

SOS factors involved in translesion synthesis

(replication-blocking lesions/2-acetylaminofluorene adducts/slippage mutagenesis/frameshift mutation hot spots/UmuDC)

RITA L. NAPOLITANO*, IAIN B. LAMBERT†, AND ROBERT P. P. FUCHS*

*Unité Propre de Recherche 9003 du Centre National de la Recherche Scientifique, Cancérogénèse et Mutagenèse Moléculaire et Structurale, École Supérieure de Biotechnologie de Strasbourg, Blvd S. Brant, 67400 Strasbourg, France; and †Biology Department, Carleton University, 1125 Colonel By Drive, Ottawa, Canada, K1S5B6

Communicated by James A. Miller, University of Wisconsin, Madison, WI, March 26, 1997 (received for review December 8, 1996)

ABSTRACT Mutations are permanent DNA sequence changes that can be induced when replication occurs on a damaged DNA template. In *Escherichia coli*, the process of translesion synthesis past a lesion that hinders replication requires the induction of SOS-controlled gene products, among which are those of the *umuDC* operon. To study translesion synthesis *in vivo*, we have constructed single-stranded vectors containing single 2-acetylaminofluorene adducts located within -1 and -2 frameshift mutation hot spots formed by short repetitive sequences. These adducts strongly hinder DNA replication as only 2–5% of the molecules give rise to progeny under non-SOS-induced conditions. Induction of the SOS response lead to a 10-fold increase in survival. Adducts present within repetitive sequences trigger the formation of misaligned primer/template replication intermediates which, upon elongation, will result in the fixation of frameshift errors (mutagenic translesion synthesis). Surprisingly we find that elongation from the nonslipped intermediate depends upon functional *umuDC*⁺ gene products, whereas elongation from the slipped intermediate is *umuDC*⁺ independent but requires another, as yet biochemically uncharacterized, SOS function. These data are discussed in terms of the different steps involved during translesion synthesis through a replication-blocking lesion.

Organisms have developed various strategies to protect their genomes from the damaging effects of endogenous and exogenous agents. Despite having robust excision repair systems that remove DNA lesions before replication, cells also possess efficient strategies for tolerating lesions persisting at the replication fork during DNA synthesis (1). Two basic strategies of lesion tolerance can be distinguished: (i) translesion synthesis (TLS) is a process during which the replication machinery reads through the lesion with an associated risk of fixing a mutation. If the lesion does not impede DNA synthesis, an unmodified replication complex can achieve TLS. However, a replication complex modified by accessory proteins encoded by genes associated with the SOS regulon is required for TLS of lesions that hinder the progression of replication; and (ii) damage avoidance designates a general strategy that facilitates replication of damaged DNA templates without the need for the polymerase to read through the lesion. This strategy takes advantage of the information contained in the complementary strand. Two models of damage avoidance have been proposed (1–3): (i) postreplication recombinational repair, in which the DNA polymerase, blocked at a lesion site, dissociates from the DNA and reinitiates replication downstream from the lesion, leaving a gap that is repaired by a recombination mechanism

involving the sister chromatid; and (ii) polymerase strand switching, in which the DNA polymerase switches temporarily from the damaged parental template to the undamaged newly synthesized strand of the sister chromatid before returning to the parental template downstream from the lesion. Both damage avoidance models are believed to be efficient and error-free. The current level of understanding of the molecular mechanisms underlying these processes is rudimentary.

To study the processes of damage avoidance, we recently have developed an approach involving double-stranded DNA constructions with a genetic strand marker located across from a single 2-acetylaminofluorene guanine adduct (dG-AAF) (4). These adducts were previously shown, using primer-template extension reactions, to be strong blocks for various DNA polymerases *in vitro*. (5–7). However, double-stranded plasmids with single adducts exhibit the same colony-forming efficiency as nonmodified control molecules when propagated *in vivo*. The presence of a strand marker across from the adduct site allowed us to measure the fraction of colonies that effectively result from a TLS event. In fact, in the absence of SOS induction, only about 0.5% of the colonies were found to result from TLS events. The induction of the SOS response increased the proportion of TLS events to about 12%. Under both SOS⁻ and SOS⁺ conditions the large majority of colonies contained only plasmids with the marker of the nonadducted strand, suggesting that the initial round of replication of the damaged plasmid molecule in these colonies occurred via a damage avoidance event (4).

The large predominance of damage avoidance events in double-stranded plasmids severely limits the study of the mechanism of TLS *per se*. Single-stranded DNA vectors containing single adducts offer a unique possibility to study TLS events because their single-stranded nature imposes TLS as the only process that can give rise to progeny *in vivo*. It is therefore possible to directly measure the toxicity, and the mutational efficiency and specificity of a given lesion. This approach was pioneered by C. Lawrence and colleagues who, using single-stranded M13 molecules with well defined UV lesions and abasic sites, demonstrated the critical role of the bacterial SOS response in promoting survival and base substitution mutagenesis (8, 9). Base substitution mutagenesis is known to strongly depend upon the expression of the SOS-controlled *umuDC* operon (10–13).

In contrast, we previously have described a frameshift mutation pathway that is SOS-inducible but *umuDC* independent (14–16). Frameshift mutations are efficiently induced by a large variety of chemical carcinogens that present an aromatic moiety. Most frameshift mutations occur within short mononucleotide or dinucleotide repeats and can be explained by slippage events during TLS (14, 17–19). Indeed, adducts formed by aromatic amines strongly stabilize misaligned primer/template replication intermediates [slipped mutagenic in-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/945733-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

Abbreviations: TLS, translesion synthesis; dG-AAF, 2-acetylaminofluorene guanine adduct; SMI, slipped mutagenic intermediate; AAF, 2-acetylaminofluorene.

intermediates (SMI), (20–22)]. These slipped intermediates are in equilibrium with their nonslipped counterparts. Elongation from nonslipped and slipped intermediates will result in error-free and mutagenic TLS, respectively (20, 23). In this paper we have determined the efficiency of elongation from slipped and nonslipped replication intermediates at two different frameshift mutation hot spots in single-stranded DNA. The results show that while elongation from the nonslipped intermediate requires functional *umuDC*, elongation from the SMI is independent of *umuDC* but depends upon another SOS function. These results will be discussed in terms of the accessory factors that are potentially required when the replication machinery performs the different biochemical steps involved in replicating through a blocking lesion.

MATERIALS AND METHODS

Oligonucleotides. Oligonucleotides with single 2-acetylaminofluorene (AAF) adducts located within the run of three Gs or the *NarI* site have been purified and characterized as described previously (20, 24). The oligonucleotides were phosphorylated using T4 polynucleotide kinase and [γ - 32 P]ATP (specific activity: 3,000 Ci/mmol). The kinased oligonucleotides were repurified by polyacrylamide gel electrophoresis (20%).

Construction and Quantification of Single-Adducted Single-Stranded Plasmids. We developed a new strategy to produce high-quality single-stranded plasmids containing a single adduct (ref. 25; Fig. 1). Briefly, double-stranded molecules containing single AAF adducts were produced and purified as described previously (20, 24). The nonadducted strand of these constructs, which contains a few uracil residues, could be selectively degraded *in vitro* thus yielding the single-stranded mono-adducted vectors (25). The single-stranded vectors were quantified, with a PhosphorImager, from the radioactivity present in the corresponding band of a dried agarose gel knowing the specific activity of the 32 P-labeled oligonucleotide (25).

Transformation, Strains, and Mutant Selection. The single-stranded plasmids were used to transform strain JM103 and its Δ *umuDC* derivative by the CaCl_2 technique. When indicated, the SOS response was induced in both strains by UV irradiation at 50 J/m² as previously described (4). After UV irradiation, the survival in strain JM103 and its Δ *umuDC* derivative was 20% and 15%, respectively.

Vectors pUC-3G.ss and pUC-Nar.ss carry a *lacZ*⁻ allele that reverts to *lacZ*⁺ by -1 and -2 frameshift mutagenesis respectively (20, 26). Wild-type and mutant transformants were quantified on indicator plates containing ampicillin, 5-bromo-4-chloro-3-indolyl β -D-galactoside, and isopropyl β -D-thiogalactoside as white and blue colonies, respectively. In addition, -1 and -2 frameshift mutants were identified by direct sequencing or by differential colony hybridization (see below) using the following specific probes: -1 mutants, 5'-AGT CAT ACC CGG ACA TGG AC-3'; and -2 mutants, 5'-TCA TCA CCG GCC ACA GAC TA-5'.

To ascertain that transformants originate from the single-stranded vectors and not from a small amount of contaminating double-stranded DNA we performed colony hybridization experiments making use of a strand marker (*Bss*HII/*Kas*I, Fig. 1) that allows unambiguous identification of any contaminating double-stranded molecules (4, 25).

RESULTS

Nomenclature of the Single-Stranded Vectors. The AAF adduct is located in two sequence contexts, a run of three Gs 5'-CCCG¹G²G³-3', or the *NarI* sequence 5'-G¹G²CG³CC-3', that were previously identified as hotspots for -1 and -2 frameshift mutations, respectively (14, 27). For both sequence

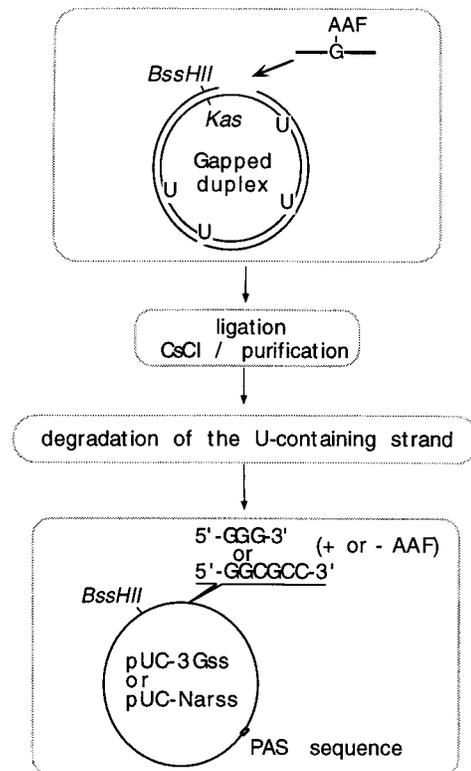


FIG. 1. Strategy for the construction of single-stranded DNA. Gapped-duplex intermediates were produced as previously described (20, 24). The gapped strand of the duplex was provided by a plasmid that was grown in a wild-type strain, whereas the complementary strand contained several uracil residues as the corresponding plasmid was propagated in a *dut*, *ung* strain (CJ 236) (25). Single AAF-adducted oligonucleotides were ligated into the gap, and the covalently closed circular plasmid was purified from CsCl equilibrium sedimentation. The uracil-containing strand was then selectively degraded using an enzymatic mixture containing uracil DNA glycosylase, and the 3' \rightarrow 5' exonuclease of exonucleaseIII and T7 DNA polymerase (25). Constructions containing the G-AAF adduct (G*) within the sequence 5'-GGG*- or 5'-GGCG*CC- were made along with the corresponding control vectors (no adduct).

contexts the AAF adduct is at G³, the most mutagenic position within each hot spot (20, 28). The single-stranded vectors with the run of three Gs were designated pUC-3G.0ss and pUC-3G.3ss for the unmodified control and the G³-AAF modified vector, respectively. Similarly, control and G³-AAF modified vectors at the *NarI* target sequence were called pUC-Nar.0ss and pUC-Nar.3ss, respectively.

Specificity of AAF-Induced Mutagenesis. Previous studies using a forward mutation assay and randomly modified double-stranded DNA have shown that approximately 90% of AAF-induced mutations are frameshifts (14, 27). The specificity of AAF-induced mutations was reinvestigated using the single-stranded vectors. In our constructions the single adduct is located in either a *SmaI* or a *NarI* restriction site. We analyzed plasmid preparations from 112 independent colonies obtained by transformation of a SOS-induced wild-type strain with pUC-3G.3ss. Among the 112 plasmids, 13 were found to be resistant to *SmaI* digestion (*SmaI*^R). Sequence analysis of the *SmaI*^R plasmids revealed a distribution of 12 frameshift (-G) mutations (CCCGGG \rightarrow CCCGG) and one G \rightarrow T targeted transversion (CCCGGG \rightarrow CCCGGT). Therefore, in agreement with our previous results, AAF adducts induce frameshift mutations at a much higher frequency than base substitutions (14, 27). For the sake of simplicity only -1 and -2 frameshift mutations were subsequently scored using the *lacZ'* detection assay described below.

Nonslipped and Slipped TLS *in Vivo*. The efficiency of TLS in a given bacterial cell is defined as the transformation efficiency of the single adducted vector relative to that of the corresponding nonmodified vector. The vectors have been engineered to allow direct determination of the induced frameshift mutation frequencies using the *lacZ'* α -complementation assay. Vectors pUC-3G and pUC-Nar carry a *lacZ'* allele that reverts to *lacZ'* by -1 and -2 frameshift mutations, respectively (20, 26). Using indicator plates containing 5-bromo-4-chloro-3-indolyl β -D-galactoside and isopropyl β -D-thiogalactoside, we define the efficiency of nonslipped and slipped TLS from the number of white and blue colonies, respectively (Table 1). The mutation frequency (i.e., the relative proportion of blue colonies divided by the total number of colonies) is a direct measurement of the polymerase error rate per TLS event *in vivo* because damage avoidance mechanisms cannot operate with single-stranded vectors.

The single-stranded vectors were introduced into both wild-type (JM103) and the corresponding Δ *umuDC* mutant strain using the CaCl_2 transformation procedure. The SOS system of these cells was either noninduced (SOS^-) or was induced by previous irradiation with UV light at a dose of 50 J/m^2 (SOS^+). Experiments were generally repeated at least three times. Average values and standard deviations for nonslipped and slipped TLS are indicated in Table 1. Despite the relatively high toxicity of the AAF adduct under certain physiological conditions (i.e., SOS^- conditions), reliable determinations of nonslipped and slipped TLS could be obtained by scoring relatively large numbers of white and blue colonies as reported in Table 1.

Survival and Polymerase Error Rate at the -1 Frameshift Mutation Hot Spot -GGG-. In these experiments the dG-AAF adduct is located at the 3' end of the run of three Gs (-GGG^{AAAF}-), the most mutagenic position for the induction of -1 frameshift mutations (20, 23). In a given strain, the toxicity conferred by the single adduct was measured by comparing the transformation efficiency of the modified vector (pUC-3G.3ss) with that of the corresponding nonmodified (pUC-3G.0ss) vector. In a wild-type strain, the survival rate increased from 2.4% to 25% when inducing the SOS response by UV irradiation. In contrast, in a Δ *umuDC* strain, the survival rate

remained low, 2.7% and 3.4%, under UV $-$ and UV $+$ conditions, respectively.

In the wild-type strain, the induced -1 frameshift mutation frequency increased from 4% to 12% when inducing the SOS response. Most surprisingly however, UV irradiation caused an increase of the mutation frequency from about 2.6% to 40% in the Δ *umuDC* strain.

Survival and Polymerase Error Rate at the -2 Frameshift Mutation Hot Spot -GGCGCC-. In vector pUC-Nar.3ss, the dG-AAF adduct is located at the third G residue within the *NarI* site -GGCG^{AAAF}CC-. It was previously demonstrated using singly adducted double-stranded vectors that this position represents the only position in the *NarI* site that induces -2 frameshift mutations (26, 28). In a wild-type strain, the survival rate of pUC-Nar.3ss increased from 4.1% to 51% when inducing the SOS response by UV irradiation. Similarly, in a Δ *umuDC* strain, the survival rate also increased from 5% to 57% under UV $-$ and UV $+$ conditions, respectively. In the wild-type strain, the induced -2 frameshift mutation frequency increased from 25% to 67% upon induction of the SOS response. However, UV irradiation of the Δ *umuDC* strain caused a further increase of the mutation frequency from $\approx 55\%$ to 96%. It should be stressed that under these conditions virtually all TLS events are mutagenic.

DISCUSSION

Mechanism of DNA Replication of the Single-Stranded Vectors. The single-stranded plasmids that we have constructed correspond to the template for lagging strand replication of the corresponding pUC double-stranded plasmid. Double-stranded pUC plasmids contain a unidirectional ColE1 origin of replication. Their lagging strand contains a DNA sequence named primosome assembly site (29, 30). This sequence is recognized by protein PriA, which directs the assembly of a ϕ X174-like primosome involving several proteins including DNA primase (DnaG) (31). The primase synthesizes primers that are extended by PolIII holoenzyme.

We believe that the majority of mutations scored in our assay are fixed during the first round of replication. Indeed, after the first round of replication, the adduct will be located in double-stranded DNA, and UvrABC-mediated nucleotide

Table 1. Translesion synthesis within DNA sequences containing AAF adducts

Measurement	Wild type		Δ <i>umuDC</i>	
	SOS^-	SOS^+	SOS^-	SOS^+
TLS at -GGG ^{AAAF} -				
Nonslipped TLS*	2.34 \pm 1.2	22.8 \pm 3.7	2.65 \pm 1.48	2.07 \pm 0.95
(Number of white clones)	(454)	(1,668)	(1,555)	(189)
Slipped TLS [†]	0.10 \pm 0.07	3.08 \pm 0.31	0.072 \pm 0.003	1.37 \pm 0.08
(Number of blue clones)	(20)	(226)	(42)	(125)
Survival	2.44	25.88	2.72	3.44
Mutation frequency, [‡] %	4.1	11.9	2.6	40
TLS at -GGCG ^{AAAF} CC-				
Nonslipped TLS*	3.06 \pm 1.48	16.7 \pm 1.0	2.19 \pm 0.45	2.31 \pm 0.78
(Number of white clones)	(329)	(1,197)	(251)	(39)
Slipped TLS [†]	1.03 \pm 0.34	34.5 \pm 0.84	2.70 \pm 0.78	54.3 \pm 1.2
(Number of blue clones)	(111)	(2,464)	(309)	(913)
Survival	4.09	51.2	4.89	56.6
Mutation frequency, [‡] %	25.2	67.4	55.2	96

Nonmodified vectors form only white colonies. The total number of blue and white colonies that have been scored to obtain the TLS values are indicated in parenthesis. The average value of three independent experiments and the standard deviation are indicated. Survival is the sum of nonslipped and slipped TLS. *Nonslipped TLS is the number of white colonies per ng of modified vector divided by the number of (white) colonies per ng formed by the corresponding nonmodified vector, expressed as a percentage. [†]Slipped TLS is the number of blue colonies per ng of modified vector divided by the number of (white) colonies per ng formed by the corresponding nonmodified vector, expressed as a percentage. [‡]The mutation frequency is the percent of slipped TLS divided by total TLS. It represents the elongation error-rate of the replication machinery per TLS event.

excision repair can efficiently excise the lesion, thus preventing mutation fixation during subsequent rounds of replication. In fact, all mutant colonies that we have analyzed contained the mutant plasmid in a pure form and not as a mixture of mutant and wild-type sequences as would be expected if mutations were formed after the first round of replication (data not shown).

Survival and Polymerase Error Rate During TLS at Single AAF Adducts. In this study, we investigated the effect on TLS of the covalent adduct formed by the strong chemical carcinogen AAF at the C8 position of guanine (dG-AAF). This lesion is known to be a strong block for *in vitro* DNA synthesis reactions (5–7). *In vivo* studies using a forward mutation assay have shown that dG-AAF induces frameshift mutations at high frequency when situated in certain DNA sequence contexts (14, 27). Studies with singly adducted double-stranded constructions have demonstrated that -1 and -2 frameshift mutations are induced at high efficiency when AAF is bound to position G³ within a run of three Gs (5'-G¹G²G³-3') and the *NarI* sequence (5'-G¹G²CG³CC-3'), respectively (20, 23, 28).

Single-stranded vectors offer a unique opportunity to study the mechanisms of translesion synthesis *in vivo* because both the efficiency of TLS (survival rate) and the error rate per TLS event can be examined under various physiological conditions that are known to affect lesion tolerance. In *Escherichia coli*, the induction of the SOS response is known to promote lesion-induced mutagenesis, a key component in this process being the *umuDC* operon (12, 13).

Within both the 5'-GGG^{AAAF}-3' and 5'-GGCG^{AAAF}CC-3' sequence contexts, a single AAF adduct strongly impairs replication of the vector *in vivo* because only $\approx 2\%$ to 5% survival was observed in the absence of SOS induction (Table 1). Because the same values are seen in SOS uninduced wild-type and Δ *umuDC* host cells, this low level of replication seems to reflect the basal level of TLS that can be mediated by the normal replication machinery, and is probably not TLS mediated by a replication complex in the presence of constitutive levels of UmuD/C proteins. Induction of the SOS response in the wild-type strain by UV irradiation caused a greater than 10-fold increase in survival in both sequence contexts (Table 1). Under SOS-induced conditions the survival of the modified vectors in the wild-type strain range between 25% and 50%, illustrating the extensive capacity of the SOS-modified replication complex to perform TLS through otherwise strongly blocking lesions. In a Δ *umuDC* strain, UV irradiation of the host did not result in an increase in survival of vectors in which the AAF adduct was situated within a run of three Gs. In contrast, survival of the pUC-Nar.3ss vector was enhanced by more than 10-fold in the UV-irradiated Δ *umuDC* strain, almost exclusively as a consequence of slipped TLS (Table 1; see below). The survival of single-stranded, M13-derived vectors containing UV lesions has previously been measured in *E. coli*. T-T cyclobutane dimers and T-T pyrimidine-pyrimidone (6–4) lesions exhibit similar survival values under SOS⁻ conditions and undergo a similar increase in survival under SOS-induced conditions as reported here for AAF adducts (8, 9). However, in striking contrast with the results described here for AAF bound to the *NarI* site, no increase in TLS was observed for cyclobutane dimer containing single-stranded DNA in a UV-irradiated Δ *umuDC* strain (32).

Frameshift Mutagenesis Within Repetitive Sequences: Elongation from Nonslipped Versus Elongation from Slipped Replication Intermediates. Frameshift mutations induced by AAF adducts within both runs of Gs and *NarI* sites can be best explained by a slippage mechanism (20, 33). We recently have shown that the *NarI* site belongs to a family of -2 frameshift mutation hot spots whose consensus sequence is 5'-GCGCX-3' (33). The key feature of the consensus sequence is a repetition of the GpC dinucleotide, the nucleotide X on the

3' side determining the relative "hotness" of the sequence. Therefore, both runs of Gs and *NarI* sequence contexts can be viewed as short repetitions of single Gs and GpC dinucleotides, respectively.

Based on experiments with singly adducted plasmids, a two-step model for frameshift mutagenesis has been proposed (20, 23, 33). The first step involves incorporation of cytosine across from the dG-AAF adduct, followed by a slippage step during which the primer template forms one or two "correct" GC base pairs at the primer terminus (Fig. 2). These SMI (20, 21, 33) extrude the adduct in a bulge containing one base for the run of Gs, or two bases for the *NarI* site (Fig. 2). The presence of the AAF adduct within the bulge has been shown by thermal stability measurements, chemical probes, and NMR to strongly stabilize the SMIs (21, 22, 34). Elongation from the SMI will fix the frameshift event (-1 and -2) in the newly synthesized strand. Provided a cytosine residue is incorporated opposite the AAF adduct, TLS within these sequences can be either error-free if it proceeds from the nonslipped intermediate or mutagenic—i.e., leading to a frameshift mutation—if it proceeds from the SMI (Fig. 2). It should be stressed that frameshift mutations do not result from a polymerase misincorporation error as do base pair substitutions, but from a polymerase miselongation step from a slipped primer/template terminus (Fig. 2).

TLS Within Repetitive Sequences: Involvement of SOS Factors. As discussed above, at repetitive sequences TLS can be error-free if it occurs from the nonslipped intermediate, or mutagenic if it proceeds from the SMI. We have previously described an SOS function that is essential for the efficient induction of -2 frameshift mutations within the *NarI* sequence context (15, 16, 35). Although the corresponding gene has not yet been isolated it was tentatively named *npf* (*NarI* processing factor) (16). Analysis of the present data in terms of elongation from slipped versus nonslipped replication intermediates allow us to define more precisely the role of both *umuDC* and *npf* factors during TLS *in vivo*.

Elongation from nonslipped replication intermediates is UmuDC dependent. Error-free elongation is achieved at both repetitive G and *NarI* sequence contexts with the same low efficiency in SOS-uninduced wild-type cells, and in SOS-induced or uninduced Δ *umuDC* cells (Table 1). At both sequence contexts, a low level of about 2.4% of TLS is measured under all these conditions, suggesting that a basal level of nonslipped TLS is achieved by the replication machinery in the absence of UmuD/C⁺ proteins. However, in the wild-type strain, the induction of the SOS response stimulated TLS (by 5- to 10-fold) at both sequence contexts indicating that UmuD/C proteins strongly promote elongation from the nonslipped intermediate. *A posteriori*, it is not surprising that the efficiency of elongation from the nonslipped primer template is sequence context independent in view of the similarities of the primer template structure (Fig. 2). Indeed, both have a dG-AAF:C "mispair" at the terminus.

Elongation from the SMI is UmuDC independent. Mutagenic TLS within the run of three Gs and the *NarI* site result from elongation of SMIs that have one and two "correct" GC base pairs, respectively, at the primer template terminus (Fig. 2). In contrast to nonslipped TLS, mutagenic elongation of both -1 and -2 SMIs appears to be largely UmuDC⁺ independent but is strongly dependent upon another UV-inducible function (Table 1). Indeed, at both sequence contexts, the induction of the SOS response increases the extent of mutagenic TLS by about 25-fold (Table 1) in both the wild-type and Δ *umuDC* cells. Moreover, mutagenic TLS appears to be ≈ 20 -fold more frequent at the *NarI* sequence as compared with the run of three Gs under both SOS⁻ and SOS⁺ conditions (Table 1). This observation suggests that the sequence context in the vicinity of the adduct, which determines the structure and/or

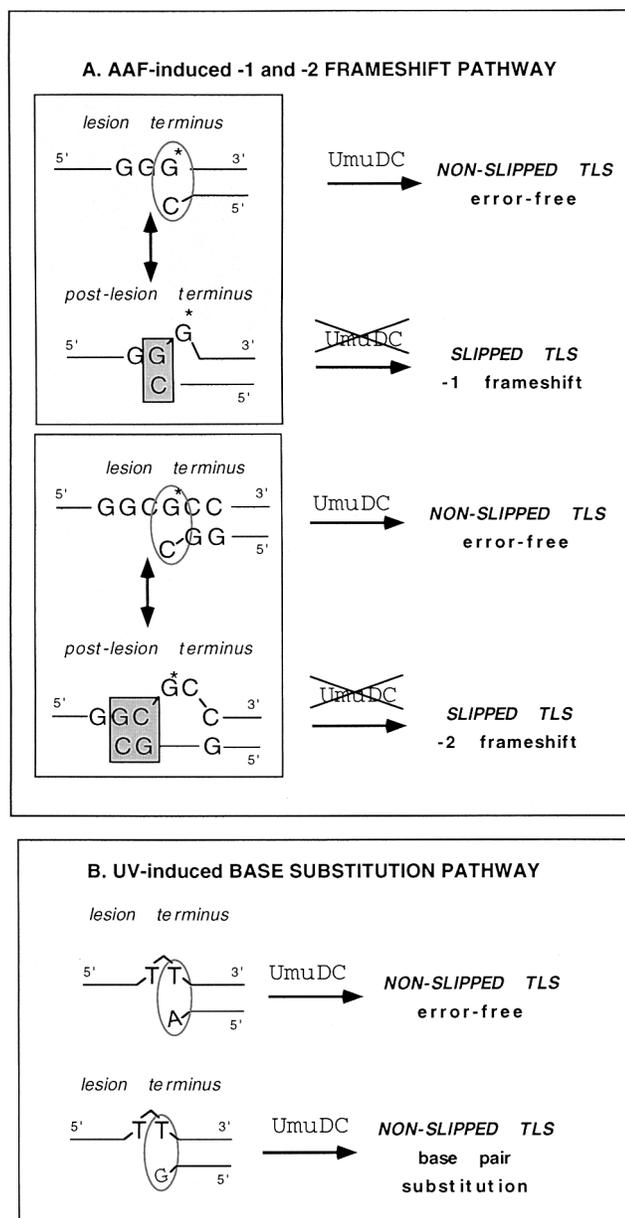


FIG. 2. Structure of replication intermediates and involvement of SOS factors. (A) AAF-induced frameshift mutagenesis. When AAF adducts are situated in repetitive sequence contexts associated with both -1 and -2 frameshifts, the TLS intermediates can assume two forms: a nonslipped intermediate or a slipped intermediate. Elongation from the nonslipped intermediate, i.e., error-free elongation, proceeds from a lesion terminus at which the terminal nucleotide of the nascent strand is situated across from the AAF lesion (circled). Elongation from the lesion terminus appears to be stimulated by UmuDC. In contrast, frameshift mutations result from the elongation of the postlesion terminus of the slipped intermediate, at which the terminal nucleotide of the nascent strand is correctly paired with residues situated 5' to the adduct on the template strand (shown in the square). G* is the AAF adducted guanine residue. (B) UV-induced base substitution mutagenesis. Both error-free and mutagenic TLS across a photoproduct involve elongation from a lesion terminus because, in the intermediate, the terminal nucleotide in the nascent strand is situated across from the lesion (circled). Both elongation events are strongly stimulated by the presence of UmuDC.

the stability of the SMI, modulates the elongation efficiency of the slipped intermediate.

Frameshift Mutagenesis: Comparison Between Double-Stranded and Single-Stranded Vectors. Using double-stranded vectors, we have previously characterized AAF-

induced -1 and -2 frameshift mutagenesis as being *umuDC* dependent and *umuDC* independent, respectively (14-16). In the present study, using single-stranded vectors, -1 frameshift mutagenesis appears to be *umuDC* independent as well. The reason for this apparent discrepancy stems from the fact that when double-stranded vectors are introduced in a $\Delta umuDC$ strain, TLS is completely abolished (i.e., lesion tolerance occurs almost entirely through damage avoidance mechanisms) when the adduct is located within the run of Gs (4) while it is maintained at the same level as in the wild-type strain when the adduct is located at the *NarI* site (T. Broschard and R.P.P.F., unpublished results). Under SOS-induced conditions, we note that all TLS events are less frequent in double-stranded versus single-stranded templates (ref. 4, T. Broschard and R.P.P.F., unpublished results; present work). In the absence of UmuDC, this phenomenon is more pronounced except for -2 mutagenesis. Therefore, it appears that the coordinated synthesis of both leading and lagging strands, a distinct feature of double-stranded DNA replication, favors the dissociation of the polymerase from the lesion site in all cases except from the slipped intermediate at the *NarI* site. Consequently, in contrast to the -2 pathway that is clearly *umuDC* independent, the complete loss of TLS (and the consequent loss of frameshifts) in the -1 pathway makes it appear to be *umuDC* dependent. This example clearly illustrates the fact that (i) mutation frequencies determined with double-stranded DNA can by no means be taken as measurements of DNA polymerase fidelity; and (ii) the genetic control of a mutation pathway determined with double-stranded vectors does not necessarily reflect the genetic control of the TLS pathway itself. Therefore, single-stranded vectors provide an unambiguous analysis of the TLS pathway itself whereas double-stranded vectors enable the study of recovery mechanisms other than TLS. In conclusion, both -1 and -2 frameshift mutagenesis induced as a consequence of AAF-mediated slippage occurs through a SOS-dependent but *umuDC*-independent pathway.

Structure of Replication Intermediates: A Comparison Between UV-Induced Base Substitution and AAF-Induced Frameshift Mutagenesis. AAF adducts and UV-induced pyrimidine dimers (cyclobutane or 6-4 photoproducts) are both strong replication blocking lesions. However, they differ in that UV lesions essentially induce base pair substitutions with little sequence context effect (for a review see ref. 1) whereas AAF adducts induce frameshift mutations with a strong effect of DNA context (14, 27). In addition, UV-light-induced base substitutions are *umuDC* dependent (1, 10) whereas AAF-induced frameshifts are not but depend upon a different SOS-controlled gene (this work, ref. 16). Despite these differences both mutation pathways can be understood in similar terms when considering the steps involved in TLS.

Let us analyze TLS in terms of the structure of the replication intermediates (Fig. 2): a replication intermediate in which the last nucleotide of the primer is located opposite the lesion in the template will be referred to as a "lesion terminus." In contrast when the last nucleotide of the primer is paired with a (nondamaged) nucleotide located 5' from the lesion in the template (such as in the SMIs) we will call it a "postlesion terminus." Elongation from a "lesion terminus" is strongly promoted by UmuDC proteins, whereas elongation from a postlesion terminus is UmuDC independent. For UV lesions, incorporation of either a correct or an incorrect nucleotide across from the lesion will yield lesion termini, the elongation of which are stimulated by UmuDC proteins (32). As a result UV mutagenesis is *umuDC* dependent.

For AAF lesion, cytosine (i.e., the "correct nucleotide") is almost always inserted across from the adduct, thus accounting for the poor capacity of AAF adducts to induce base substitutions (this work, refs. 7 and 14). Elongation from this lesion terminus is, as with the UV-lesions, *umuDC* dependent and is

error-free (Table 1). However, in repetitive sequences the lesion terminus can be converted into a postlesion terminus in the absence of a nucleotide addition step, simply by a slippage event (Fig. 2). Elongation from postlesion termini that are present in the SMI generate frameshift mutations. Although elongation from a postlesion terminus is *umuDC* independent, the whole process of TLS requires another SOS function (Table 1). Therefore, the differences between UV and AAF mutagenesis in (i) mutation specificity (base substitutions vs. frameshifts), (ii) DNA sequence context effects (no strong sequence effect vs. hot spots), and (iii) genetic requirements all can be understood within the same model of TLS.

Replication Intermediates in TLS and Genetic Requirements. Although, the core replication machinery is capable of performing a low basal level of TLS and mutagenesis, proteins encoded by the SOS regulon strongly increase the capacity of the replication machinery to achieve TLS (this work, refs. 4, 36, and 37). For UV lesions and apurinic sites, genetic and biochemical data have shown that the SOS gene products RecA, UmuD', and UmuC facilitate TLS (32, 38–41).

The first step in TLS, i.e., incorporation, appears to be efficiently performed by the polymerase itself as suggested by dNTP turnover measurements when a polymerase idles across from a variety of blocking lesions (42–44). However, to achieve complete TLS requires additional steps to be overcome: proofreading avoidance, stimulation of postlesion elongation, and/or avoidance of polymerase dissociation. It is not clear which of these steps are facilitated by RecA and UmuD/C, respectively (43, 45, 46). These proteins may only be required for elongation of lesion termini. On the other hand, during frameshift mutagenesis, the RecA and UmuD'/UmuC functions appear to be dispensable for the elongation of postlesion termini that are formed by slippage (Fig. 2). Yet an additional SOS-controlled factor was found to be involved in the whole process of TLS. Its precise role in TLS will await its isolation.

We thank S. Turban for excellent technical assistance. R.L.N. was supported by a fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil. This project was supported in part by a North Atlantic Treaty Organization Collaborative Grant. Work in the laboratory of I.B.L. is supported by the National Scientific and Engineering Research Council of Canada.

1. Livneh, Z., Cohen-Fix, O., Skaliter, R. & Elizur, T. (1993) *Crit. Rev. Biochem. Mol. Biol.* **28**, 465–513.
2. Higgins, N. P., Kato, K. & Strauss, B. (1976) *J. Mol. Biol.* **101**, 417–425.
3. Rupp, W. D., Wilde, C. E., III, Reno, D. L. & Howard-Flanders, P. (1971) *J. Mol. Biol.* **61**, 25–44.
4. Koffel-Schwartz, N., Coin, F., Veaute, X. & Fuchs, R. P. P. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 7805–7810.
5. Belguise-Valladier, P., Maki, H., Sekiguchi, M. & Fuchs, R. P. P. (1994) *J. Mol. Biol.* **236**, 151–164.
6. Belguise-Valladier, P. & Fuchs, R. P. P. (1995) *J. Mol. Biol.* **249**, 903–913.
7. Lindsley, J. E. & Fuchs, R. P. P. (1994) *Biochemistry* **33**, 764–772.
8. Banerjee, S. K., Christensen, R. B., Lawrence, C. W. & LeClerc, J. E. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8141–8145.
9. LeClerc, J. E., Borden, A. & Lawrence, C. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9685–9689.
10. Kato, T. & Shinoura, Y. (1977) *Mol. Gen. Genet.* **156**, 121–131.
11. Steinborn, G. (1978) *Mol. Gen. Genet.* **165**, 87–93.
12. Woodgate, R. & Sedgwick, S. (1992) *Mol. Microbiol.* **6**, 2213–2218.
13. Murlí, S. & Walker, G. C. (1993) *Curr. Opin. Genet. Dev.* **3**, 719–725.
14. Koffel-Schwartz, N., Verdier, J. M., Bichara, M., Freund, A. M., Daune, M. P. & Fuchs, R. P. P. (1984) *J. Mol. Biol.* **177**, 33–51.
15. Koffel-Schwartz, N. & Fuchs, R. P. P. (1989) *Mol. Gen. Genet.* **215**, 306–311.
16. Maenhaut-Michel, G., Janel-Bintz, R. & Fuchs, R. P. P. (1992) *Mol. Gen. Genet.* **235**, 373–380.
17. Refolo, L. M., Bennett, C. B. & Humayun, M. Z. (1987) *J. Mol. Biol.* **193**, 609–636.
18. Schaaper, R. M., Koffel-Schwartz, N. & Fuchs, R. P. P. (1990) *Carcinogenesis* **11**, 1087–1095.
19. Lambert, I. B., Gordon, A. J. E., Glickman, B. W. & McCalla, D. R. (1992) *Genetics* **132**, 911–927.
20. Lambert, I. B., Napolitano, R. L. & Fuchs, R. P. P. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1310–1314.
21. Garcia, A., Lambert, I. B. & Fuchs, R. P. P. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5989–5993.
22. Milhe, C., Dhalluin, C., Fuchs, R. P. P. & Lefevre, J.-F. (1994) *Nucleic Acids Res.* **22**, 4646–4652.
23. Napolitano, R. L., Lambert, I. B. & Fuchs, R. P. P. (1994) *Biochemistry* **33**, 1311–1315.
24. Koehl, P., Burnouf, D. & Fuchs, R. P. P. (1989) *J. Mol. Biol.* **207**, 355–364.
25. Napolitano, R. L. & Fuchs, R. P. P. (1997) *Chem. Res. Toxicol.*, in press.
26. Veaute, X. & Fuchs, R. P. P. (1993) *Science* **261**, 598–600.
27. Fuchs, R. P. P., Schwartz, N. & Daune, M. P. (1981) *Nature (London)* **294**, 657–659.
28. Burnouf, D., Koehl, P. & Fuchs, R. P. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4147–4151.
29. Zypursky, S. L. & Marians, K. J. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6521–6525.
30. Zypursky, S. L. & Marians, K. J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6111–6115.
31. Kornberg, A. & Baker, T. (1991) *DNA Replication* (Freeman, New York), 2nd Ed., pp. 284–286.
32. Szekeres, E. S., Woodgate, R. & Lawrence, C. W. (1996) *J. Bacteriol.* **178**, 2559–2563.
33. Koffel-Schwartz, N. & Fuchs, R. P. P. (1995) *J. Mol. Biol.* **252**, 507–513.
34. Milhe, C., Fuchs, R. P. P. & Lefevre, J.-F. (1996) *Eur. J. Biochem.* **235**, 120–127.
35. Burnouf, D. & Fuchs, R. P. P. (1985) *Biochimie* **67**, 385–392.
36. Christensen, J. R., LeClerc, J. E., Valone Tata, P., Christensen, R. B. & Lawrence, C. W. (1988) *J. Mol. Biol.* **203**, 635–641.
37. Tessman, I. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6614–6618.
38. Sommer, S., Bailone, A. & Devoret, R. (1993) *Mol. Microbiol.* **10**, 963–971.
39. Frank, E. G., Hauser, J., Levine, A. S. & Woodgate, R. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8169–8173.
40. Rajagopalan, M., Lu, C., Woodgate, R., O'Donnell, M., Goodman, M. F. & Echols, H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10777–10781.
41. Lawrence, C. L., Borden, A. & Woodgate, R. (1996) *Mol. Gen. Genet.* **251**, 493–498.
42. Villani, G., Boiteux, S. & Radman, M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3037–3041.
43. Shwartz, H., Shavitt, O. & Livneh, Z. (1988) *J. Biol. Chem.* **263**, 18277–18285.
44. Hevroni, D. & Livneh, Z. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5046–5050.
45. Fersht, A. R. & Knill-Jones, J. W. (1983) *J. Mol. Biol.* **165**, 669–682.
46. Bridges, B. A. & Woodgate, R. (1985) *Mutat. Res.* **150**, 133–139.