*Proc. Natl. Acad. Sci. USA*
Vol. 78, No. 11, pp. 6817–6820, November 1981
Biochemistry

# Base insertion and deletion mutations induced in an *Escherichia coli* plasmid by benzo[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide

(frame shift mutations/transcription terminator/phage λ/DNA sequence analysis)

HIROSHI MIZUSAWA*, CHE-HUNG LEE*, TSUYOSHI KAKEFUDA*, KEITH MCKENNEY†, HIROYUKI SHIMATAKE†, AND MARTIN ROSENBERG†

*Laboratory of Molecular Carcinogenesis and †Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

*Communicated by Elizabeth C. Miller, August 6, 1981*

**ABSTRACT** **Mutations induced by (±)*trans*-benzo-[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide (B*a*P-diol epoxide-1) were selected by using a recombinant plasmid vector system designed for the study of transcription termination signals. The plasmid contains a transcription terminator positioned between a promotor signal and the *Escherichia coli* galactokinase structural gene (*galK*). By selections for the expression of *galK* (i.e., *galK*$^-$ to *galK*$^+$), mutations are obtained in the terminator region that allow transcription from the promotor to read the *galK* gene. These mutations were characterized by direct DNA sequence of the terminator region. The DNA sequence changes caused by B*a*P-diol epoxide-1 were demonstrated for three different mutants. Two were found to be single-base-pair insertions of T·A into a cluster of consecutive T·A base pairs and the other change was a single-base-pair deletion of G·C from a cluster of consecutive G·C base pairs.**

Benzo[*a*]pyrene (B*a*P) is one of the most widely distributed environmental carcinogens. The ultimate carcinogenic form of B*a*P is most likely (±)*trans*-B*a*P-7,8-dihydrodiol-9,10-epoxide (B*a*P-diol epoxide-1), which is generated by mixed-function oxidases (1–7). B*a*P-diol epoxide-1 primarily binds covalently to deoxyguanosine residues in DNA and to a lesser extent to deoxyadenosine and deoxycytidine residues (8–13). The binding effectively inhibits the replication *in vitro* of both single- and double-stranded DNA templates (14–16) and reduces their biological activities *in vivo* (14, 17).

The mutagenic potency of carcinogenic polycyclic hydrocarbons has been demonstrated in studies of the back mutation of the *hisD* gene of *Salmonella typhimurium* and the analysis of amino acid sequences of the gene product. A one-base-pair deletion at a -G-G-G- sequence was strongly suggested (18). Recently, it was reported that the polycyclic hydrocarbon ICR170 induces G·C insertions within a G·C cluster region in yeast (19, 20).

The molecular mechanisms of B*a*P-diol epoxide-1-induced mutations have been studied (17) by modification of a plasmid (pK0482) vector carrying and expressing the *Escherichia coli* galactokinase (*galK*) gene (21). Mutations were identified as inactivation of the galactokinase activity. However, DNA sequence analysis of these mutants was not carried out because of the large size of the *galK* coding region, 1.2 kilobase pairs (kb). The mutation frequency was found to be significantly higher in AB1886(*uvrA*$^-$) than in wild type (AB1157), and no mutations were observed in AB2463(recA$^-$) (17).

The present studies probe the precise changes caused in DNA by the carcinogen B*a*P-diol epoxide-1. A plasmid vector is used that allows mutations to be readily selected and structurally analyzed (unpublished results). The vector pkG1820 (see Fig. 1) contains the *E. coli galK* gene (21); however, *galK* is not expressed from this plasmid because a signal for transcription termination ($t_o$) [originally derived from phage λ (22, 23)] is positioned between the *gal* promoter signal ($P_{gal}$) and the *galK*, structural gene. Transcription initiating at $P_{gal}$ terminates prior to *galK*, resulting in no *galK* expression. Apparently mutations that affect the function of the terminator allow transcription and, consequently, expression of the *galK* gene. Because the target region is relatively small (230 base pairs) and its exact DNA sequence is known [refs. 21 and 24; and unpublished data], mutational changes occurring in the region are easily characterized. This study presents the application of this vector system to the elucidation of B*a*P-diol epoxide-1-induced mutagenic alterations.

## MATERIALS AND METHODS

**Bacterial Strains.** The *galK*$^-$ and *amp-s E. coli* strains AB1157, AB1886(*uvrA*$^-$), or AB2463 (*recA*$^-$) were used as host strains as described (17).

**Isolation of Plasmid DNA.** Plasmid pKG1820 DNA was extracted and purified by methods as described (17). When highly purified DNA was required, the plasmid DNA was extracted by the cleared lysate method (25) and purified by two successive centrifugations in CsCl-containing ethidium bromide (26). For routine and rapid examination of the plasmid DNA, alternative methods were used (17).

**Transfection of *E. coli* with Plasmid pKG1820 DNA.** Competent cells were prepared as described (17, 27). *E. coli* AB2463 was grown in 300 ml of L broth [1% Bacto-tryptone (Difco)/0.5% Bacto-yeast extract (Difco)/0.05% NaCl/0.2% glucose, pH 7.0] at 37°C to an $OD_{660}$ of 0.6 (17). The cells were resuspended in 150 ml of cold transfection buffer (0.6% $CaCl_2$/0.3% $MnCl_2$/10 mM Tris·HCl, pH 7.5), kept for 10 min at 4°C, precipitated, and then suspended in 1.5 ml of the transfection buffer at 4°C. DNA (200 ng in 10 μl of 10 mM Tris·HCl, pH 7.5/5 mM NaCl/1 mM EDTA) was mixed with 0.1 ml of competent cells at 4°C. The mixture was kept on ice for 30 min, then heated at 40°C for 2 min, diluted with 0.9 ml of L broth, incubated for 30 min, and then plated on MacConkey agar plates containing galactose (1%) and ampicillin (50 μg/ml). The efficiency of transfection was 1–5 × $10^5$ colonies per μg of DNA.

**Mutagenesis and Screening for Mutants.** B*a*P-diol epoxide-1 was introduced into *E. coli* carrying the plasmid pKG1820—either AB1157[pKG1820] or AB1886[pKG1820](*uvrA*$^-$). These strains were grown in 5 ml of L broth overnight, precipitated, and resuspended in 5 ml of 10 mM Tris·HCl pH 7.5)/5 mM NaCl. B*a*P-diol epoxide-1 (obtained through the National

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

Abbreviations: B*a*P, benzo[*a*]pyrene; B*a*P-diol epoxide-1, (±)*trans*-B*a*P-7,8-dihydrodiol-9,10-epoxide; kb, kilobase pair(s).

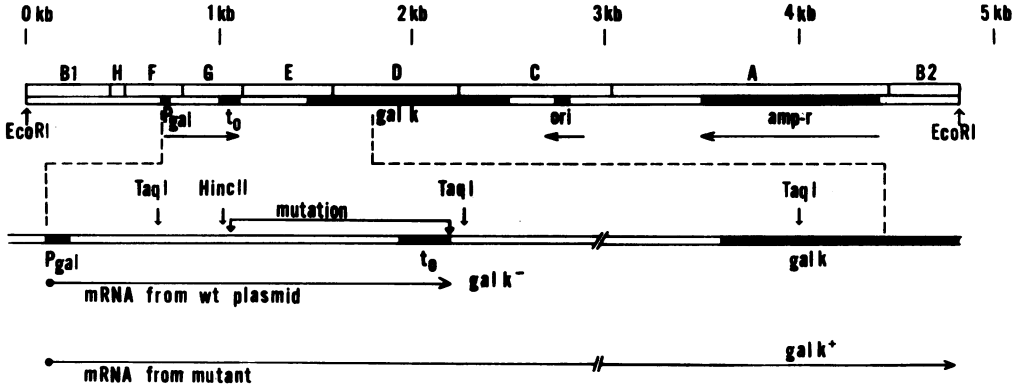


FIG. 1. Relationship between the restriction map and the functions of plasmid pKG1820. The restriction map generated by *Taq* I digestion and the genetic map are described together (*Upper*). Arrows indicate the major mRNA transcripts and orientation. The *galK* gene is not expressed in wild-type pKG1820. Transcription and translation in wild-type and mutated DNA from the $P_{gal}$ through *galK* gene are summarized in the expanded genetic map (*Lower*).

Cancer Institute depository) was dissolved in tetrahydrofuran and stored at −70°C. aliquots (1 ml) of cell suspensions were treated with B*a*P-diol epoxide-1; the final concentration of B*a*P-diol epoxide-1 used in each reaction mixture is shown in Table 1. Tetrahydrofuran was added as a control. After being mixed with B*a*P-diol epoxide-1, the cells were kept in the dark at room temperature for 2 hr. Portions of the cells were used for viable cell number determination by plating on L broth/agar/ampicillin plates. The remainder of the cells were diluted into 50 ml of fresh L broth and grown at 37°C with shaking for 18 hr. Plasmid DNA was extracted, purified and used for transfection into AB2463 (*recA*$^-$) cells. Cells transfected with wild-type pKG1820 DNA grow as white colonies on MacConkey galactose/ampicillin plates. Cells transfected with mutagen treated pKG1820 DNA grow as white colonies; however, among many white colonies, red colonies are found. Red colonies are easily detected when the total number of colonies on a single plate is about 2000. They were purified by three cycles of single colony isolation and then plasmid DNA was purified. Isolated DNA was provided to transfect AB2463 cells and was checked for *amp-r* and *galK*$^-$ to ensure that the plasmids contained the *galK*$^+$ marker.

**Sequence Determination.** As shown in Figs. 1 and 2, a restriction endonuclease *Taq* I cleavage site is located 35 base pairs downstream from the $t_0$ termination site (A·T pair marked no. 1 in Fig. 2), and a *Hin*cII restriction site is located approximately 200 base pairs upstream from the $t_0$ site (Fig. 2). The DNA fragment containing $t_0$ was isolated from mutated pKG1820 DNA after *Taq* I cleavage, labeling of the 5′ ends with [$\gamma$-$^{32}$P]ATP (ICN, 4000 Ci/mmol; 1 Ci = $3.7 \times 10^{10}$ becquerels) and polynucleotide kinase (P-L Biochemicals), and secondary cleavage with *Hin*cII. The *Taq* I–*Hin*cII fragment (230 base pairs) containing $t_0$ was purified on a 6% (wt/vol) polyacrylamide gel, and the sequence was determined by the technique of Maxam and Gilbert (28).

Table 1. Mutation frequencies of B*a*P-diol epoxide-1-treated pKG1820

| Initial strain | Exp. no. | B*a*P-diol epoxide-1, μM | Total colonies scored × $10^{-4}$ | Red colonies scored | Mutation frequency × $10^4$ |
|---|---|---|---|---|---|
| | I | 0 | 13.0 | 0 | <0.076 |
| | I | 2.5 | 13.2 | 20 | 1.5 |
| AB1886 | II | 0 | 1.08 | 0 | <0.92 |
| [pKG1820] | II | 70 | 0.39 | 1 | 2.5 |
| (*uvrA*$^-$) | III | 0 | 9.8 | 1 | 0.10 |
| | III | 70 | 2.0 | 8 | 4.0 |
| | IV | 0 | 13.5 | 1 | 0.074 |
| AB1157 | IV | 70 | 12.2 | 4 | 0.32 |
| [pKG1820] | V | 0 | 3.4 | 0 | <0.29 |
| | V | 150 | 4.0 | 3 | 0.78 |

## RESULTS

**Screening of the Mutated Plasmid pKG1820 DNA.** B*a*P treatment *in vivo* induced mutation in pKG1820 DNA decreasing the number of viable cells to 4.5% in AB1157[pKG1820] and 0.2% in AB1886[pKG1820](*uvrA*$^-$) with 150 and 70 μM B*a*P-diol epoxide-1, respectively. These results are consistent with our previous study (17) and results reported by Ivanovic and Weinstein (29). When the B*a*P-diol epoxide-1-treated AB1157[pKG1820] and AB1886 [pKG1820](*uvrA*$^-$) were plated directly onto screening plates, spontaneous reversion of the host *galK* mutation obscured the detection of mutations in the pKG1820 plasmid. To overcome this problem the B*a*P-diol epoxide-1-treated cells were grown in fresh L broth, and the pKG1820 DNA was extracted and introduced through transfection into AB2463(*recA*$^-$). In the strain AB2463 (*recA*$^-$), spontaneous reversion of the host *galK* gene was effectively suppressed (17).

To confirm that the mutations were in the plasmid, plasmid DNA was extracted from red colonies and transfected into AB2463 (*recA*$^-$). Most of the resulting colonies were red, demonstrating that the mutations were in fact in the pKG1820 DNA. All the colonies listed in Table 1 were shown to be plasmid mutants by this procedure. The mutation frequency of pKG1820 was greater than $1 \times 10^{-4}$ when pKG1820 was treated by B*a*P-diol epoxide-1 in a *uvrA*$^-$ host strain, whereas the mutation frequency of control DNA (no B*a*P-diol epoxide-1 treatment) was significantly lower than the mutagen-treated group (Table 1). When the pKG1820 was mutagen-treated in the wild-type host strain (AB1157), its mutation frequency was lower than those treated in AB1886(*uvrA*$^-$) but still higher than the control (Table 1).

**Structural Analysis of Mutations Within the Terminator Region.** Mutants were initially characterized by cleavage of the DNA with restriction endonuclease *Taq* I. The restriction patterns of almost all of the mutants were indistinguishable from that of the starting plasmid pKG1820. Only two mutants obtained in the strain AB1157 appeared to have altered restriction patterns consistent with a large DNA insertion and deletion,

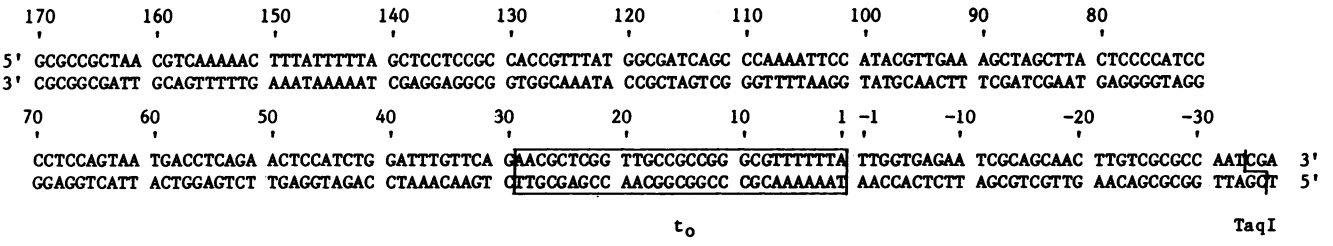


FIG. 2. DNA sequence of the terminator $t_o$ region. The orientation of transcription is left to right, and the bottom strand is transcribed. The end of the terminator is assigned as 1, where transcription terminates.

respectively. These were not further characterized.

The DNA sequence of the region surrounding the terminator signal was determined in 15 independent mutants. Three were found to have DNA sequence changes in this region. One, RB6, contained an additional T·A near the end of the terminator (Figs. 3 and 4). Another, RB17, also was found to be a single T·A insertion positioned 150 base pairs upstream of the site of mutation in RB6 (Fig. 4). The third, RB3 showed a deletion of a single G·C near the site of the mutation in RB6 (Fig. 4). In each case, the mutation occurred within a cluster of consecutive identical nucleotides; thus, which base pair was actually inserted or deleted could not be determined. None of the other mutants examined had changes in the DNA sequence in this region.

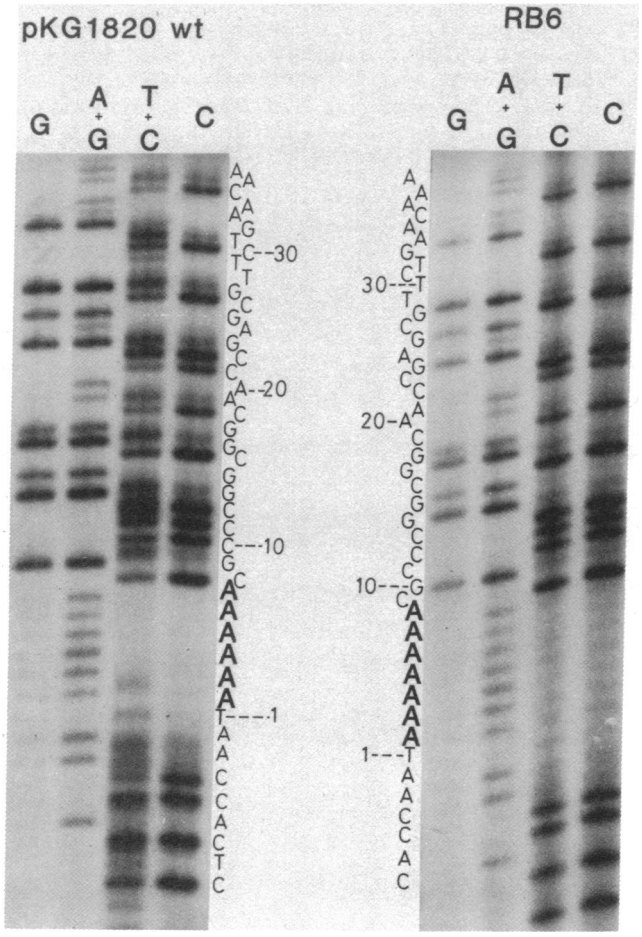


FIG. 3. DNA sequence gel at the terminator region. Autoradiogram of DNA sequence gels for wild-type (wt) and mutated pKG1820 (RB6). The numbers shown along the wild-type DNA sequences correspond to the numbers in Fig. 2.

## DISCUSSION

Mutations induced by B*a*P-diol epoxide-1 were selected and characterized with the plasmid vector pKG1820. Three independent mutations resulting from B*a*P-diol epoxide-1 were found within the terminator region of pKG1820. The fact that this region is only 230 base pairs and can be readily obtained by *Taq* I and *Hin*cII cleavage makes it reasonably simple to structurally characterize the mutants. The mutations detected in these experiments were either +1 insertions, or −1 deletions (Fig. 4), and they occurred in clusters of consecutive base pairs—i.e., deletion of G·C from $(G{\cdot}C)_3$ (RB3), insertions of T·A into $(T{\cdot}A)_6$ (RB6) or $(T{\cdot}A)_3$ (RB17). Among 15 independently selected mutants, 3 were located within the terminator region. This indicates that mutations in other regions of the vector also can lead to *galK* gene expression (e.g., promotor up-mutations, copy-number mutations, etc.). The fact that 20% of the mutants analyzed were in the region predicted (and all of them showed similar characteristics) strongly suggests that B*a*P-diol epoxide-1 preferentially induces single base-pair-deletion or -addition mutations in clusters of consecutive base pairs.

The insertions and deletions in plasmid pKG1820 DNA were produced in *E. coli* AB1886 (*uvrA*$^-$). Because the *uvrA* gene is thought to be involved in the excision repair of B*a*P-diol epoxide-1-modified bases of DNA (17, 29), the deletions and insertions observed here could have resulted from the unexcised B*a*P-diol epoxide-1-modified bases. It has been known that B*a*P-diol epoxide-1 binds preferentially to the 2-amino group of deoxyguanosine residues in DNA and, at a lower efficiency, to adenine residues (9, 10). Thus, it is possible that the deletion of G·C from a $(G{\cdot}C)_n$ cluster is due to the modification of deoxyguanosine residues in the cluster region. Although the modification of the deoxyadenosine residues by B*a*P-diol epox-

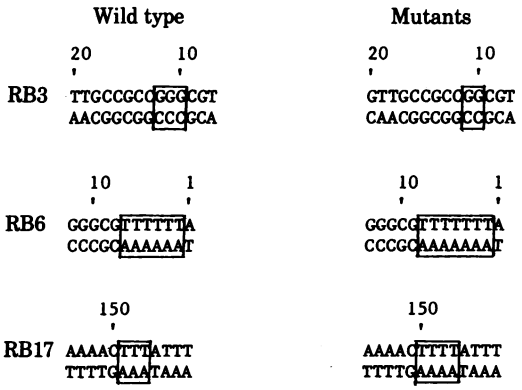


FIG. 4. Mutated DNA sequences induced by B*a*P-diol epoxide-1 treatment. The numbers shown along the DNA sequences correspond to the numbers in Fig. 2. Boxes show the location of the mutation in mutant strains RB3, RB6, and RB17 compared to corresponding wild-type DNA sequences.

ide-1 is not the predominant lesion, it may be important in mutagenesis or carcinogenesis. For example, it has been reported that the tumorigenesis of mouse skin by B*a*P-diol epoxide-1 is proportional to the binding to deoxyadenosine residues (30). An alternative explanation is that a neighboring G·C mod-modified by B*a*P-diol epoxide-1 may result in insertion at the A·T cluster region.

Similar results have been achieved in experiments with yeast, where the carcinogenic polycyclic hydrocarbon ICR170 has been shown (by analysis of the DNA sequence of the *his4* locus) to result in the insertion of one G·C to a G·C cluster (20). In addition, deletion of a single G·C from another cluster region was also suggested (18). These observations support our contention that B*a*P-diol epoxide-1 may act at clusters of consecutive nucleotides.

We thank Drs. M. Seidman and M. Bustin for valuable discussions, Drs. M. Mattern and K. Kraemer for critical reading of the manuscript, Mr. P. Morris and Miss P. Madigan for technical assistance, and Dr. H. Gelboin for support and encouragement.

1. Huberman, E., Sachs, L., Yang, S. Y. & Gelboin, H. V. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 607–611.
2. Malaveille, C., Kuroki, T., Sims, P., Grover, P. L. & Bartsh, H. (1977) *Mutat. Res.* **44**, 313–326.
3. Marquardt, H., Grover, P. L. & Sims, P. (1976) *Cancer Res.* **36**, 2059–2064.
4. Sims, P., Grover, P. L., Swaisland, A., Pal, K. & Hewer, A. (1974) *Nature (London)* **252**, 326–327.
5. Wislocki, P. G., Wood, A. W., Chang, R. L., Levin, W. & Yyagi, H. (1976) *Biochem. Biophys. Res. Commun.* **68**, 1006–1012.
6. Yang, S. K., McCourt, D. W., Roller, P. P. & Gelboin, H. V. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2594–2598.
7. Gelboin, H. V. (1980) *Physiol. Rev.* **60**, 1107–1165.
8. Jeffry, A. M., Grzeskowiak, K., Weinstein, I. B. & Nakanishi, K. (1979) *Nature (London)* **206**, 1309–1311.
9. Kakefuda, T. & Yamamoto, H. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 415–419.
10. Meeham, T., Straub, K. & Calvin, M. (1977) *Nature (London)* **269**, 725–727.
11. Osborn, M. R., Beland, F. A., Harvey, R. G. & Brooks, P. (1976) *Int. J. Cancer* **18**, 362–368.
12. Santella, R. M., Grunberger, D. & Weinstein, I. B. (1979) *Mutat. Res.* **61**, 181–189.
13. Weinstein, I. B., Jeffry, A. M., Jeannette, K. W. & Blobstein, S. H. (1976) *Science* **193**, 592–595.
14. Hsu, W-T., Lin, J. S., Harvey, R. G. & Weiss, S. B. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3335–3339.
15. Mizusawa, H. & Kakefuda, T. (1979) *Nature (London)* **279**, 75–78.
16. Moore, P. & Strauss, B. S. (1979) *Nature (London)* **278**, 664–666.
17. Mizusawa, H., Lee, C.-H. & Kakefuda, T. (1981) *Mutat. Res.* **82**, 47–57.
18. Isono, K. & Yourno, J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1612–1617.
19. Culbertson, M. R. & Underbrink, K. M. (1980) *Genetics* **95**, 833–853.
20. Donahue, T. F., Farabaugh, P. J. & Fink, G. (1981) *Science* **212**, 455–457.
21. McKenney, K., Shimatake, H., Court, D., Schmeissner, U., Brady, C. & Rosenberg, M. (1981) in *Gene Amplification and Analysis*, eds. Chirikjian, J. G. & Papas, T. S. (Elsevier/North-Holland, New York), Vol. 2.
22. Hayes, S. & Szybalsi, W. (1973) *Mol. Gen. Genet.* **126**, 275–290.
23. Howard, B. H., de Crombrugghe, B. & Rosenberg, M. (1977) *Nucleic Acids Res.* **4**, 827–842.
24. Schwarz, E., Scherer, G., Hobom, G. & Kossel, H. (1978) *Nature (London)* **272**, 410–414.
25. Clewell, D. B. & Helinski, D. R. (1970) *Biochemistry* **9**, 4428–4440.
26. Mizusawa, H., Tanaka, S., Kobayaski, M. & Koike, K. (1977) *Biochem. Biophys. Res. Commun.* **74**, 570–576.
27. Mandel, M. & Higa, A. (1970) *J. Mol. Biol.* **53**, 159–162.
28. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
29. Ivanovic, V. & Weinstein, I. B. (1980) *Cancer Res.* **40**, 3508–3511.
30. DiGiovanni, J., Romson, R. J., Linville, D., Junchau, M. R. & Slaga, T. J. (1979) *Cancer Lett.* **7**, 39–43.