Rescuing a sinking ship: The role of recombination gene products in SOS induction in *Escherichia coli*

**Abstract**

In *Escherichia coli* (E. coli) DNA damage is repaired by the process of homologous recombination (HR). There are two main types of DNA damage, double-stranded (ds) DNA breaks (DSBs) and single-stranded (ss) DNA gaps (SSGs). DSBs can arise from external DNA-damaging agents, from induction of specific endonucleases which introduce DSBs in a specific recognition site, or due to endogenous DNA damage. SSGs are formed after replication of UV irradiated E. coli cells. Both types of DNA damage, DSBs and SSGs, induce the SOS response which includes elevated expression of genes whose products are involved in DNA metabolism, inhibition of cell division and prophage induction.

DSBs are repaired by the RecBCD pathway of recombination, whereas SSGs are repaired by the RecF recombination pathway. Proteins in both recombination pathways, i.e., RecBCD and RecF, act to produce the recombinogenic RecA filament which is crucial for recombinational DNA repair and induction of the SOS response. It is known that the inactivation of some recombination gene products can lead to an impaired SOS response. Here we review the roles of recombination proteins in the formation of a RecA filament and in the induction of a SOS response.

**REGulation of THE SOS RESPONSE – THE LEXA PROTEIN AND REC A FILAMENT AS KEY PLAYERS**

In *Escherichia coli* (E. coli) there is a complex mechanism of DNA repair which is activated when DNA damage occurs or DNA replication blocked. This first line of defence in the repair of damaged DNA, the so called the SOS response, was discovered about 40 years ago (1, 2). It consists of induced expression of genes whose products are involved in inhibition of cell division, prophage induction and DNA metabolism, i.e. repair, recombination, replication and mutagenesis (3).

The regulation of the SOS system is dependent on two proteins: the LexA repressor and the activated form of the RecA protein (4) (Figure 1). When there is no DNA damage in a cell, the LexA protein exists in the form of a dimer and acts as a repressor that binds to the SOS box element in the promoter region of SOS genes (5). The SOS box is a palindromic sequence that contains an inverted consensus sequence 5’-TACTGTATATATATACAGTA-3’, with two completely conserved triplets (underlined) (3). In vitro, the RecA protein is activated in the presence of single-stranded DNA (ssDNA) and ATP or dATP. In vivo,
ssDNA regions including SSGs which appear after treatment of E. coli cells with DNA damaging agents are crucial for activation of the RecA protein (6, 7). It is known that in the presence of ssDNA and ATP, the RecA protein activates the SOS response by catalysing the proteolytic self-cleavage of the LexA repressor. This occurs between the residues Ala<sup>88</sup> and Gly<sup>89</sup> (8) which are situated within the cleavage site region (Gly<sup>73</sup>-Tyr<sup>98</sup>) in the C-terminal catalytic domain (Gly<sup>73</sup>-Asn<sup>198</sup>) (9). The crystal structure of the LexA cleavage site region revealed two conformations, cleavable (C) and non-cleavable (NC) (9). It was proposed that the RecA protein either increases the level of the C-form, binds to it and promotes cleavage (9) or it can stimulate the transition from the NC-form to the C-form of the LexA repressor (10). However, further studies are needed to elucidate the LexA-RecA cleavage mechanism. In vitro, under specific conditions (alkaline pH, in the presence of EDTA and Ca<sup>2+</sup>, Co<sup>2+</sup> or Mg<sup>2+</sup>), the auto-cleavage of the LexA repressor can occur in the absence of the RecA protein (11).

The timing and level of SOS gene expression depends on the binding strength of the LexA repressor to the SOS box. The binding strength of the LexA repressor for the SOS box depends on the sequence of the SOS box. The sequence variation of the SOS box compared with the consensus palindromic sequence can be described by the heterology index (HI) which represents the measure of LexA repressor’s binding strength (12). Using the HI and location of the putative SOS box a total of 69 genes/opernons were predicted to be regulated by the LexA repressor (12). Some SOS genes possess several SOS boxes which regulate their expression (13).

Depending on LexA binding strength to SOS boxes, there are several levels of SOS expression after UV radiation (5, 14). The first level of expression comprises SOS gene products involved in the repair of lesions arising in double-stranded DNA (dsDNA), i.e., pyrimidine dimers. These genes are yerA, yerB and yerD whose products are involved in nucleotide excision repair (NER). If the non-coding lesions are not repaired by NER, SSGs will be formed due to reinitiation of the replication machinery downstream from the non-coding lesion (15). The second level of expression includes expression of recA and other homologous recombination genes whose products are involved in the repair of SSGs and double-stranded DNA breaks (DSBs). The recombination proteins convert the SSGs into dsDNA making it a suitable substrate for the NER proteins. If the inducing signal still persists due to several lesions, RecA will be activated. The NER proteins involved in translesion DNA synthesis (TLS) will be induced. The subunits of DNA polymerase PolV involved in translesion DNA synthesis (TLS) will be induced. The subunits of DNA polymerase PolV are encoded by the umuC and umuD genes. In this step, the sfaA gene is expressed and consequently the SfaA protein inhibits cell division until the DNA damage is repaired. In the final step, if DNA damage cannot be repaired, the genes encoding colicin toxins will be expressed and will cause apoptosis and cell lysis (3, 14). Since the expression of the lexA gene is under control of the LexA repressor itself (16), the shift to uninduced state is achieved once DNA damage is repaired and when there is no more inducing signal in the cell, i.e., RecA filament (Figure 1).

**DNA recombination proteins involved in RecA filament formation**

In vivo, the RecA filament is produced by recombination proteins after processing damaged DNA. Depending on the type of DNA damage (DSBs or SSGs), there are two recombination mechanisms involved in repair. In wild type (wt) E. coli, DSBs are repaired by the RecBCD recombination pathway, whereas SSGs are repaired by the RecF recombination pathway (17). In a recBC sbcBCD mutant, DSBs are repaired by the RecF pathway. In a recB mutant lacking nuclease and RecA loading functions, i.e. recB1080, DSBs are repaired by a hybrid recombination pathway which includes components of both pathways, i.e., RecBCD enzyme and the proteins of the RecF pathway (Figure 1).

Three biochemical activities in the recombination proteins (helicase, 5'-3' exonuclease and RecA loading) are required for RecA filament formation. The RecBCD enzyme is the only component of the RecBCD recombination pathway and possesses all of these activities. In the RecF pathway, several proteins are involved in RecF filament formation, i.e., RecQ helicase, RecJ exonuclease and RecFOR complex.

The RecBCD enzyme is comprised of three subunits RecB, RecC and RecD which are encoded by the recB, recC and recD genes (18). dsDNA with blunt or nearly blunt ends is a substrate for the RecBCD enzyme (19). The mechanism of RecBCD enzyme action is dependent on the relative ratio of Mg<sup>2+</sup> and ATP. Under physiological conditions, or in vitro where Mg<sup>2+</sup> is more abundant than ATP, unwinding of dsDNA is achieved by the coordinated action of the ATP driven two helicase motor subunits (RecB and RecD) that bind and operate on opposite DNA strands (20, 21). Consequently, RecBCD enzyme acts as a bipolar helicase that has the ability to unwind dsDNA. RecD is a faster helicase and operates on the 5’ DNA strand, whereas RecB is a slower helicase and operates on the 3’ DNA strand. Therefore, a ssDNA loop is formed in front of the slower RecB helicase (21, 22, 23). The RecB subunit possesses a C-terminal nuclease domain (RecB<sup>1080</sup>) (residues Ala<sup>928</sup>-Ala<sup>1180</sup>) involved in interaction with the RecA protein and is essential for RecA loading function (24, 25, 26, 27). The crystal structure of the RecBCD enzyme revealed that unwinding of DNA strands by RecB and RecD subunits is conducted across the pin structure of the RecC subunit (28). The RecC subunit is involved in Chi activity and Chi sequence recognition (29, 30). Thus, after the 3’ DNA strand is unwound by the RecB helicase, it passes through the tunnel...
of the RecC subunit that „scans” the 3' end for the Chi sequence. Following this, the 3' end passes near the RecB-
mur domain and is frequently cleaved. On the other hand, the 5' end is distant from the RecB-
mur domain and is less frequently cleaved (28, 31). With 1009 sites in the E. coli genome, the Chi octamer sequence (5'-GCTGGTGG-3') is quite abundant and modulates activities of the RecBCD enzyme (15). Chi sequence recognition leads to pausing of the RecBCD enzyme (32), and the RecC subunit binds to the 3' end of the Chi site, preventing further cleavage of ssDNA from the 3' end. After interaction with the Chi site RecBCD translocates with reduced velocity, cleaving more frequently ssDNA from the 5' end and leaving ssDNA with the 3' end intact (31, 32, 33). In addition to this, the Chi sequence stimulates the RecBCD dependent loading of the RecA protein onto ssDNA downstream of the Chi site (34). This step in RecA filament formation is dependent on the RecB-
mur domain (26, 27). Recent publications identified RecC residues involved in Chi site recognition and suggest a role for these residues in the conformational change that converts the RecBCD enzyme from a destructive nuclease into a recombinase after it encounters a Chi sequence (35, 36). In vitro, where ATP is more abundant than Mg2++, the RecBCD enzyme unwinds dsDNA without cleaving until it reaches the Chi site. Then it introduces an endonucleolytic cut a few nucleotides from the Chi site at its 3' end. On this ssDNA, RecBCD enzyme loads RecA protein and forms a RecA filament without further DNA strand cleavage (20).

In the wt bacteria SSGs are repaired by the enzymes of the RecF recombination pathway. However, if RecBCD enzyme, as well as Exol exonuclease and SbcCD nuclease, are all inactivated by mutations, DSB repair is dependent on the proteins of the RecF recombination pathway. Exol is a 3'-ssDNA exonuclease which trims 3'ssDNA ends to form blunt dsDNA which is the substrate for the RecBCD enzyme (37, 38). SbcCD possesses dsDNA exonuclease and ssDNA endonuclease activities, and it cleaves dsDNA and especially cruciform structures (39, 40).

The proteins of the RecF machinery, i.e., RecQ, RecJ, RecF, RecO and RecR, are involved in RecA filament formation. The processing of DSBs is initiated by the RecQ protein which is a 3'-5' helicase and acts on partially dsDNA or completely dsDNA (41). In vitro, the RecQ helicase is capable of initiating homologous recombination in the presence of SSB protein and disrupting the nascent joint molecules (42). dsDNA end processing also involves the action of the RecJ protein which is a 5'-3' ssDNA exonuclease (43). In vitro, the presence of the Mg2+ ions and 7 nucleotide or longer 5' end tail are required for efficient binding of the RecJ exonuclease. Also, the DNA binding and nuclease activity of the RecJ protein are enhanced in the presence of the SSB protein (44). The result of RecQ and RecJ dsDNA end processing is a 3' overhang coated with the SSB protein. In vitro, when the SSB protein is bound to ssDNA it protects the ssDNA from RecBCD dependent cleavage (45) and obstructs the formation of secondary structures (46, 47). SSB protein competes with wt RecA for binding to ssDNA (48). The higher affinity of SSB protein for ssDNA can be overcome by addition of the RecF, RecO and RecR proteins (49). The RecF protein interacts with RecO, RecR, SSB and RecA proteins (49). It is proposed that RecO and RecR bind and modify the complex of ssDNA coated with SSB protein to promote the nucleation step of the RecA protein (50). Additionally, the RecFOR enzymatic complex recognizes ss/dsDNA junctions and loads RecA protein on SSGs coated with SSB protein (51).

In some recBCD mutants which partially affect RecBCD functions, the repair of DSBs is achieved by a hybrid recombination pathways. In the hybrid recombination pathways proteins of both recombination machineries, i.e., RecBCD and RecF, act together to produce the RecA filament. The first hybrid recombination pathway was discovered in a recD mutant where conjugation and DNA repair depends on the recf gene product (52, 53). The recD mutation inactivates the RecBCD dependent nuclease activity, such that the dsDNA is processed into a suitable recombination substrate by 5'-3'exonuclease activity of the Recf protein (54). In the recB1080, a mutant lacking nuclease and RecA loading functions of the RecBCD enzyme, DSB repair is dependent on RecB1080CD (helicase), Recf (nuclease) and RecFOR (RecA loading) (20, 55). The RecB1080CD enzyme has a substitution of Asp1080 with Ala1080 in the RecB subunit, and it is able to unwind dsDNA as well as to respond to a Chi site (54, 56). There are two additional hybrid recombination pathways which operate in a recB1080 recD double mutant and in a recC1004 single mutant. It is known that the RecD subunit has an inhibitory effect on genetic recombination and DNA repair because it inhibits the RecA loading function of the RecBCD enzyme before its interaction with a Chi site. Consequently, the inactivation of the RecD subunit results in a RecBC(D') enzyme with helicase activity and Chi-independent RecA loading function (57, 58). The hybrid recombination pathway in the recB1080 recD double mutant is primarily dependent on the recf gene product, but is also partially dependent on the recFOR gene products (59). The recC1004 mutant is a member of the recC* class of Chi recognition deficient mutants. Recombinational repair in a recC1004 mutant is achieved by a hybrid pathway and is dependent on recf and recFOR gene products (60, 61). The RecBC1004D enzyme possesses wt levels of helicase and moderately enhanced exo- and endonuclease activities (62). The presence of nuclease activity in the mutant enzyme cannot explain the fact that DNA repair in a recC1004 mutant depends on the recf gene product. In contrast to DNA repair, conjugal recombination after Hfr crosses in a recC1004 mutant is independent of recf and recFOR gene products (61).
The SOS response is dependent on recombination proteins

The importance of recombination proteins for SOS induction is obvious since the active form of RecA protein is responsible for SOS induction and exists as a RecA filament. Consequently, all the recombination proteins required for RecA filament formation are also expected to be required for SOS induction. Depending on the type of DNA damage, different recombination proteins participate in damage processing which leads to RecA filament formation and SOS induction. Therefore, inactivation of the recombination proteins or their specific biochemical activities (helicase, 5'-3' exonuclease or RecA loading) leads to altered SOS induction. If cells are exposed to DNA damaging agents that induce DSBs (e.g., ionizing radiation, DNA gyrase inhibitors, etc) SOS induction is dependent on the RecBCD enzyme (63) (Figure 1). On the other hand, if DNA damaging treatments cause SSG formation (e.g., UV radiation), SOS induction is dependent on the proteins of the RecF pathway (3) (Figure 1).

γ radiation causes DSBs as well as different types of DNA damage, e.g., base damage, ssDNA breaks. SOS induction after γ radiation in wt cells is dependent on several recombination proteins which are involved in DNA processing, i.e., recB, recO and recJ gene products (64). A recent study by these authors shows that additional ssDNA exonucleases, i.e., Exol, ExoVII, ExoX, may play a role in DNA processing needed for SOS induction after γ radiation (65). As mentioned earlier a hybrid recombination pathway operates in the recB1080 mutant. In agreement with this, it was shown that the SOS response after introduction of DSBs by the I-SceI endonuclease in the recB1080 mutant is dependent on the recJ gene product and partially dependent on recFOR gene products (66). This indicates that the processing of DSBs into a RecA filament and consequently the induction of the SOS response requires the helicase activity of the RecB1080CD enzyme, the 5'-3' nuclease activity of RecJ, and also participation of RecFOR in RecA loading. Similar results were obtained for SOS induction after γ radiation in the recB1080 mutant (Vlašić et al., unpublished data).

UV radiation causes non-coding lesions that stop the DNA replication machinery which reinitiates replication downstream from the lesion and creates a SSG (15). SSGs are repaired by proteins of the RecF recombination pathway via a RecA filament. Some of the RecA filaments which do not continue with the process of recombination can induce the SOS response. Therefore, the inactivation of the recF, recO and recR gene products causes a delay in the SOS response after UV radiation (67, 68). This indicates that soon after UV radiation, the RecA loading function of the RecFOR complex is required for the formation of a SOS inducing signal on the SSGs. SOS induction after UV radiation is not abolished in the recF, recO and recR mutants due to the presence of a functional RecBCD enzyme. It is known that during replication, unrepaired SSGs are converted to dsDNA ends (due to replication fork collapse) and are repaired by the RecBCD enzyme (69). Therefore, SOS induction in later phases after UV radiation in recF, recO and recR mutants depends on the RecBCD enzyme. Additional genetic analysis showed that SOS induction after UV radiation in a recB1080 mutant depends on the RecFOR complex, indicating that RecBCD or RecFOR dependent RecA loading on ssDNA is important for SOS induction after UV radiation (70).

In addition to exogenous DNA damage, endogenous DNA damage can also induce a SOS response. The most frequent and dangerous spontaneous type of DNA damage are collapsed replication forks (71) which arise if the replication machinery encounters a nick, break or gap in DNA (15). Some E. coli mutants exhibit constitutive SOS expression, i.e., SOS expression in the absence of external DNA damage. The constitutive SOS response can arise either due to inactivation of the proteins that are involved in DNA repair which leads to accumulation of endogenous DNA damage, or due to impaired function of proteins involved in DNA replication which leads to an increase in ssDNA. Consequently, the majority of constitutive SOS mutants have inactivated gene products that are involved in DNA recombination, repair and replication (72).

One example of a mutant with moderate constitutive SOS expression is the recB1080 mutant. The recB1080 mutation causes a constitutive SOS response which is dependent on the RecJ nuclease (66, 70). Since SOS induction after introduction of DSBs and the constitutive SOS response in recB1080 mutant have similar genetic requirements, it is likely that the constitutive SOS response in recB1080 mutant is caused by spontaneously occurring dsDNA ends, i.e., collapsed replication forks (66). A second example is a mutation which leads to an extremely high constitutive SOS response, namely recA730. The RecA730 protein has a substituted Glu198 with Lys58 (73), and it competes better with the SSB protein for ssDNA (74). Consequently, the RecA730 protein easily displaces the SSB protein from ssDNA (75), and exhibits constitutive LexA repressor cleavage activity (75). It was shown in vivo that the recA730 mutation can suppress recombination and survival deficiency after UV and gamma radiation of the recF, recO and recR mutations in both the wt and recB1080 background. This indicates that the RecA730 protein can compensate the RecA loading deficiency due to RecFOR inactivation (76, 77). On the other hand, RecA730 cannot suppress the recombination and survival deficiency after UV and gamma radiation if the defect is at the level of nuclease activity (77). The 5'-3' exonuclease activity of the RecBCD or RecJ protein is essential in achieving a high constitutive SOS response in...
The role of recombination gene products in SOS induction

Ignacija Vlašić and K. Brčić-Kostić

Period biol, Vol 116, No 2, 2014. 163

recA730 mutant. The formation of 3’ssDNA ends by the nuclease is needed for the process of RecA730 auto-loading onto ssDNA (78). Interestingly, the genetic requirements for constitutive SOS response in a recA730 background differ depending on the media used. In minimal media, the level of constitutive SOS response in the recA730 mutant is lower and is independent of the RecBCD enzyme (78, 79), whereas in rich media, it is much higher and partially dependent on the helicase function of the RecBCD enzyme (78). A possible explanation for this effect is that cells grown in the rich medium need less time to enter into stationary phase when compared to cells grown in minimal medium. Since stationary phase cells have a lower capacity for recombinational DNA repair, it is likely that DNA damage persists longer and leads to a higher SOS response. Also, DNA replication in rich medium is more extensive than in minimal medium which can lead to more frequent replication fork collapse. Consequently, recA730 mutants grown in rich medium have a higher constitutive SOS response than cells grown in minimal medium. Moreover, it was shown that this different composition of growth medium causes different SOS expression levels and different genetic requirements for DNA repair when E. coli cells were exposed to bleomycin (80). When cells were grown in rich medium with bleomycin, the SOS response was stronger and DNA repair was dependent on homologous recombination, whereas in minimal medium, bleomycin treated cells expressed a weaker SOS induction and DNA repair was independent of homologous recombination (80). A key and still unanswered question relates to the difference between the RecA filament responsible for recombination and the RecA filament involved in the SOS response? It was shown that under normal physiological conditions during logarithmic growth about 15% of the cells form RecA filaments involved in recombination (69, 71), whereas only about 0.3% cells express the SOS response (81). This indicates that only a small subpopulation of cells have induced a SOS response due to spontaneously occurring DNA damage. Some in vivo studies indicate that the ATP hydrolysis of the RecA protein is crucial for SOS induction after UV irradiation (82), whereas another study shows that ATP hydrolysis is important only for recombination and DNA repair but not for the SOS induction (83). In addition, it seems that longer RecA filaments with different, extended conformation are involved in SOS induction (79, 83). Further studies are needed to explain the differences at the structural and biochemical level between RecA filaments involved in SOS induction and recombination.

Concluding remarks

Escherichia coli possesses several recombination pathways (RecBCD, RecF and hybrid pathways) which are responsible for the repair of DNA damage. Three biochemical activities of these recombination proteins (helicase, 5’-3’ exonuclease and RecA loading) are required for the processing of DNA damage into RecA filaments which can be used for recombinational repair and for SOS induction. Although genetic requirements for the formation of RecA filaments involved in DNA repair and SOS induction are similar, there are some differences between them. Further research at the structural and biochemical level is needed to explain these differences.

Acknowledgments: This work was supported by the Croatian Ministry of Science (grant 098-0982913-2867).
REFERENCES

1. RADMAN M 1975 SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. Basic Life Sci 3A: 355-367
24. YU M, SOUAYA J, JULIN D A 1998 The 30-kDa C-terminal domain of the RecB protein is critical for the nucleolytic activity, but not the helicase activity, of the RecBCD enzyme from Escherichia coli. Proc Natl Acad Sci U S A 95(3): 981-986
27. SPIES M, KOWALCZYKOWSKI S C 2006 The RecA binding locus of RecBCD is general domain for recruitment of DNA strand exchange proteins. Mol Cell 21(4): 573-580
30. ARNOLD D A, HANNA N, KOBABASHI I, KOWALCZYKOWSKI S C 2000 A novel, 11 nucleotide variant of Chi, the recombination hot spot of Escherichia coli is a regulatory element that switches the polarity of DNA degradation by RecBCD enzyme. Genes & Dev 11(5): 571-581
31. KOWALCZYKOWSKI S C 1997 The recombination hot spot χ is a regulatory element that switches the polarity of DNA degradation by RecBCD enzyme. Genes & Dev 11(5): 571-581
33. ANDERSON D G, CHURCHILL J J, KOWALCZYKOWSKI S C 1997 Chi-activated RecBCD enzyme possesses 5´→3´ nucleolytic activity but RecBC enzyme does not: evidence suggesting that the alteration induced by Chi is not simply ejection of the RecD subunit. Genes Cells 2(2): 117-128
34. ANDERSON D G, KOWALCZYKOWSKI S C 1997 The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA by χ-regulated manner. Cell 90(4): 77-86
The role of recombination gene products in SOS induction

Ignacija Vlašić and K. Brčić-Kostić


41. UMEZU K, NAKAYAMA K, NAKAYAMA H 1990 Escherichia coli RecQ protein is a DNA helicase. PNAS 87(14): 5363-5367


47. KOWALCZYKOWSKI S C, KRUPP R A 1987 Effects of Escherichia coli SSB protein on the single-stranded DNA-dependent ATPase activity of Escherichia coli RecA protein. Evidence that SSB protein facilitates the binding of RecA protein to regions of secondary structure within single-stranded DNA. J Mol Biol 193(3): 97-113


63. CHAUDHURY A M, SMITH G R 1985 Role of Escherichia coli RecBC enzyme in SOS induction. Mol Gen Genet 201(3): 525-528

64. BREÑA-VALLE M, SERMENT-GUERRERO J 1998 SOS induction by gamma-radiation in Escherichia coli strains defective in repair and/or recombination mechanisms. Mutagenesis 13 (6): 637-641


Ignacija Vlašić and K. Brčić-Kostić

The role of recombination gene products in SOS induction

74. LAVERY P E, KOWALCZYKOWSKI S C 1990 Properties of recA441 protein-catalyzed DNA strand exchange can be attributed to an enhanced ability to compete with SSB protein. *J Biol Chem* 265 (7): 4004-4010