Mutational specificity of depurination
(bacteriophage M13mp2/transversion mutagenesis/SOS repair/base selection/noncoding lesion)

THOMAS A. KUNKEL
Laboratory of Genetics, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709
Communicated by James A. Miller, November 16, 1983

ABSTRACT The mutagenic consequences of damage to DNA produced by low pH and high temperature have been determined in a forward mutational system capable of detecting all classes of mutagenic events. When damaged single-stranded DNA from bacteriophage M13mp2 is used to transfect competent Escherichia coli cells, a 15-fold increase in mutation frequency, measured as loss of α-complementation by the lac DNS in the phage, is observed compared with an untreated DNA control transfection. The enhanced mutagenicity is largely dependent on induction of the error-prone SOS response and is proportional to the number of lethal hits introduced into the DNA. The effect is abolished by treatment of the damaged DNA before transfection with either apurinic/apyrimidinic endonuclease or alkali. Based on these observations and the rate constants for formation of the known heat/acid-produced lesions in DNA, it is concluded that the majority of the induced mutagenesis results from apurinic sites. DNA sequence analysis of 87 spontaneous and 124 induced mutants indicates that the major effect is on single base-substitution mutagenesis with a small increase in (deletion) frameshift frequency. Approximately 80% of the base-substitution mutations occur at purine positions in the viral strand, consistent with depurination as the predominant premutagenic lesion. The preference of guanine over adenine sites mutated is consistent with the preference for depurination of guanine over adenine. Transversions are observed for 57 of 79 (72%) induced base substitutions, with a strong preference for insertion of adenine residues opposite the putative apurinic site. These data in a forward mutational system provide insight into the mechanisms used by a cell to replicate DNA containing noncoding lesions.

Insight into the effects of damage to genetic information requires analysis of the interactions of cellular replication and repair processes with specific lesions in DNA. One such lesion that has received considerable attention in recent years is the abasic site (1, 2) produced by hydrolysis of the N-glycosyl bond between the sugar and base (3–5). This base loss can occur spontaneously at high frequency (3, 6) and this frequency can be increased either by base modifications resulting from certain DNA damaging agents (7–11) or enzymatically by the action of DNA glycosylases (1, 12). Apurinic/apyrimidinic (AP) sites are quite stable (4, 13), and cells have evolved mechanisms to repair these lesions (12, 14). Unrepaired AP sites have been shown to have two biological consequences, lethality (15–18) and base-substitution errors (18). These errors can be produced in vitro with purified DNA polymerases (19, 20) or can occur in vivo in Escherichia coli under conditions in which the SOS response has been induced (11, 18, 21). These results, obtained in an assay specifically designed to detect base-substitution errors that revert φX174 amber mutations, indicate that for a limited number of sites in φX174 DNA there is a strong preference for insertion of dAMP at apurinic sites (21), leading to distinctive G·C → T·A and A·T → T·A transversions.

The use of a single-base-substitution assay involving reversion of nonsense codons in an essential gene focuses on a very specific set of base changes at a limited number of sites. The inability of this system to detect all possible single-base changes as well as other classes of mutagenic events has led to the development of a forward mutational assay in a nonessential gene. This assay selects for loss of β-galactosidase “α-complementation,” encoded in the DNA of bacteriophage M13mp2 developed by Messing et al. (22). Because selection is for loss of a nonessential gene function, a wide spectrum of mutagenic events can be scored. The precise nature of the mutagenic events, whether base substitutions, frameshifs, deletions, additions, or rearrangements, can be determined by determining the sequences of the DNA of the mutants. This system is used here to determine the mutational consequences of a noncoding lesion, the apurinic/apyrimidinic site. Of particular interest are two issues, the importance of base-substitution mutations relative to other mutagenic events and the specificity of the induced base changes in a system capable of monitoring a wide spectrum of base changes at many sites.

MATERIALS AND METHODS

Bacteria and Bacteriophage. E. coli strains S90C [Δ(pro-lac), ara−, thi−, strA], S90C recA56, NR8036 [Δ(pro-lac), ara−, thi−, trpE9777], and NR8037 [Δ(pro-lac), ara−, thi−, trpE9777, umuC36::tn5] were provided by R. Dunn and B. Glickman of this Institute. E. coli strain CSH50 [ara−, thi−, Δ(pro-lac)/F'traD36, proAB, lacZΔM15] (the ΔM15 deletion spans 93 bases of the lacZ gene carried on the episome; see Fig. 1), and bacteriophage M13mp2 were obtained from J. E. LeClerc (Univ. of Rochester, Rochester, NY).

DNA Preparation. Single-stranded M13mp2 viral DNA was incubated at 70°C for various times in 30 mM potassium chloride/10 mM sodium citrate, pH 5.0. These conditions introduced one AP site per molecule in 4 min, measured by survival (18). Hydrolysis of AP sites was carried out either by incubating 10 μg of depurinated DNA for 30 min at 37°C with 20 units of apurinic/apyrimidinic endonuclease from HeLa cells [fraction VII (23)] (kindly provided by D. Mosbaugh and S. Linn, Univ. of California, Berkeley, CA) in 500 μl of 50 mM Tris-HCl, pH 7.5/10 mM MgCl2 or by incubation in 0.1 M NaOH (final pH 12.8) for 30 min at 37°C.

SOS Induction. The various strains of E. coli were grown at 37°C to a cell density of 2–4 × 108 cells/ml in YT medium (8 g of tryptone, 5 g of yeast extract, 5 g of NaCl per liter, pH 7.8), harvested by centrifugation, and then suspended in 0.9% NaCl at 10% of original volume. These cells were irradiated at 25–150 J/m2 in thin layers in plastic Petri dishes with constant gentle shaking. After removal of small aliquots for determination of cell survival, cells were diluted to 50% of their original volume in prewarmed YT medium and incubated in the dark for 45 min with vigorous aeration. These

Abbreviation: AP, apurinic/apyrimidinic.

1494
cells, or unirradiated cells grown to 2–4 × 10^6/ml in YT medium, were treated with CaCl_2 to produce competent cells as described (24).

**Transfection and Plating.** Transfections were carried out by diluting the DNA into 75 mM CaCl_2 and adding 2 vol of competent cells. The volume used was determined by the ratio of DNA molecules to cells, which was kept constant for all variables in a single experiment, typically at 20:1. The efficiency of transfection was 2 × 10^11 (50 plaques per ng of single-stranded DNA) and varied 2- to 5-fold over several experiments. After a 40 min incubation at 0°C, the mixture was heat shocked at 42°C for 3 min (or 5 min for volumes larger than 5 ml), then placed at 0°C. An amount of this mixture sufficient to yield 100–500 infective centers per plate was added to 3 ml of 0.8% agar in 0.9% NaCl containing 2 mg of 5-bromo-4-chloro-3-indolyl β-D-thiogalactoside and 0.2 ml of a logarhythmic phase culture of *E. coli* cells (CSSH50° *lacZΔM15*). This mixture was poured onto plates containing 30 ml of minimal medium solidified with 1.5% agar and containing 0.24 mg of isospropylthio-β-D-galactoside per plate. These plating conditions were specifically chosen to give intense blue color for wild-type M13mp2 plaques and to allow optimum visualization of mutant phenotypes.

**Scoring Mutants and Determination of DNA Sequences.** Inactivation of α-complementation (25) resulting from a mutation in the lac DNA in M13mp2 will give rise to mutants readily distinguished as lighter blue or colorless plaques (26). Mutants were scored after 12–15 hr of incubation at 37°C followed by 24 hr of incubation at room temperature. To eliminate false positives due to plating artifacts, mutant plaques were picked from the plate, diluted in 50 mM sodium borate buffer (pH 9.0), and mixed with an equal dilution of wild-type M13mp2 phage. Plating this mixture allows a visual comparison of phenotypes on the same plate and reduces uncertainty in identifying light blue plaques. Single-stranded DNA was prepared from the mutants scored in each of three independent experiments and sequences were determined by the chain-terminator method (27). Two individual oligonucleotides were used to prime the reactions for sequence analyses, (i) the 15-base primer from Bethesda Research Laboratories that is complementary to the coding sequence for amio acids 11–16 of the *lacZ* gene and (ii) a newly synthesized 15-base primer (New England BioLabs) complementary to the coding sequence for amino acids 46–50 of the *lacZ* gene.

**RESULTS**

The biological activity of M13mp2 viral DNA, when transfected into competent *E. coli* cells, decreases in proportion to the time of incubation at 70°C in pH 5.0 buffer (data not shown). This result is similar to that observed with *φX174* single-stranded DNA (18). It is consistent with the conclusion that, when using single-stranded DNA subject to few or no repair processes, a major biological consequence of depurination is lethality (14–16). Concomitant with this inactivation is an increase in mutation frequency (Table 1). The increase is double that of an untreated DNA control when the transfection is carried out using normal competent cells. However, if the transfection is carried out using competent cells made from bacteria UV irradiated to induce the SOS response, the mutation frequency increases to 97.8 × 10^-4, 15 times that of the untreated DNA in unirradiated cells. Both the actual mutation frequency and the relative increase above background are a function of the number of lethal hits in the DNA and the UV dose to the bacteria (data not shown). These phenomena are similar to previous observations using heat/acid-damaged *φX174* DNA (18) and imply that mutants result primarily from SOS-dependent bypass of abasic sites.

**Mutagenesis Due to AP Sites.** If apurinic (or apyrinicidic) sites are indeed responsible for the enhanced mutagenesis, then treatments that eliminate AP sites should reduce or eliminate the effect. As shown in Table 2, two independent methods of hydrolyzing AP sites reduced the enhanced mutagenesis. Hydrolysis by the highly specific (23) HeLa cell AP endonuclease (experiment 1) eliminated 95% of the enhanced mutagenicity while alkaline hydrolysis (4) reversed the effect by 85% (experiment 2). The residual mutagenesis above background may represent incomplete hydrolysis of AP sites or the presence of other lesions (3, 6, 10, 12, 28, 29). These data clearly indicate that the majority of the enhanced mutagenesis results from apurinic (and possibly apyrinicidic) sites in the DNA.

**SOS Dependence.** The mutagenic response was examined in recA° and umuC° host cells, which are defective for SOS-dependent mutagenesis (30, 31). The results (data not shown) indicate no damage-dependent increase in mutation frequency either for the recA° or umuC° mutant strains, using conditions that gave enhanced frequencies for the wild-type parent strains (32-fold and 11-fold, respectively). These data confirm the SOS dependence of the enhanced mutagenesis and further support the concept that apurinic sites, which are known to impede DNA synthesis (21, 32–34), are the responsible lesion.

**DNA Sequence Analysis.** A total of 87 spontaneous and 124 induced mutants were subjected to DNA sequence analysis. The classes of mutations observed are summarized in Table 3. The most frequent events under either condition are

<table>
<thead>
<tr>
<th>UV dose, J/m²</th>
<th>Lethal hits per DNA molecule</th>
<th>Mutation frequency (× 10^-4)</th>
<th>Relative mutation frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*E. coli* CSSH50° P° *lacZΔM15* were used for transfections. Cell survivals were 61% and 70%, respectively, for experiments 1 and 2. Phase survival was unaffected by AP endonuclease treatment and was halved by alkali treatment, in both normal and irradiated cells. Percentage reversal was calculated as follows: experiment 1, 1 – (0.8–6.0) × 100% = 95%; experiment 2, 1 – (27.0–13.8) × 100% = 85%.
Table 3. Frequency of various classes of observed mutants

<table>
<thead>
<tr>
<th>Type of event</th>
<th>No treatment (6.2 × 10⁻⁶)</th>
<th>Treated DNA, induced cells (97.80 × 10⁻⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>MF (× 10⁻⁶)</td>
</tr>
<tr>
<td>Base substitution</td>
<td>28/87</td>
<td>32.2 (42.4)</td>
</tr>
<tr>
<td>Double mutation</td>
<td>0/87</td>
<td>&lt;1.2 (&lt;1.5)</td>
</tr>
<tr>
<td>(-) Frameshift</td>
<td>9/87</td>
<td>10.3 (13.6)</td>
</tr>
<tr>
<td>(+) Frameshift</td>
<td>3/87</td>
<td>3.4 (4.5)</td>
</tr>
<tr>
<td>Deletion 93-base</td>
<td>21/87</td>
<td>24.1 (0)</td>
</tr>
<tr>
<td>Other</td>
<td>26/87</td>
<td>29.9 (39.4)</td>
</tr>
</tbody>
</table>

When the value for single-base errors (60 × 10⁻⁴) is used, it is assumed that there are 3.2 AP sites in a molecule containing 3600 purines, and the fact that (at least) 35 sites in the target are mutable (Fig. 1) is considered, the bypass frequency, calculated as described (18), is estimated to be 20%. Numbers in parentheses are percentages without considering the 93-base-deletion mutations. MF, mutation frequency.

single-base-substitution mutations. One-third of the spontaneous and two-thirds of the induced mutations are of this type. Depurination enhances base-substitution mutagenesis 31-fold above background to a frequency of 62.3 × 10⁻⁶.

A second mutation that is increased in frequency is the 93-base-deletion mutation. This event is presumably the result of recombination between the F episome containing the exact deletion (i.e., lacZAM15; see Fig. 1), and M13mp2. Further evidence that recombination is responsible is the observation that these deletion mutants have an associated adenine → guanine base change at the EcoRI site (see Fig. 1) representing the original base altered in the construction of M13mp phage. Since this deletion results from the use of F lac- containing competent cells, these mutants need not be considered further. Base substitutions thus represent an even greater percentage of the total induced mutations (79/92, 86%).

Two other classes of mutations are increased in the induced spectrum, double mutations and frameshifts. Three noninduction double mutations were observed, one of which (see Fig. 1, number 3, in the 35 promoter region) resembles a similar mutation, at this same site, in the UV-induced spectrum (26). Considering the rarity of double mutations, the existence of two similar nontandem double mutants resulting from two different mutagenic treatments suggests a common mechanism of production inherent to the structure of the DNA (35), the specificity of the replication apparatus, or both. Double mutations have also been detected twice among 32 depurination-induced amber revertants in dX174 DNA (21). They could reflect events at two different damaged sites in the same molecule or result from a targeted and an untargeted (semisemantargeted) event at some distance from the lesion. Frameshift mutations resulting from either the add}

When a frameshift occurs in a run of two or more of the same base, the exact nucleotide lost is unknown and the Δ is centered under the run. Double mutations, linked by common subscripts, have the indicated change. For example, the first double mutation, indicated with the subscript 1, contains both a C-G → A-T and a G-C → C-G transversion in the same molecule. Single-base-substitution mutations are shown above the wild-type sequence and indicate the nucleotide present in the viral strand. 0, CAP site; Δ, 35 promoter; Δ, 10 promoter; 0, transcription start; 0, operator site; 0, ribosome-binding site; 0, translation start; 0, position mutagenized to make the EcoRI site in the original mp2 construction; 0, first and last nucleotides of the 93-base deletion; 0, known mutable purine sites at which no base substitutions were observed in this study.

FIG. 1. Induced mutational spectrum for base substitutions and frameshifts in M13mp2 lac DNA. The 5' → 3' DNA sequence of the viral strand of the mutagenic target in M13mp2 is shown, from the first nucleotide after the lacI termination codon through the coding sequence for amino acid 43 of the lacZ gene. Frameshifts are shown below the sequence, indicated by Δ. When a frameshift occurs in a run of two or more of the same base, the exact nucleotide lost is unknown and the Δ is centered under the run. Double mutations, linked by common subscripts, have the indicated change. For example, the first double mutation, indicated with the subscript 1, contains both a C-G → A-T and a G-C → C-G transversion in the same molecule. Single-base-substitution mutations are shown above the wild-type sequence and indicate the nucleotide present in the viral strand. 0, CAP site; Δ, 35 promoter; Δ, 10 promoter; 0, transcription start; 0, operator site; 0, ribosome-binding site; 0, translation start; 0, position mutagenized to make the EcoRI site in the original mp2 construction; 0, first and last nucleotides of the 93-base deletion; 0, known mutable purine sites at which no base substitutions were observed in this study.
first round of replication in vivo, there is a preference for incorporation of adenine residues (47/79 or 59%). The remaining incorporation specificity is in the order thymine (28%), guanine (11%), then cytosine (1%).

**DISCUSSION**

The major goal of this work was to establish the spectrum of mutational events resulting from AP sites in DNA, to better understand the available alternatives for processing a non-coding lesion. The data reported here establish both the frequency and specificity of mutagenesis resulting from the in vivo SOS-dependent replication of unrepaird AP sites in M13mp2 DNA. Mutagenesis results primarily from an increase in single-base-substitution errors. The absence of mutants arising from the addition of a large number of bases and the negligible effect on deletion mutations indicate that, at least in this system, such events are not induced by AP sites. The spectrum of induced mutations yielded no addition frameshift mutants and only 8% deletion frameshifts. This low yield does not reflect a difference in target size for frameshifts versus other events, since frameshift mutants are detectable at most sites in the lacZ coding sequence and at many sites in the operator/promoter sequences. Thus, failure to detect larger numbers of these events indicates that they are not a major component of depurination-induced mutagenesis. The fact that only deletion frameshifts were observed suggests that looping out of the template strand containing the AP site may occur more frequently than looping out of the newly synthesized strand.

The large absolute and relative increases in base-substitution mutations (Table 3) provide a large number of mutants for analysis of specificity under conditions in which most of the mutants are induced. The small increase not dependent on SOS induction or damage to the DNA (Table 1) could result from rare bypass of AP sites under noninduced conditions, untargeted mutagenesis, or lesions other than AP sites. However, under the conditions used here to damage the DNA, other lesions (17, 28, 29, 36) are expected to be rare relative to AP sites. The results in Table 2 and the observation that 99% of the sites mutated and 82% of the total single-base substitutions are at template purines are also consistent with AP sites being primarily responsible for the observed increase. This purine site specificity is not imposed by the DNA target; a recent study of UV mutagenesis using this same target gave opposite results; eight of nine sites mutated and 32 of 35 single-base changes were at template pyrimidines (26).

Previous studies on the reversion of three amber mutations in 4X174 DNA have indicated a strong preference for insertion of adenine residues opposite apurinic sites (21, 32). The base-substitution specificity observed here for 38 sites is similar; 59% of the mutants result from incorporation of dAMP (Table 4). This specificity could result from several factors. The possibility that the base immediately preceding or following an AP site can be used as a template to determine the specificity of mutagenesis (by looping out the template or daughter strand) can explain only 25% of the observed specificity. Furthermore, if looping out is a general mechanism for bypass, one would expect to see a greater increase in frameshift mutations than actually observed. This mechanism thus seems unlikely.

Because an AP site is a noncoding lesion, incorporation specificity could reflect nucleotide substrate concentrations in vivo. Treating cells with DNA-damaging agents results in alterations in intracellular dNTP pools (37), the greatest effect being a 3- to 8-fold increase in the dATP pool. This increase could be reflected in the specificity observed here. Alternatively, purified DNA polymerases are known to incorporate dAMP opposite AP sites in vitro when equimolar dNTP concentrations are used (32, 38). Moreover, using single nucleoside triphosphates, Strauss et al. (33), found a similar DNA polymerase incorporation specificity opposite apyrimidinic sites. Thus dAMP incorporation at AP sites may not simply be a consequence of pool imbalances but may reflect inherent properties of the AP site, the replication apparatus, or both.

The transversions characteristic of adenine incorporation opposite positions of template purines are thought to be the predominant mutagenic event induced by benzo(a)pyrene (39), and aflatoxin B1 (40), which produce primarily bulky adducts in DNA. Such lesions could enhance mutagenesis through a common AP site intermediate, as suggested by the data of Schaaper et al. (11) and Drinkwater et al. (9). Alternatively, dATP may be the preferred substrate whenever base hydrogen bonding is not possible, regardless of the actual lesion. This implies that specificity is imposed by the replication apparatus, a suggestion supported by the recent observation that, in the absence of mutagen treatment, the primary base-substitution event in the lacI gene of E. coli observed on constitutive expression of error-prone replication is a G-C → T-A transversion. (J. H. Miller and K. B. Low, personal communication).

Adenine incorporation specificity has been suggested to result from preferential binding of adenine residues to the DNA polymerase in the absence of template instruction (33). Measurements with E. coli DNA polymerase I (ref. 41; binding preference, guanine > adenine > thymine > cytosine) and terminal deoxynucleotidyl transferase (ref. 42; binding preference, guanine > adenine > thymine > cytosine) do not reflect the specificity observed in this study. However, similar data are not available for other DNA polymerases, and other factors, such as exonucleolytic proofreading (32), are expected to modulate any potential correlation. This concept deserves further study.

Although incorporation of dAMP is the most frequent event, both dTMP and dGMP are incorporated at substantial frequencies. Incorporation of dTMP occurred at potential apurinic sites in 14 mutants. However, this nucleotide is incorporated opposite positions of template pyrimidines eight times, resulting in transversions not seen even once in the spontaneous spectrum. These unique mutants could represent incorporation of dTMP opposite apyrimidinic sites, although this would imply that apyrimidinic sites are produced at frequencies somewhat greater than predicted from a previous study (6). Alternatively, these mutants could result from untargeted events due to the specificity of error-prone replication.

The seven transversions resulted from incorporation of dGMP opposite potential apurinic sites. Two such transver-
sions were observed among the depurination-induced revertants of the φX174 amber 18 mutation (21). Although less frequent than mutants arising from adenine incorporation, these mutants are not rare and should not be ignored in attempting to understand what information is used to determine the mutagenic outcome of a lesion.

The distribution of mutants is not random. As in measurements of depurination-induced reversion of φX174 amber mutants (21, 32), site-specific differences are apparent here. For example, the middle guanine residue in the hotspot (Fig. 1) gave nine mutants. In contrast, no mutants were observed at the adenine residue four nucleotides away (5’ direction, Fig. 1), which is known (J. E. LeClerc and N. L. Istock, personal communication) to be mutable by the most frequent incorporation event associated with depurination (incorporation of adenine). Based on current information, five purine sites (Fig. 1) known to produce a mutant phenotype are not found in this spectrum, three of which could have resulted from the most frequently expected transversion. This apparent nonrandom distribution of mutants may reflect the relatively small number of mutants analyzed or could reflect more important factors. Of eight purine sites yielding three or more mutations, seven have purines immediately adjacent, and the hotspot is in a run of seven purines. Perhaps depurination may occur more readily in a purine-rich sequence. Differences in the rate of depurination of double- and single-stranded DNA are known (3) and guanine residues depurinate more readily than adenine residues (3). This may explain the preference for guanine versus adenine sites mutated (Table 4). Alternatively the nonrandom distribution of mutants may reflect differences in the frequency of bypass of apurinic sites (21).

AP sites are highly mutagenic under SOS-induced conditions, resulting in 0.2% mutants per lethal hit (Table 3) for a target DNA sequence of only 250 bases. This is 5- to 10-fold greater per lethal hit than the SOS-dependent mutagenesis of this same target resulting from UV irradiation of intact M13mp2 phage (26). This highly mutagenic response presumably results from two properties of AP sites. Unlike modified bases, which may retain some correct base-coding potential, AP sites are truly (base) noncoding lesions. In addition, AP sites probably distort DNA to a lesser extent and sterically constrain polymerization less than certain bulky adducts do (33, 43, 44), resulting in more frequent bypass. The estimates of AP site bypass frequency in vitro are in fact quite high for several DNA polymerases (20, 21, 32, 38), in contrast to the results obtained with bulky base modifications (33, 45). These properties of AP sites, in combination with the potentially high frequency of their production, make these lesions interesting for future studies of repair.

I would like to express appreciation to J. Motto and J. Liu for assistance in DNA sequence analysis and to M. Volkert, R. Schaeper, B. Glickman, and J. Drake of this Institute for critical evaluation of the manuscript. The gift of strains and the advice of Eugene LeClerc (University of Rochester) during the initial stages of this work were most helpful and are gratefully acknowledged.