MOLECULAR AND CELLULAR BIOLOGY, Jan. 1986, p. 277–285
0270-7306/86/010277-09$02.00/0

Vol. 6, No. 1

Sequence Specificity of Point Mutations Induced During Passage of a UV-Irradiated Shuttle Vector Plasmid in Monkey Cells

JANET HAUSER,[1] MICHAEL M. SEIDMAN,[2] KATHARINE SIDUR,[1] AND KATHLEEN DIXON[1]*

Section on Viruses and Cellular Biology, National Institute of Child Health and Human Development, [1] and Laboratory of Molecular Carcinogenesis, National Cancer Institute,[2] National Institutes of Health, Bethesda, Maryland 20205

Received 29 July 1985/Accepted 10 October 1985

A simian virus 40-based shuttle vector was used to characterize UV-induced mutations generated in mammalian cells. The small size and placement of the mutagenesis marker (the *supF* suppressor tRNA gene from *Escherichia coli*) within the vector substantially reduced the frequency of spontaneous mutations normally observed after transfection of mammalian cells with plasmid DNA; hence, UV-induced mutations were easily identified above the spontaneous background. UV-induced mutations characterized by DNA sequencing were found primarily to be base substitutions; about 56% of these were single-base changes, and 17% were tandem double-base changes. About 24% of the UV-induced mutants carried multiple mutations clustered within the 160-base-pair region sequenced. The majority (61%) of base changes were the G · C→A · T transitions; the other transition (A · T→G · C) and all four transversions occurred at about equal frequencies. Hot spots for UV mutagenesis did not correspond to hot spots for UV-induced photoproduct formation (determined by a DNA synthesis arrest assay); in particular, sites of TT dimers were underrepresented among the UV-induced mutations. These observations suggest to us that the DNA polymerase(s) responsible for mutation induction exhibits a localized loss of fidelity in DNA synthesis on UV-damaged templates such that it synthesizes past UV photoproducts, preferentially inserting adenine, and sometimes misincorporates bases at undamaged sites nearby.

Recent studies on viral and cellular oncogenes have provided strong evidence that mutations can play a fundamental role in cellular transformation and carcinogenesis (3). In addition, many heritable diseases and developmental anomalies have a mutational origin. Although numerous mutagenic agents are known to interact with DNA, the molecular mechanisms by which these agents may cause mutations in mammalian cells are not well understood. To approach this problem, we have used a UV-irradiated simian virus 40 (SV40)-based shuttle vector to analyze the sequence specificities of point mutations induced in mammalian cells by DNA damage. We reasoned that a detailed knowledge of the characteristics and spectrum of the mutations formed would provide information about the molecular mechanisms of mutation induction.

In our previous work on the replication of UV-damaged SV40 DNA in mammalian cells (1, 32, 34), we demonstrated that replication of damaged templates appears to occur by a dimer-bypass mechanism in which replication forks proceed beyond UV-induced lesions, leaving gaps (single-strand discontinuities of 50 to 150 bases) in the daughter strands. Replication is completed at the normal terminus, and relaxed circular gapped molecules are produced. These gapped molecules are then slowly converted to SV40 form I DNA. We proposed that these single-strand gaps would be filled by DNA synthesis which might include insertion of mismatched bases at the sites of UV-induced lesions in the template strands, leading to the induction of point mutations at the damage sites. An alternative mechanism for gap filling, involving recombination, would predict few mutations at the sites of UV-induced lesions but might suggest the generation of localized DNA rearrangements due to errors in the recombination process. A mechanism involving a generalized lack of fidelity in DNA synthesis occurring during gap filling would predict base changes at sites other than those of UV photoproducts. Thus, an analysis of the characteristics of UV-induced mutations should yield information on the mechanism of gap filling and mutation induction.

The analysis of induced mutations in mammalian cells has been hampered by the lack of the elegant genetic systems that have been exploited for this purpose in bacterial cells and yeast (6, 8, 9, 11, 14, 15, 19, 36). The use of shuttle vector plasmids that can replicate in both mammalian cells and bacterial cells should provide a solution to this problem. Mutagen-treated plasmid DNA can be introduced into and allowed to replicate in mammalian cells, recovered, and then analyzed for mutations in bacterial cells. In addition, plasmid vector DNA propagated in bacterial cells can be used for sequence analysis. Unfortunately, a problem encountered in the use of shuttle vectors has been the high frequency of spontaneous deletion and insertion mutations arising during passage in mammalian cells (5, 10, 21, 22, 27). However, we have shown that the recovery of these spontaneous mutations can be minimized by judicious design of the shuttle vector (28a). The vector we have used here, pZ189 (see Fig. 1; Seidman et al., in press), contains the plasmid origin of replication and β-lactamase gene from pBR327, a plasmid that lacks the poison sequences which inhibit replication of pBR322-based plasmids in mammalian cells (12). The vector contains the *supF* suppressor tRNA gene from *Escherichia coli* to provide a selectable gene for mutation analysis that can be easily sequenced. The vector also contains the early region of SV40 viral DNA to allow replication in mammalian cells. The location of the tRNA gene, between the plasmid origin of replication and the β-lactamase gene, and its small size lead to a reduced spontaneous mutation frequency (about 0.04%) and a reduced recovery of spontaneous deletions.

Here, we describe the use of this shuttle vector to analyze

* Corresponding author.

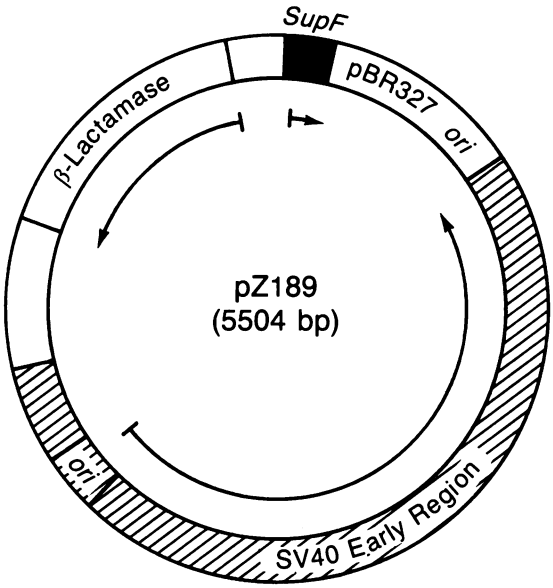

FIG. 1. pZ189 shuttle vector. The 5504-base-pair vector contains the ampicillin resistance gene (β-lactamase), the pBR327 origin of replication (ori), and the *E. coli supF* tyrosine suppressor tRNA gene from πAN7 (H. Huang, personal communication); and the origin of replication and T antigen coding region from SV40 (clockwise from SV40 nucleotides 2533 to 346). Construction and characterization of the vector have been described (28a). Arrows indicate the direction of transcription.

a series of UV-induced mutations generated in mammalian cells.

MATERIALS AND METHODS

Cells and plasmids. The TC7 clone (24) of the African Green Monkey kidney cell line CV-1 was grown in Dulbecco minimal essential medium (GIBCO Laboratories, Grand Island, N.Y.) supplemented with antibiotics and 5% fetal calf serum. *E. coli* MBM7070 (28a) has the genotype F^- *lacZ*(Am)*CA7020, lacY1 hsdR hsdM Δ(araABC-leu)7679 galU galK rpsL thi.* In the presence of isopropyl-β-D-thiogalactoside, an inducer of the *lac* operon, and 5-bromo-4-chloro-3-indoyl-β-D-galactoside, an artificial substrate for β-galactosidase, strain MBM7070 forms blue colonies if it contains an active *supF* suppressor tRNA gene and white colonies if the suppressor is inactive. The structure of the plasmid pZ189 (Seidman et al., in press) is shown in Fig. 1.

Mutation induction. Purified plasmid DNA in TBS buffer (25 mM Tris hydrochloride, 5 mM KCl, 0.9 mM Na_2HPO_4, 130 mM NaCl; pH 7.3) was irradiated with a germicidal lamp (model G15T8; maximal output at 254 nm, fluence rate of 12 J/m^2 per s; General Electric Co., Schenectady, N.Y.). Subconfluent TC7 cells growing in 100-mm plastic tissue culture dishes were transfected with the plasmid DNA (5 ng per dish) by the DEAE-dextran method (13). Forty-eight hours later, plasmid DNA was harvested from the cells by the Hirt method (7) and further purified by CsCl equilibrium density centrifugation.

Mutant selection and characterization. The bacterial strain MBM7070 was transformed with plasmid DNA by the calcium chloride procedure (29). Transformants were plated on Luria broth (LB) plates with ampicillin (50 μg/ml) and spread with 5-bromo-4-chloro-3-indoyl-β-D-galactoside (2 mg) and isopropyl-β-D-thiogalactoside (12 mg) to determine the total yield of transformed colonies and to identify colonies containing mutant plasmids. White and light blue colonies were picked, restreaked three times, and grown overnight in 5 ml of LB broth for plasmid DNA preparation. Plasmid DNA was prepared for agarose gel analysis and DNA sequencing by the alkaline extraction method (2). The tRNA region of the double-stranded supercoiled plasmid DNA was sequenced by a modification (37) of the dideoxyribonucleotide method of Sanger et al. (26) with a pBR322 *Eco*RI sequencing primer purchased from (Pharmacia P-L Biochemicals, Piscataway, N.J.) and either the Klenow fragment of DNA polymerase I or the avian myeloblastosis virus reverse transcriptase (both enzymes purchased from Boehringer Mannheim Biochemicals, Indianapolis, Ind.).

Determination of sites of UV photoproducts. The positions of UV photoproducts in the tRNA region of UV-irradiated plasmid DNA were determined by a DNA synthesis-arrest array as described by Moore et al. (17). Two primers were used, the pBR322 *Eco*RI sequencing primer that hybridizes to positions 5479 through 5494 on the vector DNA for rightward sequencing through the tRNA gene, and another primer (generously provided by M. Berman) that hybridizes to the opposite strand of the vector at positions 212 through 227 for leftward sequencing.

RESULTS

Determination of induced mutant frequency. To establish that the frequency of induced mutations is substantially higher than the frequency of spontaneous mutations in the pZ189 vector, we carried out 10 separate transfections of CV-1 cells with UV-irradiated (500 J/m^2) and unirradiated vector DNA. Vector DNA was recovered from the transfected cells 48 h later, purified, and used to transform competent MBM7070 bacterial cells. The number of transformed dark blue colonies (*supF*) and the frequency of white and light blue transformed colonies (*supF*) were then determined. In this and other experiments, irradiation of the plasmid with a UV fluence of 500 J/m^2 before transfection of the mammalian cells reduced the yield of transformed bacterial colonies to about 20% of that of the unirradiated control. We found that there was considerable fluctuation in the frequency of vector mutants arising from the individual transfections but that on the average the frequency of mutants was at least 20-fold higher when UV-irradiated vector DNA was used (Table 1). The variability in recovery

TABLE 1. Mutagenesis of unirradiated or UV-irradiated pZ189 shuttle vector in CV-1 cells

Transfection no.	No. of colonies with:			
	Unirradiated vector		UV-irradiated vector	
	Total examined	Mutants	Total examined	Mutants
1	7,063	7	1,028	9
2	7,486	2	110	5
3	10,587	6	1,675	26
4	17,666	4	139	2
5	10,666	4	118	0
6	11,973	9	245	6
7	14,647	5	15	0
8	6,700	3	141	2
9	8,540	0	281	2
10	9,573	1	1,868	21

[a] Total mutation frequencies were 0.04 and 1.1% with the unirradiated and UV-irradiated (500 J/m^2) vectors, respectively.

TABLE 2. Frequencies of spontaneous and UV-induced mutants observed

Radiation treatment	No. of colonies examined	No. (%) of mutants	Fraction with altered gel mobility
None	222,431	83 (0.037)	22/67
500 J/m^2	37,011	281 (0.76)	10/251

of plasmid from transfected cells may be a result of poor recovery of DNA in the purification method used here. In later experiments in which a modified procedure was used, less variation was seen. The frequency of spontaneous mutants found when the unirradiated vector was grown solely in *E. coli* MBM7070 was about 0.001%. Irradiation of pZ189 before transformation of the bacteria resulted in a 3 log reduction in the recovery of transformed colonies, and no increase in mutation frequency was observed. (Since our bacterial strain is *recA*$^+$, we speculate that the failure to observe an increase in mutation frequency when the UV-irradiated vector is introduced directly into bacteria may be due to a requirement for expression of the ampicillin gene before there is sufficient repair or replication of the vector for mutations to be fixed.) Thus, it appeared that it should be possible to select and characterize mutations induced in the pZ189 vector in mammalian cells.

We considered the possibility that the DNA recovered from the mammalian cells and used to transform bacteria had some unusual characteristic that caused the mutations to occur in bacteria. To rule out this possibility, we recovered DNA from a series of separate transfections and then carried out duplicate transformations with each DNA sample. All plasmid mutants recovered were sequenced. We found that, in several cases, the same rare mutant sequence was recovered multiple times from both transformations of a single-transfection sample, and the spectrum of mutations differed markedly among the various transfections samples. If the mutations had occurred in bacteria, we would not expect to find the same spectrum of mutations in the duplicate transformations and especially not the duplication of rare mutations in duplicate transformations.

Comparison of spontaneous and induced mutations. White and pale blue colonies were collected from the experiment described above and from several other similar experiments for further analysis. Vector DNA was prepared from individual clones and analyzed on agarose gels to determine the frequency of gross DNA alterations in the mutant vectors. DNA from about a third of the spontaneous mutants exhibited altered mobility on gels, whereas only about 4% of the induce mutants had DNA with altered mobility (Table 2).

The vector DNAs that appeared to migrate in gels with normal mobility were subjected to DNA sequence analysis. Since the pBR322 *Eco*RI primer we used hybridizes near the *Eco*RI site at position 5504/0 on the vector, the sequence changes from position 40 to 200 could be easily determined. The coding sequence for the mature processed tRNA extends from position 99 to 183. We sequenced UV-induced mutants from 34 separate transfections and spontaneous mutants from 10 separate transfections. Any mutants represented more than once among sequences from a single transfection were excluded from our analysis. This procedure may result in a slight underrepresentation of mutational hot spots, but it ensures that the mutants analyzed were independent.

TABLE 4. Base changes in spontaneous and UV-induced mutants

Base change	No. of mutants observed	
	Untreated	Irradiated
G · C→A · T	23	90
G · C→T · A	16	19
G · C→C · G	3	8
A · T→G · C	0	11
A · T→C · G	0	6
A · T→T · A	0	14

Characteristics of the sequence alterations in the induced mutants differed from those in the spontaneous mutants (Table 3). About 20% of the spontaneous mutants sequenced contained small deletions in the tRNA gene, whereas only 3% of the induced mutants sequenced had deletions. Most (73%) of the induced mutants had single mutations, either single-base changes or tandem double-base changes. In the induced mutants that did have multiple mutations, the majority of changes occurred within 15 bases of each other. In contrast, almost half of the spontaneous mutants had multiple mutations, and most were at distant sites (>15 bases apart). The spectrum of base changes also differed between induced and spontaneous mutants (Table 4). Interestingly, among the spontaneous mutations, only changes of G · C were found, with the G · C→A · T transition representing about half of the base changes observed. Among the induced alterations, the G · C→A · T transition was also quite prevalent, but the other transition and all four transversions were also observed.

The distributions of spontaneous and UV-induced base alterations along the sequenced portion of the vector DNA are shown in Fig. 2. All mutants that were sequenced had at least one alteration in the coding sequence for the processed tRNA except for two induced mutants that were identified as light blue colonies and had single-base changes in the region 5′ to the tRNA coding sequence. There appeared to be hot spots for UV-induced mutations at positions 123, 156, 159, 168, and 169. The spontaneous mutations appeared to be more evenly distributed within the sequenced region.

TABLE 3. Sequence alterations in spontaneous and UV-induced mutants

Radiation treatment	No. of mutants sequenced with[a]:					
	Small deletions	Single changes		Two changes		Three or more changes (>15 bases apart)
		One-base	Tandem double-base	≤15 bases apart	>15 bases apart	
None	5	6	1	2	3	6
500 J/m^2	3	54	16	13	5	5

[a] 23 spontaneous and 96 UV-induced mutants were sequenced.

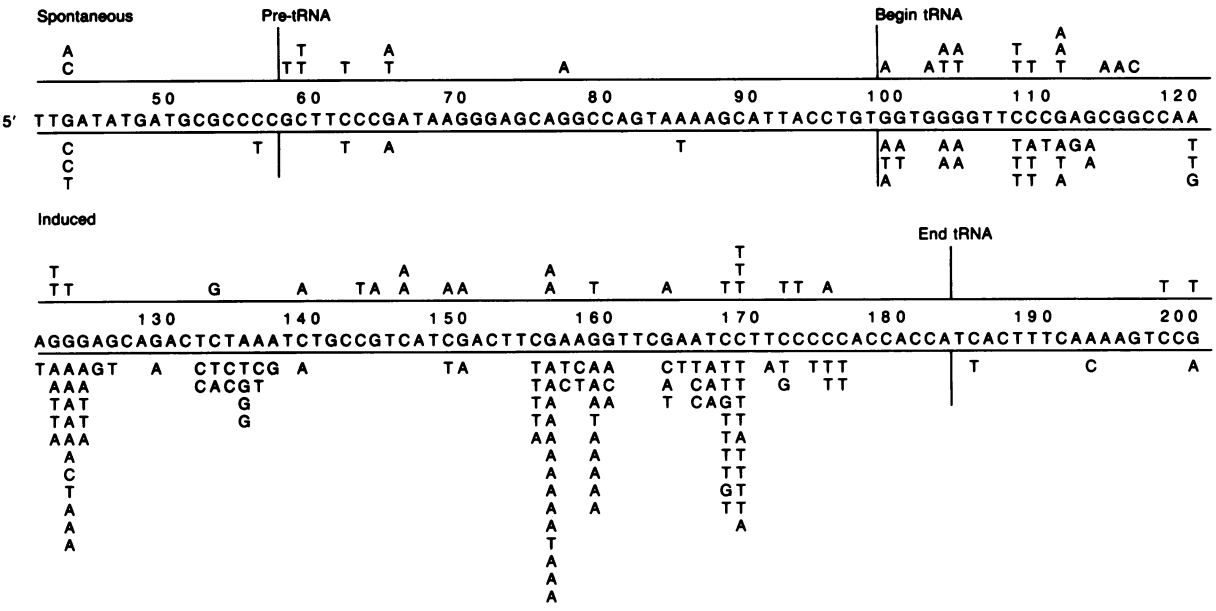

FIG. 2. Distribution of mutations in the tRNA region of pZ189. The DNA strand synthesized in the DNA sequencing reaction with the *Eco*RI rightward-sequencing primer is shown, along with the base substitutions observed in spontaneous and UV-induced mutants that were sequenced.

Determination of sites of UV-induced photoproducts. To determine whether there is a correlation between the positions and frequencies of the UV-induced mutations and the sites and frequencies of UV-induced photoproducts in the DNA, we measured the positions and relative frequencies of UV photoproducts in the tRNA region of our vector DNA. First, we used a DNA synthesis-arrest assay developed by Moore et al. (17) to identify the positions of photoproducts in the DNA template that block the progression of avian myeloblastosis virus reverse transcriptase. Supercoiled double-stranded vector DNA was irradiated at 100 or 1,000 J/m^2, annealed with primer, and incubated with avian myeloblastosis virus reverse transcriptase in the presence of labeled and unlabeled deoxyribonucleoside triphosphates. The DNA fragments synthesized were then electrophoresed in the gel system used for DNA sequencing (Fig. 3). Two different primers were used to allow determination of arrest sites on both strands of the DNA in the tRNA gene region (see Materials and Methods). DNAs from the four dideoxy sequencing reactions carried out on an unirradiated template with the same primers were electrophoresed on the same gel to serve as DNA size markers. As observed by Moore et al. (17), a distinct pattern of DNA bands was seen. This pattern did not vary significantly over a range of UV fluences from 100 to 1,000 J/m^2, although the relative amount of higher-molecular-weight DNA decreased as a function of increasing UV fluence, as expected. Presumably, these bands are generated by an arrest of the avian myeloblastosis virus reverse transcriptase at the sites of UV photoproducts in the DNA template (17). We assume that most of the UV photoproducts that cause synthesis arrest are pyrimidine-pyrimidine dimers, either cyclobutane pyrimidine dimers or 6-4 photoproducts (4, 18). The relative intensities of the bands (determined by densitometry of autoradiograms of the gels) should reflect the relative frequencies with which these photoproducts are formed at different sites along the DNA. From these data we can calculate the relative frequency with which the pyrimidine of each base pair participates in dimer formation with its neighbors (Fig. 4). There appeared to be photoproducts formed at virtually every pyrimidine-pyrimidine site; formation at TT sites appeared to be more frequent than at CT and TC, and formation at these sites was more frequent than at CC sites. These results are consistent with the known relative frequencies of formation of cyclobutane pyrimidine dimers and 6-4 photoproducts (4, 18).

Sequence specificity of UV-induced mutations. If we assume that the single mutations (single-base changes and tandem double-base changes) occur as targeted mutations opposite UV photoproducts, we can determine which DNA sequence positions are most likely to give rise to UV-induced mutations and compare these with the frequencies of UV photoproducts at these sites. We observed (Fig. 4) that UV-induced mutations occurred only at sites where UV photoproducts were formed; however, it should be noted that only 9 of the 85 base pairs of the tRNA gene sequence could not participate in dimer formation. Also, it is clear that hot spots for UV-induced mutations did not correspond to hot spots for UV photoproduct formation. In particular, A · T base pairs were underrepresented among the mutational sites. A tabulation of the types of dimer sites present and the mutations observed is presented in Table 5. Although 37% of the possible dimer sites involved T, and these sites were favored for dimer formation, only 21% of the single-base changes were at T sites, and none of the tandem double-base changes occurred at TT sites. At CC sites, double-base changes occurred preferentially. Single-base changes occurred almost twice as often at 5′-TC-3′ sequences as at 5′-CT-3′ sequences. However, it should be noted that this bias was largely due to the hot spot at position 156.

DISCUSSION

We have used an SV40-based shuttle vector to analyze UV-induced mutations occurring in mammalian cells. Mutants induced by UV radiation occur at a frequency about 20 times that of spontaneous mutations in this system; there-

fore, we can analyze the effects of UV damage on the spectrum of mutations observed. Most of the UV-induced mutations are point mutations; only about 5% are deletions. Although the design of the vector selects against deletion mutations, almost half of the spontaneous mutations are deletions. The spectra of the spontaneous and UV-induced point mutations differ; UV-induced mutations represent a broader range of base transitions and transversions, and mutations involving cytosine are favored over mutations involving thymine, despite the preference for photoproduct formation at TT sites. The G · C→A · T transition is favored over the other transition and all four transversions. We observed multiple mutations in both spontaneous and UV-induced mutants, but these appear to be clustered more closely in the UV-induced mutants. From these data, we conclude the following: (i) the pZ189 shuttle vector provides a means of analyzing induced mutations that occur in mammalian cells; (ii) the spectrum of base changes observed in the UV-induced mutants is similar to that observed by Miller et al. (16) in bacterial cells, suggesting that mutational mechanisms may be similar in the two systems; (iii) the preference for the G · C→A · T transition among UV-induced mutations suggests that the DNA polymerase(s) responsible for these mutations prefer to insert adenine opposite photoproducts in the DNA; (iv) the clustering of UV-induced multiple mutations suggests that these may arise due to a loss in replicational fidelity during the gap-filling step of DNA synthesis on UV-damaged tempates; (v) spontaneous point mutations probably occur during the error-prone repair of DNA which has been nicked or gapped during transfection.

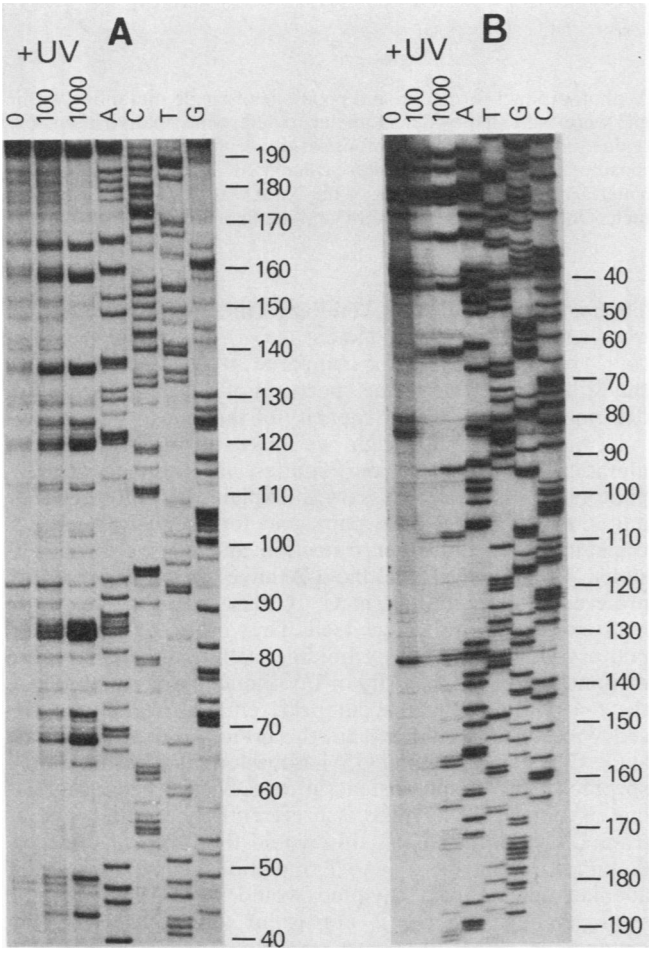

FIG. 3. Autoradiograms of DNA sequencing gels with the DNA synthesis arrest assay. Avian myeloblastosis virus reverse transcriptase was used with *Eco*RI rightward (A) and leftward (B) sequencing primers (see Materials and Methods) to determine synthesis arrest sites on both strands of UV-irradiated (100 and 1,000 J/m^2) and unirradiated (0) pZ189 DNA. The last four lanes in each panel contain the four dideoxyribonucleotide sequencing reactions. Numbers along the side refer to the base pair positions on the pZ189 plasmid. The experimental results pictured here were supplied by N. Tuteja.

The pZ189 shuttle vector has two major advantages over others designed for this purpose (5, 10, 21, 22). First, the effective target for deletion mutations is small. The tRNA marker gene itself is small and is placed between the two genes required for seletion (β-lactamase) and growth (plasmid replication function) in bacterial cells. This results in a reduction in the frequency of spontaneous deletion mutations recovered in bacteria after passage of the vector in mammalian cells. A similar vector described previously (28), in which the tRNA gene is located within a twofold-larger unselected region, exhibits about a sixfold-higher spontaneous mutation frequency. It should be noted that the formation of deletion mutations probably requires at least two separate events (21) so that the frequency of deletion mutations would not be a linear function of target size. Second, all types of point mutations, including multiple mutations, can be recovered and identified. In our system, there appears to be little limitation on mutational position within the tRNA gene or on the type of base change required. More than half of the bases in the tRNA molecule are involved in the base pairing that maintains the tRNA secondary structure (Fig. 5), and many of the unpaired bases in the loops are involved in specific interactions with the tRNA synthetase or ribosomes or in codon recognition (23). There are several bases that are invariant in all tRNAs examined (Fig. 5). Single-base changes that reduce or eliminate tRNA function have been identified at 27 different sites in the tRNA gene. Tandem double mutations involve an additional nine bases. The tRNA gene is relatively rich in runs of pyrimidines so that, in 76 of 85 of the base pairs, the pyrimidine of the base pair has the potential for participation in pyrimidine dimer formation. Although this reduces the number of nondimer sites available for assessing the contribution of nontargeted mutagenesis to UV-induced mutagenesis, it maximizes the number of sites available for assessing the types of mutational changes and their frequencies at potential dimer sites.

It is instructive to compare the spectrum of UV-induced base changes observed here to those observed in *E. coli* (Table 6). The most extensively used system for studying forward mutations in *E. coli* is the *lacI* system developed by Miller (see reference 15), in which nonsense mutations in the *lacI* gene are analyzed by genetic techniques. The data of Miller (Table 6) indicate that, as in our system, the G · C→A · T transition predominates. This bias is even more marked in the system used by Kunz and Glickman (8), in which a UV-irradiated F'Lac episome is transferred to either an unirradiated or a UV-irradiated recipient for mutation induction (Table 6). Unfortunately, the A · T→G · C transition is not easily detected in this system, so the frequency of this transition cannot be compared. Among 40 *lacI* missense mutants sequenced in the Miller study (15), however, the A · T→G · C transition occurred at about the same frequency as did each of the four transversions (Table 6). In the two phage systems studied (lambda and M13) forward mutations

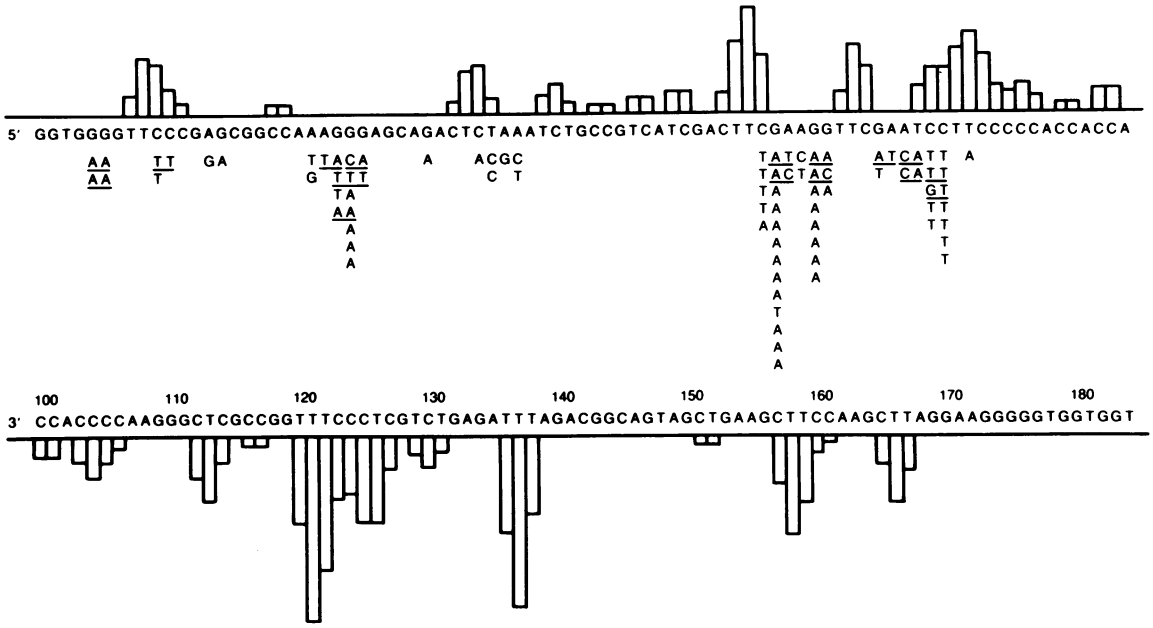

FIG. 4. Relative frequencies of participation of each pyrimidine in UV photoproduct formation and positions of single mutations within the tRNA coding sequence of pZ189. Autoradiograms of gels shown in Fig. 3 were evaluated by densitometry to determine relative intensities of individual bands. Corrections were made for increases in band intensity due to increased labeling of higher-molecular-weight DNA. From these data and the positions of the bands on the gels, the frequencies and positions of the arrest sites were calculated. The relative frequency with which each pyrimidine participates in photoproduct formation (indicated by the relative heights of the bars) was calculated, assuming that each arrest site is immediately 5' to a pyrimidine-pyrimidine dimer. Only single-base changes and tandem double-base changes (underlined) are shown here.

were characterized by direct DNA sequencing (Table 6). The two transitions were found to occur at about equal frequencies, and they were much more prevalent than any of the transversions. Since the M13 phage DNA is single stranded, each single-base change can be monitored separately; in Table 6, they have been combined for ease of comparison. In this system, the transitions C→T and T→C account for almost all of the transition mutations observed.

If, as suggested by Miller (15), the G · C→A · T transition occurs much more frequently than does the A · T→G · C transition in the *lacI* system, then it appears that the relative frequencies of base changes may depend not only on the inducing agent (in this case, UV radiation) but also on repair functions in the cell or the mode of DNA replication or both.

TABLE 5. Sequence specificity of UV-induced mutations in tRNA coding region, bases 99 through 183

Presumptive dimer sites[a]	No. of possible sites	Total no. of changes observed	
		Single	Double
Thymine sites			
XTT, TTX, TTT	8	5	0
TTC, XTC	12	1	3
CTT, CTX	7	4	1
CTC	3	1	
Cytosine sites			
XCC, CCX, CCC	24	5	18
CCT, XCT	7	13	1
TCC, TCX	12	21	3
TCT	3	2	

[a] X stands for adenine or guanine.

The results of studies by Todd and Glickman (33), in which mutagenesis in repair-proficient (*uvr*$^+$) and repair-deficient (*uvr*$^-$) bacterial strains was compared, suggest that at least in the Miller *lacI* system, the spectrum of base changes is not determined by the repair capacity of the cell.

If we assume (although we have no proof) that the mutations observed in our studies are targeted by UV photoproducts, it is necessary to explain the preference for mutations at G · C base pairs and for the G · C→A · T transition over the other transition and all four transversions. Brash and Hazeltine (4) have proposed that the preference for mutations at G · C base pairs is due to the nature of the DNA damage itself. They proposed that the 6-4 photoproducts, and not pyrimidine cyclobutane dimers, are responsible for the majority of UV-induced mutations. Since the 6-4 photoproducts occur preferentially at TC and CC sequences, this would explain the preference for mutations at G · C base pairs. Miller (15) has argued that the mutational specificity may reside instead in the DNA polymerase(s). If one assumes that adenine is preferentially inserted across from UV photoproducts, this would then explain the predominance of the G · C→A · T transitions, since insertion of adenine across from thymine would yield the wild-type sequence. This model is consistent with our data from mammalian cells. First, we observe tandem double-base changes mostly at CC sites and not at TT sites, despite the fact that single mutations are observed at these sites with about equal frequency. This observation argues that photoproducts at TT sites (almost exclusively pyrimidine cyclobutane dimers) are mutagenic, but formation of double-base changes at these sites is disfavored. This finding could be explained by a preference for insertion of adenine opposite UV photoproducts. Second, we observe a large number of mutations at cytosine sites in both CT and TC sequences.

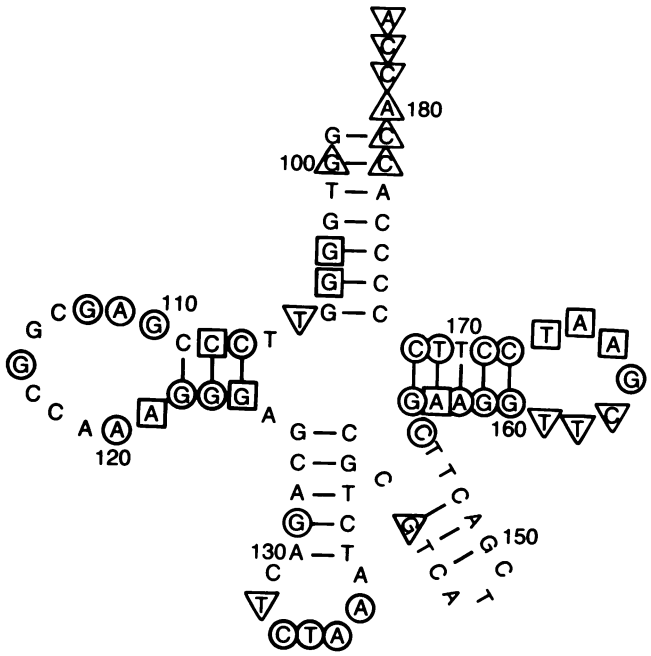

FIG. 5. Positions of single mutations within the folded structure of the tRNA gene sequence. The DNA sequence corresponding to the *supF* tRNA is shown in the normal cloverleaf structure. Positions of single mutations that were found in this study to reduce or eliminate tRNA function are indicated by circles (single-base changes) and boxes (tandem double-base changes). Positions of additional single-base mutations described in other studies (31) are indicated by triangles. Positions of invariant nucleotides (23) are indicated by upside-down triangles.

The preference observed for the TC site is due largely to the single mutational hot spot at position 156. Because 6-4 photoproducts are not detected at CT sites, mutations at these sites are probably due to pyrimidine cyclobutane dimers. The strong preference for mutations at the cytosine positions of these sites could again be explained by a preferential insertion of adenine. To explain the preference for the G · C→A · T transition over the transversions, it has been shown that, at least with DNA polymerase I from *E. coli*, purines are preferred over pyrimidines for insertion opposite UV photoproducts (20), a similar preference may exist for mammalian polymerases as well.

Another aspect of our data that needs to be explained is the presence of a significant number of multiple mutations in the UV-induced mutants. The presence of multiple mutations was also observed in the bacterial and phage systems in which mutants were characterized by DNA sequencing (8, 11, 15, 36). As discussed above, multiple mutations would not be detected in the analysis of nonsense mutations, as in the *lacI* system. It is interesting to note that, in our vector, the majority of these multiple changes are confined to short segments of the DNA, often about 15 bases. Even mutants that have multiple changes at greater distances within the tRNA gene region do not have additional mutations in a nearby nonselected region that we have also sequenced (base pairs 5320 to 5470; data not shown). It is unlikely that the multiple mutations we observe are due solely to the presence of multiple sites of UV-induced DNA damage. The UV fluence we have used (500 J/m^2) would be expected to introduce about one pyrimidine dimer per 200 bases of DNA. If none of the damage were repaired (which it is), multiple-targeted mutations would be expected to occur at a frequency of about 35% among mutants in our 160-base-pair sequence. However, these mutations would not be expected to be closely clustered, as we have observed, and it is very unlikely that we would observe such a high frequency of mutants with more than three changes. The possibility that our observation of a high frequency of multiple mutations is due to a selection bias in our system is made unlikely by the observation that 19 of 23 of the UV-induced multiple mutants we sequenced carry at least one mutation at a site that is represented among the single or tandem double mutations. Thus, these were not selected because of a requirement for more than one mutation in the tRNA gene to reduce the tRNA function sufficiently to score the mutant. The pres-

TABLE 6. Base changes (percent) of UV-induced forward mutations in *E. coli* compared with our data

Type of change	Base changes (%)						
	Our data	LacI				Lambda cI[a]	M13-lacZ[b]
		Resident F'*lac*[c]		F'*lac* transfer[d]			
		Nonsense	Sequencing	Untreated	Irradiated		
Transition							
G · C→A · T	61	63.3	60	94	95.9	47	28
A · T→G · C	7	ND[e]	7.5	ND	ND	37	41
Transversion							
G · C→T · A	13	14.3	15	2.4	1.1	4[f]	6
G · C→C · G	5	0.5	7.5	0.6	1.5	0[f]	6
A · T→C · G	4	4.7	7.5	1.2	0.4	2	5
A · T→T · A	9	16.2	2.5	2.4	1.1	10	13

[a] Data from Wood et al. (33). UV-irradiated (30 J/m^2) lambda phage were used to infect UV-irradiated bacteria (3 J/m^2). DNA sequencing was used to analyze 49 mutations.

[b] Data from LeClerc et al. (11). UV-irradiated M13mp2 phage were used to infect UV-irradiated bacteria. DNA sequencing was used to analyze 78 mutations.

[c] Data from Miller (15). Bacteria carrying the episome were irradiated with UV to ~3% survival. Forty mutants were analyzed by DNA sequencing and 1,046 nonsense mutations were analyzed by genetic techniques.

[d] Data from Kunz and Glickman (8). The episome was transferred to untreated (167 mutations analyzed) or UV-irradiated (20 J/m^2; 270 mutations analyzed) recipient bacteria from UV-irradiated donor (10 J/m^2).

[e] Not detected.

[f] One mutant reported as cytosine to purine was added to the G · C→T · A category.

ence of these multiple mutations could be explained by a localized loss of fidelity of DNA replication triggered by the presence of damage in the DNA template. According to the model that we developed to describe replication of UV-damaged SV40 DNA (see above) we would expect that these replicational errors would occur during the gap-filling process. Since the gaps left in daughter strands during replication of UV-damaged DNA appear to be between 50 and 150 bases long, loss of replicational fidelity during gap-filling would lead to the clustering of multiple mutations that we have observed. Alternatively, the multiple mutations may be due to an error-prone excision repair mechanism. The CV-1 cells are repair proficient (35) and survival data from studies with this vector (and also with SV40 itself) suggest that a large amount of UV-induced DNA damage is repaired. It is intriguing that the majority of multiple mutations lie within the distance (about 15 base pairs) known to be the size of bacterial *uvr*-mediated repair patches (25). In mammalian cells, repair patches are also thought to be quite small (30).

Miller et al. (16) recently published a genetic analysis of spontaneous suppressible nonsense mutations induced in the *lacI* region of an SV40-based shuttle vector. Ninety-three independent mutations were examined and found to result entirely from G · C→A · T or G · C→T · A changes. These changes also predominated in our collection of spontaneous mutants. In addition, we observed the G · C→C · G transversion. The presence of multiple mutations among the majority of spontaneous mutants that we observed by DNA sequencing would not be detected in the genetic analysis of nonsense mutations in the *lacI* system. Miller et al. (16) proposed that the spontaneous point mutations might result from miscoding due to depurination of guanine and deamination of cytosine. They suggested that these reactions might be enhanced in transfected DNA because of an initial localization of the DNA in lysosomes where the pH is low. Our observation that the majority of spontaneous mutants have multiple-base changes suggests that they may arise instead during repair of DNA that was damaged by nucleases during transfection (10, 21). This repair reaction appears to be highly error-prone.

In conclusion, the SV40-based shuttle vector pZ189 has been shown in this study to allow a detailed analysis of the spectrum of UV-induced mutations in mammalian cells. This vector should be useful in the further characterization of the molecular mechanisms of UV mutagenesis as well as mutagenesis by other carcinogenic agents. For example, treatment of both the host cells and the plasmid DNA with UV radiation or other DNA-damaging agents will allow the evaluation of the contribution that inducible recovery pathways (similar to the SOS pathway in bacteria) may have on the spectrum or frequency of induced mutations. In addition, since the vector can replicate in human cells, the influence on mutagenesis of repair deficiencies associated with certain human diseases (or aging) can be investigated. Further analysis of the spontaneous mutations that occur as a consequence of DNA transfection may lead to a better understanding of and appreciation for the plasticity of transfected DNA in mammalian cells and the pitfalls that may be encountered in the use of DNA transfection for characterization of normal and mutated cellular genes as well as for gene replacement therapy.

ACKNOWLEDGMENTS

We thank Arthur S. Levine for many helpful discussions during the course of this work and for critical reading of the manuscript. We also thank Daniel W. Nebert for his editorial comments. We thank our colleague Narendra Tuteja for performing the experiment shown in Fig. 3 and Michael Berman for generously providing the leftward sequencing primer. We greatly appreciate the computer analyses of our data performed by Peter Munson. We also acknowledge Eve Burrell for her expert assistance in preparation of the manuscript.

LITERATURE CITED

1. **Barnett, S. W., E. M. Landaw, and K. Dixon.** 1984. Test of models for replication of simian virus 40 DNA following ultraviolet irradiation. Biophys. J. **46**:307–321.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513.
3. **Bishop, J. M.** 1983. Cancer genes come of age. Cell **32**:1018–1020.
4. **Brash, D. E., and W. A. Haseltine.** 1982. UV-induced mutation hotspots occur at DNA damage hotspots. Nature (London) **298**:189–192.
5. **Calos, M. P., J. S. Lebkowski, and M. R. Botchan.** 1983. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. USA **80**:3015–3019.
6. **Coulondre, C., and J. H. Miller.** 1977. Genetic studies of the *lac* repressor. IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **117**:577–606.
7. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **20**:365–369.
8. **Kunz, B. A., and B. W. Glickman.** 1984. The role of pyrimidine dimers as premutagenic lesions: a study of targeted *vs.* untargeted mutagenesis in the *lacI* gene of *Escherichia coli*. Genetics **106**:347–364.
9. **Lawrence, C., R. Christensen, J. R. Christensen, and T. O'Brien.** 1983. Mechanisms of UV mutagenesis in yeast and *E. coli*. UCLA Symp. Mol. Cell. Biol. New Ser. **11**:511–520.
10. **Lebkowski, J. S., R. B. DuBridge, E. A. Antell, K. S. Greisen, and M. P. Calos.** 1984. Transfected DNA is mutated in monkey, mouse, and human cells. Mol. Cell. Biol. **4**:1951–1960.
11. **LeClerc, N., L. Istock, B. R. Saran, and R. Allen, Jr.** 1984. Sequence analysis of ultraviolet-induced mutations in M13*lacZ* hybrid phage DNA. J. Mol. Biol. **180**:217–237.
12. **Lusky, M., and M. Botchan.** 1981. Inhibition of SV40 replication in simian cells by specific pBR322 DNA sequences. Nature (London) **293**:79–81.
13. **McCutchen, J. A., and J. S. Pagano.** 1968. Enhancement of the infectivity of SV40 deoxyribonucleic acid with diethylaminoethyl dextran. J. Natl. Cancer Inst. **41**:351–357.
14. **Miller, J. H.** 1982. Carcinogens induce targeted mutations in *Escherichia coli*. Cell **31**:5–7.
15. **Miller, J. H.** 1985. Mutagenic specificity of ultraviolet light. J. Mol. Biol. **182**:45–68.
16. **Miller, J. H., J. S. Lebkowski, K. S. Greisen, and M. P. Calos.** 1984. Specificity of mutations induced in transfected DNA by mammalian cells. EMBO J. **3**:3117–3121.
17. **Moore, P. D., K. K. Bose, S. D. Rabkin, and B. S. Strauss.** 1981. Sites of termination of *in vitro* DNA synthesis on ultraviolet- and *N*-acetylaminofluorene-treated ϕX174 templates by prokaryotic and eukaryotic DNA polymerases. Proc. Natl. Acad. Sci. USA **78**:110–114.
18. **Patrick, M. H., and R. O. Rahn.** 1976. Photochemistry of DNA and polynucleotides: photoproducts, p. 35–95. *In* S. Y. Wang (ed.), Photochemistry and photobiology of nucleic acids, vol. 2. Academic Press, Inc., New York.
19. **Prakash, L., and F. Sherman.** 1973. Mutagenic specificity: reversion of iso-1-cytochrome c mutants of yeast. J. Mol. Biol. **79**:65–82.
20. **Rabkin, S. D, P. D. Moore, and B. S. Strauss.** 1983. In vitro bypass of UV-induced lesions by *Escherichia coli* DNA polymerase I: specificity of nucleotide incorporation. Proc. Natl. Acad. Sci. USA **80**:1541–1545.
21. **Razzaque, A., S. Chakrabarti, S. Joffe, and M. Seidman.** 1984. Mutagenesis of a shuttle vector plasmid in mammalian cells. Mol. Cell. Biol. **4**:435–441.
22. **Razzaque, A., H. Mizusawa, and M. M. Seidman.** 1983. Rear-

rangement and mutagenesis of a shuttle vector plasmid after passage in mammalian cells. Proc. Natl. Acad. Sci. USA **80**:3010–3014.

23. **Rich, A., and U. L. RajBhandary.** 1976. Transfer RNA: molecular structure, sequence and properties. Annu. Rev. Biochem. **45**:805–860.
24. **Robb, J. A., and K. Huebner.** 1973. Effect of cell chromosome number on simian virus 40 replication. Exp. Cell Res. **81**:120–126.
25. **Sancar, A., and W. D. Rubb.** 1983. A novel repair enzyme: UVRABC excision nuclease of *Escherichia coli* cuts a single strand on both sides of the damaged region. Cell **33**:249–260.
26. **Sanger, F., S. Nicklen, and A. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5476.
27. **Santos, E., E. P. Reddy, S. Pulciani, R. J. Feldmann, and M. Barbacid.** 1983. Spontaneous activation of a human proto-oncogene. Proc. Natl. Acad. Sci. USA **80**:4679–4683.
28. **Sarkar, S., U. B. Dasgupta, and W. C. Summers.** 1984. Error-prone mutagenesis detected in mammalian cells by a shuttle vector containing the *supF* gene of *Escherichia coli*. Mol. Cell. Biol. **4**:2227–2230.

28a. **Seidman, M. M., K. Dixon, A. Razzaque, R. Zagursky, and M. L. Berman.** 1985. A shuttle vector plasmid for studying carcinogen-induced point mutations in mammalian cells. Gene **38**:233–237.

29. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions, pp. 169–170. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Smith, C. A.** 1978. Removal of T4 endonuclease V sensitive sites and repair replication in confluent human diploid fibroblasts, p. 311–314. *In* P. C. Hanawalt, E. C. Friedberg, and C. F. Fox, (ed.), DNA repair mechanisms. Academic Press, Inc., New York.
31. **Smith, J. D.** 1972. Genetics of transfer RNA. Annu. Rev. Genet. **6**:235–256.
32. **Stacks, P. C., J. H. White, and K. Dixon.** 1983. Accommodation of pyrimidine dimers during replication of UV-damaged simian virus 40 DNA. Mol. Cell. Biol. **3**:1403–1411.
33. **Todd, P. A., and B. W. Glickman.** 1982. Mutational specificity of UV light in *Escherichia coli*: indications for a role of DNA secondary structure. Proc. Natl. Acad. Sci. USA **79**:4123–4127.
34. **White, J. H., and K. Dixon.** 1984. Gap filling and not replication fork progression is the rate-limiting step in the replication of UV-damaged simian virus 40 DNA. Mol. Cell. Biol. **4**: 1286–1292.
35. **Williams, J. I., and J. E. Cleaver.** 1978. Excision repair of ultraviolet damage in mammalian cells. Evidence for two steps in the excision of pyrimidine dimers. Biophys. J. **22**:265–279.
36. **Wood, R. D., T. R. Skopek, and F. Hutchinson.** 1984. Changes in base sequence induced by targeted mutagenesis of lambda phage by ultraviolet light. J. Mol. Biol. **173**:273–291.
37. **Zagursky, R. J., M. L. Berman, K. Baumeister, and N. Lomax.** 1985. Rapid and easy sequencing of large double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Tech. **2**:89–94.