Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector

C.W.Lawrence*, A.Borden, S.K.Banerjee and J.E.LeClerc[1]
Departments of Biophysics and [1]Biochemistry, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

Received November 6, 1989; Revised and Accepted January 8, 1990

ABSTRACT

We have investigated the mutagenic properties of an abasic site in DNA by transfecting SOS-induced and uninduced cells of *E. coli* with a single-stranded M13mp7-based vector that carries a single example of this lesion at one or other of two unique and adjacent sites. Random samples of progeny phage were sequenced to determine the nature of the replication events that occurred at and around these locations. 5% to 7% of the vectors could be replicated in SOS-induced cells, but only 0.1% to 0.7% of them gave plaques in the absence of SOS induction. In SOS-induced cells, 93% and 96% of the phage replicated resulted from the insertion of a nucleotide opposite the abasic site, while the remainder resulted from a targeted omission of a single nucleotide. At one of the sites, nucleotide insertions were 54% dAMP, 25% dTMP, 20% dGMP and 1% dCMP. At the other site they were 80% dAMP, 4% dTMP, 15% dGMP and 1% dCMP. The sequence variation in all but two of the 204 sequences analyzed was restricted to the abasic site itself. In the remaining two, a change at the abasic site was accompanied by a mutation at an immediately flanking nucleotide.

INTRODUCTION

Single-stranded vectors that carry a defined, uniquely located lesion are powerful tools for investigating the properties of the modified nucleotide and also of SOS-induced replication (1, S.K.Banerjee, A. Borden, R.B.Christensen, J.E.LeClerc and C.W. Lawrence, unpublished data). The abasic site is unusually interesting, among the various lesions that might be studied, because it is a relatively common spontaneous and mutagen-induced event (reviewed in (2)) and, in particular, because it is the clearest case of a non-informational alteration in DNA. In this latter respect, the abasic site is a benchmark against which other lesions can be measured; a comparison of mutational spectra can indicate whether the other modified nucleotides retain base-pairing specificity.

In addition, the abasic site has interest in its own right, because it reveals properties of the SOS-induced replication complex that are difficult to observe with other template modifications (3). Since abasic sites are incapable of normal hydrogen bonding, the particular nucleotide found opposite them is likely to be determined by other factors, such as polymerase imposed bias in productive nucleoside triphosphate binding, variable stabilization of dNTP's by stacking interactions with the previously added base, editing bias of 3'-5' exonuclease, variable elongation efficiency from different 3' termini, nucleotide hydrophobicity, and dNPT pool biases (2−7). Evidence indicating that several DNA polymerases show a marked tendency for the non-instructive insertion of dAMP has been obtained from a variety of experiments, using randomly depurinated or depyrimidated DNA replicated *in vivo* or *in vitro*, or using templates containing a synthetic analogue of an abasic site replicated *in vitro* (3−6, 8−11, see (2) for review). Direct evidence for this phenomenon *in vivo* has not yet been obtained, however, and similarly the possible influence of other factors discussed above has not yet been examined.

We have investigated these issues by transfecting SOS-induced or uninduced cells of a *uvrA6* strain of *E. coli* with a single-stranded DNA vector that contained a single abasic lesion at one or other of two unique and adjacent sites. The consequence of replicating these vectors was determined by sequence analysis of a random sample of progeny phage, without the aid of a reporter gene or sequence. We find that while dAMP was preferentially inserted at both of the abasic sites studied, the extent of the preference varied considerably. Similarly, the insertion frequency of dTMP and dGMP also varied at the two sites. A nucleotide was inserted opposite the abasic lesion in almost all cases; omission of a base, giving a single nucleotide deletion, occurred in only 4% and 7% of the phage. A single abasic site was extremely toxic in uninduced cells, but at least 5% to 7% could be replicated after SOS induction.

MATERIALS AND METHODS

Vector construction

Single-stranded DNA of M13mp7L1 was linearized by digestion at room temperature with 8 units of EcoRI per μg of DNA, and converted to a gapped circular species (figure 1) by annealing it at a concentration of 100 ng/μl in 25 mM NaCl with a twofold

* To whom correspondence should be addressed

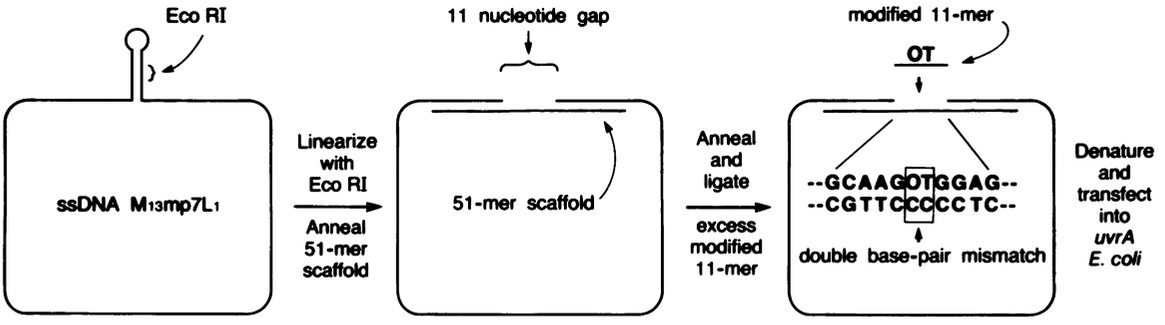

Figure 1. Method for constructing single-stranded vectors that contain a uniquely located abasic site.

molar excess of a 51-mer scaffold oligomer. Annealing was carried out at room temperature, overnight. The annealing mix was concentrated twofold by evaporation, and abasic or control T-T 11-mer ligated into the gap at 14°C overnight, using a 100-fold molar excess of oligomer and 0.1 units/μl of T4 polynucleotide ligase (Pharmacia).

Oligonucleotides

The 51-mer and the control 11-mer, with sequence 5'GCAAGTTGGAG3', were synthesized on an Applied Biosystems (Foster City, CA) model 380A synthesizer, while the uracil-containing 11-mers 5'GCAAGUTGGAG3' and 5'GCAAGTUGGAG3' were synthesized by Pharmacia (Milwaukee, WI). All oligonucleotides were deprotected by incubation in 30% ammonium hydroxide at 55°C overnight. Each of the 11-mers was purified by HPLC using a Machery-Nagel 60−7 DEAE-silica column as described (1), while the 51-mer was partially purified by electrophoresis through a 20% polyacrylamide gel. The 11-mers were phosphorylated, using T_4 polynucleotide kinase, and the kinased species purified by HPLC, using a Waters Radial Compression Nova-pak C_{18} reversed phase column as described (1). Abasic sites were introduced into the uracil-containing 11-mers by treating 12 μg of oligomer in buffer ((12), but using Tris HCl rather than HEPES) with 80 units of purified *E. coli* uracil N-glycosylase (kindly supplied by Susan Wallace, University of Vermont), at 37°C for 60 min (1 unit of enzyme releases uracil at a rate of 1 pmol/min in these conditions). Reaction mixes were extracted once with phenol and twice with chloroform, and the abasic species purified by HPLC using a Waters radial compression Nova-pak C_{18} column, and a linear gradient of 7% to 9% acetonitrile in 0.1 M triethylammonium acetate (pH 7.0) over 20 min, followed by a linear gradient to 70% acetonitrile over a further 10 minutes. The abasic oligomers had a retention time of 17 to 18 min and the parental oligomers a retention time of 22 to 24 min. Fractions containing the abasic 11-mer were subjected to four cycles of evaporation to half the initial volume followed by the addition of an equal volume of water, a final evaporation to 50 μl, and ethanol precipitation. This procedure was used because evaporation to dryness, used to collect the control 11-mer, completely cleaves the abasic material, presumably because of high pH caused by traces of triethylamine. No cleavage could be detected using the above procedure. Purified material was dissolved in water and frozen at −25°C. Under these conditions, it was stable for at least a month.

Other methods

SMH10, an F$^+$*uvrA6* derivative of AB1157 was made competent by the $CaCl_2$ method. For SOS induction, cells were irradiated at a concentration of $2-4\times10^8$ cells/ml with 4 J m^{-2} of 254 nm UV immediately before making the cells competent. To establish that the enhanced transfection efficiency in SOS-induced cells resulted from bypass of the abasic site, ligated construct was treated with the AP endonuclease activity of exonuclease III (New England Biolabs) before denaturation of the 51-mer scaffold, using buffer that contained $CaCl_2$ rather than $MgCl_2$ to inhibit the exonuclease (13), and 2.5 units of enzyme per μg of vector. Further details of the experimental procedure can be obtained from reference (1).

RESULTS

Experimental method

Single-stranded vectors containing single abasic sites were constructed in two stages, using the method described in reference (1). In the first stage one or other of the synthetic 11-mers 5'GCAAGUTGGAG3' or 5'GCAAGTUGGAG3' were treated with *E. coli* uracil glycosylase and then purified by HPLC. In the second stage, the modified 11-mer was ligated into a gapped single-stranded circular structure, made by digesting the viral DNA of M13mp7L1 with EcoRI and recircularizing the linear ssDNA by annealing it to a 51-mer scaffold oligomer (figure 1). The ends of the 51-mer are complementary to the terminal 20 nucleotides at the ends of the cut vector, while the central 11 nucleotides are complementary to the 11-mer sequence, except at the central U-T or T-U site. At this position, the 51-mer possesses a tandem C-C mismatch, which serves as a genetic label. Immediately before transfection into SOS-induced or uninduced cells of SMH10, an F$^+$ *uvrA6* Δ(*pro lac*) derivative of AB1157, the ligated construct was heated at 90°C for 2 min, to denature the 51-mer, and the material diluted 100-fold to prevent reannealing. Exposure of ^{32}P-labelled abasic 11-mer to this temperature for up to 5 minutes failed to produce any detectable cleavage at the abasic site (data not shown). Exposure of the modified oligomers to alkali, however, efficiently cleaved them to yield ^{32}P-labelled fragments of the expected size (figure 2), showing that these 11-mers possess a lesion at the expected site.

Replication efficiencies of abasic vectors in SOS-induced and uninduced cells

Matched samples of vector containing either unmodified oligomer or 11-mer with an abasic site were transfected into competent cells of SMH10 that had been irradiated with 4 J m^{-2} UV or not irradiated. In one series of experiments the abasic site replaced the sixth nucleotide of the 11-mer, and in another series it replaced the seventh nucleotide. In a third series of experiments the 11-mer

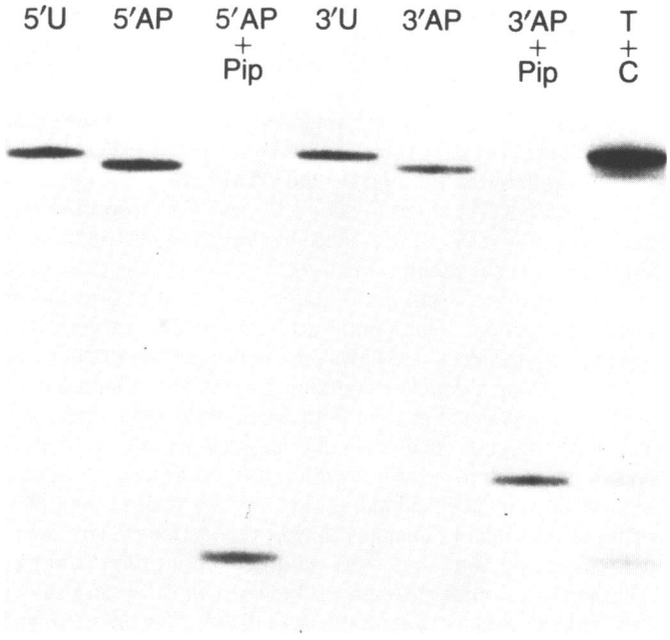

Figure 2. Polyacrylamide gel electrophoresis analysis of 11-mers abasic at the position of the 6th nucleotide (5′ site) or the 7th nucleotide (3′ site), treated with 1 M piperidine at 90°C for 30 min. Each oligomer was labelled at the 5′ end with ^{32}P. Lanes 1 and 4 contain 11-mer with uracil at the 5′ or 3′ sites, respectively. Lanes 2 and 5 contain 11-mer with an abasic site in one of these two positions, while lanes 3 and 6 contain such oligomers treated with piperidine. Lane 7 contains the T + C reactions from a Maxam and Gilbert sequencing procedure, which gives size standards for the cleaved material.

Table 1. Replication efficiency, percent of the T-T control, of vectors carrying a single abasic site (O) or a single uracil (U)

SOS	% vectors replicated				
induction	T-T	O-T	T-O	U-T	T-U
−SOS	100	0.6	0.3	0.7	0.1
+SOS	100	6.9	5.0	nd	nd

Data are the average of between 2 and 4 experiments, and 5 to 8 plates/table entry in each experiment. Average numbers of plaques per plate in the T-T controls were 126 (−SOS) and 144 (+SOS). Data exclude progeny of uncut M13mp7L1, or those (containing GG) derived from the undenatured construct. Such phage, detected by sequence analysis, constitute 2% to 3% of all constructs. Cells were grown in the *ung*$^+$ strain, SMH10.
nd, not determined.

Table 2. Replication efficiency, percent of the T-T control, of vectors carrying a single uracil, grown in *ung*$^-$ and *ung*$^+$ strains without SOS induction

Strain	% vectors replicated		
	T-T	U-T	T-U
ung$^-$	100	90	54
ung$^+$	100	3	1

Results are the average of two experiments. Data from progeny of uncut M13mp7L1 and G-G phage were not excluded. Strains used were SMH77, an F′*lacZ* ΔM15 *pro*+, Δ(*pro-lac*) *leu*+ derivative of AB1157, and SMH99, an *ung* derivative of SMH77, prepared by P1 transduction using a lysate from BW280 (*ung*-1::Tn10).

contained uracil at either the sixth or seventh position, and the vector was transfected into uninduced cells only. As shown in Table 1, vectors carrying a single abasic site were replicated with

Table 3. Effect of the AP endonuclease activity of exonuclease III on the replication efficiency of vectors carrying a single abasic site, transfected into SOS-induced cells

Treatment	% vectors replicated		
	T-T	O-T	T-O
−AP endo	100	6.4	6.9
+AP endo	105	0.7	0.0

Data are the average of two experiments. Results from the abasic vectors have been corrected for the number of plaques found in uninduced cells, and are expressed as% of the T-T control.

only very low efficiency in uninduced cells. The small number of progeny phage obtained were nevertheless likely to have resulted from the replication of vectors containing an abasic site, rather than from those carrying the unmodified parental 11-mer, because uracil containing vectors were also replicated with very low efficiency (Table 1). This observation implies that uracil is removed very efficiently from the DNA by the endogenous uracil glycosylase, and that this occurs before replication. Such an interpretation is supported by the finding that the efficiency with which uracil containing vectors were replicated was much increased in SMH99, an *ung*$^-$ strain, compared to the *ung*$^+$ strain SMH77 (Table 2). In SOS-induced cells, vectors carrying a single abasic site were replicated with an efficiency of between 5% and 7%, an increase of over 10-fold (Tables 1 and 3). The SOS-dependent increase in replication efficiency concerned abasic site containing vectors, rather than any contaminating species, because it could be abolished by treating the ligated construct with exonuclease III. This enzyme possesses an AP endonuclease activity which cleaves the phosphodiester chain 5′ to the site of base loss. Following denaturation, the vectors are therefore linearized, and such a species transfects with an efficiency that is 100- to 1000-fold lower than for the circular vector (unpublished data).

Nucleotide sequence of progeny phage

Nucleotide sequence was determined in the progeny of 104 phage carrying an abasic site at the position of the sixth nucleotide (designated OT), and of 100 that carried this lesion at the position of the seventh nucleotide (designate TO), all obtained from SOS-induced cells. Nucleotide sequence was determined in a region starting about 25 nucleotides on the 3′ side of the 11-mer insertion site and extending 80 to 100 nucleotides to the 5′ side of this location.

With only two exceptions, the sequence of all phage analyzed was entirely normal at sites other than that occupied by the abasic site, within the region examined. In the two exceptions, a nucleotide substitution was found immediately adjacent to the abasic site (Table 4). It cannot be eliminated that all or some of the single nucleotide deletions arose by insertion of dAMP at the abasic site, and deletion at the site of the immediately adjacent T, but the simpler interpretation of deletion at the abasic site itself is more probable.

As shown in Table 4, the most likely consequence of replication past the template abasic site was the insertion of a nucleotide at this site: failure to insert a nucleotide, giving a single base deletion, occurred in only 7% (OT) or 4% (TO) of the phage. At both sites, the addition of dAMP to the nascent strand, giving a viral strand T, was the most common insertion, but its frequency varied greatly at the two locations: 80% at the 3′ site (TO) but only 54% at the 5′ site (OT), a difference which is significant ($p < 0.001$). Similarly, the two sites differed with respect to the

Table 4. Numbers of mutant and normal sequences in SOS-induced cells transfected with vectors carrying a single abasic site (O)

	O-T	T-O
TT	52	77
AT	25[a]	0
CT	19	0
GT	1	0
TA	0	4
TC	0	14
TG	0	1
−1 nuc	7	4[b]
other	0	0
total	104	100

[a]One of these was A-A (OT→AA).
[b]One of these was A (TO→A).

Table 5. Nucleotide insertion frequency (%) opposite template abasic sites

Nucleotide inserted	*In vivo* 3′ terminal nucleotide				*In vitro* 3′terminal nucleotide					
	A[a]	C[a]	V[b]	V[c]	A[d]	C[d]	T[e]	T[f]	V[g]	V[h]
dAMP	54	80	56	81	88	80	89	71	65	74
dGMP	20	15	15	13	8	12	7	15	15	17
dCMP	1	1	–	6	2	4	<1	3	–	9
dTMP	25	4	29	–	2	5	4	9	20	–

[a]Present data, Table 4.
[b]Kunkel (1984) depurinated G.
[c]Kunkel (1984) depurinated A.
[d]Randall *et al.* (1987), *Drosophila* DNA polymerase α.
[e]Takeshita *et al.* (1987), structure II AMV reverse transcriptase.
[f]Takeshita *et al.* (1987), structure II calf thymus DNA polymerase α.
[g]Hevroni and Livneh (1988) depurinated G, *E. coli* DNA polymerase III holoenzyme.
[h]Hevroni and Livneh (1988) depurinated A, *E. coli* DNA polymerase III holoenzyme.
The data of Hevroni and Livneh are for nucleoside triphosphate turnover, rather than for insertion frequency.
V = various.

next most favored insertions: at the 5′ site, a template A, followed closely by a template C, were the next most common events, but at the 3′ site, a template C was over threefold more abundant than a template A. A single example of template G occurred at each site.

DISCUSSION

We have determined the identity of the nucleotide sequences that result from replication past a single abasic site, and also the frequency with which such templates are replicated, by transfecting SOS-induced and uninduced cells of *uvrA6* strain of *E. coli* with a single-stranded M13-based vector that carries this lesion at one or other of two unique sites. We find that the most common response to the presence of the abasic site in the template was the addition of dAMP to the end of the nascent chain, a result that is in qualitative agreement with the conclusions from earlier work using randomly depurinated or depyrimidated DNA, replicated *in vivo* or *in vitro*, and from studies using templates containing a synthetic analogue of an abasic site, replicated *in vitro* (3−6, 8−11, Table 5). It has been suggested that such a bias towards dAMP insertion arises from a polymerase preference for binding dATP (3) and more specifically, from a polymerase imposed constraint that favors productive binding of dATP, that is binding that leads to polymerization (5). Although our data agree qualitatively with earlier results, they differ from them quantitatively in several ways. In particular, the extent of the preference for dAMP insertion varies at the two sites studied: at the 5′ (OT) site 54% of phage resulted from dAMP insertion, while at the 3′ (TO) site the frequency was 80%, a difference that is significant ($p < 0.001$). Similarly, the two sites differ in relative preference for dGMP and dTMP insertion (20% and 25%, respectively, at the 5′ site; 15% and 4%, respectively, at the 3′ site, $p = 0.02-0.01$). Since the incoming nucleotide stacks against a terminal adenine when the 5′ site is abasic, and against terminal cytosine when the 3′ site is abasic, it is possible that different stacking interactions account for the varying base-insertion preferences. However, a kinetic analysis of nucleotide insertion, using a template containing a synthetic analogue of an abasic site replicated *in vitro* with *Drosophila* α polymerase (5, Table 5) suggests that stacking interactions are unlikely to contribute much to insertion preference, compared to variation in productive nucleotide binding. Possibly the relative importance of these two factors changes during replication *in vivo* with a natural abasic site and SOS-modified replication complex. Alternatively, other differences between *in vitro* and *in vivo* experiments may be important: specifically, the *in vitro* study excluded the effects of editing and elongation. Whether the 3′ to 5′ editing exonuclease is active in SOS-induced replication is not yet known, but if it is, editing efficiency may be sequence dependent. Similarly, the efficiency with which 3′ termini can be elongated may also depend on sequence.

Although an abasic lesion cannot participate in normal base-pairing, DNA polymerase nevertheless recognizes with high efficiency that its template site is an appropriate place for nucleotide insertion: as shown in Table 4, failure to add a base at this site, giving a single nucleotide deletion, occurs in only 4% and 7% of the phage, a frequency similar to that observed previously (10) with randomly depurinated DNA. The mechanism responsible for maintaining the accuracy of polymerase translocation along the template, and therefore for avoiding such deletions, is not known but perhaps depends either on contacts between the protein and the deoxyribose moiety of the nucletoide or on contacts with a large enough set of nucleotides such that loss of a single contact is of little consequence. It is also evident from Table 4 that the influence of an abasic lesion is confined strictly to its location: substitutions or deletions occurred only at the abasic site itself in 202 of 204 sequences analyzed. In the remaining two sequences, changes targeted at the abasic site were accompanied by substitutions at an immediately flanking site.

Our results also show that in most cells a single abasic site is likely to constitute a lethal event because it appears to block replication with high efficiency. In uninduced cells, only 0.3% to 0.7% of the vectors could be replicated completely. Replication efficiency rose to about 5% to 7% when SOS dependent functions were induced, suggesting therefore that this lesion is substantially more toxic than the *cis-syn* T-T cyclobutane dimer (1), the *trans-syn* T-T cyclobutane dimer (S.K.Banerjee, A.Borden, R.B.Christensen, J.E.LeClerc and C.W.Lawrence, unpublished data) or the 6,4 pyrimidine-pyrimidone dimer (J.E.LeClerc, A.Borden, S.K.Banerjee, and C.W.Lawrence, unpublished data). It is possible, however, that our estimates of abasic vector replication efficiency are underestimates, because the endogenous AP endonuclease activity of endonuclease VII, which is active on single-stranded DNA (14), may have linearized some of the constructs before replication, which would constitute a non-repairable lethal event. Finally, a comparison between the nucleotide sequence changes caused by the abasic sites and those

produced by the UV lesions above, all of which are located at the same vector site and within the same sequence, shows that each possesses a characteristic and different mutation spectrum. The lesions produced by SOS-dependent mutagens should not therefore be regarded as homogeneously 'non-instructive'. Instead, their mutagenic potential is likely to depend on a variety of interactions between template, nucleotide and polymerase and include, at least in some cases, correct base-pairing.

ACKNOWLEDGEMENTS

We thank Susan Wallace (University of Vermont, Burlington) for the generous gift of purified *E. coli* uracil glycosylase. This work was supported by grants No. GM32885 and GM27817 from the National Institutes of Health and grant No. DE-FG02-88ER60626 from the United States Department of Energy.

REFERENCES

1. Banerjee,S.K., Christensen,R.B., Lawrence,C.W. and LeClerc,J.E. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8141–8145.
2. Loeb,L.A. and Preston,B.D. (1986) *Ann. Rev. Genet.* **20**, 201–230.
3. Strauss,B., Rabkin,S., Sagher,D. and Moore,P. (1982) *Biochimie* **64**, 829–838.
4. Kunkel,T.A., Schaaper,R.M. and Loeb,L.A. (1983) *Biochemistry* **22**, 2378–2384.
5. Randall,S.K., Eritja,R., Kaplan,B.E., Petruska,J. and Goodman,M.F. (1987) *J. Biol. Chem.* **262**, 6864–6870.
6. Takeshita,M., Chang,C.-N., Johnson,F., Will,S. and Grollman,A.P. (1987) *J. Biol. Chem.* **262**, 10171–10179.
7. Kypr,J. (1988) *J. Theor. Biol.* **135**, 125–126.
8. Schaaper,R.M., Kunkel,T.A. and Loeb,L.A. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 487–491.
9. Sagher,D. and Strauss,B. (1983) *Biochemistry* **22**, 4518–4526.
10. Kunkel,T.A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1494–1498.
11. Hevroni,D. and Livneh,Z. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5046–5050.
12. Lindahl,T. (1980) In Grossman,L. and Moldave,K. (eds.) pp. 284–295. *Methods in Enzymology* Vol. **65**, Academic Press, New York.
13. Rogers,S.G. and Weiss,B. (1980) In Grossman,L. and Moldave,K. (eds.) pp. 201–211. *Methods in Enzymology* Vol. **65**, Academic Press, New York.
14. Friedberg,E.C., Bonner,T., Radany,E.H. and Love,J.D. (1981) In Boyer,P.D. (ed.) pp. 251–279. *The Enzymes* Vol. **XIV part A**. Academic Press, New York.