Relative Probability of Mutagenic Translesion Synthesis on the Leading and Lagging Strands during Replication of UV-Irradiated DNA in a Human Cell Extract

David C. Thomas,‡ Dinh C. Nguyen,‡ Walter W. Piegorsch,§ and Thomas A. Kunkel*,‡

Laboratory of Molecular Genetics and Statistics and Biomathematics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Received July 27, 1993; Revised Manuscript Received August 27, 1993®

ABSTRACT: We have previously demonstrated mutagenic bypass of pyrimidine dimers during SV40 origin-dependent replication of UV-irradiated DNA in human cell extracts [Thomas, D. C., & Kunkel, T. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 7744–7748]. Here we use two vectors having the origin of replication on opposite sides of a *lacZα* reporter gene to examine the relative probability of mutagenic translesion synthesis on the leading and lagging strands. Although replication of both vectors is inhibited by UVB irradiation in a dose-dependent manner, the covalently closed DNA products of replication contain T4 endonuclease sensitive sites, indicating that bypass of cyclobutane pyrimidine dimers occurred. At fluences of 70 and 100 J/m^2, the mutant frequencies obtained with both vectors are substantially higher than with control DNAs. Sequence analysis of mutants obtained with both vectors reveal three types of mutations at frequencies significantly above those obtained from replication of undamaged DNA. These are C → T transitions, accounting for about two-thirds of the mutants, a small number of CC → TT substitutions, and complex mutations. Comparing the distribution of C → T substitutions in the two spectra permits an estimation of the probability of mutagenic translesion replication of the same sequence when replicated as the leading or lagging strand. The data suggest that the overall average UV-independent C → T substitution probability per phenotypically detectable dipyrimidine site is the same during leading and lagging strand replication. However, statistically significant differences are observed when the distribution of C → T substitutions is considered. The substitution probability at some positions is higher when replicated as the lagging strand than when replicated as the leading strand, while at other sites the opposite is observed. Thus the fidelity of leading and lagging strand translesion synthesis varies by position.

DNA replication is an asymmetric process. There are different protein requirements for synthesis of the continuously replicated leading strand and the discontinuously replicated lagging strand (Kornberg & Baker, 1992). Our laboratory is interested in whether such asymmetry might lead to differences in error rates for leading and lagging strand replication. To examine this issue in human cell extracts, we developed a strategy (explained in detail below) using the well-characterized SV40[1] origin-dependent replication system (Kelly, 1988; Stillman, 1989; Hurwitz et al., 1990). Our first applications of this strategy examined the relative rates of base substitution and minus-one nucleotide errors on the two strands during replication of undamaged DNA substrates (Roberts et al., 1991, 1993).

The issue is also relevant to replication of damaged DNA. Our first attempt to examine the fidelity of mutagenic translesion replication has been prompted by observations on ultraviolet light-induced mutations in mammalian cells. Substantial information on this subject has been obtained using shuttle vectors [for review, see Sarasin (1989)]. In excision repair-competent cells, where the transcribed strand of the *hprt* gene is preferentially repaired, most UV light-induced mutations occur on the nontranscribed strand (Vrieling et al., 1989, 1991; McGregor et al., 1991; Menichini et al., 1991). However, when repair-deficient cell lines were tested, most mutations were observed on the transcribed strand (McGregor et al., 1991; Menichini et al., 1991; Vrieling et al., 1989, 1991; Dorado et al., 1991). This bias was suggested to result from a difference in the fidelity of leading and lagging strand synthesis (Menichini et al., 1991; Vrieling et al., 1991). Furthermore, evidence that the transcribed strand of the *hprt* gene may be the leading strand for replication (Sykes et al., 1988) is consistent with the possibility that translesion synthesis could be more error-prone on the leading strand.

Using SV40 origin-dependent replication mutagenesis assays, we (Thomas & Kunkel, 1993) and others (Carty et al., 1993) have recently observed mutagenic synthesis past pyrimidine dimers in UV-irradiated DNA. Here we extend the analysis to examine the relative rates of mutagenic translesion replication of the same nucleotide sequences when copied by leading and lagging strand replication proteins.

EXPERIMENTAL PROCEDURES

Materials. Escherichia coli strains and other materials have been described (Roberts & Kunkel, 1993). Bacteriophage M13mp2SV containing the SV40 origin of replication on either side of the *lacZα* complementation gene (ori left and ori right) are described in Roberts et al. (1991), except that both constructs now contain the wild-type *lacZα* coding sequence encoding a blue-plaque phenotype.

UV Irradiation, Replication Reactions, and Product Analyses. These were performed as described (Thomas & Kunkel, 1993). Some samples were treated with T4 endonuclease V

* To whom correspondence should be sent: phone, (919)-541-2644; Fax, (919)-541-7613.

‡ Laboratory of Molecular Genetics.

§ Statistics and Biomathematics Branch.

® Abstract published in *Advance ACS Abstracts,* October 1, 1993.

[1] Abbreviations: UV, ultraviolet; SV40, simian virus 40; T antigen, SV40 large tumor antigen; ESS, T4 endonuclease V sensitive sites; PCNA, proliferating cell nuclear antigen.

prior to electrophoresis as described (Thomas & Kunkel, 1993), and the number of enzyme-sensitive sites (ESS) was calculated from the zero term of the Poisson distribution.

Determination of Mutant Frequencies and Sequence Analysis. The mutant frequencies for various DNA samples were determined by electroporation of excision repair proficient *E. coli* cells to score forward mutants among total plaques on indicator plates (Roberts et al., 1991). The assay scores errors in the wild-type *lacZα* gene. Correct replication produces DNA that yields dark blue plaques upon transfection. Errors are scored as lighter blue or colorless plaques. Since the assay measures loss of a function (α-complementation of β-galactosidase activity) that is not essential for phage production, mutations at different sites can be recovered and scored (Roberts & Kunkel, 1993). For a description of detectable sites for scoring errors relevant to this study, see the legend to Figure 1.

Statistical Analysis of Mutational Spectra. When represented as numbers of mutations for a given category or genomic location, mutational spectra may be viewed as contingency tables and analyzed by statistical methodology for categorical data (Piegorsch & Bailer, 1993).[2] Usually when two mutational spectra are compared, categorical data tables are sparse in data content; i.e., they contain many zeros and ones. Piegorsch and Bailer (1993) suggest an approximate measure for sparseness that is the number (N) of total mutants divided by the degrees of freedom in the table. For degrees of freedom, denote the number of different mutant categories as R and the number of different spectra for comparison as T (here $T = 2$). The spectra then comprise an $R \times T$ contingency table, and the degrees of freedom are $(R-1)(T-1)$. When $N/\{(R-1)(T-1)\}$ is less than about 10, Piegorsch and Bailer (1993) suggest that the sparseness in the table is extreme. In this case, the hypergeometric test of Skopek and Adams (1987) should be used to compare the spectra. A low P-value from that test indicates that the spectra exhibit significant heterogeneity in their patterns of mutations.

For data tables that are not sparse, one can employ more standard methods, such as the usual chi-square test for $R \times T$ tables or the closely related likelihood ratio test. Piegorsch and Bailer (1993) review the various options in this case. They also note that it is possible with nonsparse tables to identify which mutant positions or sites contribute significantly to any spectral heterogeneity. As presented originally by Gabriel (1966), the method essentially removes rows of data corresponding to sites of interest one or more at a time, until the data no longer exhibit heterogeneity. [However, one must adjust for post-hoc multiplicity; see Piegorsch and Bailer (1993).] The test begins by assessing heterogeneity for the entire table of data using a likelihood ratio test. If the statistic indicates no heterogeneity, the procedure stops. If, however, spectral heterogeneity is observed, the statistic is recomputed after data rows of interest are removed to indicate if the rows contribute significantly to the observed heterogeneity. Application is described in more detail by Piegorsch and Bailer (1993).

RESULTS

Strategy To Determine Mutagenic Translesion Replication Probabilities on the Leading and Lagging Strands. Examining the relative probability of mutagenic translesion replication on the leading and lagging strands requires comparing results with two vectors. Both are 7.4 kilobase pairs and contain the *lacZα*-complementation target sequence, comprising nucleotides −84 to +170, where +1 is the first transcribed nucleotide. They also contain an SV40 origin, but on opposite sides of the *lacZα* gene. For a reference point to describe the distance from the origin to the *lacZα* target, we use the transition point from discontinuous to continuous DNA synthesis (nucleotide 5210 in SV40 DNA; Hay & DePamphilis, 1982). Given the locations of the origins in the M13 vectors, the shortest distance one of the two replication forks must proceed from the origin in order to fully replicate the *lacZα* target is 594 base pairs for the ori left vector and 403 base pairs for the ori right vector. In both vectors, the other fork must synthesize more than 10 times as many base pairs to replicate the mutational target. Previous studies (Edenberg & Huberman, 1975; Li & Kelly, 1985) have indicated that the rate of replication fork movement is similar in both directions from the origin. Thus, with the ori left vector, the (+) viral strand is inferred to be replicated as the lagging strand (Figure 1). For the ori right vector, the (+) strand is inferred to be replicated as the leading strand.

Fidelity measurements with these two vectors thus allow an estimation of the fidelity of replication of the same sequence by either the leading or lagging strand apparatus, so long as the strand on which the error was made can be assigned. Here this is possible because UV irradiation-dependent mutations result from mutagenic bypass of cyclobutane pyrimidine dimers (Thomas & Kunkel, 1993; Carty et al., 1993).

Replication of UV-Irradiated M13mp2SV Ori Left and Ori Right DNA in a HeLa Cell Extract. Ori left and ori right DNAs were simultaneously irradiated with UVB light at fluences ranging from 0 to 140 J/m^2. These were replicated in a HeLa cell extract in the presence of T antigen. Replication was inhibited by UV irradiation in a dose-dependent manner similar to that observed previously (Thomas & Kunkel, 1993). An aliquot of each reaction was digested with restriction endonuclease *Dpn*I either with or without T4 endonuclease V, and the products were resolved on an agarose gel. At the highest UV dose, the amount of covalently closed monomer-length product was diminished to less than 2% of the unirradiated control [data not shown, but see Thomas and Kunkel (1993)]. The covalently closed circular DNA products were increasingly sensitive to the endonuclease with each higher fluence (data not shown), indicating the presence of pyrimidine dimers in the replicated DNA. Results were indistinguishable with both vectors, suggesting that the replication apparatus performed trans-dimer synthesis with similar efficiency on both substrates.

Replication of both Ori Left and Ori Right DNAs Is Mutagenic. The *lacZα* mutant frequencies of control and replicated ori left and ori right DNAs were determined by transfection of an appropriate *E. coli* strain (Table I). For both vectors, unincubated DNA, with or without irradiation, and incubated DNA, with or without irradiation but minus T antigen, served as controls. The values ranged from 2.8×10^{-4} to 6.8×10^{-4} (Table I), similar to previously measured values (Roberts & Kunkel, 1988; Thomas et al., 1991; Thomas & Kunkel, 1993). Thus, UV-irradiated DNA not subject to replication in the extract is not mutagenic upon transfection into excision repair proficient, non-SOS-induced *E. coli* cells. Likewise, the mutant frequencies for undamaged, replicated DNA were not substantially higher than control values, reflecting the high fidelity of the replication apparatus with undamaged DNA. However, mutant frequencies for UV-

[2] Reprints of this article may be obtained from W. Piegorsch (current address: Department of Statistics, University of South Carolina, Columbia, SC 29208).

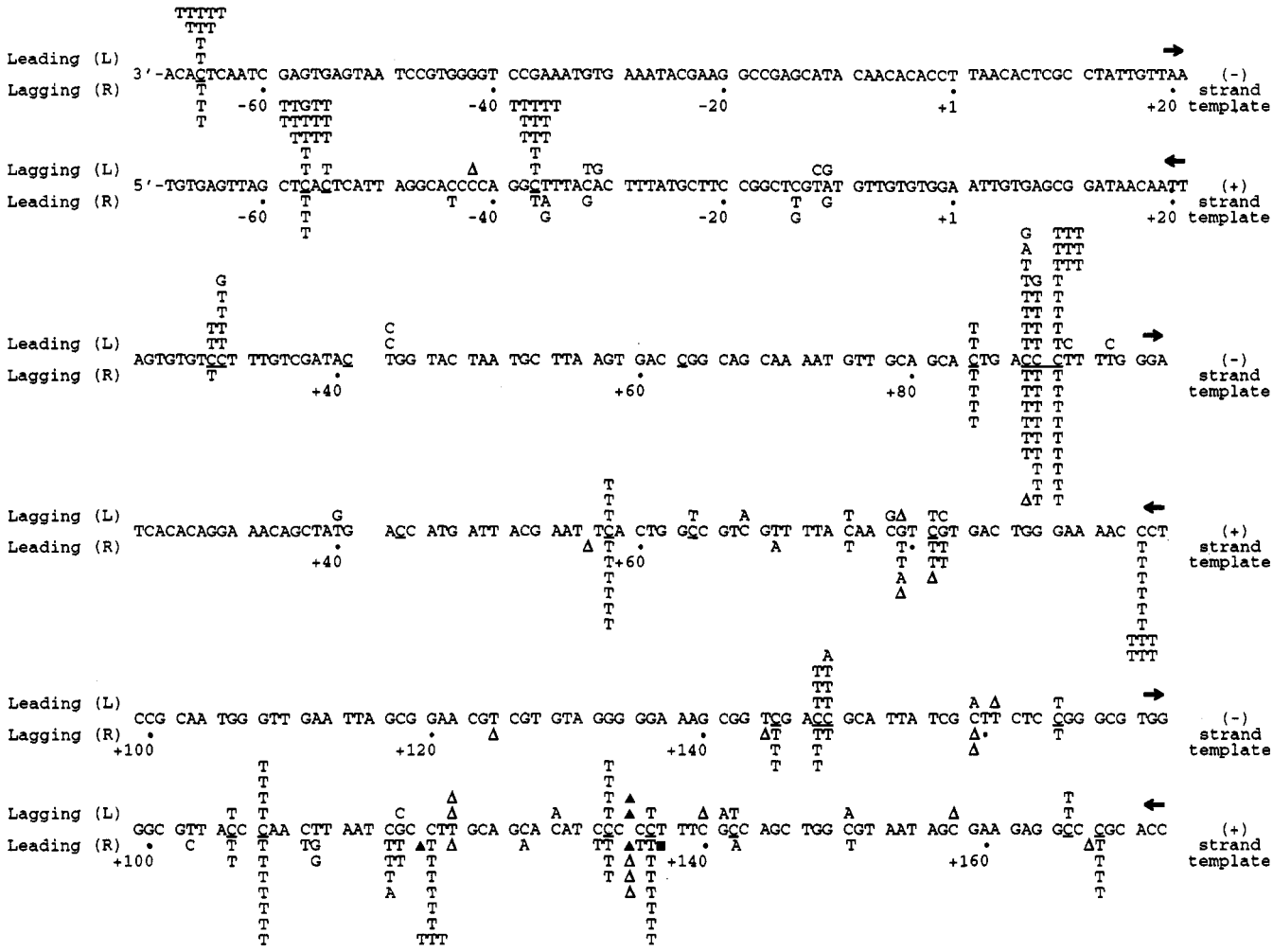

FIGURE 1: Spectra of single-base mutations produced by replication of UV-irradiated M13mpSV ori left and ori right DNA in HeLa cell extracts. Three lines of double-stranded primary wild-type DNA sequence are shown. The direction of synthesis is indicated by arrows. As described in the text, the (–) and (+) strand templates are inferred to be copied as the leading or lagging strand depending on the DNA substrate used, whether ori left (L) or ori right (R). Position +1 is the first transcribed base of the *lacZα* gene. Only single-base mutations are shown. Base substitutions are shown as letters directly above or below the wild-type base. Substitutions at dipyrimidine sequences are placed above or below the strand containing the dipyrimidine rather than the opposite dipurine-containing strand, since these are inferred to represent errors resulting from replication past cyclobutane pyrimidine dimers (see text). Substitutions at nondipyrimidine locations are arbitrarily placed above or below the (+) strand. The loss of a base is indicated by an open triangle (Δ), and the addition of a base is indicated by a closed triangle (▲). When a frameshift occurs at positions of iterated or reiterated bases, it is not possible to distinguish which base was lost or added; thus the symbol is centered under or over the run. In one mutant, an A was added between the nucleotides at positions 136 and 137, as indicated by a closed square (■). Extensive prior use of this reporter gene in M13mp2 and M13mp2SV ori left vectors (Roberts & Kunkel, 1993) has revealed 27 dipyrimidine sites (underlined nucleotides) at which C·G → T·A transitions can be scored. This study and two previous ones (Roberts et al., 1991, 1993) have shown that the M13mp2SV ori right vector is consistently slightly more sensitive for detecting α-complementation mutants, such that to date there are 35 (+) dipyrimidine sites in this vector at which C·G → T·A transitions can be scored. The eight additional sites are positions –44, –14, 96, 111, 117, 120, 132, and 135.

irradiated, replicated DNA were severalfold higher at two different doses (Table I). Photoreactivation of ori left DNA irradiated at 70 J/m^2 prior to replication restored the mutant frequency to near background levels (Thomas & Kunkel, 1993), suggesting that most of the mutants were cyclobutane dimer-induced. At both doses, the ori right DNA yielded higher mutant frequencies than the ori left vector. This difference partly reflects an increased number of phenotypically detectable sites in the ori right vector (see below).

Sequence Analysis of Mutants. Mutants obtained from replication of DNAs irradiated at 70 and 100 J/m^2 were analyzed by DNA sequence analysis. Since the mutant frequencies were similar, the sequencing results for the two doses were combined. Totals of 290 and 255 mutants were analyzed from the ori left and ori right reactions, respectively (Table II). For comparison, a breakdown of the mutants obtained from replication of undamaged ori left DNA [from Thomas et al. (1991)] is also included. An equivalent spectrum for mutants from undamaged, replicated ori right DNA was not generated, since the frequencies for the two vectors were similar to each other and to the control values in Table I.

For replication of UV-irradiated ori left and ori right DNAs, single-base substitutions account for over 80% of the observed changes, and about 90% of these are at dipyrimidine sites (Table II). All but 5% of these are at a template C, and nearly all were C → T transitions. When the UV-induced ori left and ori right collections are compared to the undamaged DNA collection, the frequencies for total base substitutions, substitutions at dipyrimidines and template C, and C → T transitions are significantly above the undamaged values. This general specificity extends the results obtained in earlier studies (Thomas & Kunkel, 1993; Carty et al., 1993) and is consistent with the interpretation that most of these mutations arise as a result of mutagenic bypass of cytosine-containing pyrimidine dimers.

Table I: Mutant Frequencies after Replication in Extracts

reaction conditions	UV fluence (J/m^2)	plaques scored total	plaques scored mutant	mutant frequency ($\times10^{-4}$)
unincubated DNA				
ori left	0	64 975	21	3.2[a]
	100	24 121	10	4.1[b]
ori right	0	ND[c]	ND	
	100	27 663	19	6.8
reactions without T antigen				
ori left	0	19 993	10	5.0[b]
	70	14 367	4	2.8[b]
ori right	0	15 442	12	7.8
	100	17 009	11	6.4
replicated DNA				
ori left	0	21 637	12	5.5
	70	71 917	170	24
	100	48 915	120	25
ori right	0	14 306	14	9.8
	70	37 611	153	41
	100	21 981	102	46

[a] Data taken from Roberts and Kunkel (1988). [b] Data taken from Thomas and Kunkel (1993). [c] ND = not determined.

Table II: Sequence Analysis of Mutants

mutation	ori left[a] no.[b]	ori left[a] frequency ($\times10^{-5}$)	ori left + UV no.	ori left + UV frequency ($\times10^{-5}$)	ori right + UV no.	ori right + UV frequency ($\times10^{-5}$)
total	78	74	290	240	255	430
substitutions	37	35	120	99	125	210
non-pyr-pyr site	10	9.4	12	9.9	12	20
pyr-pyr site	27	26	108	89	113	190
template C	21	20	104	86	107	180
template T	6	5.7	4	3.3	6	10
C·G → T·A	12	11	96	79	104	175
others	15	14	12	9.9	9	15
frameshifts	8	7.6	9	7.4	16	27
CC → TT	0	≤0.9	5	4.1	6	10
insertions	0	≤0.9	4	3.3	1	1.7
complex	1	0.9	2	1.7	4	6.7
no change	28		150		103	

[a] Data taken from Thomas et al. (1991). [b] The total number of mutants for ori left includes four deletions.

All other classes of mutations combined make up less than 20% of the mutants detected in the target (Table II). Single-base frameshifts, both plus-one base pair and minus-one base pair, were observed in both collections. At least with the ori left reactions, however, their frequency is not above background. The frequency of frameshifts in the ori right collection is about 3-fold above the ori left background frequency, and all but one of these occur at dipyrimidines. Four of the frameshifts occur at a run of five C's (positions 132–136, Figure 1). This is where three such errors were also observed in the spontaneous collection. Thus, this site may be an active sequence for frameshifts during replication of undamaged DNA. Even if some of the remaining frameshifts are UV-induced, the frequency is low compared to C → T transitions.

Several CC → TT tandem double mutations were observed in both collections (Table II). Though comprising only 4% of the mutants, these have never been seen in previous collections with undamaged DNA but have been noted in other UV-induced collections (Brash et al., 1991; Carty et al., 1993). Their frequency is at least 5–10-fold above background, and thus they are likely to be UV-induced. We also recovered a few insertions and complex mutations not detected previously with undamaged DNA.

Mutagenic Translesion Replication on the Leading and Lagging Strands. For analysis of UV-induced mutations on the leading and lagging strands, only C → T transitions were considered, since these are the most likely to be UV-dependent on the basis of signal-to-noise ratios (Table II). All (–) strand C → T transitions generated in the ori left reactions and all (+) strand C → T transitions from the ori right reactions (Figure 1) were grouped together to constitute leading strand errors. Likewise, all (–) strand C → T transitions from ori right reactions and (+) strand C → T transitions from ori left reactions were combined and were inferred to represent lagging strand errors. Given the number of detectable sites for this substitution error, the overall average frequency of C → T transitions per detectable dipyrimidine site is identical for leading and lagging strand synthesis (Table III). Nonetheless, among the 13 dipyrimidine positions available for detecting C → T transitions with both vectors, and at which mutations were scored more than once with either vector, some sites show a higher frequency with one vector or the other (Figure 1 and Table IV). At some positions (e.g., positions 145 and 89), lagging strand errors predominate, while at others (e.g., positions 136 and 168) leading strand errors predominate.

Table III: C → T Substitution Specificity on the Leading and Lagging Strands

strand	detectable sites[a]	no. of mutants	frequency ($\times10^{-5}$)	frequency per site ($\times10^{-5}$)
leading strand	35	114	63	1.8
lagging strand	27	87	48	1.8

[a] An explanation of detectable sites for C → T transitions in the ori left and ori right vectors is provided in the legend to Figure 1.

Table IV: Comparison of C → T Substitutions by Position

template position[a]	lagging mutants	lagging frequency ($\times10^{-6}$)	leading mutants	leading frequency ($\times10^{-6}$)	ratio of leading to lagging
145	3	50	0	≤8	≥6.2:1
89	9	150	4	33	4.5:1
84	4	67	2	17	3.9:1
–57	15	120	3	50	2.4:1
88	6	100	6	50	2.0:1
148	3	50	3	25	2.0:1
–66	10	83	3	50	1.7:1
90	9	150	14	120	1.2:1
133	4	33	3	50	1:1.5
108	7	120	4	33	1:3.6
58	3	25	6	100	1:4.0
168	0	≤8	4	67	≤1:8.0
136	1	8	7	120	1:14

[a] Fourteen mutants at position –37, 13 on the lagging strand and one on the leading strand were not included in this analysis, since this site was observed to have an unusually high background mutant frequency for undamaged, replicated ori left DNA.

Table V: C → T Substitution Specificity at Dipyrimidine Sites

strand	site considered[a]	detectable sites	no. of mutants	mutants, detectable site	Y = T	Y = C
leading strand	5′-C-Y-3′	12	8[b]	0.67	1	7
	5′-Y-C-3′	12	47	3.9	27	20
	5′-Y-C-Y-3′	10	57	5.7	37[c]	20
lagging strand	5′-C-Y-3′	10	11[b]	1.1	4	7
	5′-Y-C-3′	10	39	3.9	26	13
	5′-Y-C-Y-3′	6	24	4.0	11[c]	13

[a] Y represents a pyrimidine residue. [b] The fourteen mutants at position –37, 13 on the lagging strand and one on the leading strand were not included in this analysis, since this site was determined to have an unusually high background mutant frequency for undamaged, replicated ori left DNA. [c] These mutations were flanked by both a T and a C residue; i.e., the 5′ to 3′ sequence was either TCC or CCT.

When the specificity of C → T substitutions was examined by location within the dipyrimidine sequence (Table V), the mutated C residue was found to preferentially be the 3′ nucleotide, for both leading and lagging strand errors. Overall, there does not appear to be a strong preference for either a T or C as the flanking (nonmutated) member of the dimer.

DISCUSSION

Previous studies from this (Thomas & Kunkel, 1993) and another laboratory (Carty et al., 1993) had suggested that pyrimidine dimers were bypassed during SV40 origin-dependent replication of UV-irradiated DNA in human cell extracts and that this bypass is mutagenic. The primary objectives of the present study were to define extensively the spectrum of errors resulting from this translesion replication and to examine whether mutagenic replication of the same damaged template sequence was different depending on whether it was copied as the leading or lagging strand.

Error Specificity during Trans-Dimer Replication. Sequence analysis of mutants from replication of damaged DNA indicates that mutations occur at frequencies significantly above those for replication of undamaged DNA only for C → T transitions, CC → TT substitutions, and complex mutations (Table II). This does not necessarily imply that other types of errors (e.g., one-base frameshifts or substitutions at TT dimers) do not result from trans-dimer replication, only that these were not observed here at frequencies above those obtained with undamaged DNA using this forward mutation assay.

As noted earlier (Thomas & Kunkel, 1993), and here extended to a much larger mutant collection, the error specificity generated during trans-dimer replication in cell extracts is quite similar to that observed *in vivo* in eukaryotic cells (Keyse et al., 1988; Armstrong & Kunz, 1990) and in skin tumors (Schaaper et al., 1987; Brash et al., 1991). This includes 11 occurrences of the unusual tandem CC → TT substitution (Table II), which is considered a signature mutation resulting from UV irradiation (Brash et al., 1991) or oxidative stress (Reid & Loeb, 1993). Eleven mutants comprising eight different mutations were also recovered from replication of damaged DNA. These had either insertions or complex deletions/replacements involving several nucleotides (Table II). Since the frequency of these events was higher than for replication of undamaged DNA, they may reflect aberrant replication of damaged DNA. Possible models to explain simple and complex frameshifts have invoked kinetic and/or structural barriers to continued synthesis [for reviews, see Kunkel (1990) and Ripley (1990)]. UV photoproducts are one such barrier, as they are well-known to inhibit synthesis by purified DNA polymerases (Moore et al., 1981; O'Day et al., 1992) and replication of double-stranded DNA (Thomas & Kunkel, 1993; Carty et al., 1993; this study).

The most frequent mutation in the UV spectra is the C → T substitution. The 7–16-fold increase in mutant frequency for this event (Table II) suggests that 85%–94% of these are UV-dependent. These are inferred to result from insertion of dAMP opposite the position of a cytosine in a cytosine-containing dipyrimidine sequence. Statistical comparison of either UV-induced spectrum (ori left + UV or ori right + UV) with the previous ori left spectrum obtained with undamaged DNA (Table II) indicates strongly significant differences. The sparseness measure (see Experimental Procedures) for both analyses is large enough (approximately 52 for the ori left + UV vs ori left comparison and approximately 47 for the ori right + UV vs ori left comparison) that the spectra form nonsparse tables. Thus, from Piegorsch and Bailer (1993), we conclude that the Gabriel (1966) analysis is acceptable. The resulting *P*-values for both comparisons are less than 0.0001.

The results in Table V indicate a preference for errors at cytosines that are at 3′ rather than 5′ positions in the dimer. This is not a target size bias, since the frequencies are expressed per detectable site. It could reflect a difference in initial damage distribution, a difference in the efficiency of translesion replication of the two types of dimers, or a difference in the coding potential and/or the rate of deamination (see below) of a 3′ versus 5′ cytosine in a dimer.

Mutagenic Dimer Bypass during Leading and Lagging Strand Replication. Interpreting the spectra observed with the two vectors (Figure 1) in terms of leading and lagging strand replication requires several assumptions. One is that the strand which templates errors can be assigned. We have shown that the replicated fraction of UV-irradiated DNA contains cyclobutane dimers and that pretreatment of the irradiated DNA with photolyase prior to replication largely reverses the inhibition of replication and the mutagenicity (Thomas & Kunkel, 1993; this study). This suggests that the sites of UV-induced mutations are cyclobutane dimers and thus identifies the template strand. C → T substitutions are the only class of mutants comprising a large number of occurrences and showing a substantial UV irradiation-dependent component. We therefore compare leading and lagging strand error rates only for this mutation.

Another assumption is that one knows which of the two replication forks emanating from the bidirectional origin was responsible for translesion replication. Previous studies (Edenberg & Huberman, 1975; Li & Kelly, 1985) have indicated that the rate of fork movement is similar in both directions from the SV40 origin. The two vectors used here contain 7396 base pairs, and the *lacZα*-complementation target comprises nucleotides −84 to +170. From the base pair in the origin where the switch from lagging to leading strand replication occurs (Hay & DePamphilis, 1982), the closest replication fork in the ori left vector must replicate only 594 base pairs to completely synthesize the *lacZα* target. If it does so before the other fork replicates the 6802 base pairs needed to copy past the *lacZα* sequence, then in the ori left vector the (+) viral strand is replicated as the lagging strand (Figure 1). Similarly, the ori right vector contains the origin to the right of the target, and the closest replication fork need replicate only 402 base pairs to complete the target. Thus, with this vector the (+) viral strand is assumed to be replicated as the leading strand.

A third assumption is that other parameters are constant and do not influence our interpretations. Here we compare error rates on the same primary DNA sequence when replicated as the leading or lagging strand, so explanations for any observed differences are distinct from well-known sequence context effects on fidelity. The two DNA substrates were irradiated in parallel and subsequently handled in the same manner, so we assume that damage distribution in the reporter gene was similar. Furthermore, reactions with the two vectors were performed in parallel under the same conditions, using the same reagents and extract. Thus, any observed effects are suggested to reflect the fidelity of synthesis catalyzed by the DNA polymerases/accessory proteins responsible for leading and lagging strand replication.

As noted in Table III, the overall fidelity for C → T transitions is similar on the two strands. This suggests that neither mode of replication is more or less error-prone for this

particular error when results are averaged over the target sequence. Nonetheless, there is a difference in the distribution of errors for leading versus lagging strand replication (Table IV). Several sites yielded more errors when replicated as the lagging strand, while other sites yielded more errors when replicated as the leading strand. In order to determine if the difference in the overall pattern in Table IV is statistically significant, we first calculated the sparseness measure $N/\{(R-1)(T-1)\}$, which here is equal to 9.83. This is just below the recommended cutoff of 10 given by Piegorsch and Bailer (1993). Thus, a hypergeometric test of the two spectra (Adams & Skopek, 1987) was performed. This yielded a P-value of 0.005, suggesting that the overall variation is highly significant. Next, we assessed the contributions of individual template positions using the Gabriel (1966) test. Although this test should be viewed with caution since there is evidence of marginal sparseness, it nonetheless corroborates the hypergeometric indication of significant leading-versus-lagging heterogeneity ($P = 0.004$). Finally, when we used the Gabriel method for post-hoc evaluation of positions that contribute to the heterogeneity, positions 168 and 136 were found to contribute most significantly to the difference, while when only one of the sites showing the greatest difference (e.g., sites 136, −57, or 145) is removed, the significant heterogeneity in the table is retained. These analyses suggest that the site specificity of mutagenic trans-dimer synthesis is different during leading and lagging strand replication of the same sequence. Thus, the mutagenic risk posed by replicating past a lesion at a key location in a critical gene may not be constant but may depend on its position relative to surrounding origins. Any circumstance wherein new or alternative origins are used, e.g., replication from additional origins during rapid replication in embryonic tissues, may alter the mutagenic potential of a DNA damaging agent.

Evidence has accumulated that UV-dependent C → T substitutions result from incorporation of dAMP opposite uracil-containing templates resulting from deamination of cytosine in cyclobutane dimers (Tessman et al., 1992; Jiang & Taylor, 1993). Regarding this possibility, the differences in distribution of C → T substitutions in the present study can be explained in four ways. First, the rate of deamination may be different in the two vectors. This seems unlikely since the ori left and ori right substrates used here were processed similarly. Second, the C → T substitutions may result from deamination, and the observed differences represent random experimental fluctuation. Although this seems unlikely given the statistical analysis of significance, it nonetheless cannot be excluded. Third, a template uracil within a dimer may have a different coding potential or probability of being bypassed when replicated by a leading-versus-lagging strand DNA polymerase/accessory protein complex.

Finally, C → T substitutions may result from misincorporation of dAMP opposite cytosine rather than uracil. The potential for such errors may differ for leading-versus-lagging strand replication proteins in a site-specific manner for any of several possible reasons. One possibility is that replication of the two strands is performed by different DNA polymerases or perhaps the same polymerase but with a different complement of accessory proteins, thus yielding differences in misinsertion rates and/or ability to extend versus proofread mispaired intermediates at dimer sites. That replication on the leading strand is highly processive may be relevant for bypassing dimers in particular sequence contexts, thus fixing misincorporations via continued synthesis from the mispair in the dimer. In comparison, lagging strand replication is discontinuous, may involve more than one DNA polymerase and/or one or more switches between enzymes, and requires both RNA primer synthesis and eventual replacement. Thus, the added complexity of lagging strand replication provides ample opportunities for nonrandom rates of mutagenic trans-dimer replication.

In vivo observations (Sarasin & Hanawalt, 1980; Berger & Edenberg, 1986; Vrieling et al., 1991) have led several authors to speculate that the probability of replication past UV photoproducts may differ on the two strands. Recently, O'Day et al. (1992) have demonstrated that DNA polymerase δ, in the presence of the accessory protein PCNA, is capable of replication past cis–syn and trans–syn TT dimers. Given the evidence for the involvement of DNA polymerase δ and PCNA in SV40 origin-dependent replication of undamaged DNA (Weinberg & Kelly, 1989; Lee et al., 1989; Tsurimoto et al., 1990), it is possible that they are responsible for mutagenic translesion replication of templates harboring C-containing cyclobutane pyrimidine dimers in the replication system used here. Synthesis by human DNA polymerase α, the enzyme responsible for at least some synthesis on the lagging strand, is inhibited by pyrimidine dimers (Moore et al., 1981), suggesting that either additional accessory factors are needed for this polymerase to perform trans-dimer synthesis or that another DNA polymerase bypasses the dimer. Since the Hela cell extract has already been fractionated and replication has been reconstituted *in vitro* from purified fractions (Kelly, 1988; Stillman, 1989; Hurwitz, et al., 1990), it may eventually be possible to fully characterize the process of replication past UV photoproducts on both strands.

ACKNOWLEDGMENT

We thank R. Stephen Lloyd for his generous gift of T4 endonuclease V and Roel M. Schaaper and Kenneth R. Tindall for critical comments on the manuscript.

REFERENCES

Adams, W. T., & Skopek, T. R. (1987) *J. Mol. Biol. 194*, 391–396.
Armstrong, J. D., & Kunz, B. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9005–9009.
Berger, C. A., & Edenberg, H. J. (1986) *Mol. Cell. Biol. 6*, 3443–3450.
Brash, D. E., Rudolph, J. A., Simon, J. A., Lin, A., McKenna, G. J., Baden, H. P., Halperin, A. J., & Ponten, J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10124–10128.
Carty, M. P., Hauser, J., Levine, A. S., & Dixon, K. (1993) *Mol. Cell. Biol. 13*, 533–542.
Dorado, G., Steingrimsdottir, H., Arlett, C. F., & Lehmann, A. R. (1991) *J. Mol. Biol. 217*, 217–222.
Edenberg, H. J., & Huberman, J. (1975) *Annu. Rev. Genet. 9*, 245–284.
Gabriel, K. R. (1966) *J. Am. Stat. Assoc. 61*, 1081–1096.
Hay, R. T., & DePamphilis, M. L. (1982) *Cell 28*, 767–779.
Hurwitz, J., Dean, F. B., Kwong, A. D., & Lee, S.-H. (1990) *J. Biol. Chem. 265*, 18043–18046.
Jiang, N., & Taylor, J.-S. (1993) *Biochemistry 32*, 472–481.
Kelly, T. J. (1988) *J. Biol. Chem. 263*, 17889–17892.
Keyse, S. M., Amaudruz, F., & Tyrrell, R. M. (1988) *Mol. Cell. Biol. 8*, 5425–5431.
Kornberg, A., & Baker, T. (1992) *DNA Replication*, W. H. Freeman & Co., New York.
Kunkel, T. A. (1990) *Biochemistry 29*, 8003–8011.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 362–387.
Lee, S.-H., Eki, T., & Hurwitz, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7361–7365.
Li, J. J., & Kelly, T. J. (1985) *Mol. Cell. Biol. 5*, 1238–1246.

McGregor, W. G., Chen, R.-H., Lukash, L., Maher, V. M., & McCormick, J. J. (1991) *Mol. Cell. Biol. 11*, 1927–1934.
Menichini, P., Vrieling, H., & van Zeeland, A. A. (1991) *Mutat. Res. 251*, 143–155.
Moore, P. D., Bose, K. K., Rabkin, S. D., & Strauss, B. S. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 110–114.
O'Day, C. L., Burgers, P. M. J., & Taylor, J.-S. (1992) *Nucleic Acids Res. 20*, 5403–5406.
Piegorsch, W. W., & Bailer, A. J. (1993) Technical Report No. 8270, National Institute of Environmental Health Sciences, Research Triangle Park, NC.
Reid, T. M., & Loeb, L. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3904–3907.
Ripley, L. S. (1990) *Annu. Rev. Genet. 24*, 189–213.
Roberts, J. D., & Kunkel, T. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7064–7068.
Roberts, J. D., Thomas, D. C., & Kunkel, T. A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3465–3469.
Roberts, J. D., Nguyen, D., & Kunkel, T. A. (1993) *Biochemistry 32*, 4083–4089.
Sarasin, A. R. (1989) *J. Photochem. Photobiol. 3*, 143–155.
Sarasin, A. R., & Hanawalt, P. G. (1980) *J. Mol. Biol. 138*, 299–319.
Schaaper, R. M., Dunn, R. L., & Glickman, B. W. (1987) *J. Mol. Biol. 198*, 187–202.
Stillman, B. (1989) *Annu. Rev. Cell Biol. 5*, 197–245.
Sykes, R. C., Lin, D., Hwang, S. J., Framson, P. E., & Chinault, A. C. (1988) *Mol. Gen. Genet. 212*, 301–309.
Tessman, I., Liu, S.-K., & Kennedy, M. A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1159–1163.
Thomas, D. C., & Kunkel, T. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7744–7748.
Thomas, D. C., Roberts, J. D., Sabatino, R. D., Myers, T. W., Tan, C.-K., Downey, K. M., So., A. G., Bambara, R. A., & Kunkel, T. A. (1991) *Biochemistry 30*, 11751–11759.
Tsurimoto, T., Melendy, T., & Stillman, B. (1990) *Nature 346*, 534–539.
Vrieling, H., van Rooijen, M. L., Groen, N. A., Zdzienicka, M. Z., Simons, J. W. I. M., Lohman, P. H. M., & van Zeeland, A. A. (1989) *Mol. Cell. Biol. 9*, 1277–1283.
Vrieling, H., Venema, J., van Rooijen, M. L., van Hoffen, A., Menichini, P., Zdzienicka, M. Z., Simons, J. W. I. M., Lohman, P. H. M., Mullenders, L. H. F., & van Zeeland, A. A. (1991) *Nucleic Acids Res. 19*, 2411–2415.
Weinberg, D. H., & Kelly, T. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9742–9746.