen et al., 2001; de Massy, 2003; Hillers et al., 2003; Thompson et al., 1999). Defects in many recombination functions result not only in DNA repair defects, but also in more general genomic instability. Much of this is associated with stalled replication forks (Bjergbaek et al., 2002; Chakraverty et al., 1999; Kolodner et al., 2002; Myung et al., 2001; Myung et al., 2002; Osborn et al., 2002; Thompson et al., 2002; van Gent et al., 2001; Venkitaraman, 2001). Regulation determines which pathway is used to correct a double strand break in DNA in eukaryotes (Haber, 2000). Humans with mutations conferring a hyperrec phenotype have an increased risk
for cancer (Bishop et al., 2003). Mitotic recombination is regulated, at least in part, by the mismatch repair system in eukaryotes (Chen et al., 1998; Datta et al., 1996). In many cases, the operative regulatory mechanisms are not well defined.

The pattern is not limited to eukaryotes. Mutations in some bacterial recombination proteins reduce homologous recombination and associated repair processes, but can produce large increases in illegitimate recombination (Hanada et al., 2000; Hanada et al., 1997; Lovett et al., 2002; Lovett et al., 1995). Much of this genomic instability is again associated with stalled replication forks (Bierne et al., 1994; Cox et al., 2000; Hanada et al., 1997; Hyrien, 2000; Kowalczykowski, 2000; Kuzminov, 1999). The potential deleterious consequences of unregulated recombination are as obvious in bacteria as they are in eukaryotes.

The recombinases of the RecA family represent an obvious target for regulation. The highly conserved bacterial RecA protein is found in virtually all bacteria (Brendel et al., 1997; Roca et al., 1990; Roca et al., 1997). The only bacteria found to date that lack a gene encoding RecA protein are certain endosymbionts that have undergone dramatic genome size reductions (Moran et al., 2000; Tamas et al., 2002). True structural and functional homologues of RecA have been found in bacteriophage (Jiang et al., 1993), in archaea (Sandler et al., 1996; Seitz et al., 1998; Seitz et al., 2001), and in eukaryotic cells (Baumann et al., 1998; Bishop, 1994; Gupta et al., 1997; New et al., 1998; Ogawa et al., 1993; Shinohara et al., 1992; Sung, 1994). These proteins typically promote the central steps of recombination and recombinational DNA repair in the pathways that utilize them. Recombination in general is a complex process now known to involve dozens of individual proteins. This review focuses on those proteins that directly affect or modulate RecA protein function in some manner.

THE RecA PROTEIN

The biochemistry of RecA protein has recently been reviewed in detail (Cox, 2006; Cox, 2007). A short introduction is provided here so that readers can approach the information about regulation without the need to read other articles. The RecA protein of *Escherichia coli* (EcRecA) is the prototype of this class of protein and it has three very different roles in the cell. First, it promotes the central steps of recombination, aligning and pairing two DNA molecules, and then promoting a strand switch followed by branch migration (Cox, 1999; Cox, 2002; Cox, 2003; Lusetti et al., 2002). In *vitro*, the RecA protein catalyzes a DNA strand exchange reaction (Figure 1) that serves as the major experimental model for its recombination activities. Second, RecA itself has a regulatory function. RecA has an activity generally referred to as its coprotease function, facilitating the autocatalytic cleavage of the LexA repressor (Little, 1991) as required for induction of the SOS response (Walker et al., 2000). The RecA coprotease activity also facilitates the autocatalytic cleavage of the UmuD protein to the activated UmuD, a component of DNA polymerase V (Pham et al., 2002; Tang et al., 1999). The affected proteins (LexA, UmuD, and a few others) cleave themselves only when they are bound to an activated RecA filament. Third, RecA functions to directly facilitate replicative bypass of DNA lesions by DNA polymerase V during SOS, utilizing a mechanism that is gradually yielding to investigation (Dutreix et al., 1989; Echols et al., 1991; Pham et al., 2002; Schlacher et al., 2006; Schlacher et al., 2005; Schlacher et al., 2006; Tang et al., 1999).

EcRecA protein is a 352-residue polypeptide chain (Mr 37,842). The structure of RecA protein was elucidated by Story and Steitz in 1992 (Story et al., 1992; Story et al., 1992), and multiple structures of RecA and its homologues have appeared since (Conway et al., 2004; Datta et al., 2003; Datta et al., 2003; Datta et al., 2000; Krishna et al., 2006; Qian et al., 2005; Rajan et al., 2004; Wu et al., 2004; Wu et al., 2005; Xing et al., 2004;
FIGURE 2 RecA protein structure. The figure is available in color in the online publication. (A) A filament is shown, based on the structure of Story and Steitz (Story et al., 1992; Story et al., 1992), with 24 subunits (four turns of the filament). (B) One monomeric subunit of the larger filament is colored differently in panel a, and enlarged here in a ribbon diagram. The core domain is in blue, with bound ADP shown as a space-filling model. The N- and C-terminal domains are shown in green and purple, respectively. (C) The sequence of the C-terminal 25 amino acid residues (disordered in all of the published crystal structures) is shown. Arrows indicate the locations of C-terminal truncations used in a recent characterization of C-terminus function (Eggler et al., 2003; Lusetti et al., 2003; Lusetti et al., 2003).

Xing et al., 2004). Unfortunately, none of these structures includes bound DNA, and much remains to be elucidated about the DNA binding sites on this protein. The monomer structure (Figure 2) has a large core domain, and two smaller domains at the N- and C-termini. The core domain is the prototype of a motif called the RecA fold, found to be common to a range of other proteins including motor proteins such as the F1 ATPase, multiple helicases, DNA transport proteins, and certain other transporters (Abrahams et al., 1994; Amano et al., 1994; Blanchet et al., 1997; Bird et al., 1998; Egelman, 2000; Gomis-Ruth et al., 2001; Yu et al., 1997). In the active RecA filament, ATP is bound at the subunit-subunit interface (Cox et al., 2006; VanLoock et al., 2003). A depiction of the RecA filament, based on the Story and Steitz structure, is shown in Figure 2. There are six RecA monomers per helical turn of the filament, corresponding to the 18 bp per turn extended...
conformation of the DNA bound within the filament. The helical groove providing access to the filament interior can be seen between the monomers colored gray.

In principle, any activity of EcRecA may be regulated, but most work has focused on the recombinational function of RecA. The DNA strand exchange reaction has three distinct phases (Cox, 1999; Cox, 2002; Cox, 2003; Lusetti et al., 2002). RecA protein first binds to the single-stranded or gapped DNA substrate, producing a right-handed helical filament containing one RecA monomer for every 3 nucleotides or base pairs of DNA. Next, a homologous duplex DNA is aligned with the single strand already within the filament and a nascent hybrid DNA product is formed. Extension of this nascent hybrid DNA occurs in the final reaction phase, in which the capacity of the filament to promote a strand switch is augmented by ATP hydrolysis.

ATP hydrolysis (dATP readily substitutes) renders the DNA strand exchange reaction unidirectional (Jain et al., 1994; Shan et al., 1996), allowing it to bypass substantial structural barriers in the DNA substrates (Kim et al., 1992; Rosselli et al., 1991; Shan et al., 1996), and permits strand exchange reactions involving 4 DNA strands (Kim et al., 1992; Shan et al., 1996). Models for how RecA protein couples ATP hydrolysis to these activities have been described (Cox, 2003; Cox, 2007). RecA protein will promote replication fork regression, a reaction that may be relevant to RecA function at an arrested replication fork. This reaction requires ATP hydrolysis.

Much of the regulation characterized to date is directed at the assembly and disassembly of RecA filaments. RecA filaments assemble on DNA in discrete nucleation and extension phases, with the former generally being rate-limiting (Figure 3). Nucleation occurs more rapidly on single-stranded DNA (ssDNA) than on double-stranded DNA (dsDNA). Nucleation on ssDNA is greatly slowed if SSB is bound to the ssDNA. Once nucleation is achieved (often with the aid of regulatory proteins), filament growth occurs primarily on the 3′-proximal end (5′ to 3′ direction) (Register et al., 1985; Shan et al., 1997). RecA binding here is a simple reversible association of ATP-liganded RecA monomers. The reported rates of 5′-to-3′ filament growth range from 120 to 1200 subunits min\(^{-1}\) (Galletto et al., 2006; Joo et al., 2006; Shivashankar et al., 1999; van der Heijden et al., 2005). Dissociation occurs primarily on the 5′-proximal end. The release of RecA monomers at this end is not part of a simple reversible binding process, as dissociation is coupled to ATP hydrolysis—(a fact that is not entirely accounted for in recent attempts to measure dissociation constants and/or filament growth at the disassembly end in the presence of ATP using single-molecule approaches (Galletto et al., 2006; Joo et al., 2006). Dissociation from ssDNA occurs at a rate of \(\sim 70\) monomers per minute (Arenson et al., 1999), rising to 120 monomers per minute when the filament is bound to dsDNA (Cox et al., 2005; Shivashankar et al., 1999). The monomer-monomer interface should be identical on both filament ends. When ATP analogues that cannot be hydrolysed are substituted for ATP, the filaments should be able to grow in either direction, as has been observed (Galletto et al., 2006).

The DNA inside a RecA filament is in an unusual conformation (Cox, 2006; Nishinaka et al., 1997). It is extended by 50% to 60%, and underwound to \(\sim 18\) base pairs per turn. The resulting strain might facilitate both the exchange of DNA strands and the homology search that precedes it (Cox, 2006; Klapstein et al., 2004). Each RecA monomer in the filament covers three nucleotides or DNA base pairs.

**FIGURE 3** RecA protein filament formation on DNA. The figure is available in color in the online publication. RecA protein filaments are generally formed in two distinct phases, nucleation and extension. Nucleation occurs most readily on single strand DNA. Extension proceeds largely in the 5′ to 3′ direction, and can continue readily into any adjacent duplex DNA. Direct binding to duplex DNA is observed if the slow nucleation step is overcome, and certain RecA mutant proteins are proficient at the rapid nucleation of filament formation on duplex DNA.
Once formed, filaments that hydrolyse ATP are dynamic, continuing to grow and shorten at the appropriate ends. When bound to a DNA circle, they might contain breaks where dissociation at a disassembly end is rapidly compensated for by association to the trailing assembly end. Again, the filaments on ssDNA and dsDNA differ, with the ATP hydrolysis turnover rate per monomer being ~30 ATP molecules per minute on the former and 20 per minute on the latter.

The properties of RecA filaments formed on different DNA substrates are diagnostic of different filament states (Cox, 2006; Haruta et al., 2003). As RecA binds to ssDNA in the presence of ATP, it undergoes a transition between states, designated O and A, respectively. The A state on ssDNA is converted to a state, designated P, on dsDNA. The capacity of RecA protein to pair DNAs and promote DNA-strand exchange improves in the progression O to A to P, and many filament properties change.

**REGULATION OF RecA FUNCTION**

The activity of RecA protein, and presumably all related recombinases, is regulated on at least three levels. First, recA gene expression is controlled within the SOS regulon. This aspect of RecA function will not be considered here. Second, RecA protein is subject to autoregulation. Its activities are suppressed, to degrees that vary with conditions, by the C-terminus (and perhaps other parts) of the protein. Third, the activity of RecA protein is modulated by a growing array of other proteins.

**AUTOREGULATION OF RecA PROTEIN FUNCTION**

In most of the published RecA crystal structures, the C-terminal 25 amino acids (Figure 2) are disordered. Recent work has shown that this part of the RecA polypeptide, and particularly the C-terminal 17 amino acid residues, acts as a kind of autoregulatory flap (Eggler et al., 2003; Lusetti et al., 2003a, 2003b). The first studies of the RecA C-terminus indicated that DNA binding (particularly dsDNA binding) and some DNA pairing functions were enhanced when all or large segments of the C-terminal 25 amino acid segment were removed (Benedict et al., 1988; Tateishi et al., 1992). An extensive follow-up demonstrated that removal of 17 amino acid residues enhances a wide range of RecA activities. The wild type RecA protein requires the addition of 6 to 8 mM Mg ion for optimal DNA strand exchange function. The deletion mutant no longer requires free Mg$^{2+}$ ion for optimal strand exchange activity (Lusetti et al., 2003a). The same truncated RecA protein binds more quickly to duplex DNA (Lusetti et al., 2003b), rapidly displaces SSB on single-stranded DNA (Eggler et al., 2003), and promotes the cleavage of LexA protein more rapidly when bound to duplex DNA (Lusetti S. L., unpublished data). This C-terminal peptide appears to affect virtually every RecA function, with almost all activities being more robust when it is absent. As such, it is a logical interaction point for other proteins that modulate RecA function.

**REGULATION OF RecA PROTEIN BY OTHER PROTEINS**

**Overview**

Classically, the regulation of RecA has focused on loading of the protein onto single-stranded DNA. There are two major pathways for RecA loading, and these in turn define the two major recombination pathways of *E. coli*, laid out over three decades ago by Clark and coworkers (Clark et al., 1994; Horii et al., 1973; Kowalczykowski et al., 1994; Kuzminov, 1999).

During conjugation, the major pathway for recombination utilizes the RecBCD helicase/nuclease, and is called the RecBCD pathway. This pathway is generally required to process DNA ends at the sites of double strand breaks. As detailed later, RecBCD processes a DNA end to create a long single-strand extension on the 3'-ending strand, upon which RecA protein can be loaded. The RecA then aligns this single-stranded DNA with homologous sequences in a duplex, and promotes a DNA strand invasion as a first step in the repair of a double strand break (Figure 4).

**FIGURE 4** The DNA strand invasion reaction. The single strand extension in the top DNA has a 3' terminus. The single-stranded region serves as a loading point for RecA protein filament formation. The RecA then pairs this single strand region with a complementary strand in another duplex, leading to the strand invasion reaction shown.
The RecF, RecO, and RecR proteins (often abbreviated RecFOR) have been highlighted as prototype mediator proteins (Beernik et al., 1999; Sancar et al., 1993), necessary to load RecA protein onto SSB-coated DNA (Sandler, 2001; Umezlu et al., 1993; Umezlu et al., 1994). The RecFOR recombination pathway appears to complement the RecBCD pathway. The recF, recO, and recR genes form an epistatic group tailored more for the repair of DNA gaps than double strand breaks (Asai et al., 1994; Courcelle et al., 1997; Horii et al., 1973; Kolodner et al., 1985; Wang et al., 1984). The RecFOR pathway plays a secondary role during conjugational recombination. When the RecBCD pathway is removed by mutation of one of the genes encoding the RecBCD helicase/nuclease, the recombinational defect is suppressed by mutations in sbcA, sbcB, or sbcCD (Bidnenko et al., 1999). During conjugation, the RecFOR pathway supplants the RecBCD pathway only in a recBCD− sbc(A, B, or CD) genetic background (Clark et al., 1994; Horii et al., 1973; Kowalczykowski et al., 1994; Kuzminov, 1999). However, the major function of recombination in bacteria is the repair of stalled or collapsed replication forks (Cox et al., 2000; Kowalczykowski, 2000; Kuzminov, 1999). In this context, the RecFOR and RecBCD pathways appear to be equally important (Cox, 1999; Galitski et al., 1997; Steiner et al., 1998).

The RecBCD and RecFOR functions seem sufficient to target RecA filaments to the locations requiring them, yet recent work has shown that RecA regulation is much more complex. The RecF, RecO, and RecR proteins modulate both the assembly and disassembly of RecA filaments (Shan et al., 1997), and RecF protein in particular may have multiple functions (Rangarajan et al., 2002; Sandler, 1996). Several additional proteins play important roles. The DniI protein stabilizes RecA filaments (Lusetti et al., 2004). The RecX protein blocks RecA filament extension, leading to net disassembly in certain contexts (Drees et al., 2004). Certain helicases, such as UvrD, act to disrupt RecA filaments (Lovett et al., 1995; Mendonca et al., 1993; Morel et al., 1993; Petranovic et al., 2001). The PsiB (Bagdasarian et al., 1992; Bailone et al., 1988) and RdgC (Moore et al., 2003; Ryder et al., 1996) proteins have been implicated by genetic studies in the suppression of RecA activity by mechanisms that are still largely undefined. The study of proteins that regulate RecA function is still in its early stages. Recent studies have highlighted the growing complexity of the network of proteins that regulate RecA protein.

Surprisingly, the RecBCD and RecFOR functions that load RecA protein are not as ubiquitous in the eubacteria as is RecA protein itself (Rocha et al., 2005; Sandler, 2001). The RecFOR proteins are more widespread than the RecBCD enzymes (Rocha et al., 2005). In a survey of the genomes of 117 bacterial species, the recB and recC genes are missing in 77 of them (Rocha et al., 2005). An orphan variant of the recD gene is present in 30 of these species (Rocha et al., 2005). In some species, the homologous addAB system replaces recBC, but both RecBC and AddAB are absent in at least 33 species, including Mycoplasma, Helicobacter, Campylobacter, Desulfovibrio, Ralstonia, Bdelovibrio, Nitrosonomas, Bacteroides, Wolinella, Corynebacterium, and Streptomyces species (Rocha et al., 2005). In contrast, RecFOR is entirely lacking in only nine species. However, these three proteins are not always found together. RecR protein is the most widespread, missing in all but the nine that have no RecFOR components (Rocha et al., 2005). RecF is missing 29 times, and RecO 17 times. Many of the Mycoplasma have RecO and RecR, but no RecF or RecBCD/AddAB. The c-proteobacteria group tends to have RecR protein alone. The Buchnera endosymbionts (which lack RecA protein) and a few Mycoplasma species appear to lack all of the known systems for loading RecA protein onto DNA (Rocha et al., 2005).

SSB

The single-strand DNA binding protein (SSB) of Escherichia coli was first reported in 1972 (Sigal et al., 1972). It was the second example of a protein in this class, with the gene 32 protein of bacteriophage T4 being the first (Alberts et al., 1970). Like RecA recombinases, SSB-class proteins are ubiquitous and essential to DNA metabolism in all organisms. The eukaryotic equivalent is the heterotrimeric replication protein A (RPA) protein. In E. coli, the ssb gene is essential for cell viability (Curth et al., 1996; Lohman et al., 1994; Meyer et al., 1990). Each of these proteins features a DNA binding motif called a oligonucleotide/oligosaccharide binding fold (OB fold). The E. coli SSB polypeptide (M, 18,843) includes one OB fold in a large amino-terminal domain. EcsSB functions as a homotetramer (Molineux et al., 1975; Weiner et al., 1975). Most bacterial SSB proteins are similar. Exceptions are the SSB proteins of Deinococcus radiodurans and Thermus sp., with polypeptides that encompass two OB folds and function as homodimers (Dabrowski et al., 2002; Eggington et al., 2004).
SSB plays a complicated role in RecA reactions. RecA filament nucleation is inhibited, and under some conditions blocked entirely, if SSB is allowed to coat the DNA prior to RecA addition (Bork et al., 2001; Kowalczykowski et al., 1987; Lavery et al., 1990; Shan et al., 1997; Umezu et al., 1994). This inhibition of binding nucleation is overcome in the bacterial cell by the mediator proteins, RecO and RecR (Bork et al., 2001; Hobbs et al., 2007; Morimatsu et al., 2003; Shan et al., 1997; Umezu et al., 1994). However, WT E. coli RecA protein does not bind well to secondary structure in ssDNA, and addition of SSB after RecA protein disrupts the secondary structure and allows RecA to form a contiguous filament on the DNA (Kowalczykowski et al., 1987).

SSB is sometimes viewed as a DNA-binding polypeptide that is otherwise inert, but the reality is more complex. SSB interacts directly and specifically with a wide range of proteins involved in DNA metabolism, including PriA (Cadman et al., 2004), uracil-DNA glycosylase (Acharya et al., 2002; Handa, 2001 #4619), exonuclease I (Genschel et al., 2000), the χ subunit of DNA polymerase III (Gulbis et al., 2004; Witte et al., 2003), RecO protein (Hobbs et al., 2007; Kantake et al., 2002; Umezu et al., 1994), TopB (DNA topoisomerase III) (Butland et al., 2005), the nucleases RecJ and SbcB (Butland et al., 2005), the RecG helicase (Butland et al., 2005), and the RecQ helicase (Butland et al., 2005). SSB has its own conserved acidic C-terminus (Figure 5), and most of the protein-protein interactions occur at the C-terminus (Acharya et al., 2002; Cadman et al., 2004; Curth et al., 1996; Genschel et al., 2000; Gulbis et al., 2004; Handa et al., 2001; Kinebuchi et al., 1997; Ma et al., 2004; Richard et al., 2004; Savvides et al., 2004; Witte et al., 2003; Yuzhakov et al., 1999). Instead of an inert DNA coating that must be displaced for DNA metabolism to proceed, SSB may instead be viewed as an organizational scaffold where DNA metabolic complexes are assembled. Making the issue of SSB function more intriguing, the protein exhibits multiple single-stranded DNA binding modes that depend on salt concentration (Bujalowski et al., 1987; Bujalowski et al., 1988; Lohman et al., 1994). Potential effects of these different SSB binding modes on reactions in which SSB is a participant has not been properly assessed in most cases.

In addition to its effects on RecA filament formation, SSB plays a role in DNA strand exchange, binding to the displaced single-strand product and facilitating its release from the filament (Lavery et al., 1992).

**The RecBCD Helicase/Nuclease**

The *E. coli* recB and recC genes were identified in early screens for mutations that produced defects in standard recombination assays (Clark et al., 1965; Emmerson, 1968). The recD gene was discovered only two decades later (Amundsen et al., 1986), after a third integral subunit was recognized as a required component. The RecB (M, 133,973 (Finch et al., 1986)), RecC (M, 128,860 (Finch et al., 1986)), and RecD (M, 66,973 (Finch et al., 1986)) proteins function as a heterotrimeric complex.

RecBCD is part of a broader family of eubacterial enzymes called either RecBCD or AddAB. The RecBCD helicase/nuclease both unwinds and degrades duplex DNA from one end. The recent elucidation of the structure of the RecBCD enzyme bound to DNA has capped several decades of careful investigation and rendered the pathway in three dimensions (Singleton et al., 2004). The RecB subunit possesses a 3’ to 5’ helicase activity, and the RecD subunit has a 5’ to 3’ helicase activity (Dillingham et al., 2003; Taylor et al., 2003; Yu et al., 1998). The nuclease function, sequestered in a separate domain of RecB, generally degrades the 3’-ending strand more efficiently at first. When RecBCD encoun-

---

**FIGURE 5** The SSB C-terminus. The single strand DNA binding protein of *E. coli* has an acidic C-terminus, similar to other bacterial SSB proteins. The C-terminus is a site for interaction with many other proteins involved in DNA metabolism. Deletion of the eight C-terminal amino acid residues eliminates many of these interactions.
ters a copy of an 8 base pair sequence called Chi (5'-GCTGGTGG-3'), a site in the RecC subunit binds to chi on the 3'-ending strand and prevents its further degradation. The 5'-ending strand is then degraded, leading to the creation of a long 3' single strand extension at the end. Finally, RecBCD actively loads RecA protein onto the single-strand segments it creates (Anderson et al., 1997; Arnold et al., 2000; Churchill et al., 1999), utilizing an interaction site in the RecB subunit (Churchill et al., 2000). RecBCD is not a dedicated RecA regulatory function. However, its dynamic interaction with RecA makes it an important factor in RecA biochemistry in vivo, and it provides one of the key pathways that direct assembly of RecA protein filaments. Given the tendency of RecA filaments to disassemble in the 5' to 3' direction, it will be interesting to learn if RecBCD not only triggers the loading of RecA onto the 3' single strand extensions that it creates, but also stabilizes the 5'-proximal ends of the assembled filaments. Alternatively, other proteins such as DinI protein (described below), might serve as the stabilizing factor.

The RecFOR Proteins

Mediator proteins are nearly as ubiquitous as recombinases, ensuring the targeted assembly of recombinase filaments. In E. coli, the RecF, RecO, and RecR proteins function in this capacity, and perhaps have other functions as well. Given the non-ubiquitous distribution of the recFOR genes in eubacteria (Rocha et al., 2005), it is not clear that this class of mediator is present in all bacterial species. However, RecFOR provides a good prototype for investigation of bacterial mediator function.

The RecF Protein

The E. coli recF gene was discovered (Horii et al., 1973) as a UV sensitive, recombination-deficient mutant in a recBC sbcBC background. Although it has never been implicated in replication, the gene is contained in an operon that also includes the dnaA, dnaN, and gyrB genes. The sequenced recF gene encodes a 357-amino acid polypeptide (Mr 40,519 (Blanar et al., 1984)). The RecF protein has been purified and characterized in vitro (Griffin et al., 1990; Madiraju et al., 1992; Madiraju et al., 1991; Morimatsu et al., 2003; Umezu et al., 1993; Umezu et al., 1994; Webb et al., 1995; Webb et al., 1997; Webb et al., 1999). It binds to ssDNA with an apparent stoichiometry of 1 RecF monomer per 15 nucleotides (Madiraju et al., 1991). In the presence of ATP, the RecF protein also binds to dsDNA (Madiraju et al., 1992; Webb et al., 1995). The protein contains a consensus nucleotide-binding fold (Walker A box). The protein binds ATP, and has a weak DNA-dependent ATPase activity (kcat about 1.0 min\(^{-1}\)) (Webb et al., 1995; Webb et al., 1999). ATP hydrolysis leads to RecF dissociation from DNA (Webb et al., 1999).

The structure of the RecF protein from Deinococcus radiodurans has recently been elucidated by Koroleva and coworkers (2007). The RecF protein exhibits extensive structural similarity with the head domain of the eukaryotic Rad50 protein (Figure 6), but lacks the long coiled-coil domain of Rad50 (Koroleva et al., 2007). RecF belongs to the ATP-binding cassette (ABC) ATPase family of proteins, and it has the Walker A, Walker B, and signature motifs characteristic of that family. ATP binding triggers RecF dimerization (Koroleva et al., 2007). Mutations in the Walker A (K36R) (Webb et al., 1999), Walker B (D300N) (Koroleva et al., 2007), and the signature (S268R) (Koroleva et al., 2007) motifs all disrupt ATP hydrolytic activity.

RecF protein has been proposed to direct RecA loading to the boundaries of single strand gaps in duplex DNA (Hegde et al., 1996; Morimatsu et al., 2003). However, a report that RecF binds specifically to the duplex-ssDNA junction at DNA gaps (Hegde et al., 1996) has not been confirmed (Webb et al., 1999).

The RecO Protein

The E. coli recO gene was identified by Kolodner and colleagues (Kolodner et al., 1985). It is situated in an operon with the rnc gene, which encodes ribonuclease III, and the era gene, which encodes a GTP-binding protein with sequence similarities to the yeast RAS proteins (Ahnn et al., 1986). The sequenced recO gene encodes a protein with 242 amino acids (Mr 27,260) and includes a Walker A box (Morrison et al., 1989; Takiff et al., 1989). The purified protein binds to both ssDNA and dsDNA and behaves as a monomer in solution (Luisi-DeLuca, 1995; Luisi-DeLuca et al., 1994; Umezu et al., 1993; Umezu et al., 1994). Binding or hydrolysis of ATP has not been reported. The protein forms a functional complex with the RecR protein, as described below.

RecO protein promotes an ATP-independent renaturation of complementary DNA strands (Luisi-DeLuca et al., 1994). This reaction is enhanced when RecO forms a complex with the SSB protein (Kantake et al., 2002).
The renaturation reaction promoted by RecO-SSB is inhibited by the RecR protein, suggesting the presence of a RecR-SSB competition for RecO complex formation (Kantake et al., 2002). The RecO-SSB DNA renaturation activity is reminiscent of the reaction promoted by the eukaryotic Rad52 protein (Mortenson et al., 1996; Reddy et al., 1997; Sugiyama et al., 1998), and could function in some RecA-independent pathways for recombination and recombinational DNA repair.

The structure of the RecO protein from Deinococcus radiodurans has been elucidated (Leiros et al., 2005; Makharashvili et al., 2004). The protein contains three domains that include an OB fold, a helical bundle, and a zinc-finger motif moving from the N- to C-terminus (Figure 7). Sites for interactions with DNA and other proteins have been suggested (Leiros et al., 2005; Makharashvili et al., 2004) and some of these have been tested, but more work is needed to define these complexes.

**RecR Protein**

The recR gene was identified in 1989 by Mahdi and Lloyd (Mahdi et al., 1989; Mahdi et al., 1989). It is co-transcribed with the dnaX gene and shares an operon with a small open reading frame of unknown function called orf-12 or ybaB. The recR gene encodes a 201 amino acid protein (M, 21,965). The sequence includes two putative DNA-binding motifs (helix-turn-helix and zinc finger) (Alonso et al., 1993). The purified protein has been examined in vitro, where it was studied in concert with the RecF and RecO proteins (Shan et al., 1997; Umezu et al., 1993; Umezu et al., 1994; Webb et al., 1995; Webb et al., 1997). There is no indication that the E. coli RecR protein alone binds directly to DNA, although the RecR proteins from D. radiodurans (Lee et al., 2004) and B. subtilis (Alonso et al., 1993) do. The E. coli RecR protein is a dimer in solution (Umezu et al., 1994). The structure of the D. radiodurans RecR protein (44% identity with EcRecR) has recently been determined (Lee et al., 2004). The protein crystallizes as a tetrameric ring with a central hole large enough to accommodate a molecule of dsDNA (Lee et al., 2004) (Figure 8). Koroleva and colleagues (2007) have suggested that the EcRecR protein may require a not-yet-identified clamp-loading function to effect binding to DNA.

**Interaction of RecF, O, and R Proteins**

Several lines of evidence indicate that these three proteins function at the same stage of recombination, and tie them to a role in displacing SSB and modulating RecA filament assembly. The phenotypes of mutations in the three genes are very similar, defining them as an epistatic group (Clark et al., 1994; Smith, 1989). Mutations in all three genes are partially suppressed by recA441 and recA803 mutations (Wang et al., 1993). In vitro, the same RecA441 (previously tif) and RecA803 proteins exhibit an enhanced capacity to displace SSB
and bind ssDNA (Lavery et al., 1988; Madiraju et al., 1992). In addition, a gene in bacteriophage λ called ninB or orf (described further below) has been identified which can replace recF, recO, and recR functions in lambda recombination (Sawitzke et al., 1992; Sawitzke et al., 1994). In vivo, mutant bacteria missing any of the recFOR functions exhibit a delayed activation of the SOS response that might reflect slow formation of the RecA filaments required to facilitate LexA cleavage (Madiraju et al., 1988; Whitby et al., 1995). E. coli strains in which SSB is overexpressed exhibit a recFOR-like phenotype (Moreau, 1988), again suggesting that these proteins function together in overcoming the barrier to RecA filament nucleation represented by SSB.

FIGURE 7  The structure of the RecO protein of *Deinococcus radiodurans* (Leiros et al., 2005; Makharashvili et al., 2004). The figure is available in color in the online publication. The structure is oriented with the amino-terminal OB fold shown at bottom, the helical bundle in the middle, and the C-terminal zinc finger domain at top. A 90° rotation of the structure is shown in panel B.

FIGURE 8  The structure of the RecR protein of *Deinococcus radiodurans* (Lee et al., 2004). The figure is available in color in the online publication. The protein is arranged as a tetrameric ring, with a space in the middle large enough to accommodate a double-stranded DNA molecule. Parts of the A and C, and B and D, subunits overlap extensively.
A more detailed examination of the literature, however, shows that the roles of these proteins are not always confluent. In particular, RecF protein is an apparent outlier. RecF activity can be toxic to the cell at least in some contexts. In a strain lacking the function of PriA protein, the additional loss of RecO is about 10 times more deleterious than the loss of either RecF or RecR (Grompone et al., 2004). The effect of RecO loss is moderated in recOrecR or recOrecF strains, suggesting that RecF and RecR are doing something deleterious to the cell in the absence of RecO. An earlier report identified recF mutants as the most deleterious in the presence of a priA mutation (Sandler et al., 1996). The apparent discrepancy may be explained by some extra steps taken in the former study (Grompone et al., 2004) to avoid the appearance of suppressors in the very sick recOpriA strains. Both studies, however, draw a clear distinction between the effects of recF and recO mutations in the priA background. The RecF protein, but not RecO or RecR, is needed for the activation of mutagenic translesion DNA synthesis (TLS) (Rangarajan et al., 2002). This work suggests that RecF may work with RecOR in some processes and independently in others. As already noted, the recFOR genes do not always occur together in bacterial genomes (Rocha et al., 2005; Sandler, 2001). In Bacillus subtilis (which has all three genes), RecF protein recruitment to repair foci is preceded by the appearance of RecO protein by several minutes (Kidane et al., 2004). Overexpression of RecF protein in E. coli reduces SOS induction, UV resistance, and viability at 42°C (Sandler et al., 1993). The overexpression of the RecOR proteins suppresses many of the effects of either RecF overexpression (Sandler, 1994) or a recF null mutation (Sandler et al., 1994). These varied results suggest that the current pictures of RecFOR and RecF function require expansion.

The functional distinction between RecF and RecO proteins is also quite evident in vitro (Shan et al., 1997; Umezu et al., 1993; Umezu et al., 1994; Webb et al., 1995; Webb et al., 1997). RecR protein forms alternative complexes with RecF and RecO protein (Bork et al., 2001; Morimatsu et al., 2003; Shan et al., 1997; Webb et al., 1995; Webb et al., 1997). The RecR TOPRIM domain is the interaction site for both RecF and RecO (Honda et al., 2006). RecF protein generally interferes with RecOR function (Bork et al., 2001; Hobbs et al., 2007), and strongly competes with RecO for the formation of complexes with RecR (Bork et al., 2001).

The EcRecOR complex stimulates RecA protein binding to ssDNA coated with SSB, in a process that is not further stimulated by RecF protein under most conditions (Bork et al., 2001; Hobbs et al., 2007; Shan et al., 1997; Umezu et al., 1993; Umezu et al., 1994). Early models indicated that RecOR does not displace SSB, but instead binds to it to form a RecO-RecR-SSB complex that facilitates RecA nuclease (Shan et al., 1997; Umezu et al., 1994). RecO and RecR proteins remain associated with the RecA filament after it is formed (Shan et al., 1997; Umezu et al., 1994). In addition to stimulating nuclease of RecA filament formation on SSB-coated ssDNA, the EcRecOR complex prevents a net end-dependent dissociation from linear ssDNA (Shan et al., 1997). It is not yet clear whether the evident stability of RecA filaments formed when RecOR is present reflects an actual stabilization of the 5'-proximal ends of RecA filaments, a rapid re-nucleation of RecA at RecOR complex sites after the previous RecA filament dissociation commences, or both. The EcRecO and EcRecR proteins are not active independently in these processes. Both must be added to see a stimulation of EcRecA filament loading.

The formation of a RecOR-SSB complex might obviate a direct interaction of RecOR with ssDNA. However, RecOR functions surprisingly slowly in the loading of RecA onto SSB-coated ssDNA. When RecOR concentrations are optimized, the lag in RecA loading can be reduced by more than five-fold. However, a lag of about 10 minutes remains (Hobbs et al., 2007). The slow step in the activation of RecOR is not the binding of SSB, but the displacement of SSB to permit the formation of a direct RecOR-ssDNA complex (Hobbs et al., 2007). The only condition that eliminates the 10 minutes lag is to add RecOR to the ssDNA prior to the SSB. RecOR is functional as a RecA nucleation site only when this complex with ssDNA is formed (Hobbs et al., 2007). When SSB is present on the ssDNA, the pathway leading to the formation of a RecOR-ssDNA complex functional in RecA loading involves an interaction with the SSB C-terminus. When a deletion mutant of SSB lacking the eight C-terminal amino acids replaces the wild type SSB, RecOR function is largely eliminated (Hobbs et al., 2007).

Does RecF protein contribute to RecOR function? RecF inhibits the RecOR-mediated loading of RecA onto SSB-coated ssDNA under most conditions (Bork et al., 2001; Hobbs et al., 2007; Shan et al., 1997; Umezu et al., 1993; Umezu et al., 1994). There is no evidence
yet reported for an interaction between the RecO and RecF proteins. With one exception (Morimatsu et al., 2003) discussed below, there are no reports of a stimulatory effect of RecF on any RecA activity. However, the presence of RecO and RecR appears to nullify a strong inhibitory effect of RecF on RecA-mediated reactions in vitro (Umezu et al., 1993). Optimized in vitro reconstitution of several steps of one major pathway for recombination-dependent replication restart requires the presence of the RecOR proteins (Xu et al., 2003). RecF protein reduces the stimulation provided by RecOR in this reaction, probably as a result of its negative effects on RecA loading (Xu et al., 2003).

The RecF and RecR proteins form a complex in an ATP and DNA-dependent fashion (Webb et al., 1995). The RecFR complex binds primarily to dsDNA, and the complex is stable enough to halt RecA filament extension (Webb et al., 1997). The RecR protein stimulates the RecF ATPase, but reduces the rate of RecF transfer from one DNA to another (Webb et al., 1995; Webb et al., 1999).

It is always possible that the failure to detect the formation or activity of a RecFOR complex could reflect a failure to find the right reaction conditions. One recent study has provided evidence that RecF, O, and R can act together to facilitate RecA protein filament formation on SSB-coated DNA gaps (Morimatsu et al., 2003). In this study, an enhancement of RecOR-mediated loading of RecA onto SSB-coated ssDNA by RecF protein was observed when DNA substrates contained single strand gaps bordered by short duplex regions (Morimatsu et al., 2003). The positive effects of RecF protein depended on the inclusion of a large excess of SSB protein relative to available SSB binding sites (Morimatsu et al., 2003). The reason for the excess SSB requirement is not clear. The extra SSB may be titrating some factor that affects the formation of the proposed RecFOR complexes. No physical evidence for a RecFOR complex has appeared.

**The DinI and RecX Proteins**

These two proteins are related in the sense that they have opposing activities, each antagonizing the function of the other.

**The RecX Protein**

The RecX protein of *Escherichia coli* (Mr 19,293) is encoded by the *E. coli* version of a widespread bacterial gene often found just downstream or even overlapping the recA gene (De Mot et al., 1994; Papavinasasundaram et al., 1997; Sano, 1993; Vierling et al., 2000; Yang et al., 2001). In a few cases, the gene is found in another region of the chromosome (Stohl et al., 2001). In *E. coli*, the recX gene is just downstream of the recA gene, and expressed from the recA promoter via a 5% to 10% transcriptional read-through of a hairpin sequence separating the two genes (Pages et al., 2003). The first indications that recX was involved in RecA regulation appeared in the early 1990s (De Mot et al., 1994; Sano, 1993). In some bacterial species, RecX protein is necessary to overcome deleterious effects of overexpression of RecA protein, implying that RecX is a negative modulator of RecA expression or function (Papavinasasundaram et al., 1998; Sano, 1993; Sukchawalit et al., 2001; Vierling et al., 2000). Deletion of the gene in *E. coli* produces no clear phenotype (Pages et al., 2003), although overexpression of the recX gene can reduce the induction of the SOS response (Stohl et al., 2003). When purified, both the *Mycobacterium* RecX (Venkatesh et al., 2002) and the *E. coli* RecX protein (Stohl et al., 2003) inhibit the ATPase and strand exchange activities of RecA protein in vitro. The RecX protein binds deep within the major helical groove of an AMPNP-stabilized RecA filament (VanLoock et al., 2003).

The RecX protein acts as a RecA inhibitor by blocking the extension of RecA filaments during assembly, almost certainly by capping the filament (Drees et al., 2004a). When RecA filaments have been formed on circular ssDNAs, there is generally no net dissociation and ATP hydrolysis proceeds at a constant steady state. There are generally breaks in the filaments where dissociation at a disassembly end can occur, but the resulting ends are quickly filled in by growth of the trailing filament assembly end (Figure 3). When RecX protein is added at relatively low concentrations (about one RecX per 20–100 bound RecA monomers), a net disassembly of the RecA filaments occurs that takes 10–15 min to complete. Detailed characterization of this phenomenon led to the filament capping model (Drees et al., 2004a) (Figure 9). The moderating effect of RecA C-terminal deletions provides evidence that the RecA C-terminus plays a significant role in the RecX-RecA interaction (Drees et al., 2004b).

**The DinI Protein**

The *E. coli* DinI protein is a small (81 amino acids; Mr 8,818) polypeptide that was identified a decade ago in a search for new LexA-regulated genes (Lewis...
The DinI and RecX Proteins Each Antagonize the Function of the Other

DinI protein stabilizes the RecA filament, and RecX destabilizes it (Lusetti et al., 2004). With sufficient DinI protein present, a challenge with low concentrations of RecX has no effect on RecA filaments (Lusetti et al., 2004). If DinI protein is added after the RecX protein challenge, the filaments recover. High concentrations of RecX (nearly stoichiometric with RecA protein) do displace the DinI protein and destabilize the RecA filament. In addition to its capping function, the RecX
FIGURE 10  DinI regulation of SOS. The figure is available in color in the online publication. (A) The DinI protein binds all along the length of the RecA filament, stabilizing it and preventing dissociation. The function of the RecA filament is subtly affected by DinI (Lusseti et al., 2004). DNA strand exchange with at least some DNA substrates proceeds normally. The autocatalytic cleavage of LexA protein is reduced somewhat, but the autocalytic cleavage of the UmuD protein is essentially blocked (Y asuda et al., 1998; Y asuda et al., 2001). (B) The DinI protein is induced early in the SOS response (the kinetics shown here are adapted from (V oloshin et al., 2001)). This induction pattern can explain the observed temporal regulation of the SOS response. At early stages in SOS, DinI blocks cleavage of UmuD protein, and thus blocks the activation of the mutagenic DNA polymerase V. The non-mutagenic processes of DNA repair thus predominate early in the SOS response. As DinI levels recede, cleavage and activation of UmuD protein can occur, leading to the formation of active DNA polymerase V. This initiates the mutagenic phase of the SOS response.

protein appears to compete with DinI for binding sites in the RecA filament groove. This idea meshes well with the binding of RecX within the RecA filament groove as seen in the EM studies of Egelman and colleagues (VanLoock et al., 2003). The replacement of RecX with DinI is a relatively slow process, occurring over 10 minutes or more (Lusetti et al., 2004). In effect, RecA-RecX and RecA-DinI define two RecA filament forms that are functionally distinct.

The RecF Protein Antagonizes RecX Function

The interaction of RecX with other proteins in the RecA regulatory web is not limited to its antagonistic relationship with DinI. RecX protein inhibits the loading of RecA protein to SSB-bound ssDNA under most conditions, even when the RecOR proteins are included in the reaction. However, when RecF is present, the inhibitory effects of RecX protein disappear. RecF blocks RecX function, likely via a direct interaction between the two proteins (Lusetti et al., 2006). To completely neutralize the effects of RecX, RecF is generally required at concentrations equivalent or somewhat greater than that of RecX protein. This provides an alternative mechanism for RecF to affect RecA filament formation in a positive way. By interfering with the function of a protein that blocks RecA filament extension, RecF can facilitate the extension phase of RecA filament formation (Lusetti et al., 2006), complementing its postulated role in facilitating the nucleation phase.

The PsiB and RdgC Proteins

These proteins are implicated in the modulation of RecA filament formation and/or function, but have either not been purified or characterized minimally in vitro. Both have the potential for new insights into modes of RecA regulation.

The PsiB Protein

The PsiB protein is encoded by a number of conjugative plasmids (Delver et al., 1997; Sarno et al., 2002). PsiB proteins are generally small, ~140 amino acid residues giving a molecular mass of 15–16 kDa. PsiB inhibits the induction of SOS when conjugation is taking place, presumably by interacting with the RecA protein or blocking RecA filament formation (Bagdasarian et al., 1986). The psiB gene is transferred early in conjugation and transcribed transiently (Bagdasarian et al., 1986). None of the known PsiB proteins have been studied in vitro.

When bacterial cells are starved, there is a transient induction of the SOS response that leads to genome-wide hypermutation. This is called adaptive mutation, and it requires RecA function both directly and for the SOS induction (McKenzie et al., 2000; McKenzie et al., 2001). In this context, DinI protein has little effect on the SOS response, but PsiB is a potent inhibitor (in cells with an F plasmid) (McKenzie et al., 2000). These results again suggest that PsiB is an inhibitor of RecA function (and reinforce the impression that DinI is not an inhibitor. Also of interest, the IncN plasmid pKM101 (broad host range) encodes both a PsiB protein and a RecX protein (Delver et al., 1997).
The RdgC Protein

RdgC is a fairly abundant DNA binding protein that appears to affect the function of RecA and RecFOR. The rdgC gene is located near sbeC and sbeD on the E. coli chromosome, encoding a protein of Mr 33,862 (Ryder et al., 1996). A deletion of the rdgC gene by itself has little obvious effect. However, the same deletion in a recBCsbeBC background is viable only if the RecA and RecF proteins are functional (hence, recombination-dependent growth or rdg) (Ryder et al., 1996). The RdgC protein is also important in a priA background. The poor viability of priA mutants is suppressed by certain mutations in dnaC such as dnaC212. These allow the DnaC protein to circumvent PriA in the pathway by which the DnaB helicase is loaded onto a repaired fork structure (Sandler et al., 2000; Sandler et al., 1999). Introducing an rdgC deletion into a priAdnaC212 background confers a slow growth phenotype (Moore et al., 2003). Suppressors of several types arise rapidly in these strains. First, the slow growth is suppressed by mutations that eliminate RecF, RecO or RecR function (Moore et al., 2003). This suggests that an inappropriate loading of RecA protein onto ssDNA causes the slow growth. It also suggests that the RdgC protein has a role in preventing the inappropriate loading or function of RecA protein. The slow growth is also suppressed by certain mutations in the ssb gene (Moore et al., 2003). The characterized ssb mutants are R97C and Δ115–144. These ssb alterations could define interaction points for Rec(F)OR. The RdgC protein is present at about 1000 copies per cell (compared to about 800 tetramers of SSB) (Moore et al., 2003), and in vitro the protein binds to both ssDNA and dsDNA (Moore et al., 2003).

In vitro, RdgC protein is a potent inhibitor of RecA protein function. It competes with RecA for DNA binding sites, particularly on dsDNA (Drees et al., 2006). When bound to the duplex substrate of DNA strand exchange, RdgC effectively sequesters the dsDNA and prevents its use by the RecA filaments bound to the homologous duplex (Drees et al., 2006). It is not yet clear how this might affect RecA in the cell, since RdgC does not appear to interact with RecA protein. To affect RecA function in vivo, RdgC would need to be targeted to sites of RecA assembly. The RdgC protein is interesting on its own. It is present in the cell at a level comparable to SSB, and it uniformly coats dsDNA (Drees et al., 2006). It might be considered a double-stranded DNA binding protein (DSB?), a kind of duplex DNA counterpart to SSB. Rdg C crystallizes as a dimeric ring, with head to head and tail to tail organization. The 30Å central hole provides a likely DNA binding site (Briggs et al., 2007).

The UvrD Helicase

In every organism, there appear to be helicases that antagonize recombination functions. This is perhaps best characterized in yeast. The yeast Srs2 helicase negatively modulates recombination (Aboussekhra et al., 1989; Aguiler et al., 1988). Additional work suggests that Srs2 actively removes recombination complexes or structures in vivo (Chanet et al., 1996; Kaytor et al., 1995; Milne et al., 1995; Schild, 1995). The yeast Sgs1 protein is a helicase from the RecQ family (Gangloff et al., 1994). A combination of srs2 and sgs1 null mutants results in a near-lethal slow growth phenotype (Fabre et al., 2002; Klein, 2001; Lee et al., 1999; Mankouri et al., 2002). Mutations in a number of recombination functions, including Rad51, Rad52, Rad55, and Rad57, alleviate the defect in the sgs1srs2 double mutant (Fabre et al., 2002; Klein, 2001). The work indicates that Sgs1 and Srs2 are involved in the removal of toxic recombination intermediates, and can partially substitute for each other. In vitro, the Srs2 helicase will disrupt Rad51 protein filaments and interfere with their DNA strand exchange activity (Krejci et al., 2003; Veauke et al., 2003).

In bacteria, parallels are evident but less developed. The RuvA and B proteins displace RecA filaments in vitro (Adams et al., 1994), but the genetics provides only limited support for this role in vivo. There may be better candidates, and null mutants of the uvrD gene have a phenotype implicating the UvrD protein in recombinational DNA repair (Mendonca et al., 1995). UvrD is induced as part of the SOS response, and has a putative role in dismantling RecA filaments.

The EcUvrD protein (or E. coli DNA helicase II; Mr 81,859) is encoded by a gene located at about 84 minutes on the E. coli chromosome. UvrD is homologous to the somewhat smaller E. coli Rep helicase, and can even form heterodimers with Rep (Wong et al., 1993). Deletion of both the rep and uvrD genes in E. coli is lethal (Washburn et al., 1991). As Rep has been shown to remove proteins from the DNA (Yancey-Wrona et al., 1992), it has been proposed that Rep might serve to remove proteins in the path of the replication fork (DNA synthesis is slowed 50% in cells lacking Rep function (Lane et al., 1975)). UvrD could in principle play a
similar role in recombination and repair, and genetic studies suggest that it does.

UvrD plays a role in many aspects of DNA metabolism, although its detailed molecular function remains somewhat enigmatic. This helicase is involved in both methyl-directed mismatch repair (Lahue et al., 1989; Modrich, 1989) and the DNA excision repair mediated by the UvrABC exculinuclease (Kumura et al., 1985). A role in chromosomal replication is suggested by the constitutive induction of the SOS response that is observed in many uvrD mutant cells (George et al., 1994; Ossanna et al., 1989), presumably because replication forks are stalling in these strains. Cells lacking UvrD function have a defect in recombinational DNA repair (Howard-Flanders et al., 1981; Lloyd, 1983; Mendonca et al., 1993). At the same time, uvrD− strains have a hyperrecombination phenotype, with large increases in illegitimate recombination (Lovett et al., 1995; Washburn et al., 1991). Certain alterations of UvrD can suppress the phenotypes of ruvB and recF (Lovett et al., 1995) mutations, further suggesting a complex involvement in recombination processes. Perhaps most important, UvrD has a demonstrated anti-recombinase function in vitro that may involve the destabilization of recombination intermediates, the complexes that form them, or both (Lovett et al., 1995; Morel et al., 1993; Petranovic et al., 2001). In strains lacking the RecBCD pathway, recQ and uvrD null mutations are synthetically lethal (Mendonca et al., 1995).

The EcUvrD protein is a 3′ to 5′ helicase (Matson, 1986) (as is Rep and the yeast Srs2 helicase) and unwinds duplex DNA best when there is a 3′ single strand extension upon which to bind and initiate. However, UvrD also exhibits significant unwinding activity even when initiating the reaction at a nick (Runyon et al., 1990) or blunt end, and this capability could be important for repair systems designed to address strand breaks. In addition, UvrD will unwind RNA-DNA hybrids in a reaction more robust than the unwinding of DNA (Matson, 1989), perhaps suggesting a role in replication fork repair on the lagging strand. The UvrD helicase functions as a dimer (Ali et al., 1999; Maluf et al., 2003), although the protein binds well to single-stranded DNA as a monomer (Maluf et al., 2003; Mechanic et al., 1999; Velankar et al., 1999).

Finally, and significantly, Petit and collaborators have demonstrated that the UvrD helicase dismantles RecA filaments in vitro (Veauté et al., 2005). Filament disassembly was monitored by electron microscopy in this study (Veauté et al., 2005). In the same series of experiments (Veauté et al., 2005), Rep helicase did not promote RecA filament disassembly, indicating that this may be a specialized role of the UvrD protein.

CONCLUDING REMARKS

The known regulatory protein network affecting the bacterial RecA recombinase is in an expansion phase. The proteins already assigned to the network are incompletely understood, and new proteins are being added. Obvious questions for future investigation are highlighted throughout the text above. Prospects for successful biochemical reconstitution of recombinational DNA repair depend upon a better understanding of these regulatory processes. The mechanistic issues are important in part because a very similar situation exists in eukaryotes. The stakes are substantial, especially in human cells where recombination represents a significant cancer-avoidance system.

There is an intriguing possibility that additional regulatory proteins remain to be discovered in bacteria, even in E. coli. In every genome, the functions of many genes remain unknown. In E. coli, a significant number of uncharacterized open reading frames exist in operons where other genes are dedicated to DNA metabolism. Some recent additions to the list of RecA regulatory proteins (e.g., RecX, DinI) are encoded by genes that generate quite subtle phenotypes when mutated, and were not identified in classic genetic screens for recombination functions. In spite of the limited genetic effects, these proteins have important regulatory functions. Genomics provides some new paths for relevant gene discovery. Interesting new candidates for RecA regulatory functions already exist. For example, mutations in the E. coli mgsA (raA) gene produce a phenotype that partially overlaps with that of recA mutants, and the gene is orthologous to a eukaryotic gene (yeast MGS1) that plays a role in maintaining genome stability (Shibata et al., 2005). The MgsA protein could be a regulator and/or a function that complements RecA at stalled replication forks. The relatively slow action of RecOR and the sometimes confusing results with RecF protein suggest that one or more additions to the RecFOR protein family may remain to be found. Genetic studies in Bacillus subtilis may provide a guide. A mutation designated recL is epistatic to the recFOR genes in Bacillus, producing the RecFLOR complementation group (Carrasco et al., 2001; Fernandez et al., 1999; Petit et al., 2002). The gene
corresponding to recL has not been identified. There are actually five separate mutations in B. subtilis that affect recombination but have not been assigned to genes: recH, recL, recP, recV (formerly recD) and rec3 (J. Alonso, personal communication). Some or all of these might have a regulatory role with respect to RecA. It is not clear even in Bacillus that the search for recombination functions has been saturated. It is also not clear how many of the Bacillus genes will turn out to have E. coli homologues, or how many additional E. coli recombination functions may be present.

ACKNOWLEDGMENTS

Work described from the author’s laboratory was supported by grant GM52725 from the National Institutes of Health of the United States of America. The author thanks Sergey Korolev for sharing his group’s RecF structural results prior to publication and for permitting the use of structural data prior to release for use in the generation of Figure 6. The author also thanks past and present members of the Cox laboratory for numerous discussions and insights that contributed to the work described here.

REFERENCES


Yu, M., Souaya, J., and Julin, D.A. 1998. The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from Escherichia coli. *Proc Natl Acad Sci USA* 95:981.
