Molecular Microbiology (1991) 5(6), 1295–1299 ADONIS 0950382X91001434

MicroReview

The RecA protein as a recombinational repair system

M. M. Cox
Department of Biochemistry, 420 Henry Mall, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA.

Summary

The *Escherichia coli* RecA protein plays a central role in homologous genetic recombination, recombinational repair, and several other processes in bacteria. *In vitro*, an extended filament involving thousands of RecA monomers promotes a reaction in which individual DNA strands switch pairing partners (DNA strand exchange). This reaction has been extensively studied as a paradigm for the central steps in recombination. Because the strand-exchange reaction is relatively simple and isoenergetic, the complexity of the RecA system that carries it out has led to controversy about the functional significance of many fundamental properties of RecA. Filamentous protein structures involving thousands of RecA monomers, which hydrolyse 100 ATPs per base pair of heteroduplex DNA formed, are hard to rationalize in the context of recombination between two homologous DNAs. The thermodynamic barriers to strand exchange are much too small. These molecular features of the system can be easily rationalized, however, by shifting the focus to DNA repair.

Introduction

The RecA system evolved over 1.5 billion years ago and is highly conserved among bacteria (Roca and Cox, 1990). In *E. coli*, identified functions of RecA protein include homologous recombination, recombinational repair, the induction of the SOS response to DNA damage, and SOS mutagenesis (Rupp and Howard-Flanders, 1968; Roberts *et al.*, 1978; West *et al.*, 1981; Radding, 1982; 1988; Dutreix *et al.*, 1989; Smith and Wang, 1989; Woodgate *et al.*, 1989). This protein has been studied extensively *in vitro*, often with emphasis on experimental systems that mimic its putative activities in recombination. The complexity and energetic cost of the system revealed in these studies, however, suggests that more attention should be given to the recombinational repair function.

Received 22 November, 1990. Tel. (608) 262 1181; Fax (608) 262 3453.

DNA-repair systems deal with a plethora of chemical insults to cellular chromosomes. Even in the absence of external DNA-damaging agents, the spontaneous lesions (oxidative damage, cytosine deamination, base loss, etc.) that occur in a typical mammalian genome each day number at least in the tens of thousands (Friedberg, 1985; Richter *et al.*, 1988). Much of this damage is repaired successfully and accurately because DNA consists of two complementary strands. A lesion in one strand can be excised and replaced by replicating the information in the undamaged strand.

Under some circumstances, however, the second strand is absent or also damaged. These circumstances include lesions in single-stranded gaps (e.g. lesions left behind by replication bypass), double-strand breaks, and double-strand crosslinks. When these types of damage are present, the information required for accurate repair must come from a separate, homologous DNA molecule. The repair mechanism utilizes the cellular homologous recombinational system in processes that are often called recombinational repair. Evidence is accumulating that lesions requiring recombinational repair are very common (Friedberg, 1985; Richter *et al.*, 1988; Roca and Cox, 1990). Many studies also indicate that these types of lesions are highly recombinogenic (Cole, 1971; Rupp *et al.* 1971; Konrad, 1977). For example, the loss of function mutations that produce hyper-rec phenotypes in *E. coli* are generally those that would tend to increase the number of unrepaired single- and/or double-strand breaks (*polA*, *lig*, *dam*, *dut*) (Konrad, 1977). If recombinational repair is compromised by a mutation that eliminates RecA function, each of these mutants is rendered non-viable (Weinstock, 1987).

In *E. coli*, the most important pathways for recombinational repair clearly involve RecA (Rupp and Howard-Flanders, 1968; West *et al.*, 1981; Kushner, 1987; Weinstock, 1987; Cheng *et al.*, 1988; Smith and Wang, 1989). The putative pathway for post-replication repair (West *et al.*, 1981) involves the RecA-mediated DNA strand-exchange activity (Fig. 1). The same RecA activity is envisaged as a central part of the repair of double-strand breaks and crosslinks (Sinden and Cole, 1978; Cheng *et al.*, 1988; Smith and Wang, 1989). A wild-type *E. coli* strain

can survive conditions that produce 65 psoralen crosslinks per cell, whereas a *recA13 uvrA6* mutant cell cannot survive one crosslink (Cole, 1971). In fact, sensitivity to DNA-damaging agents is the prototypical phenotype of RecA mutants, and RecA mutations render the cell more sensitive to ultraviolet light than mutations in excision repair genes such as *uvrA*. Even in the absence of external DNA-damaging agents, only about 50% of the cells in a *recA* null mutant strain are viable and 10% do not contain DNA (Capaldo *et al.*, 1974). This indicates that levels of spontaneous DNA damage of types requiring recombination repair are high enough to incapacitate one cell out of two in the absence of RecA. If this interpretation is correct, then the repair function of RecA is much more important to cell survival than the recombination function. Nevertheless, it is the recombination function that is usually highlighted, especially in reports dealing with the activities of RecA protein *in vitro*.

RecA has been extensively studied in the context of a recombination system. The simple thesis of this short review is that RecA is a repair system first and a recombination system second. This is by no means a novel viewpoint, and this theme can be seen in many *in vivo* studies (see Radding, 1982; Smith and Wang, 1989; Roca and Cox, 1990) and in several *in vitro* studies of strand exchange through structural barriers (Das Gupta and Radding, 1982; Livneh and Lehman, 1982; Bianchi and Radding, 1983). The idea also echoes the arguments of Campbell (1984), Bernstein (Bernstein *et al.*, 1985), Heffron (Nickoloff *et al.*, 1989), and others, namely that DNA repair is the most important function of homologous genetic recombination. My purpose here is to summarize some key *in vitro* properties of the system and evaluate them with respect to the recombination and repair functions identified *in vivo*.

Recombination and recombinational repair are closely related activities, and RecA protein-mediated DNA strand exchange is central to both processes. However, the potential thermodynamic barriers to strand exchange and the perception of what is mechanistically reasonable change when RecA is presented as a repair system rather than a recombination system. Recombination is usually associated with the generation of genetic diversity. In this role, an element of chance is acceptable and perhaps even desirable. There is no real need for anything but spontaneous or random branch migration following the creation of a crossover. As drawn, the substrates and products of the DNA strand-exchange reaction in Fig. 1A are equivalent in terms of base pairs and the reaction is isoenergetic. If recombination is the primary mission of this system and represents the selective pressure for its evolution and maintenance, then many of the key properties of the RecA system observed *in vitro*, in particular the filamentous structure and a large investment in ATP hydrolysis, become hard to rationalize, particularly in view of the fact that some eukaryotic proteins promote DNA strand exchange but do not form filaments or hydrolyse ATP (Hsieh *et al.*, 1986; Ganea *et al.*, 1987; Kolodner *et al.*, 1987; Halbrook and McEntee, 1989; Moore and Fishel, 1990). In contrast, the recombinational repair process places a variety of special demands on the DNA strand-exchange reaction, in particular a requirement for bypassing DNA lesions that may take a variety of molecular forms. When a lesion is present in a region of single-stranded DNA, as in Fig. 1B, a complementary strand must be provided before the lesion can be repaired; an element of chance is no longer acceptable. In many instances, cell survival will depend upon promoting strand exchange beyond the particular base pair containing the lesion. *In vitro*, RecA protein readily promotes strand exchange past mismatches (Das Gupta and Radding, 1982), pyrimidine dimers (Livneh and Lehman, 1982), and short heterologous inserts (Bianchi and Radding, 1983).

Putting recombinational repair ahead of recombination is not a trivial semantic exercise. It has very important implications for the major outstanding questions about structure, mechanism, and the use of chemical energy in RecA-mediated DNA strand exchange. It also has a large but often unperceived effect on the choice of questions and experimental approaches within the field. For example, the substrates used for RecA-mediated strand-exchange reactions *in vitro* rarely include the DNA lesions or other barriers likely to be encountered *in vivo* (with exceptions, some of which are cited above). The relative importance of genetic exchange versus DNA repair in the cell will not be addressed further here (see Radding, 1978; Kushner, 1987; Weinstock, 1987; Smith and Wang, 1989; Roca and Cox, 1990). Instead, the focus will be on the molecular properties of the RecA system that alone, I believe, provide compelling evidence that the primary mission of RecA protein is DNA repair. At least three arguments can be found in the data derived from the past decade's detailed physical and enzymological characterization of RecA protein and the reactions it promotes.

RecA filaments

The active species in RecA-mediated DNA strand exchange is a helical filament that can include thousands of RecA monomers. There are generally a few thousand RecA monomers in a bacterial cell, and this number can increase 20-fold upon ultraviolet induction (Weinstock, 1987). This represents a substantial cellular investment just in terms of protein biosynthesis. Is this enormously complex system required for recombination? No; it is easy to imagine simpler enzymatic pathways. For example, an enzyme analogous to the integrase class site-specific recombinases (Craig, 1988) could align homologous

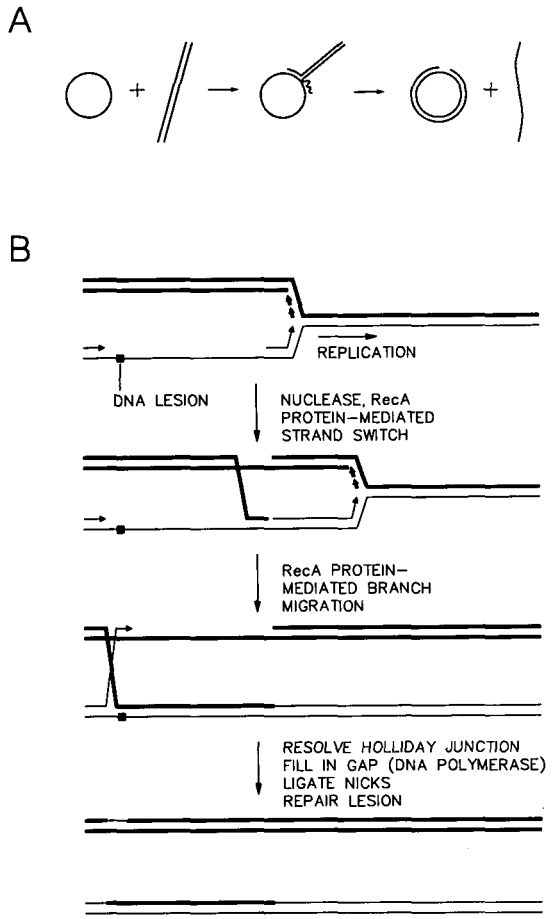

Fig. 1. RecA-mediated DNA strand exchange.
A. A simple strand-exchange reaction involving a circular single-stranded DNA and a homologous linear duplex DNA. These substrates are those most commonly employed *in vitro* and are generally derived from bacteriophages.
B. A strand-exchange pathway for post-replication repair. This model is adapted from that of West *et al.* (1981).

sequences and create Holliday junctions. After some spontaneous branch migration, this junction could be resolved to recombinant products. A few copies of these enzymes would suffice to promote recombination during conjugation, etc. The extended filament is demonstrably unnecessary to facilitate pairing or any other step in strand exchange. It is clear that most of the eukaryotic proteins isolated to date that promote strand exchange do not form filaments. In addition, some can promote strand exchange when added in surprisingly small amounts (Heyer and Kolodner, 1989). Thus it would be unreasonable to expect that the evolutionarily more sophisticated bacterial cell would retain a much more expensive system to generate genetic diversity.

Is the filament required for recombinational repair? Yes; if the strand-exchange reaction is to be promoted any distance at the high efficiency needed in the repair mode, it is necessary to exclude the many DNA-binding proteins present in any cell from its path. Any one of these proteins would act as a major barrier to a spontaneously migrating branch point. The filamentous design contributes greatly to the efficiency of strand exchange merely by excluding other proteins from the region where strand exchange is required.

I note that this is not necessarily the only function of the RecA filament. RecA protein is also involved in the induction of the SOS response (facilitating cleavage of the *lexA* repressor) and in SOS mutagenesis, although how the filamentous form of RecA contributes to these functions is not understood.

DNA binding

Genetic recombination is generally viewed as an exchange of genetic information between two duplex DNA molecules, and model-building traditionally revolves around the adventures of four DNA strands. Any critical examination of RecA properties, however, reveals a system designed to operate with three strands rather than four. RecA does not bind to duplex DNA and initiate exchanges at random locations. Instead, the molecular design efficiently targets the protein to genomic locations at which repair is likely to be required, i.e. at single-strand gaps and lesions. Many separate experimental facts contribute to this picture.

(i) RecA binds to single-stranded DNA almost exclusively at pH values above 7.0 (Radding, 1982; Cox and Lehman, 1987; Roca and Cox, 1990). In the cell, therefore, RecA binding would be expected to be localized at the single-strand gaps created by replication bypass of lesions (Fig. 1). Another possibility is binding to single-stranded DNA ends created by the action of RecBCD, as discussed later.

(ii) The apparent exclusion of RecA from duplex DNA is explained by the pathway for duplex DNA binding (Kowalczykowski *et al.*, 1987; Pugh and Cox, 1987; 1988). The rate-limiting step in binding is nucleation. Once the first monomer (or other binding unit) is bound, the filament is extended rapidly. A key feature of the nucleation step is that it involves an extension and underwinding of the DNA characteristic of the entire nucleoprotein filament once it is assembled. Nucleation is therefore facilitated by structural perturbations that make the DNA more 'unwindable'. A single-strand gap is best, but binding (nucleation) is also facilitated to varying degrees by A/T-rich sequences (Kowalczykowski *et al.*, 1987), stretches of DNA in the Z conformation (Blaho and Wells, 1987; Kim *et al.*, 1989), DNA ends (Lindsley and Cox, 1990a), and, notably, DNA lesions (Lu *et al.*, 1986; Kojima *et al.*, 1990).

(iii) Recent physical studies indicate a capacity for binding three, but not four, DNA strands in the major filament groove (Takahashi *et al.*, 1989; Müller *et al.*, 1990).

(iv) RecA promotes exchanges involving four strands as well as with three, but the end of the second homologous duplex DNA must overlap the single-stranded gap of the DNA within the filament, i.e. four-strand exchanges are generally initiated in three-strand regions. An overlap of 50–100 base pairs stimulates a four-strand exchange by at least two orders of magnitude (Conley and West, 1990; Lindsley and Cox, 1990b).
(v) Alignment of two DNAs occurs much more readily with three strands (one single strand + one duplex) than with four (Lindsley and Cox, 1990b).
(vi) It can be argued that RecA is really only the second enzyme in the overall recombination pathway and that it is designed to act on single-stranded substrates provided by the action of RecBCD (Smith, 1987). It can be counter-argued that RecBCD itself is designed primarily for a repair function. Sensitivity to DNA-damaging agents is a major phenotype of *recB* and *recC* mutations. RecBCD enzyme binds DNA efficiently only at DNA ends. DNA ends will occur only at double-strand breaks, which is one of the types of damage requiring recombinational repair. Double-strand breaks are one of the prominent outcomes of oxidative DNA damage (Richter *et al.*, 1988).

ATP hydrolysis

RecA is a DNA-dependent ATPase with a relatively low turnover number of 25–30 per minute. ATP is hydrolysed uniformly by RecA monomers throughout each filament (Brenner *et al.*, 1987). Although some studies suggest a role for ATP hydrolysis in strand exchange, the molecular function has not been clearly defined. The real controversy revolves around the fact that about 100 ATPs are hydrolysed for every base pair of heteroduplex DNA formed in a typical RecA-mediated DNA strand-exchange reaction. This apparent inefficiency has led many workers to discount the relevance of ATP hydrolysis to strand exchange.

One enlightening result is that of Menetski *et al.* (1990), demonstrating that three-strand exchange reactions involving homologous DNA substrates do not strictly require ATP hydrolysis, occurring to a substantial degree under some conditions with ATP[γS]. As outlined elsewhere (Roca and Cox, 1990), this result still leaves four potential functions for ATP hydrolysis in DNA strand exchange, most of which are particularly relevant to repair: (i) to render the reaction unidirectional (the ATP[γS] reaction does not go to completion), (ii) to provide the energy required for branch migration to bypass significant structural barriers in the DNA (lesions, etc.), (iii) to permit four-strand exchanges, and (iv) to recycle RecA filaments by promoting dissociation.

It is worth noting that whereas the dissociation of RecA monomers occurs largely from only one of two filament ends and requires ATP hydrolysis (Lindsley and Cox, 1990a), this accounts for only a very small fraction of ATP hydrolytic events. Much evidence indicates that dissociation of RecA is not required for DNA strand exchange (see Roca and Cox, 1990). Current ideas for how ATP hydrolysis might be used to facilitate unidirectional strand exchange past DNA lesions and other barriers, and a molecular rationalization for the apparent low efficiency of ATP hydrolysis in this process, are reviewed elsewhere (Roca and Cox, 1990).

Here my primary intention is to point out that the apparently wasteful use of ATP by RecA is not unreasonable when viewed in the context of repair. Cells maintain elaborate and redundant layers of defences to protect and repair cellular chromosomes. One theme that can be seen in many repair systems is that chemical energy conservation is not a high cellular priority when the integrity of the DNA is at stake. The methyl-directed mismatch repair system of *E. coli* provides a good example (Modrich, 1987). In this process, a DNA strand is believed to be degraded and replaced between a hemi-methylated GATC sequence and the mismatch. Since the mismatch may be over a thousand base pairs from the nearest GATC sequence, this process represents an investment of thousands of high-energy bonds (in DNA synthesis) to repair one mismatch. A similar situation is seen in the direct repair of O^6-alkyl-guanine (Walker, 1985). Here repair occurs by transfer of the O^6-methyl or ethyl group to a cysteine on the O^6-alkylguanine-DNA alkyltransferase. The enzyme is irreversibly inactivated by this alkyl transfer. Hence, a 19000 M_r protein is used up to repair one lesion.

As energetically inefficient as it seems to be, the RecA ATPase is not a major drain on cellular metabolism. One thousand RecA monomers operating at V_{max} would utilize less than 0.1% of the energy invested in protein biosynthesis in the same cell (Cox, 1990). This ATP hydrolysis will occur only when RecA is bound to DNA, and as noted above this binding will generally take place only when a suitable single-strand gap or lesion is present to facilitate nucleation. If RecA uses ATP hydrolysis to ensure efficient and unidirectional strand exchange past DNA lesions of all kinds (for purposes of repair) then the energy may well be a worthwhile investment in genome maintenance.

Concluding remarks

DNA damage represents a challenge to cell survival during every cell generation, and provides a good explanation for the early evolution and selective maintenance of this complex system. According to this viewpoint, the occasional exchange and recombination of genetic material between cells is a process that evolved later to take advantage of the properties of RecA.

The perception of a system's primary function inevitably

affects each investigator's approach to it. If generating genetic diversity through recombination is the primary function of the RecA system, then the filament, ATP hydrolysis, and some of the DNA-binding properties are perverse molecular details that must be explained away. If repair is the primary function then these aspects of the molecular design of the system are more readily understood. For the sake of argument, I will offer a rather extreme view and suggest that the implicit emphasis on recombination in most *in vitro* work to date has actually been detrimental to a full appreciation of mechanism and molecular design in the RecA system. RecA, of course, is a complex system and a final evaluation of the functional significance of these and other properties will require a better understanding of the role of RecA in SOS induction and SOS mutagenesis.

The repair theme can be extended to eukaryotes, and much evidence exists, especially in yeast, that eukaryotic general recombination systems are really repair systems in disguise (Szostak *et al.*, 1983; Roca and Cox, 1990). A change in conceptual emphasis from recombination to repair could well facilitate the resolution of many questions in the RecA system and provide an enlightening perspective for the study of recombination itself.

Acknowledgements

Results cited from the author's laboratory were supported by National Institutes of Health Grant GM 32335.

References

Bernstein, H. *et al.* (1985) *Science* **229:** 1277–1281.
Bianchi, M.E., and Radding, C.M. (1983) *Cell* **35:** 511–520.
Blaho, J.A., and Wells, R.D. (1987) *J Biol Chem* **262:** 6082–6088.
Brenner, S.L. *et al.* (1987) *J Biol Chem* **262:** 4011–4016.
Campbell, A. (1984) *Cold Spring Harbor Symp Quant Biol* **49:** 839–844.
Capaldo, F.N. *et al.* (1974) *J Bacteriol* **118:** 242–249.
Cheng, S. *et al.* (1988) *J Biol Chem* **263:** 15110–15117.
Cole, R.S. (1971) *J Bacteriol* **107:** 846–852.
Conley, E.C., and West, S.C. (1990) *J Biol Chem* **265:** 10156–10163.
Cox, M.M. (1990) In *The Biology of Nonspecific DNA–Protein Interactions.*Revzin, A. (ed.). Boca Raton: CRC Press, pp. 171–196.
Cox, M.M., and Lehman, I.R. (1987) *Annu Rev Biochem* **56:** 229–262.
Craig, N. (1988) *Annu Rev Genet* **22:** 77–105.
Das Gupta, C., and Radding, C.M. (1982) *Proc Natl Acad Sci USA* **79:** 762–766.
Dutreix, M. *et al.* (1989) *J Bacteriol* **171:** 2415–2423.
Friedberg, E.C. (1985) *DNA Repair.* New York: Freeman.
Ganea, D. *et al.* (1987) *Mol Cell Biol* **7:** 3124–3130.
Halbrook, J., and McEntee, K. (1989) *J Biol Chem* **264:** 21403–21412.
Heyer, W.D., and Kolodner, R.D. (1989) *Biochemistry* **28:** 2856–2862.
Hsieh, P. *et al.* (1986) *Cell* **44:** 885–894.
Kim, J. *et al.* (1989) *J Biol Chem* **264:** 21848–21856.
Kojima, M. *et al.* (1990) *Nucl Acids Res* **18:** 2707–2714.
Kolodner, R. *et al.* (1987) *Proc Natl Acad Sci USA* **84:** 5560–5564.
Konrad, E.B. (1977) *J Bacteriol* **130:** 167–172.
Kowalczykowski, S.C. *et al.* (1987) *Proc Natl Acad Sci USA* **84:** 3127–3131.
Kushner, S.R. (1987) In Escherichia coli *and* Salmonella typhimurium. *Cellular and Molecular Biology.* Vol. 2. Neidhart, F.C. *et al.* (eds). Washington, D.C.: American Society for Microbiology, pp. 1044–1053.
Lindsley, J.E., and Cox, M.M. (1990a) *J Biol Chem* **265:** 9043–9054.
Lindsley, J.E., and Cox, M.M. (1990b) *J Biol Chem* **265:** 10164–10171.
Livneh, Z., and Lehman, I.R. (1982) *Proc Natl Acad Sci USA* **79:** 3171–3175.
Lu, C. *et al.* (1986) *Proc Natl Acad Sci USA* **83:** 619–623.
Menetski, J.P. *et al.* (1990) *Proc Natl Acad Sci USA* **87:** 21–25.
Modrich, P. (1987) *Annu Rev Biochem* **56:** 435–466.
Moore, S.P., and Fishel, R. (1990) *J Biol Chem* **265:** 11108–11117.
Müller, B. *et al.* (1990) *J Mol Biol* **212:** 97–112.
Nickoloff, J.A. *et al.* (1989) *J Mol Biol* **207:** 527–541.
Pugh, B.F., and Cox, M.M. (1987) *J Biol Chem* **262:** 1326–1336.
Pugh, B.F., and Cox, M.M. (1988) *J Mol Biol* **203:** 479–493.
Radding, C.M. (1978) *Annu Rev Biochem* **47:** 847–880.
Radding, C.M. (1982) *Annu Rev Genet* **16:** 405–437.
Radding, C.M. (1988) In *Genetic Recombination.* Kucherlapati, R., and Smith, G.R. (eds). Washington, D.C.: American Society for Microbiology, pp. 193–229.
Richter, C. *et al.* (1988) *Proc Natl Acad Sci USA* **85:** 6465–6467.
Roberts, J.W. *et al.* (1978) *Proc Natl Acad Sci USA* **75:** 4714–4718.
Roca, A.I., and Cox, M.M. (1990) *Crit Rev Biochem Mol Biol* **25:** 415–456.
Rupp, W.D., and Howard-Flanders, P. (1968) *J Mol Biol* **31:** 291–304.
Rupp, W.D. *et al.* (1971) *J Mol Biol* **61:** 25–44.
Sinden, R.R., and Cole, R.S. (1978) *Proc Natl Acad Sci USA* **75:** 2373–2377.
Smith, G.R. (1987) *Annu Rev Genet* **21:** 179–201.
Smith, K.C., and Wang, T.C. (1989) *BioEssasy* **10:** 12–16.
Szostak, J. *et al.* (1983) *Cell* **33:** 25–35.
Takahashi, M. *et al.* (1989) *J Mol Biol* **205:** 137–147.
Walker, G.L. (1985) *Annu Rev Biochem* **54:** 425–457.
Weinstock, G.M. (1987) In Escherichia coli *and* Salmonella typhimurium. *Cellular and Molecular Biology.* Vol. 2. Neidhart, F.C. *et al.* (eds). Washington, D.C.: American Society for Microbiology, pp. 1034–1043.
West, S.C. *et al.* (1981) *Nature* **294:** 659–662.
Woodgate, R. *et al.* (1989) *Proc Natl Acad Sci USA* **86:** 7301–7305.