RecBCD Helicase/Nuclease

Deana A Arnold, *University of California, Davis, California, USA*

Stephen C Kowalczykowski, *University of California, Davis, California, USA*

The RecBCD enzyme of *Escherichia coli* participates in several aspects of DNA recombination and repair. It is essential to the main pathway of genetic homologous recombination, where it contributes to the exchange of genetic material between homologous DNA molecules (i.e. conjugal recombination), and to the recombinational repair of potentially lethal chromosomal double-stranded breaks.

Introduction

The *Escherichia coli* RecBCD enzyme is a multifunctional protein complex (330 kDa) containing three subunits, the products of the *recB*, *recC*, and *recD* genes. This enzyme displays four distinct activities: nuclease, helicase, ATPase, and site-specific recognition of the DNA regulatory sequence chi (crossover hotspot instigator, χ). Originally identified as exonuclease V, the RecBCD enzyme is responsible for the seemingly disparate functions of DNA degradation and repair of the bacterial chromosome. The former function is achieved by the combined action of its helicase and nuclease activities, whereas a recombinationally activated form accomplishes the latter, after interaction of RecBCD enzyme with χ. The RecBCD enzyme is a principal component of the main pathway for homologous genetic recombination in *E. coli*, referred to as the RecBCD pathway. Structural or functional analogues of the RecBCD enzyme are present in both Gram-negative and Gram-positive bacteria, and interspecies complementation studies suggest that the mechanism of homologous recombination via the RecBCD pathway is conserved, at least among bacteria.

Genetics/Effects of Null Mutations

Although the cellular level of RecBCD enzyme is low (~ 10 copies per cell), the absence of this enzyme has a profound effect on DNA metabolism in *E. coli*. Mutations that inactivate RecB or RecC proteins lower viability approximately 4-fold, reduce conjugal recombination 100- to 1000-fold, and sensitize the bacterium to DNA-damaging agents such as ultraviolet (UV) and gamma irradiation. These cells also lack detectable levels of ATP-dependent nuclease and so are susceptible to infection by bacteriophages. These findings emphasize the many roles of RecBCD enzyme in *E. coli*, and underscore the importance of the RecB and RecC subunits to the most basic functions of the holoenzyme.

In striking contrast, however, strains containing null mutations in the *recD* gene display wild-type cell viability, resistance to UV and gamma irradiation, and are recombination proficient. Like the *recB⁻* and *recC⁻* strains, these cells do not display nuclease activity, making them sensitive to infection by bacteriophages. The biochemical basis for this phenotype will be discussed later.

Helicase Activity

Under optimal conditions, the RecBCD enzyme is a highly processive helicase, unwinding an average of 30 kilobases per binding event at a rate of 1000–1500 base pairs per second (**Figure 1**) (Roman *et al.*, 1992). A blunt double-stranded DNA (dsDNA) end is the preferred substrate for initiation of unwinding, but single-stranded DNA (ssDNA) tails of less than 25 nucleotides can also serve as initiation sites. RecBCD enzyme displays a high affinity for dsDNA ends, about 50-fold higher than that for ssDNA ends, arguing that the relevant substrate *in vivo* is a blunt or nearly blunt dsDNA end. At the dsDNA end, the RecB subunit associates with the strand terminating with a 3′ hydroxyl, and the RecC and RecD subunits associate with the strand terminating with a 5′ phosphate, with translocation occurring in the 3′ to 5′ direction relative to the 3′-hydroxyl end. Thus, the requirement for a blunt end may be due to the need for all three subunits to interact simultaneously with the appropriate DNA termini.

The amino acid sequence of the RecB subunit displays homology to other known helicases, such as Rep and UvrD proteins. The purified RecB subunit itself possesses both DNA-dependent ATPase and weak 3′ to 5′ helicase activities. The RecB and RecC subunits can be reconstituted to produce the RecBC enzyme, which is an active helicase with a lower affinity for dsDNA ends than the RecBCD holoenzyme.

During unwinding, either ssDNA loop-tails, or twin-loops of ssDNA are formed, which extend from the RecBCD enzyme complex (Taylor and Smith, 1980). At physiological temperature (37°C), the loops grow at a rate of about 100 nucleotides per second. The presence of the
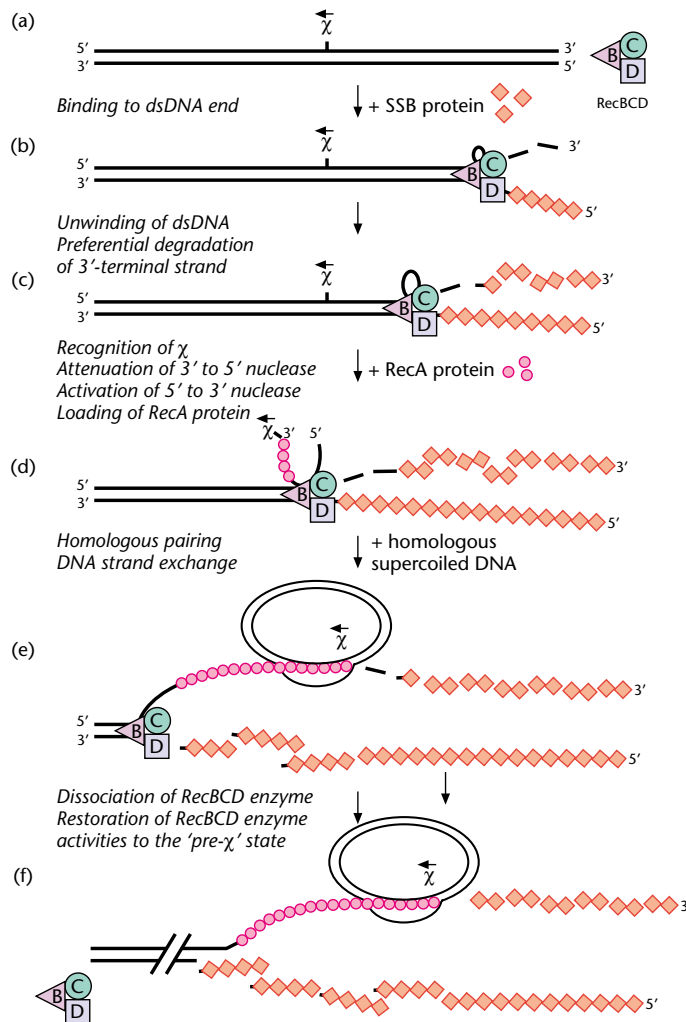
Figure 1 Initiation of homologous recombination by the coordinated activities of RecBCD enzyme and RecA protein. An *in vitro* model depicting recombination between a linear χ-containing double-stranded DNA (dsDNA) and a supercoiled plasmid is shown. (a–c) First, RecBCD enzyme enters at a dsDNA end and unwinds the duplex while preferentially degrading the strand corresponding to the 3′ terminus at the point of entry; SSB protein binds the single-stranded DNA (ssDNA) produced. (d,e) Upon recognition of χ, the 3′ to 5′ nuclease activity is attenuated and a weaker 5′ to 3′ nuclease activity is activated on the opposite strand. Following the interaction with χ, RecBCD enzyme facilitates the loading of RecA protein (to the exclusion of SSB protein) onto the ssDNA produced by continued translocation and unwinding of the DNA molecule. (e,f) This RecA protein–ssDNA filament then invades a homologous duplex DNA molecule, producing a recombination intermediate known as a joint molecule.

loop-tail or twin-loop structures implies that there are at least two translocating domains in the holoenzyme: a more rapidly moving domain(s) containing the helicase activity and a slower domain(s) that translocates along the ssDNA produced. Thus, as the dsDNA substrate is processed, ssDNA loop(s) form between these domains. *E. coli* single-stranded DNA-binding protein (SSB) favours production of the loop-tail structure with the loop being created from the DNA strand terminating with a 3′ hydroxyl at the point of entry of the RecBCD enzyme, presumably by binding to the complementary strand and disrupting its contact with the slower translocating domain(s) (Anderson and Kowalczykowski, 1998).

The DNA helicase activity requires the nucleotide cofactor adenosine triphosphate (ATP) and magnesium ion, with 1.7–3 molecules of ATP being hydrolysed per base pair unwound. Helicase activity is inhibited by the ssDNA produced during processing of a dsDNA molecule. Binding of SSB protein to the ssDNA products alleviates this inhibition (**Figure 1**) (Anderson and Kowalczykowski, 1998).

Nuclease Activity

In addition to being a helicase, the RecBCD enzyme is also a potent nuclease, functioning to protect the cell from invasion by infecting viral DNA. The RecBCD enzyme degrades both dsDNA and ssDNA, but the activity toward ssDNA is much lower than that for dsDNA.

The dsDNA nuclease activity is coincident with translocation by RecBCD enzyme. Thus, although this degradation of dsDNA is formally defined as an 'exonuclease' activity, the cleavage is actually endonucleolytic and the requirement for a dsDNA end pertains to helicase function. Degradation during unwinding is asymmetric, with the 3′ terminal strand (relative to the dsDNA entry site) being degraded much more vigorously than its complement (**Figure 1a–c**). Hence the enzyme degrades dsDNA primarily in a 3′ to 5′ direction (Dixon and Kowalczykowski, 1993). SSB protein, in addition to binding the potentially inhibitory ssDNA produced by the helicase activity as discussed above, also moderates the 5′ to 3′ nuclease activity, lowering the frequency of cutting by the enzyme (Anderson and Kowalczykowski, 1998).

In vitro, the ratio of magnesium ion concentration to ATP concentration greatly influences the observed dsDNA exonuclease activity. As the ratio of magnesium ion concentration to ATP concentration rises, the frequency of endonucleolytic cleavages increases, resulting in the production of increasingly smaller ssDNA fragments. When this ratio is very low, nuclease activity is barely detectable and the enzyme behaves essentially as a helicase. The seemingly complex effects of magnesium ion concentration are the result of competing requirements for this cofactor. Magnesium is required in the form of a complex with ATP for helicase activity, but in its free form for nuclease activity. Hence, when magnesium ion concentration exceeds the ATP concentration, the resultant excess free magnesium ion concentration increases the dsDNA nuclease activity, but not the helicase activity. The free intracellular magnesium ion concentration is $1–2\,mM$, consistent with the high levels of RecBCD-dependent nuclease activity present *in vivo*.

Bacteriophages that infect *E. coli* have developed strategies to circumvent the destructive nuclease activities of the RecBCD enzyme. For example, the λ phage produces the Gam protein early in its lytic cycle; this protein specifically binds to RecBCD enzyme, preventing degradation of the linear λ phage genome. A different defence is utilized by T4 phage; rather than binding to the enzyme, T4 gene *2* protein caps the ends of the T4 genome, blocking entry by RecBCD enzyme. The N15 phage does not produce an inhibitory protein; instead it forms hairpin loops at the ends of its linear genome, thereby blocking binding of the RecBCD enzyme (Malinin *et al.*, 1992). As discussed below, even *E. coli* itself has developed a means to protect broken chromosomal DNA from degradation

by RecBCD enzyme, a mechanism which also serves to initiate repair of these DNA lesions.

This voracious nuclease activity seems at odds with the fact that RecBCD enzyme plays a principal role in promoting DNA repair and homologous recombination, processes that, by their very nature, require the preservation of DNA. Resolution of this apparent inconsistency is found in the regulatory aspects of the recombination hotspot χ.

Modification of the Nuclease Activities of RecBCD Enzyme by χ

χ is a DNA locus that stimulates the frequency of genetic recombination in its vicinity. This recombination hotspot was originally discovered as a mutation in λ phage that protected the phage genome from degradation by RecBCD enzyme (Lam *et al.*, 1974). As shown in **Figure 2**, the sequence of χ is the octamer 5′-GCTGGTGG-3′, and most single base mutations within the octamer reduce χ activity (Smith *et al.*, 1981). Recombination in the vicinity of a χ site is stimulated by 5- to 10-fold over background levels. Key features need to be emphasized to understand the nature of this recombination hotspot. First, the stimulation is highly polar, with the region of enhanced recombination extending downstream of the 5′ end of the χ sequence. Enhancement of recombination downstream of χ decreases by a factor of two for every 2.2–3.2 kb, returning to background levels 10 kb downstream when no heterologous regions intervene (Myers *et al.*, 1995). Second, all recombination stimulated by this site requires the activity of the RecBCD enzyme, whose nuclease activity would seemingly destroy its own substrate. Numerous genetic and biochemical studies have shown that the increase in recombination is due to a direct
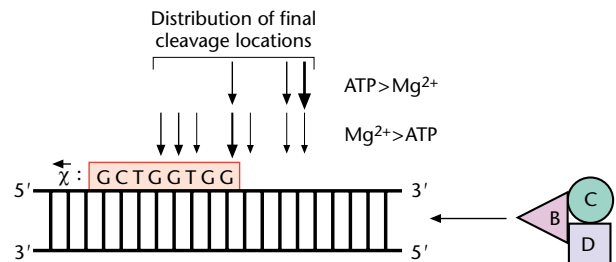

Figure 2 Orientation dependence of χ recognition. For recognition to occur, the RecBCD enzyme must approach the χ site from the 3′ side, as depicted. Recognition results in attenuation of the 3′ to 5′ nuclease activity of the enzyme with the final cleavage event occurring near the 3′ end of χ. Arrows denote the positions at which the final cut may occur, with corresponding thicknesses indicating the relative frequencies at each position. Adapted from Kowalczykowski *et al.* (1994) and Taylor and Smith (1995).

interaction between the χ sequence and the RecBCD enzyme, and that this stimulation only occurs if the enzyme approaches from the 3′ side (**Figure 2**) (Taylor and Smith, 1995). The interaction with χ elicits several changes in enzyme function that are manifest in an overall decrease in nuclease activity, which accounts for the protection of DNA observed *in vivo*. Thus, χ is a regulator of RecBCD enzyme and, hence, of genetic recombination.

The sequence of events preceding and succeeding interaction with χ are illustrated in **Figure 1**. As described above, the RecBCD enzyme translocates through a dsDNA molecule, simultaneously unwinding the dsDNA and degrading primarily the ssDNA strand that is 3′ terminated at the entry site. Upon reaching a χ site in the correct orientation, the enzyme pauses and several enzymatic changes occur. Although the helicase activity of the modified enzyme is relatively unaffected by this encounter, degradation of the strand corresponding to the 3′ end is attenuated at least 500-fold (Dixon and Kowalczykowski, 1993). The last nonspecific cleavage event occurs with a high probability near χ, thereby generating a 3′ hydroxyl end within a few base pairs of the χ site (**Figure 2**). This attenuation of nuclease activity is manifest until the enzyme dissociates from the DNA, explaining the elevated recombination frequency downstream of χ sites. But, in addition to the attenuation of 3′ to 5′ nuclease activity, a weaker 5′ to 3′ nuclease activity is activated on the 5′ terminal strand (**Figure 1e**) (Anderson and Kowalczykowski, 1997a). This switch in the polarity of nuclease activity produces a dsDNA molecule with an ssDNA tail retaining χ at the 3′ terminus (**Figure 1f**). Further, the RecBCD enzyme facilitates the loading of the RecA DNA strand exchange protein onto this 3′ terminal ssDNA tail, which is the preferred substrate for RecA protein-dependent strand invasion of a homologous recipient dsDNA (**Figure 1d**–**e**) (Dixon and Kowalczykowski, 1991; Anderson and Kowalczykowski, 1997b). The cooperative 5′ to 3′ polymerization of RecA protein onto this ssDNA substrate yields an invasive 3′ end that is completely coated with RecA protein (**Figure 1f**). Thus, χ coordinates the activities of RecBCD enzyme and RecA protein such that stimulation of homologous pairing initiates at the 3′ end generated by the overall attenuation of nuclease activity at χ and propagates downstream from χ.

The precise biochemical basis for these changes is not known, but they are probably the result of a modification of the RecD subunit. Both genetic and biochemical evidence demonstrates that the RecBC enzyme, without the RecD subunit, mimics many, but not all, characteristics of the χ-modified RecBCD enzyme. This view explains why the level of conjugal recombination in $recD^-$ strains is similar to that observed in wild-type strains, but is not stimulated by χ. The recombination proficiency of these strains can be largely attributed to the lack of RecBCD nuclease activity. However, unlike recombination in a wild-type strain, it is also dependent on a 5′ to 3′ ssDNA

nuclease, the RecJ protein. Together, the nuclease-deficient RecBC enzyme and the RecJ protein produce a 3′ terminal ssDNA substrate appropriate for RecA protein-promoted strand invasion and, hence, comprise an effective substitute for the χ-modified RecBCD enzyme.

A special class of *recBCD* mutants, the *recC** class, specifically affects the ability of the enzyme to respond to χ. The *recC** strains display moderate to wild-type levels of recombination and resistance to DNA-damaging agents, but they lack stimulation of recombination at χ. Unlike the *recD* mutants, however, these strains possess ATP-dependent nuclease activity. At least one member of this class, the RecBC[1004]D enzyme, recognizes a novel sequence variant of χ, known as χ* (5′-GCTGGTGCTCG-3′) (Handa *et al.*, 1997). These findings suggest that the RecC subunit is responsible for χ recognition.

Role in Conjugal Recombination

During conjugation, a piece of chromosomal DNA is delivered from a donor cell to a recipient cell. After transfer, genetic markers may be exchanged between the transferred DNA and the genome of the recipient cell. Since this exchange occurs by recombination, it is referred to as conjugal recombination. During conjugation, DNA enters the recipient cell as a single-stranded molecule and is largely converted to linear dsDNA by DNA replication. RecBCD enzyme can load at the ends of this dsDNA molecule, both unwinding and degrading the duplex DNA until it encounters a χ site. As described above, this encounter modifies the enzyme from its destructive form to its recombination-promoting form. Approximately 80% of conjugal recombination events are believed to occur at the ends of transferred DNA via the RecBCD pathway, and nearly all of these are χ-mediated, correlating well with biochemical studies.

Role in Recombinational DNA Repair

Alkylating agents, gamma radiation and ultraviolet light cause damage to DNA; whether directly or indirectly, these lesions can result in the production of dsDNA breaks, which are potentially lethal to cells. Repair of this type of DNA damage is dependent on the RecBCD pathway and on the availability of an additional intact homologous chromosome. The processing of such dsDNA breaks by RecBCD enzyme initiates the recombination-dependent repair of these breaks, and is termed recombinational DNA repair. This repair is facilitated by the abundant presence of χ sites, occurring approximately once every 4.6 kb, making it one of the most over-represented octamers in the *E. coli* genome. Interestingly, the orientation of these sequence elements is not random in the

chromosome: 75.5% of χ sites in *E. coli* are oriented toward the origin of replication, *oriC* (**Figure 3a**).

This biased orientation of χ sites accommodates another aspect of cell survival: namely, the repair of detached arms of replication forks. Replication fork collapse is common in *E. coli*, contributing to the lowered viability of *recB* and *recC* null mutants. Collapse of a replication fork can be caused by an ssDNA nick in either template strand, or by proteins bound to the template DNA that block progression of the fork. Such an interruption is potentially lethal, ultimately resulting in a double-strand DNA break and creating an entry site for the RecBCD enzyme. Recombi-

national repair of the break allows restoration and resumption of replication (**Figure 3**). Thus, a major function of the *E. coli* dsDNA break repair pathway is the reassembly of collapsed replication forks, and the preferential orientation of χ sites facilitates this repair through its interaction with the RecBCD enzyme (Kuzminov, 1995). As shown in **Figure 3c**, a double-stranded break resulting from the collapse of a replication fork provides a dsDNA end at which RecBCD enzyme can initiate unwinding and degradation in the direction of the origin of replication, *oriC*. The preferential orientation of χ sites toward *oriC* favours recognition by, and χ-dependent modification of,
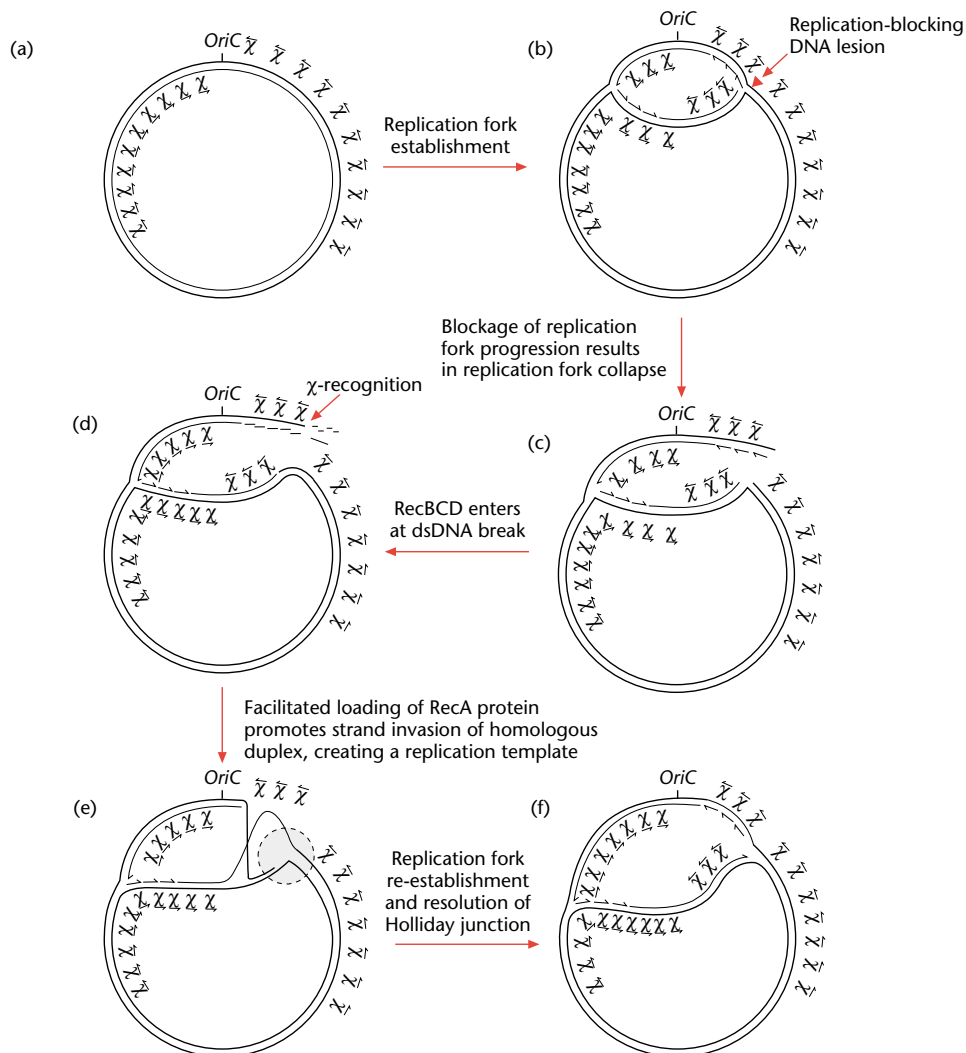

Figure 3 Recombinational repair of a collapsed replication fork. (a) Unreplicated *E. coli* chromosome. (b) Replication is bidirectional with two replication forks initiating at *oriC* and progressing in opposite directions around the chromosome; a lesion in the DNA template blocks progression of one replication fork. (c) The blockage causes detachment of one arm and collapse of the replication fork resulting in a dsDNA break. (d) RecBCD enzyme enters at the dsDNA end, degrading until reaching a χ site; enzymatic modification of RecBCD enzyme occurs and the facilitated loading of RecA protein follows. (e) RecA protein promotes strand invasion of the ssDNA substrate into the homologous duplex, recreating a template for replication as shown in the shaded circle. (f) The replication fork is reassembled and replication of the chromosome resumes. The Holliday junction formed by the strand exchange event is resolved by the specific resolvases of *E. coli*.

RecBCD enzyme (**Figure 3d**), thereby promoting recombination between the broken and intact replication arms and restoring the replication fork (**Figure 3e, f**).

Summary

The multifunctional RecBCD enzyme serves many functions in *E. coli*. Its potent nuclease activities degrade foreign DNA, such as viral genomes, that might otherwise infect and destroy the cell. Subsequent to the occurrence of dsDNA breaks, it would probably do the same to the bacterial genome itself, were it not for the highly over-represented χ site. Acting as a 'self-recognition' tactic, interaction with the χ site promotes a modification of RecBCD activities, effectively changing a destructive nuclease into a genome-protecting recombination enzyme. In addition to formation and preservation of the invasive 3′ terminal ssDNA with χ at its terminus, the χ-activated RecBCD enzyme stimulates loading of the RecA DNA strand-exchange protein onto this strand, prompting its invasion into a homologous dsDNA molecule. This ability of RecBCD enzyme to promote recombination provides a means to repair DNA damage to the *E. coli* chromosome, such as the dsDNA breaks formed by radiation, chemical mutagens or the collapse of replication forks.

References

Anderson DG and Kowalczykowski SC (1997a) The recombination hot spot χ is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme. *Genes and Development* **11**: 571–581.

Anderson DG and Kowalczykowski SC (1997b) The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a χ-regulated manner. *Cell* **90**: 77–86.

Anderson DG and Kowalczykowski SC (1998) SSB protein controls RecBCD enzyme nuclease activity during unwinding: a new role for looped intermediates. *Journal of Molecular Biology* **282**: 275–285.

Dixon DA and Kowalczykowski SC (1991) Homologous pairing in vitro stimulated by the recombination hotspot, Chi. *Cell* **66**: 361–371.

Dixon DA and Kowalczykowski SC (1993) The recombination hotspot χ is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell* **73**: 87–96.

Glickman BW (1979) *rorA* mutation of *Escherichia coli* K-12 affects the RecB subunit of exonuclease V. *Journal of Bacteriology* **137**: 658–660.

Handa N, Ohashi S, Kusano K and Kobayashi I (1997) χ*, a χ-related 11-mer sequence partially active in an *E. coli recC** strain. *Genes to Cells* **2**: 525–536.

Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, and Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiological Reviews* **58**: 401–465.

Kuzminov A (1995) Collapse and repair of replication forks in *Escherichia coli*. *Molecular Microbiology* **16**: 373–384.

Lam S T, Stahl MM, McMilin KD and Stahl FW (1974) Rec-mediated recombinational hot spot activity in bacteriophage lambda. II. A mutation which causes hot spot activity. *Genetics* **77**: 425–433.

Malinin A, Vostrov AA, Rybchin VN and Svarchevski AN (1992) Structure of ends of linear plasmid N15. *Molekuliarnaia Genetika, Mikrobiologia, i Virusologa* **28**: 19–22.

Myers RS, Stahl MM and Stahl FW (1995) χ recombination activity in phage λ decays as a function of genetic distance. *Genetics* **141**: 805–812.

Roman LJ, Eggleston AK and Kowalczykowski SK (1992) Processivity of the DNA helicase activity of *Escherichia coli* RecBCD enzyme. *Journal of Biological Chemistry* **267**: 4207–4214.

Smith GR, Kunes SM, Schultz DW, Taylor A and Triman KL (1981) Structure of Chi hotspots of generalized recombination. *Cell* **24**: 429–436.

Taylor A and Smith GR (1980) Unwinding and rewinding of DNA by the recBC enzyme. *Cell* **22**: 447–457.

Taylor AF and Smith GR (1995) Strand specificity of nicking of DNA at Chi sites by RecBCD enzyme. Modulation by ATP and magnesium levels. *Journal of Biological Chemistry* **270**: 24459–24467.

Further Reading

Eggleston AK and West SC (1997) Recombination initiation: easy as A, B, C, D… chi? *Current Biology* **7**: R745–R749.

Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, and Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiological Reviews*. **58**: 401–465.

Kuzminov A (1996) *Recombinational Repair of DNA Damage*. Austin, TX: RG Landes.

Myers RS and Stahl FW (1994) Chi and the RecBCD enzyme of *Escherichia coli*. *Annual Review of Genetics* **28**: 49–70.

Smith GR (1989) Homologous recombination in *E. coli*: multiple pathways for multiple reasons. *Cell* **58**: 807–809.

Smith GR (1991) Conjugal recombination in *E. coli*: myths and mechanisms. *Cell* **64**: 19–27.

Stahl FW and Stahl MM (1977) Recombination pathway specificity of Chi. *Genetics* **86**: 715–725.