Extending the gap and loading RecA:
the presynaptic phase plays a pivotal role in modulating lesion tolerance pathways

Luisa Laureti[1,*], Lara Lee[1], Gaëlle Philippin[1], Michel Kahi[1], Vincent Pagès[1,*]

[1]Cancer Research Center of Marseille: Team DNA Damage and Genome Instability | CNRS, Aix Marseille Univ, Inserm, Institut Paoli-Calmettes, Marseille, France
[*]vincent.pages@cnrs.fr
[*]Correspondence may also be addressed to: luisa.laureti@inserm.fr

ABSTRACT

During replication, the presence of unrepaired lesions results in the formation of single stranded DNA (ssDNA) gaps that need to be repaired to preserve genome integrity and cell survival. All organisms have evolved two major lesion tolerance pathways to continue replication: Translesion Synthesis (TLS), potentially mutagenic, and Homology Directed Gap Repair (HDGR), that relies on homologous recombination. In *Escherichia coli,* the RecF pathway repairs such ssDNA gaps by processing them to produce a recombinogenic RecA nucleofilament during the presynaptic phase. In this study, we show that the presynaptic phase is crucial for modulating lesion tolerance pathways. Indeed, impairing either the extension of the ssDNA gap (mediated by the nuclease RecJ and the helicase RecQ) or the loading of RecA (mediated by the RecFOR complex) leads to a decrease in HDGR. We suggest a model where defects in the presynaptic phase delay the formation of the D-loop and increase the time window allowed for TLS. We indeed observe an increase in TLS independent of SOS induction. In addition, we revealed an unexpected synergistic interaction between *recF* and *recJ* genes, that results in a *recA* deficient-like phenotype in which HDGR is almost completely abolished.

INTRODUCTION

Genome of all living organisms is constantly damaged and some lesions can impair replication, potentially leading to mutagenesis and genome instability. In the presence of an unrepaired replication-blocking lesion, a single-stranded DNA (ssDNA) gap is formed, due to the initiation of the next Okazaki fragment (on the lagging strand) or to a repriming event (on the leading strand) (1). In order to complete replication and preserve cell survival, the ssDNA gap is filled in by one of the two lesion tolerance pathways**,** identified both in prokaryotes and eukaryotes: 1) Translesion Synthesis (TLS), which employs specialized DNA polymerases able to replicate damaged DNA, with the potential to introduce mutations (2–4); 2) Damage Avoidance (DA) pathways, which use the information of the sister chromatid to bypass the lesion in a non-mutagenic way through homologous recombination mechanisms (5). In *Escherichia coli*, we have shown that in non-stressed conditions Homology Directed Gap Repair (HDGR), a DA mechanism that relies on the recombinase RecA, is the major lesion tolerance pathway employed by cells while TLS pathway is a minor one (6, 7). We have also identified another lesion tolerance strategy, named Damaged Chromatid Loss (DCL) that promotes cell proliferation at the expense of losing the damaged chromatid; in particular when HDGR mechanism is impaired, cells can survive by replicating only the undamaged chromatid (6).

In *E. coli* there are two well-known RecA dependent recombinational pathways: the RecBCD pathway, involved in the repair of double strand breaks, and the RecF pathway (also known as the RecFOR pathway), involved in the repair of ssDNA gap as well as in SOS induction (8). Historically, the RecF pathway has received less attention because it was thought to be a “back-up” recombinational pathway; however, in the last decades, it became clear that this pathway plays an important role during replication of a damaged DNA.

The RecF pathway is evolutionary conserved in all bacteria and functional orthologs

are also found in Eukaria. This pathway encompasses several proteins and is divided into three distinct phases: i) the presynaptic phase in which the damaged ssDNA gap is processed by different proteins (*i.e.* SSB, RecF, RecO, RecR, RecJ, RecQ) to promote the formation of an active RecA nucleofilament; ii) the synaptic phase in which homology pairing and strand exchange by the RecA nucleofilament occur to produce a D-loop and iii) the postsynaptic phase involving the processing of the recombinational intermediates by the resolvase RuvABC or by the helicase RecG (9). Some of these proteins have already been well characterized *in vitro* (10–12), but their role *in vivo* and their impact on lesion tolerance pathways is not yet fully elucidated. The presynaptic phase is of particular interest since the ssDNA gap is the common substrate for both HDGR and TLS, hence how the ssDNA gap is processed to produce a RecA nucleofilament might influence the choice of lesion tolerance pathways.

Several genetic analyses indicate that the *recF, recO* and *recR* genes belong to the same epistasis group as the double mutants behave like the single mutants, at least for the UV sensitivity and the delay in SOS induction (9). According to *in vivo* and *in vitro* studies, it was proposed that two complexes, RecOR and RecFR, are formed whose function at a ssDNA gap is to disassemble the filament of single strand binding (SSB) proteins in order to load the recombinase RecA and promote homologous recombination (13, 14). *In vitro*, high amounts of the RecOR complex are sufficient to displace SSB proteins and load RecA, promoting strand exchange in the absence of RecF (15, 16). On the other hand, RecF was shown to bind to gapped substrates, preferentially on the 5' double strand-single strand (ds-ss) DNA junction (14, 17) and often colocalize with the replisome (18). The UV sensitivity phenotype of the *recFOR* deficient strains is presumably due to a defect in RecA loading that leads to recombinational repair deficiency. Accordingly, a RecA mutant (*i.e.,* the *recA730* allele) able to load itself on a ssDNA partially restores UV resistance of the *recFOR* deleted strains.

The *recJ* and *recQ* genes have also been associated with the presynaptic phase; however, compared to the *recFOR* deficient strains, no strong UV sensitivity phenotype was observed when *recJ* or *recQ* genes were deleted. Courcelle laboratory was the first to propose that the products of these genes were involved in processing the nascent lagging DNA strand during replication blockage (19). The helicase RecQ (3'->5') would help the exonuclease RecJ (5'->3') to enlarge the ssDNA at the 5' end of the gap to lengthen the substrate for RecA filament formation. *In vitro* data showing that SSB protein is able to stimulate the activity of RecQ and of RecJ corroborate the hypothesis that both RecQ and RecJ are recruited at the ssDNA gap formed upon lesion encountering (12). Recently, we have shown that in the presence of an unrepaired lesion, RecJ is required to enlarge the ssDNA gap not only in the lagging strand but also in the leading strand (20), thus providing further evidence of the *in vivo* role of RecJ in the presynaptic phase.

In the present study, we have explored the role of the genes involved in the presynaptic phase of the RecF pathway, their genetic interaction and their effect on lesion tolerance pathways. We confirmed that, regarding lesion bypass, the *recF, recO* and *recR* genes belong to same epistasis group, likewise *recJ* and *recQ* genes showed to be epistatic. Defect in RecA loading (mediated by the RecFOR complexes) or in enlarging the ssDNA gap (mediated by RecQ and RecJ) results in HDGR impairment that is partially compensated by TLS without induction of the SOS response. Interestingly, our data revealed a strong synergistic interaction between *recF* and *recJ* (as well as *recF* and *recQ*). Indeed, loss of these two genes results in a *recA* deficient-like phenotype in which HDGR is almost completely abolished, indicating that both the size of the ssDNA gap and the loading of RecA play an important role in the formation of an efficient D-loop structure, thus favoring the non-mutagenic HDGR mechanism over TLS pathway.

MATERIAL AND METHODS

Bacterial strains and growth conditions

All *E. coli* strains used in this work are listed in Supplementary Table 1. They are either derivative of strains FBG151 and FBG152 (7), that carry a plasmid that allows the expression of the *int–xis* genes after IPTG induction, or derivative of strains EC1 and EC2 (this study), in which the *int–xis* genes have been inserted into the chromosome of *E. coli*. Strains were grown on solid and in liquid Lysogeny Broth (LB) medium. Gene disruptions were achieved by the one-step PCR method (21) or by P1 transduction using the Keio collection (22).

Following the site-specific recombination reaction, the lesion is located either in the lagging strand (FBG151 or EC1 derived strains) or in the leading strand (FBG152 or EC2 derived strains). Antibiotics were used at the following concentrations: ampicillin 50 or 100 μg/ml; tetracycline 10 μg/ml, kanamycin 100 μg/ml, chloramphenicol 30 μg/ml. When necessary IPTG and X-Gal were added to the medium at 0.2mM and 80 μg/ml, respectively.

Plasmids

pVP135 expresses the integrase and excisionase (*int–xis*) genes from phage lambda under the control of a *trc* promoter that has been weakened by mutations in the -35 and the -10 region. Transcription from Ptrc is regulated by the *lac* repressor, supplied by a copy of $lacI^q$ on the plasmid. The vector has been modified as previously described (7).

pVP146 is derived from pACYC184 plasmid where the chloramphenicol resistance gene has been deleted. This vector, which carries only the tetracycline resistance gene, serves as an internal control for transformation efficiency.

pVP141-144 and pGP1, 2 and 9 are derived from pLDR9-attL-lacZ as described in (7). pLL1, pLL2c and pLL7 are derived from pVP141 and contain several genetic markers as previously described (6). All these plasmid vectors contain the following characteristics: the ampicillin resistance gene, the R6K replication origin that allows plasmid replication only if the recipient strain carries the *pir* gene, and the 5' end of the *lacZ* gene in fusion with the *attL* site-specific recombination site of phage lambda. The P'3 site of *att*L has been mutated (AATCATTAT to AATTATTAT) to avoid the excision of the plasmid once integrated. These plasmids are produced in strain EC100D pir-116 (from Epicentre Biotechnologies, cat# EC6P0950H) in which the pir-116 allele supports higher copy number of R6K origin plasmids. Vectors carrying a single lesion for integration were constructed as previously described following the gap-duplex method (7). A 13-mer oligonucleotide, 5′-GCAAGTTAACACG-3′, containing no lesion or a TT(6-4) lesion (underlined) in the *Hinc*II site was inserted either into the gapped-duplex pLL1/2c leading to an out of frame *lacZ* gene (to measure HDGR) or into the gapped-duplex pGP1/2 leading to an in frame *lacZ* gene (to measure TLS0). A 15-mer oligonucleotide 5'-ATCACCGGCGCCACA-3' containing or not a single G-AAF adduct (underlined) in the *Nar*I site was inserted into the gapped-duplex pLL1/7 (to measure HDGR) or into the gapped-duplexes pVP141-142 or pVP143-144 to score respectively for TLS0 Pol V-dependent and for TLS-2 Pol II-dependent. A 13-mer oligonucleotide, 5′-GAAGACCTGCAGG, containing no lesion or a dG-BaP(-) lesion (underlined) was inserted into the gapped-duplex pVP143/pGP9 leading to an in frame *lacZ* gene (to measure TLS).

Monitoring HDGR and TLS events

The protocol for lesion integration assay is described in details in (23). Cells were diluted and plated before the first cell division using the automatic serial diluter and plater EasySpiral Dilute (Interscience) and were counted using the Scan 1200 automatic colony counter (Interscience). No differences were observed when we used FBG or EC derivative strains.

Following the integration of the pLL1/2c vector (TT6-4 lesion) or pLL1/7 vector (AAF lesion), sectored blue/white colonies represent HDGR events. Following integration of the pVP141/142, pVP143/144, pGP1/2, pVP143/pGP9 vectors (AAF, TT6-4 or BaP lesion, respectively), sectored blue/white colonies represent TLS events. The relative integration efficiencies of lesion-carrying vectors compared with their lesion-free homologues, normalized by the transformation efficiency of pVP146 plasmid in the same electroporation experiment, allow the overall rate of lesion tolerance to be measured. For each lesion we combine the results obtained with the different plasmids to measure HDRG and TLS events (for BaP lesion we only measure TLS events).

The data in every graph represent the average and standard deviation of at least three independent experiments of a lesion inserted in the leading (or in the lagging) orientation. Tolerance events (Y axis) represent the percentage of cells able to survive in presence of the integrated lesion compared to the lesion-free control.

RESULTS

We have set up a genetic assay that allows to insert a single replication-blocking lesion into the chromosome of *E. coli,* exploiting the site-specific recombination system of the bacteriophage lambda. Using cognate lesion-containing plasmids, we can directly monitor either TLS or HDGR

mechanisms by a colorimetric assay based on the *lacZ* reporter gene (6, 7, 24). For this study, we used two blocking lesions: the UV-induced thymine-thymine pyrimidine (6-4) pyrimidone photoproduct (TT6-4) and the chemical guanine adduct, the N-2-acetylaminofluorene (G-AAF). *E. coli* possesses three TLS polymerases, involved in the bypass of different types of DNA lesions: Pol II (encoded by *polB*), Pol IV (encoded by *dinB*) and Pol V (encoded by *umuDC*) (2). The expression of these TLS polymerases is strongly increased upon SOS induction (2). UV lesions are exclusively bypassed by Pol V while the G-AAF lesion when positioned in a particular sequence context (*i.e.*, the *Nar*I hotspot sequence) can be bypassed either by Pol V (named TLS0 events) or Pol II (named TLS-2 events, frameshift -2) (25). The *E. coli* strains used in this study are deficient for the mismatch repair system (*mutS*) to prevent corrections of the genetic markers of the integration plasmids, as well as for nucleotide excision repair (*uvrA*), to avoid excision of the lesion and to focus on lesion tolerance events (see also (6, 7)). In our genetic assay, we can insert the lesion either on the lagging or in the leading orientation regarding the replication fork direction. Since we did not observe any significant difference between the two strand orientations, we will present in the manuscript the results obtained for a lesion inserted in the leading strand, while lagging strand results will be presented in Suppl Fig 1.

Loading of RecA: role of the mediators

In a previous study, we have shown that in the presence of a TT6-4 lesion the single mutants *recF* or *recO* but also the double mutant *recF recO* strongly reduced HDGR (Fig 1) (23). In the absence of its mediators (*i.e.* RecF, RecO or RecR), RecA loading can still occur, but randomly and with a slower kinetics (26, 27). Since HDGR relies on the recombinase RecA (6), this mechanism is necessarily affected. We measured a very low level of Pol V dependent TLS event (<0.5%) because there is essentially no Pol V expression in non-SOS induced cells (28). In addition, Pol V requires interaction with the tip of a RecA nucleofilament formed by RecFOR complex which is not present in the *recF* strain (29, 30).

In the presence of a G-AAF lesion, deletion of either *recF, recO* or *recR* resulted again in a strong decrease in HDGR events, accompanied this time by a significant increase in Pol II dependent

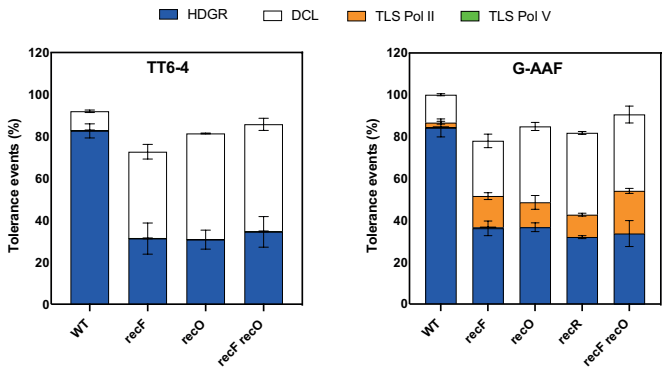


Figure 1. Deficiency in RecA loading promotes replication through TLS polymerases.
The graphs represent the partition of lesion tolerance pathways (i.e. HDGR, TLS and DCL) for the single mutants *recF, recO* and *recR* or the double mutant *recF recO* in the presence of a TT6-4 lesion* (left panel) or a G-AAF lesion (right panel). All the strains are compared to our parental strain (WT in the graphs) which is *uvrA* and *mutS* deficient. *The data for the TT6-4 lesion have been previously published (23).

TLS events (Fig 1). The double mutant *recF recO* showed no further decrease in the level of HDGR compared to the single mutants, neither a significant increase in the level of Pol II dependent TLS (Fig 1). Altogether our genetic data indicate that the *recF, recO* and *recR* genes belong to the same epistasis group for lesion bypass and that the three genes play an important role in HDGR, which is expected due to their role in mediating RecA loading.

In a previous study, we have shown that affecting HDGR favors TLS: mutant alleles of RecA, that can form a nucleofilament but are impaired for homology search and strand exchange (*i.e.* are impaired in the D-loop formation), strongly favor lesion bypass by TLS polymerases (31). However, in this particular context, SOS response was constitutively activated thus the amount of TLS polymerases in the cells was significantly increased shifting the balance towards TLS pathway (31). Interestingly, in the *recF* (as well as *recO* or *recR*) deficient strain, SOS response is not induced, and yet, the level of Pol II TLS at a G-AAF lesion increases to a level similar to the one observed after SOS activation (either by deleting *lexA* or by UV-induction) (Suppl Fig 2). Our data point out that TLS can be regulated by factors other than the SOS response: defect in RecA loading is indeed sufficient to strongly increase the level of Pol II dependent TLS events, without activating the SOS response. To further confirm this, we monitored Pol IV activity on a benzo(a)pyrene (BaP): we show that the level of TLS in a *recF* deficient strain is again increased to

the level observed upon SOS induction by UV-irradiation of the cells (Suppl Fig 2).

Effect of the ssDNA gap extension on lesion bypass

We have recently shown that the size of the ssDNA gap can modulate lesion tolerance pathways. Lack of the nuclease RecJ results in a shorter RecA-covered ssDNA gap, thus homology search and strand exchange are less efficient, impacting HDGR mechanism on one hand and favoring TLS pathway on the other hand (Fig 2 and (20)). While we have previously shown this for the G-AAF lesion, we confirmed that deletion of *recJ* affects HDGR mechanism also for the TT6-4 lesion (Fig 2).

The helicase RecQ has been proposed to help RecJ in enlarging the ssDNA gap at the 5' end (19). To confirm this role, we deleted the *recQ* gene and monitor lesion tolerance pathways in the presence of both G-AAF and TT6-4 lesions. We also observed a decrease in HDGR, but less important than when we deleted *recJ* or *recF* (Fig 2). This might indicate that even if the helicase activity of RecQ is required to provide the substrate for the nuclease activity of RecJ, it is not essential or that another helicase, yet to be identified, can partially compensate for the loss of RecQ. In the presence of a G-AAF lesion, we measured an increase in Pol II dependent TLS similar to the *recJ* strain (Fig 2).

Deletion of both *recJ* and *recQ* in the presence of either a G-AAF or a TT6-4 lesion showed the same level of lesion tolerance pathways (for both HDGR and TLS) as the single mutant *recJ*. Hence, from our genetic analyses, *recJ* and *recQ* appeared to belong to the same epistasis group for lesion bypass, however RecJ plays a more critical role in HDGR mechanism.

It is interesting to note that in the *recJ, recQ* and *recJ recQ* mutants we observe a slight increase in Pol V TLS events (Fig 2 and Suppl Fig 3). In these strains, unlike in the *recF* deficient strain, RecA loading and nucleofilament formation occurs efficiently allowing the contact with Pol V that is required for its activation (29). This increase is only measured with the G-AAF lesion most likely because the low physiological level of Pol V is a limiting factor in the presence of the strong blocking lesion TT6-4. Nevertheless, the result with the G-AAF lesion further confirms that impairing the extension of the ssDNA gap favors replication by TLS polymerases.

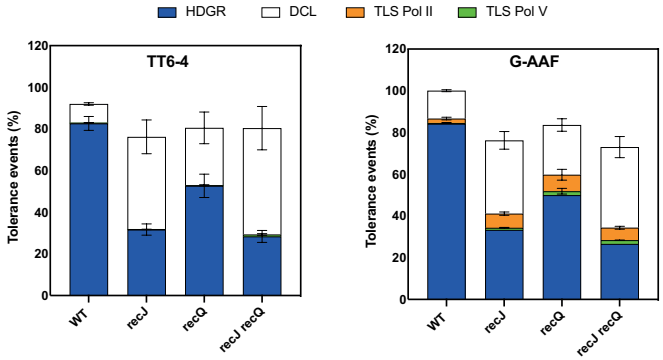


Figure 2. The size of the ssDNA gap affects the efficiency of HDGR mechanism favouring TLS.
The graphs represent the partition of lesion tolerance pathways (i.e. HDGR, TLS and DCL) for the single mutants *recJ** and *recQ* as well as for the double mutant *recJ recQ* in the presence of a TT6-4 lesion (left panel) or a G-AAF lesion (right panel). All the strains are compared to our parental strain (WT in the graphs) which is *uvrA* and *mutS* deficient. *The data for *recJ* in the presence of the G-AAF lesion have been previously published (20).

Synergistic effect between *recF* and *recJ* genes

From our data, it appears that both RecA loading and the size of the ssDNA gap are important factors to modulate the partition and the efficiency of lesion tolerance pathways. To further investigate the presynaptic phase, we decided to analyze the genetic interaction between *recF* and *recJ*, by deleting both genes and monitoring lesion bypass in the presence of a G-AAF and TT6-4 lesion. Interestingly, for both lesions the double mutant showed a drastic drop in HDGR events (resulting in less than 10% of HDGR), compensated mainly by damaged chromatid loss (Fig 3). Similar results were obtained for the double mutant *recF recQ* (Fig 3). These data revealed that *recF* and *recJ* (or *recQ*) genes are synergistic for HDGR and not epistatic as it was expected since they are part of the same recombinational RecF pathway. No synergistic effect was observed for TLS events: the additional loss in HDGR is compensated by an increase in DCL, not by a further increase in TLS. This is probably due to the fact that in the *recF* strain, the level of TLS already reaches that of a SOS induced strain (Suppl Fig 2) and cannot increase further. Likewise, no significant increase in Pol IV dependent TLS was observed when we deleted both *recF* and *recJ* genes in the presence of a BaP lesion (Suppl Fig 2).

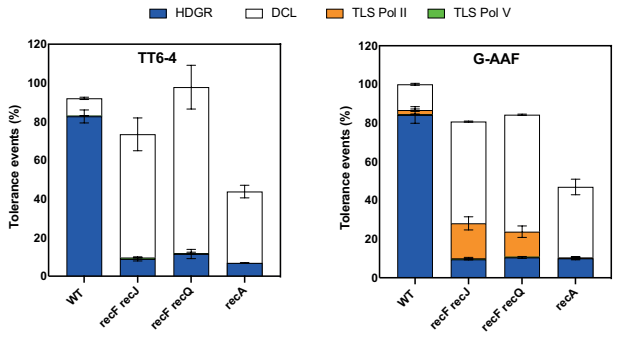


Figure 3. Synergistic effect of *recF* and *recJ* results in loss of HDGR mechanism.
The graphs represent the partition of lesion tolerance pathways (i.e. HDGR, TLS and DCL) for the double mutant *recF recJ* and *recF recQ* in the presence of a TT6-4 lesion (left panel) or a G-AAF lesion (right panel). All the strains are compared to our parental strain (WT in the graphs) which is *uvrA* and mutS deficient. The data for the *recA* deficient strain have been already published (6).

A similar phenotype, *i.e.* a strong loss of HDGR, was previously observed in a *recA* deficient strain (see Fig 3 and also (6)). However, in the absence of RecA, we do not observe an increase in TLS. This might be explained by the fact that in the absence of RecF, RecA loading can still occur (but less efficiently) and this is sufficient to help stabilizing the ssDNA gap favoring replication by Pol II TLS. Similarly, a better stability of the ssDNA gap could explain why we did not observe a drastic decrease in cell survival for the double mutant *recF recJ* as in a *recA* deficient strain.

Altogether these results indicate that both the loading of RecA and the size of the ssDNA gap are equally important for an efficient HDGR. Indeed, the simultaneous deficiency of these two processes condemns cells to survive mostly on damaged chromatid loss.

DISCUSSION

In this study we investigated the role of the presynaptic phase of the RecF pathway on lesion bypass. The presynaptic phase is involved in the processing of the ssDNA gap formed upon encountering a lesion, and ends with the formation of a RecA nucleofilament. Once formed, the nucleofilament will engage in homology search and strand pairing, resulting in a D-loop structure that promotes HDGR and prevents TLS. The use of our single lesion assay allows to finely explore what occurs in the presence of a single replication blocking lesion and to better decipher lesion bypass mechanisms.

The presynaptic phase of the RecF pathway comprises several proteins that were expected to belong to the same epistasis group. While we proved this for the *recF-recO-recR* genes as well as for *recQ-recJ*, we unveiled a synergistic effect between *recF* and *recJ* (and *recF-recQ*), that results in a drastic loss of HDGR mechanism reminiscent of a *recA* deficient phenotype. We also showed that a defect in RecA loading and nucleofilament formation, as in a *recF* strain, or a defect in ssDNA gap extension, as in *recJ* strain, not only strongly affects HDGR mechanism but also promotes replication by Pol II or Pol IV TLS in the absence of SOS induction (*i.e.* without increasing the amount of TLS polymerases).

In non-stressed conditions, tolerance of highly blocking lesions relies almost exclusively on HDGR while TLS is a minor pathway (6). We have already shown that several factors contribute to regulate the balance between HDGR and TLS. In bacteria, TLS is positively regulated by the SOS response that increases the amount of the TLS polymerases in the cell, but TLS can also be modulated by HDGR efficiency: RecA mutants impaired in D-loop formation show a defect in HDGR and an increase in TLS (31). Similarly, the proximity of lesions on opposite strands generates overlapping ssDNA regions that prevent HDGR, leading again to an increase in TLS (20). In this work, we show that it is specifically the presynaptic phase of the RecF pathway that is critical for lesion tolerance pathway choice. Indeed, impairing the presynaptic phase will lead to a delay in the formation of an efficient D-loop and therefore increases the time window allowed for TLS to occur, leading to a significant increase in TLS.

Based on these data, we propose a model that illustrates how the choice of lesion tolerance pathways is directed during the presynaptic phase of the RecF pathway (Fig 4). In the presence of an unrepaired lesion, the replicative polymerase temporarily stalls before a repriming event allows replication to continue, leaving a ssDNA gap. SSB protein will immediately cover the ssDNA gap to protect it and to orchestrate the recruitment of several proteins. SSB was indeed shown to directly interact with most of the proteins of the presynaptic phase as well as with the TLS polymerases (32, 33). The SSB-coated ssDNA gap can be regarded as a platform for protein recruitment for the processing of the ssDNA gap and for lesion bypass. The 5' ds-ssDNA junction of the ssDNA gap

appears to be a shared substrate for both RecQJ and RecFR complexes. Whether and which factors regulate the timing of RecQJ and RecFR recruitment is not known yet, however, we suggest that RecQ and RecJ are the first to be recruited at the 5' junction of the nascent DNA. We based this hypothesis on the fact that RecQ and RecJ were shown to directly interact with SSB in the presence of a ssDNA gap and this interaction stimulates their catalytic activity to enlarge the ssDNA gap (12, 34–36). Besides, our data demonstrate that the processing of the ssDNA gap is equally important as the loading of RecA for an efficient HDGR mechanism. Finally, since RecA inhibits RecJ activity (37), it suggests that the extension of the gap must come before the loading of RecA. We have previously estimated the size of the ssDNA gaps to be in the range of 1.8-3.5 Kb (20) which correlates with the processivity established *in vitro* for RecJ (35). Once the complex RecQJ falls off, RecFR complex can have access to the 5' ds-ssDNA junction and this also allows the recruitment of the RecOR complex to the ssDNA gap where RecO directly interacts with SSB (38, 39). Finally, RecA loading and filament formation can start. Since it occurs preferentially in the direction 5'->3' (40), it will inhibit a new RecQ-RecJ recruitment. Once the RecA nucleofilament on the ssDNA gap is formed it will engage in homology search and strand invasion, forming the so-called D-loop structure required for the HDGR mechanism, that we showed to be the major lesion tolerance pathway employed by cells under non-stressed conditions (6, 7).

Our data show that defect in the genes of the presynaptic phase affects HDGR mechanism favoring TLS pathway. Indeed, in the absence of either RecJ or RecQ, the size of ssDNA gap in not enlarged, thus impairing the efficiency of HDGR mechanism. Likewise, deletion of the RecA mediators (*recF, recO* or *recR*) causes a significant delay in RecA nucleofilament formation, as shown *in vitro* where nucleation time goes from 2-10 min in the presence of RecFOR to 10-60 min in its absence (26). In both cases, we observed an increase in TLS pathways. Absence of RecF, other than causing a delay in RecA filament formation, could allow RecJ to be recruited again and further extend the gap, a too long ssDNA gap can also affect HDGR mechanism. Indeed, in *recF* strain it was observed a RecJ dependent degradation of the nascent DNA that caused a delay in replication resumption (41, 42). Finally, the concomitant deficiency of RecF and RecJ results in a drastic decrease of HDGR mechanism, indicating that the slow formation of a small RecA-covered ssDNA strongly inhibits HDGR.

From our model, we can evince that the presynaptic phase, that includes the different steps to achieve the formation of the RecA nucleofilament, is crucial in modulating lesion tolerance pathways and in turn the outcome of lesion bypass (non-mutagenic versus mutagenic). The presynaptic phase can be considered as the "time window" where TLS can take place. Once the D-loop is formed, the recombinogenic RecA filament engages in HDGR process, thus closing the window that was allowed for TLS. Any delay in the presynaptic phase will extend this time window leading to an increase in TLS.

ACKNOWLEDGEMENT

We thank Jean-Hugues Guervilly and Florencia Villafañez for critical reading of the manuscript.

FUNDING

This work was supported by the Agence Nationale de la Recherche (ANR) Grant [GenoBlock ANR-14-CE09-0010-01]; and the Fondation pour la Recherche Médicale Equipe FRM-EQU201903007797.

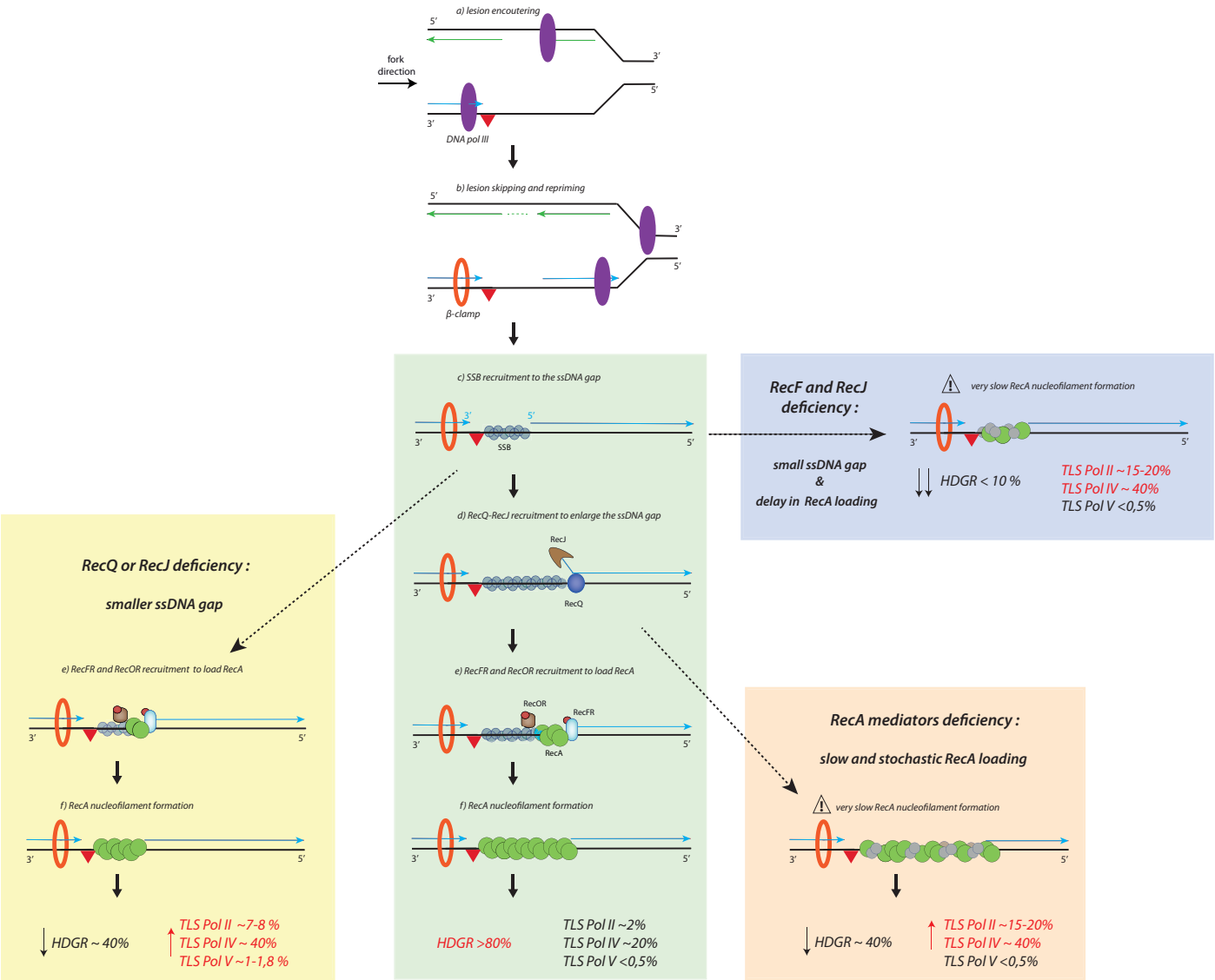


Figure 4. The presynaptic phase of the RecF pathway

Here is represented a scheme of the presynaptic phase during lesion bypass in the presence of a replication-blocking lesion (red triangle) in the leading strand. For simplicity after the repriming event (b), we have drawn only the molecular mechanisms occurring on the ssDNA gap.

During replication, unrepaired lesions can temporarily block the replicative polymerase (a). The polymerase can skip the lesion and restart replication downstream, leaving a ssDNA gap (b). SSB protein is immediately recruited to the ssDNA gap (c) to protect it from undesired degradation, but also to orchestrate the recruitment of the proteins involved in the processing and the filling in of the ssDNA gap. The helicase RecQ and the nuclease RecJ are loaded at the 5' ds-ssDNA junction and their concerted action enlarge the size of the gap (d). The processivity of RecJ-RecQ drops and they are replaced by the RecFR complex that is also recruited at the 5' junction (e). The action of RecFR and of the RecOR complex at the ssDNA gap is to dismantle SSB and load the recombinase RecA in the 5'->3' direction. The presynaptic phase ends with the formation of the RecA nucleofilament along the ssDNA gap. At this point the synaptic phase can take place, leading to homology search and strand exchange that forms the well-known D-loop structure, leading to HDGR mechanisms.

In the absence of RecQ and in particular RecJ (yellow panel), the size of the ssDNA gap is not enlarged and the efficiency of homologous recombination is significantly affected, in favor of the TLS polymerases. In the absence of the RecF, RecO or RecR (orange panel), RecA loading occurs randomly and the kinetics of filament formation is slower, resulting in a decrease in HDGR mechanism and in an increase in Pol II and Pol IV TLS. The combination of RecF and RecJ deficiency (blue panel) has a drastic effect on HDGR similar to the one observed in the recA deficient strain, since the ssDNA gap is small and the loading of RecA is strongly delayed.

REFERENCES

1. Marians,K.J. (2018) Lesion bypass and the reactivation of stalled replication forks *Annual review of biochemistry* **87**, 217-238.
2. Napolitano,R., Janel-Bintz,R., Wagner,J. and Fuchs,R.P. (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *The EMBO journal* **19**, 6259-6265.
3. Sale,J.E. (2013) Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harbor perspectives in biology* **5**, a012708.
4. Henrikus,S.S., van Oijen,A.M. and Robinson,A. (2018) Specialised DNA polymerases in Escherichia coli: roles within multiple pathways *Current genetics* **64**, 1189-1196.
5. Prado,F. (2014) Homologous recombination maintenance of genome integrity during DNA damage tolerance *Molecular & Cellular Oncology* **1**, e957039.

6. Laureti,L., Demol,J., Fuchs,R.P. and Pagès,V. (2015) Bacterial proliferation: keep dividing and don't mind the gap *PLoS genetics* **11**, e1005757.
7. Pagès,V., Mazón,G., Naiman,K., Philippin,G. and Fuchs,R.P. (2012) Monitoring bypass of single replication-blocking lesions by damage avoidance in the Escherichia coli chromosome *Nucleic acids research* **40**, 9036-9043.
8. Spies,M. and Kowalczykowski,S.C. (2005) Homologous recombination by RecBCD and RecF pathways *The bacterial chromosome* 389-403.
9. Kuzminov,A. (1999) Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiology and molecular biology reviews : MMBR* **63**, 751-813.
10. Handa,N., Morimatsu,K., Lovett,S.T. and Kowalczykowski,S.C. (2009) Reconstitution of initial steps of dsDNA break repair by the RecF pathway of E. coli *Genes & development* **23**, 1234-1245.
11. Morimatsu,K., Wu,Y. and Kowalczykowski,S.C. (2012) RecFOR proteins target RecA protein to a DNA gap with either DNA or RNA at the 5' terminus: implication for repair of stalled replication forks. *The Journal of biological chemistry* **287**, 35621-35630.
12. Morimatsu,K. and Kowalczykowski,S.C. (2014) RecQ helicase and RecJ nuclease provide complementary functions to resect DNA for homologous recombination. *Proceedings of the National Academy of Sciences* **111**, E5133-42.
13. Sandler,S.J. (2001) Rec FOR Protein *e LS*
14. Morimatsu,K. and Kowalczykowski,S.C. (2003) RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Molecular Cell* **11**, 1337-1347.
15. Umezu,K., Chi,N.W. and Kolodner,R.D. (1993) Biochemical interaction of the Escherichia coli RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein. *Proc Natl Acad Sci U S A* **90**, 3875-3879.
16. Sakai,A. and Cox,M.M. (2009) RecFOR and RecOR as distinct RecA loading pathways. *The Journal of biological chemistry* **284**, 3264-3272.
17. Hegde,S.P., Rajagopalan,M. and Madiraju,M.V. (1996) Preferential binding of Escherichia coli RecF protein to gapped DNA in the presence of adenosine (gamma-thio) triphosphate. *J Bacteriol* **178**, 184-190.
18. Henrikus,S.S., Henry,C., Ghodke,H., Wood,E.A., Mbele,N., Saxena,R., Basu,U., van Oijen,A.M., Cox,M.M. and Robinson,A. (2019) RecFOR epistasis group: RecF and RecO have distinct localizations and functions in Escherichia coli *Nucleic acids research* **47**, 2946-2965.
19. Courcelle,J. and Hanawalt,P.C. (1999) RecQ and RecJ process blocked replication forks prior to the resumption of replication in UV-irradiated Escherichia coli *Molecular and General Genetics MGG* **262**, 543-551.
20. Chrabaszcz,É., Laureti,L. and Pagès,V. (2018) DNA lesions proximity modulates damage tolerance pathways in Escherichia coli *Nucleic acids research* **46**, 4004-4012.
21. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645.
22. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006.0008.
23. Laureti,L., Lee,L., Philippin,G. and Pagès,V. (2017) A non-catalytic role of RecBCD in homology directed gap repair and translesion synthesis. *Nucleic Acids Research* **45**, 5877-5886.
24. Pagès,V. and Fuchs,R.P. (2018) Inserting Site-Specific DNA Lesions into Whole Genomes. *Methods Mol Biol* **1672**, 107-118.
25. Pagès,V. and Fuchs,R.P.P. (2002) How DNA lesions are turned into mutations within cells? *Oncogene* **21**, 8957-8966.
26. Bell,J.C. and Kowalczykowski,S.C. (2016) RecA: Regulation and Mechanism of a Molecular Search Engine. *Trends in Biochemical Sciences* **41**, 491-507.
27. Wu,H.Y., Lu,C.H. and Li,H.W. (2017) RecA-SSB interaction modulates RecA nucleoprotein filament formation on SSB-wrapped DNA *Scientific reports* 1-13.
28. Sommer,S., Boudsocq,F., Devoret,R. and Bailone,A. (1998) Specific RecA amino acid changes affect RecA-UmuD'C interaction. *Mol Microbiol* **28**, 281-291.
29. Fujii,S., Isogawa,A. and Fuchs,R.P. (2006) RecFOR proteins are essential for Pol V-mediated translesion synthesis and mutagenesis. *The EMBO journal* **25**, 5754-5763.
30. Raychaudhury,P. and Marians,K.J. (2019) The recombination mediator proteins RecFOR maintain RecA* levels for maximal DNA polymerase V Mut activity. *J Biol Chem* **294**, 852-860.
31. Naiman,K., Pages,V. and Fuchs,R.P. (2016) A defect in homologous recombination leads to increased translesion synthesis in E. coli *Nucleic acids research* **44**, 7691-7699.
32. Shereda,R.D., Kozlov,A.G., Lohman,T.M., Cox,M.M. and Keck,J.L. (2008) SSB as an organizer/mobilizer of genome maintenance complexes *Critical reviews in biochemistry and molecular biology* **43**, 289-318.
33. Bianco,P.R. (2017) The tale of SSB. *Prog Biophys Mol Biol* **127**, 111-118.
34. Cheng,K., Xu,H., Chen,X., Wang,L., Tian,B., Zhao,Y. and Hua,Y. (2016) Structural basis for DNA 5-end resection by RecJ *Elife* **5**, e14294.
35. Han,E.S. (2006) RecJ exonuclease: substrates, products and interaction with SSB *Nucleic Acids Research* **34**, 1084-1091.
36. Shereda,R.D., Bernstein,D.A. and Keck,J.L. (2007) A central role for SSB in Escherichia coli RecQ DNA helicase

function *Journal of Biological Chemistry* **282**, 19247-19258.
37. Corrette-Bennett,S.E. and Lovett,S.T. (1995) Enhancement of RecA strand-transfer activity by the RecJ exonuclease of Escherichia coli *The Journal of biological chemistry* 1-5.
38. Umezu,K. and Kolodner,R.D. (1994) Protein interactions in genetic recombination in Escherichia coli. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J Biol Chem* **269**, 30005-30013.
39. Ryzhikov,M., Koroleva,O., Postnov,D., Tran,A. and Korolev,S. (2011) Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein *Nucleic acids research* **39**, 6305-6314.
40. Bianco,P.R. and Kowalczykowski,S.C. (1998) RecA Protein - eLS - Bianco - Wiley Online Library *eLS*
41. Courcelle,J., Carswell-Crumpton,C. and Hanawalt,P.C. (1997) recF and recR are required for the resumption of replication at DNA replication forks in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 3714-3719.
42. Chow,K.H. and Courcelle,J. (2004) RecO acts with RecF and RecR to protect and maintain replication forks blocked by UV-induced DNA damage in Escherichia coli. *J Biol Chem* **279**, 3492-3496.

Extending the gap and loading RecA: the presynaptic phase plays a pivotal role in modulating lesion tolerance pathways

Luisa Laureti[1,*], Lara Lee[1], Gaëlle Philippin[1], Michel Kahi[1], Vincent Pagès[1,*]

___

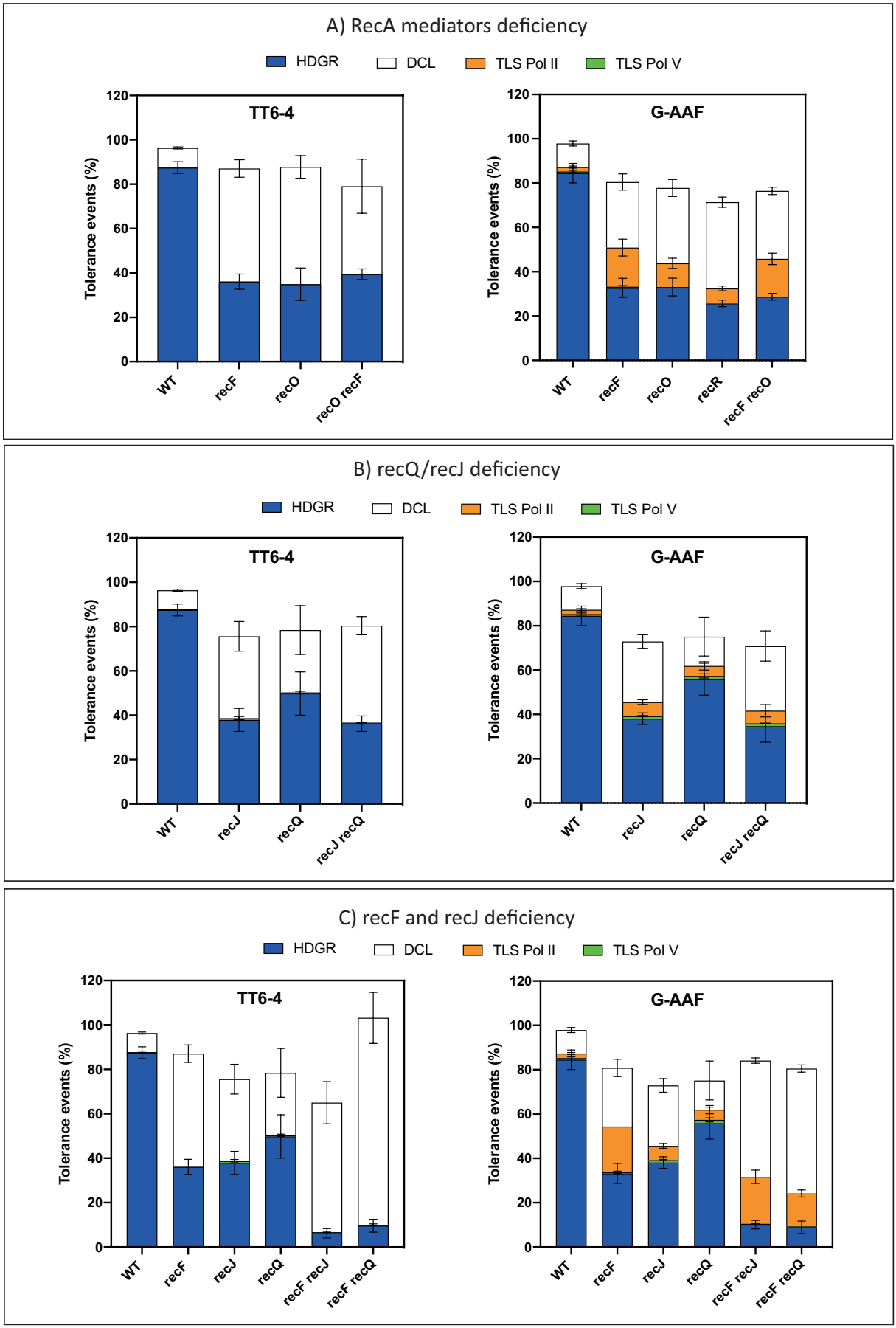
Supplementary data

Supplementary Table 1 : *E. coli* strains used in this study

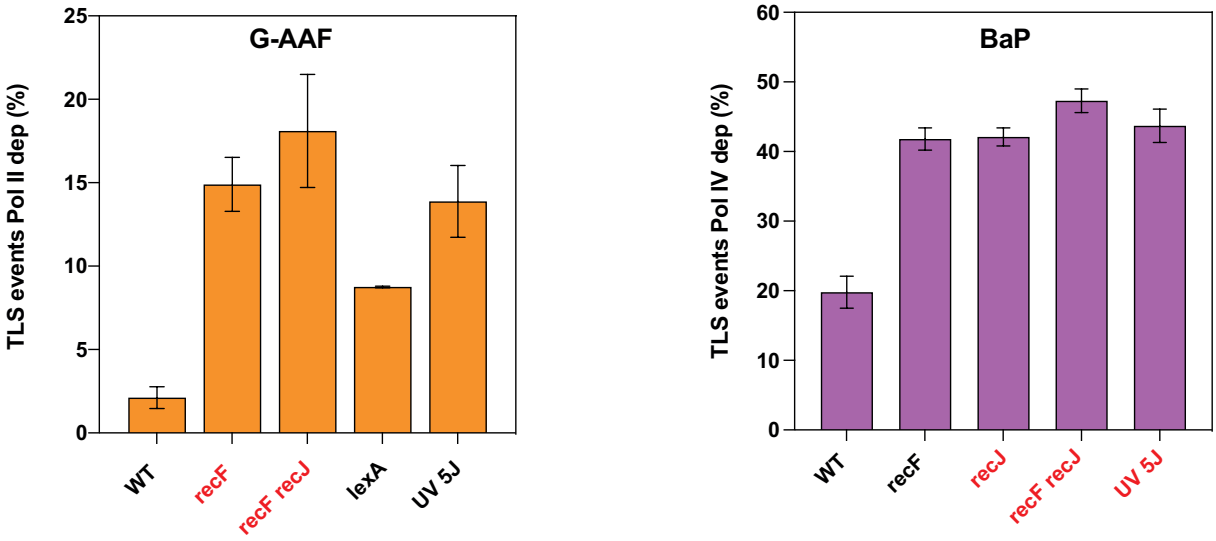
Name	Short name	Genotype
EVP22/23	Parental strain	FBG151/152 *uvrA*::frt *mutS*::frt
EVP146/147	*recO*	FBG151/152 *uvrA*::frt *mutS*::frt *recO::*frt
EVP195/196	*recQ*	FBG151/152 *uvrA*::frt *mutS*::frt *recQ* ::frt
EVP602/603	*recO recF*	FBG151/152 *uvrA*::frt *mutS*::frt *recO::*frt *recF::*Kan
EVP347/348	*recJ*	FBG151/152 *uvrA*::frt *mutS*::frt *recJ::frt*
EVP498/499	*phrB recF*	FBG151/152 *uvrA*::frt *mutS*::frt *phrB* ::frt *recF* ::frt
EVP588/589	*recJ recF*	FBG151/152 *uvrA*::frt *mutS*::frt *recJ*::frt *recF* ::frt
EVP615/616	*recQ recF*	FBG151/152 *uvrA*::frt *mutS*::frt *recQ* ::frt *recF* ::frt
EVP963/964	*recQ recJ*	FBG151/152 *uvrA*::frt *mutS*::frt *recQ* ::frt *recJ::frt*
EVP993/994	*recR*	FBG151/152 *uvrA*::frt *mutS*::frt *recR*::frt
EVP989/990	*phrB recF recO*	FBG151/152 *uvrA*::frt *mutS*::frt *phrB* ::frt *recF* ::frt *recO* ::frt
EC13/14	Parental strain	EC1/EC2 *uvrA*::frt *mutS*::frt
EC39/40	*recO*	EC1/EC2 *uvrA*::frt *recO::*frt *mutS*::frt
EC49/50	*recO recF*	EC1/EC2 *uvrA*::frt *recO::*frt *mutS*::frt *recF* ::frt
EC51/52	*recF*	EC1/EC2 *uvrA*::frt *recF::*frt *mutS*::frt

FBG151 = MG1655 ΔlacIZ ΔattBλ::attRλ-3'lacZ-aadA (lesion inserted in the lagging orientation compared to replication)
FBG152 = MG1655 ΔlacIZ ΔattBλ::3'lacZ-attRλ-aadA (lesion inserted in the leading orientation compared to replication)
EC1 = MG1655 ΔlacI-lacZ::frt ΔattBλ::attRλ-3'lacZ-aad intergenic(aqpZ-ybjD)::high pTRC λ int-xis lacIq
EC2 = MG1655 ΔlacI-lacZ::frt ΔattBλ::3'lacZ-attRλ-aad intergenic(aqpZ-ybjD)::high pTRC λ int-xis lacIq

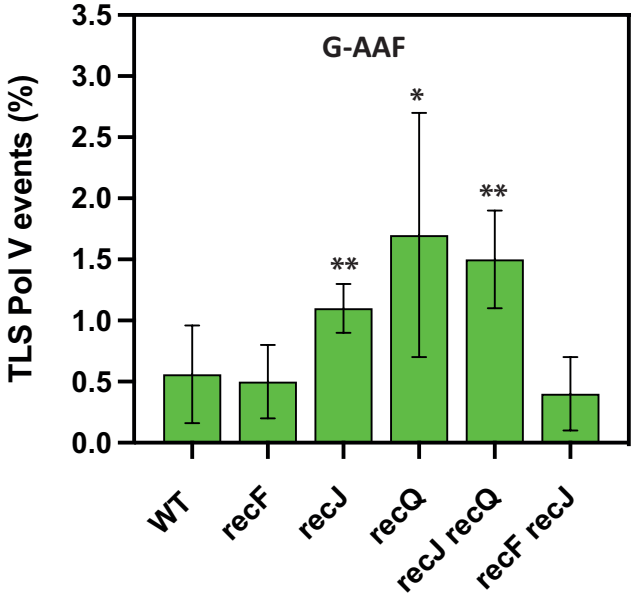
Supplementary Figure 1



Supplementary Figure 1: The graphs represent the partition of lesion tolerance pathway (*i.e.* HDGR, TLS and DCL) in the presence of a TT6-4 lesion (on the left side) or a G-AAF lesion (on the right side) inserted in the lagging orientation. In panel **A)** are represented the data for the single mutants *recF, recO, recR* and the double mutant *recF recO*. In panel **B)** are represented the data for the single mutants *recJ, recQ,* and the double mutant *recJ recQ*. In panel **C)** are represented the data for the double mutants *recF recJ* and *recF recQ*. All the mutant strains are compared to our parental strain (WT in the graphs) which is *uvrA* and *mutS* deficient.



Supplementary Figure 2: The graphs represent the percentage of Pol II dependent TLS events in the presence of a G-AAF lesion (left panel) or the percentage of Pol IV dependent TLS events in the presence of a BaP lesion (right panel). All the mutant strains are compared to our parental strain (WT in the graphs) which is *uvrA* and *mutS* deficient. SOS induction is obtained either by deleting the *lexA* repressor gene or by UV induction (using 5J). The data in red have been obtained in this study, while the data in black have been previously published in (Naiman et al., 2014; Naiman et al., 2016).



Supplementary Figure 3: The graph represents the percentage of Pol V dependent TLS events in the presence of a G-AAF lesion inserted either in the leading or in the lagging orientation (we pulled together the data since no difference was observed). There is a slight but statistically significant increase for the single mutants *recJ* and *recQ* as well for the double mutant *recJ recQ*. T-test was performed to compare values from the mutant strain to our parental strain (WT). *P < 0.05; **P < 0.005.