Styring, S., & Rutherford, A. W. (1987) *Biochemistry 26*, 2401-2405.
Tamura, N., & Martin, G. (1985) *Biochim. Biophys. Acta 809*, 245-257.
Teo, B.-K., & Lee, P. A. (1979) *J. Am. Chem. Soc. 101*, 2815-2832.
Teo, B.-K., Antonio, M. R., & Averill, B. A. (1983) *J. Am. Chem. Soc. 105*, 3751-3762.
Velthuys, B., & Kok, B. (1978) *Biochim. Biophys. Acta 502*, 211-221.
Vincent, J. B., & Christou, G. (1986) *FEBS Lett. 207*, 250-252.
Vincent, J. B., Christmas, C., Huffman, J. C., Christou, G., Chang, H.-R., & Hendrickson, D. N. (1987a) *J. Chem. Soc., Chem. Commun.*, 236-238.
Vincent, J. B., Chang, H.-R., Folting, K., Huffman, J. C., Christou, G., & Hendrickson, D. N. (1987b) *J. Am. Chem. Soc. 109*, 5703-5711.
Witt, H. T., Saygin, Ö., Brettel, K., & Schlodder, E. (1987) in *Progress in Photosynthesis Research* (Biggins, J., Ed.) Vol. I, pp 523-531, Martinus Nijhoff, Dordrecht.
Yachandra, V. K., Guiles, R. D., McDermott, A., Britt, R. D., Dexheimer, S. L., Sauer, K., & Klein, M. P. (1986a) *Biochim. Biophys. Acta 850*, 324-332.
Yachandra, V. K., Guiles, R. D., Sauer, K., & Klein, M. P. (1986b) *Biochim. Biophys. Acta 850*, 333-342.
Yachandra, V. K., Guiles, R. D., McDermott, A. E., Cole, J. L., Britt, R. D., Dexheimer, S. L., Sauer, K., & Klein, M. P. (1987) *Biochemistry 26*, 5974-5981.

# Mechanisms of Mutagenesis by the Vinyl Chloride Metabolite Chloroacetaldehyde. Effect of Gene-Targeted in Vitro Adduction of M13 DNA on DNA Template Activity in Vivo and in Vitro†

J. Steven Jacobsen and M. Zafri Humayun*
*Department of Microbiology and Molecular Genetics, University of Medicine and Dentistry of New Jersey—New Jersey Medical School, 185 South Orange Avenue, Newark, New Jersey 07103*
*Received May 23, 1989; Revised Manuscript Received August 28, 1989*

ABSTRACT: 2-Chloroacetaldehyde (CAA), a metabolite of the carcinogenic industrial chemical vinyl chloride, reacts with single-stranded DNA to form the cyclic etheno lesions predominantly at adenine and cytosine. In both ethenoadenine and ethenocytosine, normal Watson–Crick hydrogen-bonding atoms are compromised. We have recently shown that CAA adduction leads to efficient mutagenesis in *Escherichia coli* predominantly at cytosines, and less efficiently at adenines. About 80% of the mutations at cytosines were C-to-T transitions, and the remainder were C-to-A transversions, a result similar to that of many noninstructional DNA lesions opposite which adenine residues are preferentially incorporated. It is widely believed that noninstructional lesions stop replication and depend on SOS functions for efficient mutagenesis. We have examined the effects of in vitro CAA adduction of the *lacZα* gene of phage M13AB28 on in vivo mutagenesis in SOS-(UV)-induced *E. coli*. CAA adduction was specifically directed to a part of the *lacZ* sequence within M13 replicative form DNA by a simple experimental strategy, and the DNA was transfected into appropriate unirradiated or UV-irradiated cells. Mutant progeny were defined by DNA sequencing. In parallel in vitro experiments, the effects of CAA adduction on DNA replication by *E. coli* DNA polymerase I large (Klenow) fragment were examined. Our data do not suggest a strong SOS dependence for mutagenesis at cytosine lesions. While adenine lesions remain much less mutagenic than cytosine lesions, mutation frequency at adenines is increased by SOS. SOS induction does not significantly alter the specificity of base changes at cytosines or adenines. The in vitro replication results show that these lesions create kinetic pause sites rather than absolute replication blocks. The simplest interpretation of the specificity of base changes is that the mechanisms of base incorporation opposite these lesions are analogous to those opposite the canonical noninstructional lesions. The high efficiency of mutagenesis opposite cytosine lesions without the aid of induced levels of SOS functions suggests that all DNA lesions lacking normal Watson–Crick base-pairing ability may nevertheless not block replication. In addition, the relatively nonmutagenic bypass of adenine lesions focuses attention on the need to understand mechanisms of error avoidance in the absence of normal base-pairing information.

Following the description of the formation of cyclic DNA lesions by the mutagen glyoxal (Shapiro, 1969), a large number of chemical carcinogens have been shown to induce such lesions. For example, 2-chloroacetaldehyde (CAA),[1] a metabolite of vinyl chloride (VC), predominantly reacts with unpaired adenine and cytosine residues to form 1,$N^6$-ethenoadenine (εA) and 3,$N^4$-ethenocytosine (εC) which are the final stable lesions derived from the dehydration of the initially

† This work was supported by USPHS Grant CA47234. M.Z.H. is the recipient of USPHS Research Career Development Award CA00907.
* Author to whom correspondence should be addressed.

[1] Abbreviations: CAA, chloroacetaldehyde; VC, vinyl chloride; A, adenine; C, cytosine; T, thymine; G, guanine; ds, double stranded; ss, single stranded; polI(k), *Escherichia coli* DNA polymerase I large (Klenow) fragment; εC, 3,$N^4$-ethenocytosine; εC·$H_2O$, hydrated 3,$N^4$-ethenocytosine [3,$N^4$-(hydroxyethano)cytosine]; εA, 1,$N^6$-ethenoadenine; εA·$H_2O$, hydrated 1,$N^6$-ethenoadenine [1,$N^6$-(hydroxyethano)adenine]; UV, ultraviolet.

0006-2960/90/0429-496$02.50/0 

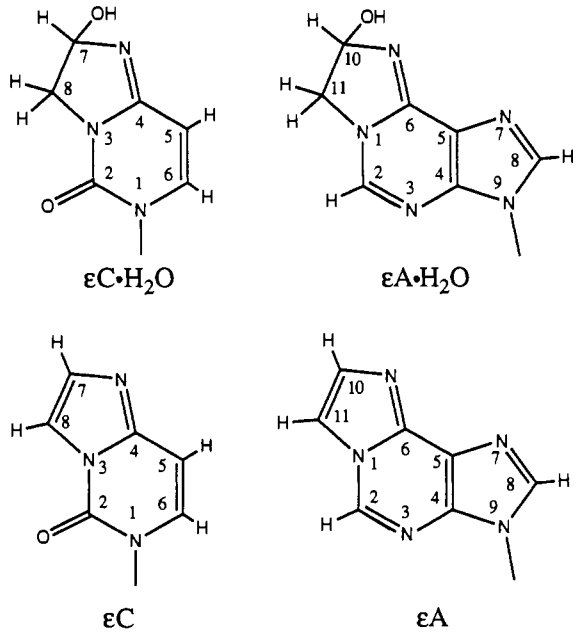


FIGURE 1: Structures of εC and εA and their respective hydrated intermediate forms, εC·$H_2O$ and εA·$H_2O$.

formed hydroxyethane intermediates εA·$H_2O$ and εC·$H_2O$ (Figure 1). However, the biological impact of cyclic DNA lesions, especially their in vivo mutagenic potential, has been largely unexplored. Using a high-efficiency mutation detection system based on the *lacZα* sequence of coliphage M13, we have recently provided evidence that cytosine lesions inflicted by CAA (likely ethenocytosines) are highly mutagenic (Jacobsen et al., 1989).

Mutagenic DNA lesions are usually classified into misinstructional and noninstructional categories on the basis of purported mechanisms of mutagenesis. Misinstructional lesions, exemplified by DNA uracil (a product of cytosine deamination) and $O^6$-methylguanine, are assumed to cause mutations by simple miscoding and require only the normal DNA replication machinery. Noninstructional lesions, exemplified by abasic sites, UV photoproducts, and bulky chemical adducts, are believed to block DNA replication due to either a lack of template instruction or steric hindrance. Bypass of these lesions, at least in *Escherichia coli*, is presumed to require the intervention of induced genetic functions (the SOS pathway). DNA sequence analysis of mutations induced by a few examples of noninstructional lesions suggests that, irrespective of the chemical diversity of these SOS-dependent lesions, adenines are most frequently incorporated opposite these lesions, followed by thymines and, much less frequently, other bases [see, e.g., Kunkel (1984)]. Observation of a similar preference for adenine misinsertion opposite noninstructional lesions in defined in vitro systems has led to the hypothesis that this preference may be an inherent property of many DNA polymerases confronted with a site devoid of accessible template function [see, e.g., Rabkin and Strauss (1984)].

The specificity of misinsertion opposite CAA-induced cytosine lesions (adenine, 80%; thymine, 20%), as well as other considerations, have led us to previously suggest that εC lesions might act as noninstructional lesions. Paradoxically, even though εA lesions were present at about the same level as εC lesions, mutations at adenine lesions were an order of magnitude less frequent than at cytosines. Nevertheless, A-to-T transversions predominated the few mutations at adenines, suggesting that these lesions were also acting as noninstructional lesions. Since noninstructional lesions should show a marked SOS dependence for mutagenesis and be able to act as replication stop sites, CAA-induced lesions should show these properties. In this paper, we describe in vivo and in vitro experiments to examine these issues.

## Materials and Methods

*Bacterial and Phage Strains.* The suppressor-less *E. coli* K12 strains KH2A (*uvrA*$^-$) and KH2AM (KH2A carrying the *mucAB*$^+$ plasmid pGW270) were constructed as described by Sambamurti et al. (1988). The amber-less phage M13AB28 was constructed as described (Sambamurti et al., 1988). AB28 contains the *E. coli lacZα*-complementing fragment and yields dark blue plaques on suppressor-less *E. coli* cells on appropriate indicator plates.

*DNA Preparation and CAA Adduction.* The preparation of AB28 ss DNA was as described (Refolo et al., 1987). Gapped duplex DNA was constructed as described in detail elsewhere (Jacobsen et al., 1989). In this DNA, a 178 nucleotide long segment of the minus strand (within the *lacZα* gene) has been removed so as to leave the plus strand sequence between bases 6236 and 6413 in a single-stranded conformation (see Figure 2). The preparation of CAA-treated DNA, adduct characterization, and quantitation have been described (Jacobsen et al., 1989). Briefly, the modification procedure involved a 15-min exposure of gapped duplex DNA at room temperature to 71.6 mM CAA, followed by DNA isolation and incubation at 80 °C for 60 min to dehydrate the etheno lesions. This procedure resulted in an average of 2.1 and 3.4 εC and εA lesions, respectively, per DNA molecule. Mock-treated DNA (0 CAA) was prepared by the same protocol, including the post-"modification" 80 °C/60-min incubation.

*SOS Induction and Transfection Procedures.* Procedures for transfection of CAA-treated and mock-treated (no CAA) DNAs into unirradiated *E. coli* cells, plating of infectious centers, and determination of the total number of plaques, as well as the number of *lacZ*$^-$ plaques, were as described (Jacobsen et al., 1989). Transfection efficiency of mock-treated gapped DNA averaged $10^6$ plaque-forming units/μg and was close to that observed for untreated AB28 supercoiled ds DNA. For examining the effect of SOS induction, *E. coli* KH2AM (or KH2A) cells were irradiated with ultraviolet (UV) light to induce the expression of SOS functions. Parameters for UV irradiation were as described (Sambamurti et al., 1988; Bennett et al., 1988).

*Mutant Selection, DNA Isolation, and DNA Sequence Analysis.* Mutant phage were identified as those exhibiting a light blue or colorless plaque phenotype in comparison to the dark blue phenotype of wild-type AB28. Plaque phenotypes were confirmed essentially as described (Kunkel, 1984). Mutant phage ss DNA preparation and sequencing procedures were as described (Refolo et al., 1987).

*In Vitro DNA Replication Analysis.* Phage M13mp8 (Messing & Vieira, 1982) ss DNA or the AB28 gapped duplex DNA was subjected to CAA modification essentially as described for gapped duplex DNA (Jacobsen et al., 1989). Adducted mp8 DNA was incubated for 1 h at either 0 or 65 °C before use. Procedures for analyzing the template characteristics of CAA-treated and mock-treated (no CAA) ss DNA were essentially as described previously (Jacobsen et al., 1987). Briefly, an appropriate oligonucleotide primer was annealed to the ss DNA template, and elongation was initiated at 37 °C by *E. coli* DNA polymerase large "Klenow" fragment [polI(k)] in the presence of either 8 mM $MgCl_2$ or 0.5 mM $MnCl_2$. Elongation products were radiolabeled by a 15-min pulse (2.14 μM [α-$^{32}$P]dATP, 200 Ci/mmol, and 5 μM each of dGTP, dTTP, and dCTP) followed by a 15-min chase (50 μM of each dNTP). Primer elongation was terminated by

Table I: Effect of in Vitro CAA Treatment of Phage M13 AB28 Gapped Duplex DNA on Survival and Mutagenesis in SOS-Induced and Uninduced Cells[a]

| host strain[b] | initial CAA exposure[c] (mM) | plaques/μg of DNA (×10⁶) [% survival] | | % mutagenesis | |
|---|---|---|---|---|---|
| | | −SOS | +SOS | −SOS | +SOS |
| (A) KH2AM | 0 | 5.2 [100] | 4.5 [100] | 0.41 | 1.07 |
| | 71.6 | 1.4 [27.0] | 2.2 [49.2] | 4.92 | 12.6 |
| (B) KH2A | 0 | 7.7 [100] | 3.3 [100] | 0.20 | 0.55 |
| | 71.6 | 2.6 [33.3] | 1.1 [31.6] | 1.66 | 5.43 |

[a] Data shown are averages of two transfections. [b] *E. coli* strain KH2AM is *uvrA*⁻, *mucAB*⁺; KH2A is *uvrA*⁻, *mucB*⁻ (see text). [c] After CAA exposure, DNA was isolated and subjected to incubation for 1 h at 80 °C for conversion of hydrated lesions into the stable dehydrated lesions as described (Jacobsen et al., 1989). The number of adducts per gapped duplex molecule is estimated as 2.1 and 3.4, respectively, for εC and εA [see Jacobsen et al. (1989)].

adding a commercially available (New England Biolabs) formamide-containing "stop" buffer, and the labeled elongation products were fractionated by electrophoresis on high-resolution 8% polyacrylamide–urea gels. Analysis of the replication-blocking potential of CAA lesions was studied by using the same CAA-modified gapped duplex used for in vivo mutagenesis. The single-stranded portion of the DNA was replicated with DNA polI(k) by extending the 3′-hydroxyl at the *Bgl*I site (right boundary of the target, see Figure 2). The elongation products were radiolabeled by a 5-min pulse, and replicate samples chased for 1, 3, 30, or 90 min. Elongation was terminated by rapid freezing on dry ice, and the products were digested with *Mst*II, which cuts the DNA once approximately 75 nucleotides 3′ to the *Bgl*I site. The resulting labeled DNA products were fractionated on high-resolution gels as above alongside standard dideoxy sequence ladders.

## Results

In order to examine the effects of SOS induction on CAA-induced mutagenesis in *E. coli*, we have devised a convenient strategy using the phage M13 *lacZ* based experimental system for the detection and characterization of forward mutations (Jacobsen et al., 1989). This "gapped" duplex strategy, by exploiting the known single-strand preference of CAA, permits the specific adduction of a part of the *lacZα* gene sequence. Furthermore, since the lesions are in the ss DNA, mutations can be unambiguously assigned as those occurring at G, A, T, or C. Plating on appropriate *E. coli* lawn cells and media containing a *lac* inducer and a chromogenic substrate for β-galactosidase permits the identification of wild-type (blue plaques) and mutant (colorless or light blue plaques) phage. The construction of phage M13AB28 as well as the basis for phenotypic selection have been described elsewhere (Sambamurti et al., 1988).

*Effect of CAA Adduction and SOS Induction on Survival and Mutagenesis.* Table IA shows that CAA treatment of DNA under conditions yielding an average of 2 εC and 3 εA lesions within the target sequence enhances mutagenesis by an order of magnitude over the background in unirradiated cells. Upon SOS induction, the fold enhancement in mutagenesis remains at about an order of magnitude over the corresponding background. Since SOS induction increases both background and CAA-induced mutagenesis by about 2.6-fold, there is very little Weigle mutagenesis. The relatively high background mutation frequency is probably the result of DNA manipulations required for construction of the gapped duplex DNA molecules and post-"modification" heat treatment. SOS induction does not significantly enhance survival of CAA-treated DNA. It is difficult to assess the significance of the marginal survival effect observed (27% to 49% in Table IA), because replicate transfections often show a 2–5-fold variation in survival [e.g., Kunkel (1984), Refolo et al. (1987), and Sambamurti et al. (1988)].

Table II: Base Changes in Apparent Wild-Type Plaques Obtained by Transfection of CAA-Treated DNA[a]

| mut no.[b] | base no. | sequence change from | to | mut no.[b] | base no. | sequence change from | to |
|---|---|---|---|---|---|---|---|
| 1 | 6322 | C | A | 14 | 6396 | C | T |
| | 6388 | C | A | 15 | 6380 | G | A |
| 2 | 6391 | C | A | 16 | 6258 | G | A |
| 3 | 6249 | C | A | 17 | 6364 | A | T |
| | 6346 | C | T | 18 | 6238 | C | T |
| | 6347 | C | T | 19 | 6316 | A | T |
| | 6383 | C | T | | 6346 | C | T |
| 4 | 6251 | A | C | 20 | 6396 | C | T |
| 5 | 6242 | G | T | 21 | 6238 | C | T |
| 6 | 6344 | C | T | 22 | 6292 | C | T |
| 7 | 6340 | C | T | 23 | 6347 | C | T |
| | 6344 | C | T | 24 | 6238 | C | T |
| 8 | 6239 | G | A | | 6246 | C | A |
| | 6338 | C | A | 25 | 6238 | C | A |
| 9 | 6340 | C | T | | 6289 | C | T |
| 10 | 6238 | C | T | 26 | 6251 | A | T |
| 11 | 6313 | G | C | | 6367 | A | T |
| | 6340 | C | T | | 6405 | G | C |
| 12 | 6341 | A | G | 27 | 6401 | A | T |
| | 6361 | C | A | | | | |
| 13 | 6249 | C | T | | | | |
| | 6289 | C | T | | | | |

[a] All identified nucleotide changes were at sequences within the single-stranded portion of the gapped duplex molecule (Figure 2). [b] Mutant No.

The gene products of UmuD, UmuC, and RecA are believed to be required for SOS-dependent mutagenesis in *E. coli*. MucA and MucB (MucAB) are inducible plasmid encoded functional homologues to cellular UmuDC proteins and are frequently used in short-term assays to amplify the mutagenic response to certain DNA-damaging agents. For example, McCann et al. (1975) used an excision-deficient *mucAB*⁺ (pKM101) strain of *Salmonella* to detect the mutagenic activity of CAA. Table IB shows that fold enhancements in CAA mutagenesis in SOS-induced as well as uninduced KH2A cells lacking MucAB remain close to 1 order of magnitude. These results suggest that CAA mutagenesis does not require induced levels of SOS functions.

*DNA Sequence Analysis of Background and CAA-Induced Mutations.* Figure 2 represents the DNA sequence of the single-stranded gap region and indicates the nucleotide changes (lower case letters) detected in 30 background mutants (shown above the sequence) and in 49 CAA-induced mutants (shown below the sequence) obtained under SOS-induced conditions. We have previously observed that the gene-targeted adduction strategy employed results in a large number of phenotypically silent (*lacZ*⁺) base substitutions (Jacobsen et al., 1989). To determine the true CAA-induced mutation frequency in SOS-induced cells, 40 plaques displaying a blue (*lacZ*⁺) phenotype were randomly picked from the same plates from which mutant plaques had been picked. Sequence analysis of the 40 apparent *lacZ*⁺ phage revealed that 27 (67.5%) contained one or more base substitution events (Table II).

The data in Figure 2 and Table II are summarized in Table III to show the classes of mutations as well as the specificity of the base changes. Table IIIB shows that, out of 50 CAA-induced *lacZ*⁻ mutants sequenced, 49 (98%) suffer one or more substitutions. In addition, 3 of the 49 also contain a single

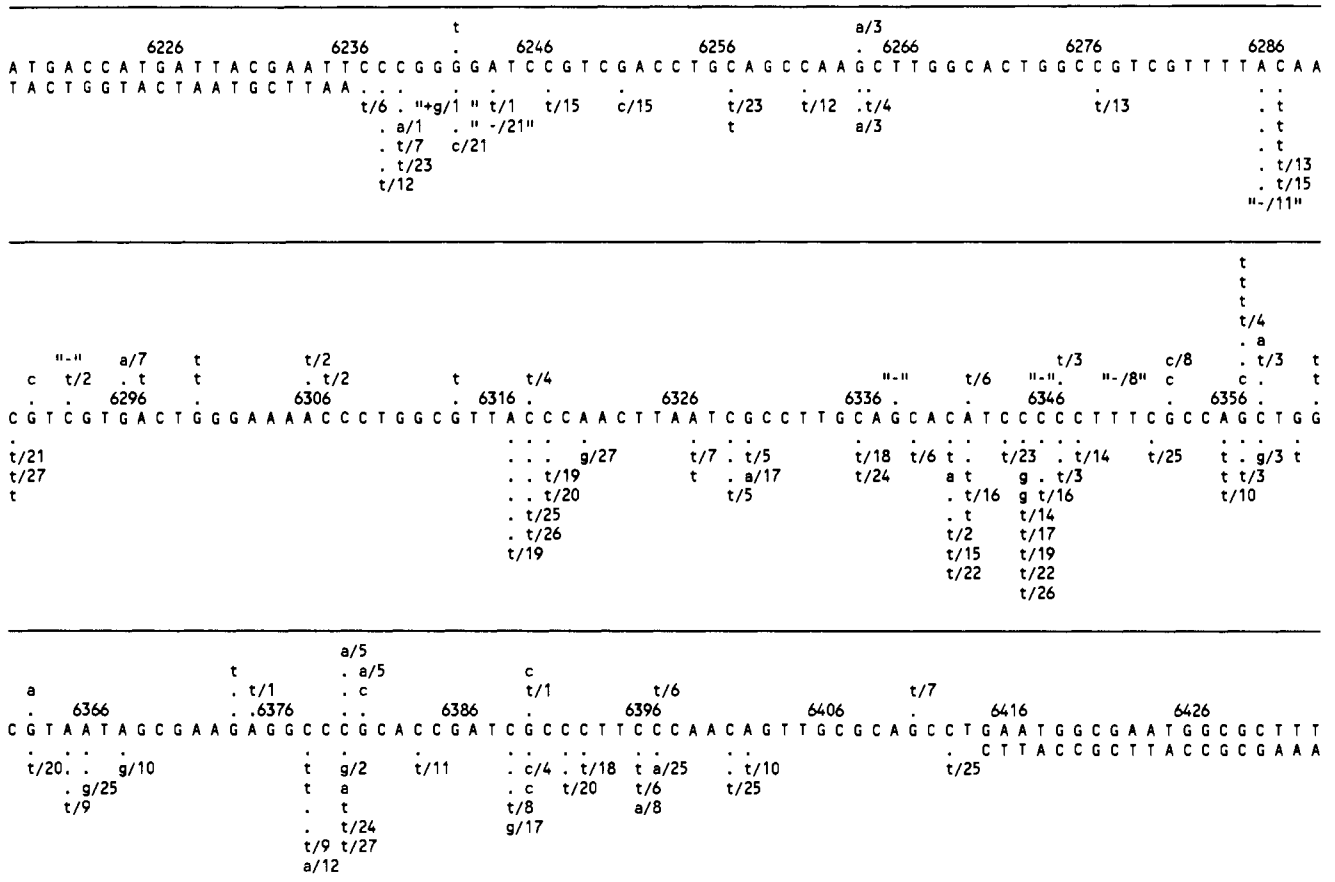


FIGURE 2: Part of the phage AB28 *lacZ* DNA sequence indicating background and CAA-induced mutations. The plus strand is shown on top (written 5′ to 3′). The single-stranded portion of the gapped duplex molecule was created by the procedures described elsewhere (Jacobsen et al., 1989). All mutants sequenced were derived from the transfection of heat-treated (80 °C) gapped DNA into SOS(UV)-induced KH2AM cells. Background mutations are shown above the sequence. Multiple base changes occurring in the same mutant are identified by a common number (e.g., base-change a/3 at position G6263 and change t/3 at C6346 both occur in the same mutant; similarly, the changes t/19 at A6316, t/19 at C6318, and t/19 at C6344 all occur in the same mutant). Dashes enclosed in double-quote marks ("–") indicate the four 1 bp deletions (positions C6292, G6337, and C6345 and a T in TTT6348–6350). CAA-induced mutants are shown below the sequence by use of conventions similar to those for background mutants. A plus enclosed in double-quote marks ("+") indicates the single 1 bp insertion (positions GGGG6239–6242). Three CAA-induced mutations (C to A at 6195, part of mutant a/17; A to T at 6146, part of mutant t/28; and C to T at 6147, part of mutant t/28) mapping upstream of the initiating codon at 6217 are not shown but are included in the data in Table III. These three mutations are arbitrarily depicted as changing plus-strand bases.

Table III: Summary of CAA-Induced Base Changes in Apparent Wild-Type (*lacZ*$^+$) Phage

| DNA treatment (plaque phenotype) | ttl no. seqd[a] | no seq chng[b] | −126 bp mut[c] | mut with base subst[d] | base substitution events[e] | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | events | | | | | | A to | | | C to | | | G to | | |
| | | | | | ×1 | ×2 | ×3 | ×4 | ×6 | ttl | c | g | t | a | g | t | a | c | t |
| (A) background (*lacZ*$^-$)[f] | 34 | 4 | 1 | 26 | 19 | 5 | 2 | 0 | 0 | 35 | 0 | 0 | 4 | 2 | 0 | 6 | 4 | 6 | 13 |
| (B) CAA induced (*lacZ*$^-$)[f] | 50 | 1 | 0 | 49 | 22 | 16 | 7 | 3 | 1 | 93 | 0 | 3 | 13 | 7 | 5 | 55 | 2 | 4 | 4 |
| (C) background (*lacZ*$^+$)[g] | 20 | 20 | | | | | | | | | | | | | | | | | |
| (D) CAA induced (*lacZ*$^+$)[h] | 40 | 13 | 0 | 27 | 16 | 9 | 1 | 1 | 0 | 41 | 1 | 1 | 5 | 8 | 0 | 20 | 3 | 2 | 1 |
| (E) CAA induced (*lacZ*$^-$ and *lacZ*$^+$)[i] | | | 0 | 76 | 38 | 25 | 8 | 4 | 1 | 134 | 1 | 4 | 18 | 15 | 5 | 75 | 5 | 6 | 5 |

[a] Total number sequenced. [b] Number showing no sequence changes between bases 6130 and 6430. [c] 126 bp deletions, representing recombinational background, are analogous to the "93 bp deletions" in M13mp2 (Kunkel, 1984; LeClerc et al., 1984). The polylinker sequence in M13AB28 adds 33 bp to the size of the deletion. [d] Number suffering base substitutions. [e] ×1, ×2, ×3, ×4, ×6, and ttl represent either one, two, three, four, six, or total substitution events, respectively; sequence changes in the plus strand are shown. [f] Data for phenotypically detected (*lacZ*$^-$) mutations (background and CAA induced; 80 °C treatment). [g] Analysis of 20 randomly picked blue (*lacZ*$^+$) plaques resulting from transfection of mock-treated, 80 °C incubated gapped duplex DNA (see text). [h] Analysis of 40 randomly picked blue (*lacZ*$^+$) plaques resulting from transfection of CAA-treated, 80 °C incubated gapped duplex DNA (see text). [i] Pooled data (sum of rows B and D) on base changes observed as a result of CAA treatment of DNA.

frame-shift event. These include two 1 bp deletions (at A6243 and A6285; see Figure 2) and a G insertion in the GGGG run at 6239–6242. The 34 background *lacZ*$^-$ mutants (Table IIIA) are comprised of fewer base substitutions (76%). One mutant has a single substitution at G6352 and a 1 bp deletion at a TTT run at 6348–6350 (mutations marked with /8 in Figure 2). Three frame-shift mutants arise by the loss of a single base pair (C6292; G6337; and a C in the CCCCC run at 6343–6347; Figure 2). The data in Table IIID show that CAA-induced silent mutations have a specificity similar to that of phage with a *lacZ*$^-$ phenotype. Sequence analysis of 20 randomly picked *lacZ*$^+$ plaques from control DNA plates showed that all 20 had the wild-type sequence (Table IIIC). These data show that, under SOS conditions, CAA mostly induces base substitutions, a result similar to that obtained in unirradiated cells (Jacobsen et al., 1989).

Table IV: Analysis of CAA Enhancement of Mutagenesis with and without SOS Induction at A, C, G, and T Residues[a]

| | rel mutation frequency ($10^{-3}$) | | | |
|---|---|---|---|---|
| | background | | CAA | |
| | -SOS | +SOS | -SOS | +SOS |
| all mutations | 4 | 10.7 | 575 | 801 |
| at A | 0.37 | 1.22 | 34 | 137 |
| at C | 0.73 | 2.45 | 520 | 568 |
| at G | 2.9 | 7.03 | 21 | 96 |
| at T | | | | |

[a] No mutations were detected at thymines; the "-SOS" data are from Jacobsen et al. (1989). The "+SOS" background mutation frequency is from Table I. CAA-induced "unselected" mutation frequency was derived from the data in Table IIID (*lacZ*$^+$, CAA induced) as follows:

$$\frac{\text{no. of plaques with base changes (27)}}{\text{no. of plaques sequenced (40)}} = 0.675$$

It should be noted, however, that the so-called "unselected" mutants (blue plaques) are in fact "reverse-selected" in the sense that light blue and colorless plaques (together constituting 12.6% of all plaques; Table I) are omitted from this collection. Therefore, the true total mutation frequency (MF) is derived as follows:

$$\text{MF} = \text{unselected MF (0.675)} + \text{selected MF (0.126)} = 0.801$$

Relative mutation frequency (RMF) is the fractional mutation frequency contributed by mutations occurring at G, A, T, or C to the MF, as shown for CAA-induced adenine targeted mutations:

$$\text{RMF} = \frac{\text{no. of A-targeted substitutions (23)}}{\text{total substitutions (134)}}[\text{MF (0.801)}]$$

(Note that substitution frequency should be used in the place of MF in this calculation; however, since an overwhelming majority of CAA-induced mutations are substitutions, substitution frequency is taken to equal MF.) There are no significant specificity differences in the selected and silent mutations: e.g., the fraction of C-targeted mutations in the selected pool (Table IIIB) is 67/93 = 0.72, and for the unselected mutations (Table IIID), it is 28/41 = 0.68. Therefore, using the pooled sequence data simply serves to increase the number of total mutations without a significant effect on the fraction of the mutations occurring at each of the four bases. RMF values can also be calculated specifically for selected mutations (MF = 0.126; Table I), or specifically for silent mutations (MF = 0.675). Since the true mutation frequency is determined by the number of mutants in the unselected as well as the selected pool as shown above, RMF figures for the pooled data are likely to be the most representative.

*Specificity of CAA-Induced Base Substitution Mutations.* Table IIIE shows that the majority of CAA-induced base substitution mutations are targeted to cytosines (71%), with fewer occurring at adenine (17%) and guanine (12%) residues, and with none occurring at thymines. The corresponding figures for background substitutions are C (23%), A (11%), and G (66%). Table IIIE also shows that, of the 95 CAA-induced base changes affecting cytosines, 75 are C-to-T (79%) mutations, followed by 15 C-to-A (16%) and 5 C-to-G (5%) mutations. Of the 23 adenine-targeted substitutions, 13 (78%) are A-to-T, 3 are A-to-G, and 1 is an A-to-C substitution. These results, in comparison to those in uninduced cells (Jacobsen et al., 1989), show that SOS induction does not significantly alter the specificity of base changes induced by CAA.

*Efficiency of CAA-Induced Mutagenesis.* Relative mutation frequency calculations (Table IV) show that CAA treatment increases overall mutagenesis as well as cytosine-specific mutagenesis to a similar extent in SOS-induced and uninduced cells. Substitutions at A and G residues are enhanced nearly 4-fold with SOS induction. However, in comparison to mutation fold enhancement at cytosine (200–600-fold) and at adenine (100-fold), the increase at guanine (6–10-fold) is modest.

Table V: Efficiency of Mutagenesis of CAA-Induced Cytosine Lesions with and without SOS Induction

| | -SOS[a] | +SOS[b] |
|---|---|---|
| (A) surviving fraction | 0.300 | 0.490 |
| (B) mutant fraction[c] | 0.575 | 0.801 |
| (C) wild-type fraction (1 − B)[c] | 0.425 | 0.199 |
| (D) corrected wild-type fraction (C × A)[d] | 0.128 | 0.098 |
| (E) corrected mutant fraction (B × A)[d] | 0.173 | 0.392 |
| (F) no. of mutants (selected and unselected) with base substitutions | 69 | 76 |
| (G) no. of mutants with single C-targeted substitutions[e] | 48 | 24 |
| (H) mutant fraction with single C-targeted events [E × (G/F)] | 0.120 | 0.124 |
| (I) mutagenic efficiency at single C-targeted events (H/0.271)[f] | 0.443 | 0.457 |

[a] "-SOS" values are taken from indicated sources in Jacobsen et al. (1989). [b] "+SOS" values are taken from sources indicated in the present paper. [c] Fraction of survivors with (B) or without (C) base changes. [d] Wild-type (D) and mutant (E) fractions corrected for survival. [e] These figures were obtained by counting the number of mutants suffering only one C-targeted mutation [Figure 4 of Jacobsen et al. (1989); Figure 2 and Table II]. [f] At an average of 2 $\epsilon$C lesions per molecule, the Poisson fraction of DNA molecules suffering 0, 1, 2, and more than 2 $\epsilon$C lesions are 0.135, 0.271, 0.271, and 0.323, respectively.

As considered further under Discussion, mutagenic cytosine lesions induced by CAA are probably $\epsilon$C (or $\epsilon$C·$H_2O$). From the average number of $\epsilon$C lesions per DNA molecule, and the frequency of single-base substitution events at template C residues, it can be argued that the efficiency of mutagenic bypass of single $\epsilon$C lesions is over 40% (Table V) in both SOS-induced and uninduced cells. These calculations should be viewed as first approximations until the availability of data with site-specific $\epsilon$C lesions at a number of sites. Similar calculations for events at A residues (not shown) suggest that mutagenesis at single $\epsilon$A lesions is inefficient (3%) in the absence of SOS induction but increases (to about 13%) in SOS-induced cells. An interesting feature of the analysis in Table V is that the corrected wild-type fractions (Table VD; 0.128 for -SOS and 0.098 for +SOS) are close to the Poisson fraction of molecules bearing zero $\epsilon$C lesions (0.135). Since at 3 $\epsilon$A lesions (on the average per DNA molecule) the Poisson fraction of DNA molecules with zero $\epsilon$A lesions is less than 0.05, these calculations imply that $\epsilon$A lesions make only a small contribution to mutagenesis and contribute little to lethality.

*Effect of DNA Sequence Context on the Frequency and Specificity of CAA-Specific Substitution Mutations at Template C Residues.* In order to analyze the effect of nucleotide sequence on CAA-induced mutagenesis at cytosines, we have plotted the normalized relative mutation frequency for each of the 14 NCN triplet sequences found (out of 16 possible NCN sequences; Figure 3) in the ss DNA "gap". The salient points derived from this analysis can be summarized as follows. (1) Mutagenesis is detected at all triplet nucleotide sequence environments represented within the ss DNA target (although not necessarily at every occurrence of the same triplet sequence), supporting the notion that CAA probably reacts randomly with all non-base-paired C residues [see Jacobsen et al. (1987)]. (2) Normalized data for C-to-N events (Figure 3) or C-to-T and C-to-A events (data not shown) are not significantly affected (Student's *t* test) by SOS induction. (3) In general, 5′-C-Pu-3′ sequences are associated with higher frequencies of mutagenesis as compared to 5′-C-Py sequences.

*Effect of CAA Modification of Template DNA on Primer Elongation by polI(k).* In order to determine whether CAA adducts stop or attenuate DNA synthesis, M13mp8 ss DNA was subjected to CAA modification, and DNA replication on the modified template was initiated by primer extension in the

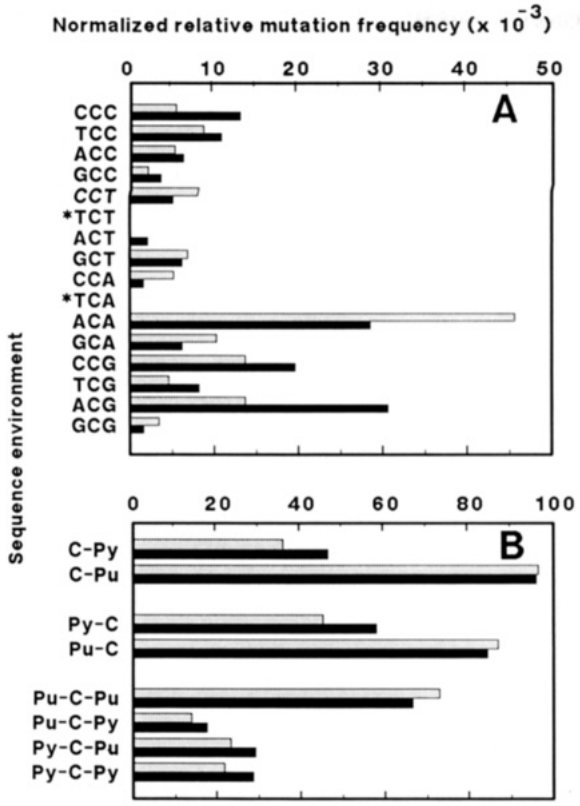


FIGURE 3: Effect of SOS induction and nucleotide sequence environment on CAA-induced base substitution mutations at C lesions. Normalized relative mutation frequencies (RMF) at each of the 14 NCN triplets found within the "gap" region (out of a possible 16) are depicted. The normalized RMF was calculated from the RMF for all mutations occurring in a particular triplet and the number of occurrences of the triplet within the ss DNA gap sequence, as illustrated for the triplet 5′-ACA (+SOS conditions) in eq 1 and 2.

$$\text{RMF(5′-ACA)} = \frac{\text{C mutations in 5′-ACA (14)}}{\text{all events at C (93)}}[\text{RMF(C) (0.568)}] \ (=0.0855) \quad (1)$$

$$\text{normalized RMF(5′-ACA)} = \frac{\text{RMF(5′-ACA) (0.0855)}}{\text{no. of 5′-ACA triplets (3)}} \ (=0.0285) \quad (2)$$

The RMF(C) value is from Table IV. Grey bars represent −SOS data, and dark bars, +SOS data. The sequences 5′-TCT or 5′-TCA (asterisks) are not represented in the 178-base ss DNA gap, and therefore, no data are available for these sequence contexts. Panel A, effect of each individual base 5′ or 3′ to C; panel B, effect of purine (Pu) or pyrimidine (Py) flanking the C. In the above analysis, 28 unselected and 67 selected C-targeted events distributed over 51 sites are analyzed. The rationale for pooling the sequence data is given in the footnote for Table IV. Because of the inclusion of the unselected mutations, the pool is taken to represent most of the base changes induced by CAA within the target sequence. It should be noted that, with or without the "normalization" (eq 2 above), the major conclusions regarding mutational hot spots as well as the lack of SOS effect remain the same.

presence of *E. coli* DNA polI(k), divalent cation, and nucleotide precursors (Rabkin & Strauss, 1984). Labeled elongation products were analyzed on DNA sequencing gels. Preliminary results using this approach had shown that CAA-induced DNA damage creates pause sites at A and C residues (Jacobsen et al., 1987). It was observed that replication terminated predominantly one nucleotide before (3′ to) the lesion, and less frequently opposite the lesion. Figure 4 shows that the substitution of $Mn^{2+}$ for $Mg^{2+}$ (known to be mutagenic; Sirover & Loeb, 1976) causes qualitative changes in replication stops at A and C lesions. At certain sites, there is a shift in the replication pause site from one nucleotide before

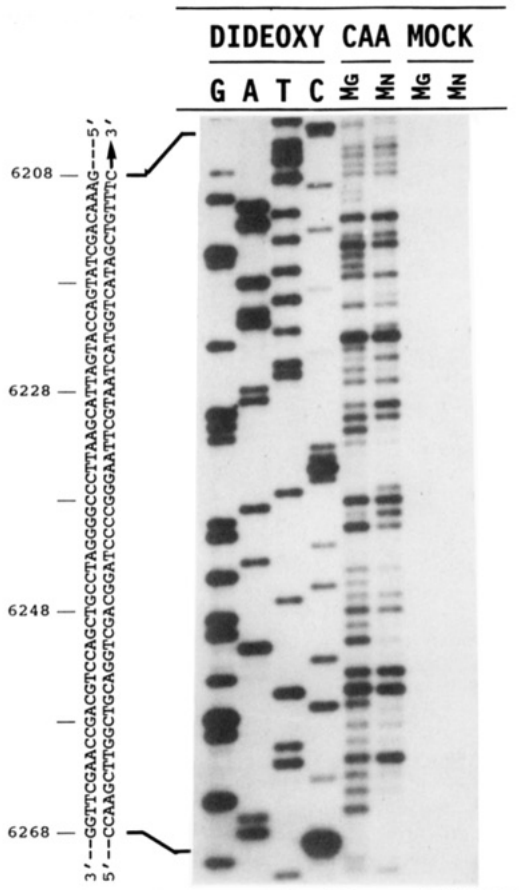


FIGURE 4: In vitro replication analysis of CAA-modified M13mp8 ss DNA template in presence of either $Mg^{2+}$ or $Mn^{2+}$. The lanes G, A, T, and C represent standard Sanger (dideoxyribonucleotide) sequence ladders. The relevant part of M13mp8 *lacZ* sequence (primer strand indicated by arrow) is shown. CAA-modified or mock-modified (no CAA control) ss DNA template was replicated with DNA polI(k) by primer extension in the presence of either 8 mM $MgCl_2$ (Mg) or 0.5 mM $MnCl_2$ (Mn) as described under Materials and Methods. The series of bands in CAA-modified ss DNA lanes represent polymerase pause sites. In mock-treated DNA lanes, there are no visible bands, indicating lack of chain termination. (With mock-treated templates, high molecular weight DNA products traveling at the top of the gel are observed; data not shown.) The pause sites in CAA-treated DNA correspond to chain termination events opposite or, more frequently, one nucleotide before (3′) template C or A residues. CAA-treated DNA replicated in the presence of $MnCl_2$ shows fewer pause sites in comparison to the DNA replicated with $MgCl_2$. In the presence of $MnCl_2$, several bands are apparently completely bypassed (e.g., band one nucleotide 3′ to template C-6264 at the bottom of the CAA/Mg lane), while at others, there is a "shift-up" effect, suggesting that terminations occurred more frequently opposite a template C or A residue (e.g., compare terminations opposite template C-6236 in the CAA/Mg and CAA/Mn channels).

the lesion to a position opposite the lesion. At other sites, a complete bypass of the lesion is apparent. These effects, which appear to be influenced by the sequence context, are observed at several A and C residues. These results suggest that CAA adduction specifically creates pause sites at adenine and cytosine residues and that polI(k) can incorporate bases opposite CAA-induced adenine and cytosine lesions in vitro.

To determine whether CAA-induced lesions introduced absolute replication blocks to polymerase or kinetic (temporary) pause sites, the ss DNA portion of CAA-modified gapped duplex DNA (used in mutagenesis studies described above) was replicated by elongating the 3′-OH end of the complementary strand as described (see Materials and Methods and legend to Figure 5). Figure 5 shows that, with $Mg^{2+}$ as a cofactor, polymerase pause sites are observed initially (1–3

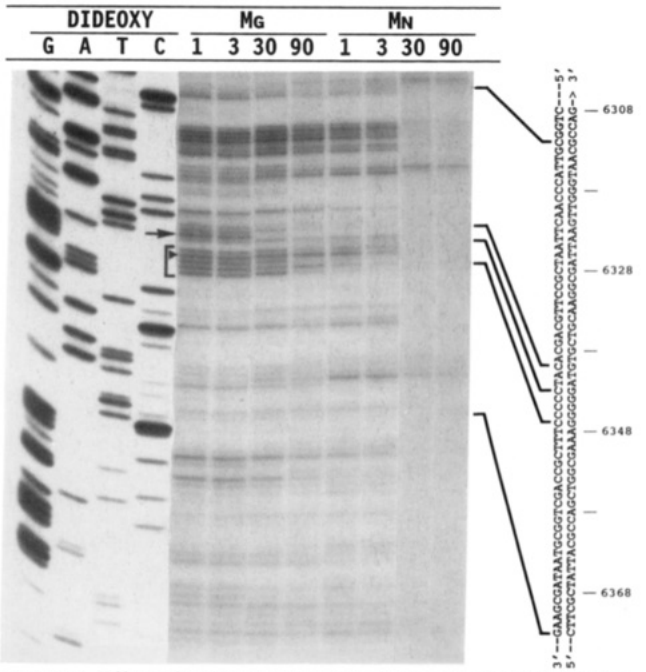


FIGURE 5: Effect of time of incubation on the in vitro polI(k) pause sites on CAA-treated ss DNA template contained within the gapped duplex. For these experiments, the ss DNA portion of CAA-modified gapped duplex DNA was replicated from the 3′-hydroxyl terminus present opposite template G6414 in the presence of polI(k), nucleotide precursors, and either $Mg^{2+}$ (Mg) or $Mn^{2+}$ (Mn) as described; see Materials and Methods. The numbers 1, 3, 30, and 90 indicate the number of minutes the elongation reaction was allowed (following an initial 5-min pulse) before termination. G, A, T, and C are standard Sanger sequence ladders. A relevant part of the ss DNA sequence is shown (primer strand identified by an arrow). The gradual decrease in relative intensity of bands corresponding to pause sites with time is observable at several sites in the Mg lanes and is much more pronounced in the Mn lanes for any given time interval. For an explanation of the arrow, right-pointing triangle, and bracket, see text.

min) at almost all A and C residues. Some of the pause sites disappear at longer time intervals (30–90 min), indicating that these lesions temporarily stall DNA synthesis and are not absolute blocks. The figure also suggests that $Mn^{2+}$, as compared to $Mg^{2+}$, increases the efficiency of (i.e., decreases the time required for) in vitro translesion synthesis.

The data in Figure 5 suggest that bypass synthesis may occur more readily at some sequences than at others. The example marked by the arrow corresponds to a template C residue in the sequence 5′-ACA-3′ (6339–6341; see Figure 3), known to be a mutational hot spot in vivo. (This effect was also observed at a second 5′-ACA-3′ site at 6285–6287; data not shown.) Despite this apparent correlation between in vitro bypass and in vivo mutagenesis, in vitro translesion bypass may not accurately predict in vivo mutagenesis. Figure 5 shows that equivalent pause sites are initially formed at all five bases in the CCCCC run at 6343–6347 (marked by the bracket). However, at 90 min ($Mg^{2+}$) the blocks at 6343, 6345, and 6347 (corresponding to the first, third, and fifth bracketed bands, respectively, from the top) are diminished in comparison to those at 6344 and 6346 (corresponding to the second and fourth bracketed bands, respectively). Nevertheless, in vivo mutations were observed at position 6344 (closed triangle, Figure 5; also see Figure 2), but not at the lesions that appear to be more readily bypassed in vitro. This observation raises the possibility that guanine may be incorporated opposite C lesions at certain sites in vivo, resulting in error avoidance. In the above analysis, we have assumed that equivalent pause sites imply equivalent (random) adduction, an assumption that remains to be tested. Finally, caution is required in extrapolating in vitro replication results obtained with a model polymerase such as polI(k) to in vivo mutagenesis, which is undoubtedly much more complex.

## Discussion

*Mutagenic DNA Lesions Induced by CAA.* The preponderance of evidence, reviewed briefly by Jacobsen et al. (1989), indicates that the mutagenic lesions induced by CAA are etheno derivatives, notably ethenocytosine. It is conceivable that a small fraction of the C-to-T mutations, but not the approximately 20% mutations changing C to A or G, are caused by simple deamination of cytosine during the in vitro modification procedures. However, the following considerations suggest that the cytosine deamination to yield DNA uracil is unlikely to play a significant role in the observed mutagenesis. (1) Hall et al. (1981), who specifically looked for cytosine deamination during CAA treatment, did not find significant deamination. (2) The postlabeling procedure used by us can detect the formation of deoxyuridine at significant levels, but none was detectable by this method. (3) In order to increase the sensitivity of detection, we have examined deamination of deoxycytidine 5′-monophosphate subjected to the various modification and incubation conditions used here (data not shown). CAA treatment by itself resulted in little deamination above the background. The postmodification 80 °C/1-h incubation, as expected, did result in detectable deamination, which, however, constituted a small fraction of $\epsilon$C formed during the procedures. No detectable deamination occurred after a 0 °C/1-h incubation. (4) Jacobsen et al. (1989) have shown that the frequency and specificity of mutations obtained by transfection of DNA subjected to 0 °C/1-h incubation do not significantly differ from those of 80 °C/1-h incubated DNA. (5) The background mutations analyzed here (Table IIIA) and elsewhere (Jacobsen et al., 1989) were obtained by transfecting mock-treated DNA subjected to the 80 °C/1-h incubation. Nevertheless, C-to-T mutations are a smaller fraction of background mutations as compared to mutations at guanines. Even if it is unlikely that simple cytosine deamination could account for a significant fraction of mutagenesis, random-adduction experiments of the type described here always leave open the possibility that other, as yet unidentified, primary or secondary cytosine lesion(s) may be responsible for the observed mutagenesis. Nevertheless, this work does establish that in vitro CAA treatment of ss DNA results in a large increase in SOS-independent mutagenesis specifically at cytosines and that this in vivo mutagenesis correlates with a CAA dose-dependent formation of $\epsilon$C lesions in DNA.

*Role of Cytosine Lesions in Mutagenesis Induced by Vinyl Chloride.* Vinyl chloride, a major industrial chemical produced in large quantities around the world, is a well-known carcinogen. Nevertheless, the DNA lesions responsible for the genotoxicity of vinyl chloride are not known. Vinyl chloride is known to be metabolized in vivo to 2-chloroethylene oxide, which can spontaneously and rapidly rearrange to CAA (Guengerich et al., 1979). The major DNA adduct induced by vinyl chloride (via chloroethylene oxide) in vitro and in vivo is *N*-7-(2-oxoethyl)guanine (oxet-G; Laib et al., 1981). In vitro reaction of metabolically activated vinyl chloride with DNA (Laib et al., 1981) and RNA (Laib et al., 1977, 1978) also yields $\epsilon$C and $\epsilon$A; reaction with DNA also yields $N^2$,3-ethenoguanine (Oesch & Doerjer, 1982). After the original demonstration of $\epsilon$C and $\epsilon$A in the livers of rats subjected to low-level vinyl chloride exposure over a period of time (Green & Hathway, 1978), this finding could not be confirmed until recently. Since induction of $\epsilon$C and $\epsilon$A lesions in DNA

probably occurs only in single-stranded regions, these lesions are expected to be present at very low levels, making their detection and quantitation in genomic DNA difficult. However, very recently, Eberle et al. (1989), using a specific monoclonal antibody, have detected and quantitated εC and εA in liver and lung DNA of rats exposed to vinyl chloride.

In bacterial reversion systems, 2-chloroethylene oxide induces mostly G·C to A·T transitions, and fewer A·T to T·A transversions (Barbin et al., 1985a), and this mutagenesis is largely SOS-independent (Barbin et al., 1985c). Therefore, both the mutational specificity and the lack of SOS dependence of 2-chloroethylene oxide are strikingly similar to those of CAA [this work and Jacobsen et al. (1989)] and can be readily explained by postulating that the significant mutagenic lesions formed by chloroethylene oxide exposure were in fact the same cytosine lesions as those formed by CAA in vitro. However, since the reversions could have arisen by DNA lesions at either a guanine or a cytosine, it is important to evaluate whether the oxet-G adduct could lead to this specificity. Because of a lack of in vitro miscoding properties, oxet-G is believed to be nonmutagenic (Barbin et al., 1985b). Since guanine N7 adduction by bulky chemicals can lead to mutagenesis [e.g., Sambamurti et al. (1988)], it is premature to dismiss a mutagenic role for oxet-G. Any hypothesis implicating the guanine adduct (e.g., an abasic site intermediate), however, must account for both the mutational specificity as well as the SOS independence in bacterial test systems.

*Possible Mechanisms of Mutagenesis by εC.* The specificity of mutagenesis at cytosines suggests that C lesions are treated as noninstructional lesions by the polymerase. Such a hypothesis is consistent with the expected loss of normal hydrogen-bonding capability in εC (Figure 1) and with the observation that DNA containing hydrated as well as dehydrated εC lesions yields similar patterns of mutagenesis (Jacobsen et al., 1989). This hypothesis must, however, account for the SOS-independent, highly efficient mutagenic bypass of cytosine lesions and the observation that these lesions cause temporary pause sites in vitro rather than absolute blocks expected of "true" replication-blocking lesions such as aflatoxin $B_1$ (Refolo et al., 1985; Jacobsen et al., 1987). The apparent inconsistency can be explained by assuming that εC lesions do not impede the progression of the polymerase because their structural properties permit the maintenance of the integrity of the helix in the absence of proper hydrogen bonding. Two caveats need to be considered here: First, the *E. coli* DNA polymerase utilized in mutagenic DNA replication past the DNA lesions in the gapped DNA in vivo is not known and can be either polIII, polI, or even polII. Therefore, the observed SOS independence may have other explanations. Second, even though inferences on in vivo mutagenic mechanisms have been made in the literature on the basis of simple in vitro replication systems using model polymerases such as polI(k), the validity of such extapolation remains to be investigated.

Finally, it is possible that εC has multiple miscoding properties which account for the various types of cytosine-targeted mutations observed. However, until convincing miscoding schemes become available, it may be useful to consider these lesions to be a special class of highly mutagenic noninstructional lesions.

*Mutagenesis and Error Avoidance at εA Lesions.* Jacobsen et al. (1989) have previously suggested that εA lesions exist in an equilibrium between an "unfavorable" and a "favorable" configuration. In the unfavorable state, they act as mutagenic noninstructional lesions responsive to SOS functions. This state can be imagined as a helix-distorting conformation. While the present data do show that mutagenesis at adenines (unlike that at cytosines) is SOS-responsive, the facile, error-free, bypass of a majority of these lesions is difficult to explain.

It is possible that a significant fraction of εA lesions are spontaneously or enzymatically reversed before replication in vivo. This might also account for the observation that in vitro replication pause sites are found at both adenines and cytosines (Figures 4 and 5), and yet adenine lesions do not appear to cause significant lethality. However, such a hypothesis must account for the following observations: (1) the extensive literature on these derivatives, which were extensively used as enzyme substrate analogues for adenine (Leonard, 1984), suggests that these are stable lesions; (2) repair enzymes are generally believed not to act on ss DNA and therefore may not be able to eliminate εA lesions before replication; (3) in the present system (as for other ss DNA phages), a strand discontinuity caused by excision repair is expected to be lethal, whereas εA lesions appear to contribute little to lethality. In order to account for the error-free bypass of these lesions, we have previously suggested that εA lesions in the alternative "favorable" conformation are capable of templating the incorporation of thymine at a high efficiency (Jacobsen et al., 1989). Whether such "templating" involves stabilization of an εA·T "base pair" by novel hydrogen bonds, or repesents a novel helix-stabilizing mechanism involving stacking and other hydrophobic forces, remains to be investigated.

*Effect of Sequence Environment on CAA-Induced Mutagenesis.* The availability of the DNA sequences of a large number of phenotypically selected as well as silent mutations presented in this paper and elsewhere (Jacobsen et al., 1989) offers an opportunity to examine the effects of sequence context on the mutational process. There is at present no evidence that the sequence environment influences the reactivity of CAA with unpaired adenine or cytosine. [This is distinct from the decreased reactivity of adenine or cytosine in base-paired regions such as the stems of hairpins; see Jacobsen et al. (1987)]. Even though interpretation of such data is not unambiguous, Figure 5 shows that initial pause sites for a model polymerase at many cytosine sites are equivalent, implying random adduction. The distribution of mutations is distinctly nonrandom (Figures 2 and 3), a phenomenon echoed as variability in the bypass efficiency of these lesions during in vitro replication by a model polymerase (Figure 5). The in vivo observations presented here imply that the mutagenic potential of a DNA lesion is strongly modulated by the sequence environment. In that sense, an identical chemical lesion can in fact give rise to a number of distinctly different premutagenic lesions.

## Acknowledgments

We thank X. Luo and J. Callahan for assistance with DNA sequencing and K. Sambamurti for strains and helpful discussions.

**Registry No.** CAA, 107-20-0.

## References

Barbin, A., Besson, F., Perrard, M.-H., Be're'ziat, J.-C., Kaldor, J., Michel, G., & Bartsch, H. (1985a) *Mutat. Res. 152*, 147–156.

Barbin, A., Laib, R. J., & Bartsch, H. (1985b) *Cancer Res. 45*, 2440–2444.

Barbin, A., Tenenbaum, L., Toman, Z., Radman, M., & Bartsch, H. (1985c) *Mutat. Res. 152*, 157–159.

Bennett, C. B., Luo, X., Refolo, L. M., & Humayun, M. Z. (1988) *Mutat. Res. 202*, 223–234.

Eberle, G., Barbin, A., Laib, R. J., Ciroussel, F., Thomale, J., Bartsch, H., & Rajewsky, M. F. (1989) *Carcinogenesis 10*, 209-212.
Green, T., & Hathway, D. E. (1978) *Chem.-Biol. Interact. 22*, 211-224.
Guengerich, F. P., Crawford, W. M., Jr., & Watanabe, P. G. (1979) *Biochemistry 18*, 5177-5182.
Hall, J. A., Saffill, R., Green, T., & Hathway, D. E. (1981) *Carcinogenesis 2*, 141-146.
Jacobsen, J. S., Refolo, L. M., Conley, M. P., Sambamurti, K., & Humayun, M. Z. (1987) *Mutat. Res. 179*, 89-101.
Jacobsen, J. S., Perkins, C. P., Callahan, J. T., Sambamurti, K., & Humayun, M. Z. (1989) *Genetics 121*, 213-222.
Kunkel, T. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1494-1498.
Laib, R. J., & Bolt, H. M. (1977) *Toxicology 8*, 185-195.
Laib, R. J., & Bolt, H. M. (1978) *Arch. Toxicol. 39*, 235-240.
Laib, R. J., Gwinner, L. M., & Bolt, H. M. (1981) *Chem.-Biol. Interact. 37*, 219-231.
LeClerc, J. E., Istock, N. L., Saran, B. R., & Allen, R. (1984) *J. Mol. Biol. 180*, 217-237.
Leonard, N. J. (1984) *CRC Crit. Rev. Biochem. 15*, 125-199.
McCann, J., Simmon, V., Streitweiser, D., & Ames, B. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3190-3193.
Messing, J., & Vieira, J. (1982) *Gene 19*, 269-278.
Oesch, F., & Doerjer, G. (1982) *Carcinogenesis 3*, 663-665.
Rabkin, S. D., & Strauss, B. S. (1984) *J. Mol. Biol. 178*, 569-594.
Refolo, L. M., Conley, M. P., Sambamurti, K., Jacobsen, J. S., & Humayun, M. Z. (1985) *Proc. Acad. Natl. Sci. U. S.A. 82*, 3096-3100.
Refolo, L. M., Bennett, C. B., & Humayun, M. Z. (1987) *J. Mol. Biol. 193*, 609-636.
Sambamurti, K., Callahan, J., Luo, X., Perkins, C. P., Jacobsen, J. S., & Humayun, M. Z. (1988) *Genetics 120*, 863-873.
Shapiro, R. (1969) *Ann. N.Y. Acad. Sci. 163*, 624-630.
Sirover, M. A., & Loeb, L. A. (1976) *Science 194*, 1434-1436.

# Proton NMR Studies of Transforming and Nontransforming H-*ras* p21 Mutants

Ilme Schlichting, Jacob John, Mathias Frech, Pierre Chardin,[‡] Alfred Wittinghofer, Herbert Zimmermann,[§] and Paul Rösch*

*Department of Biophysics, Max-Planck-Institut for Medical Research, Jahnstrasse 29, D-6900 Heidelberg, West Germany*

*Received April 18, 1989; Revised Manuscript Received August 24, 1989*

ABSTRACT: One- and two-dimensional nuclear magnetic resonance spectroscopy (1D and 2D NMR) and site-directed mutagenesis were used to study the influence of mutations on the conformation of the H-*ras* oncogene product p21. No severe structural differences between the different mutants, whether they were transforming or nontransforming, could be detected. Initially, selective incorporation of 3,5-deuterated tyrosyl residues into p21 and 2D NMR were used to identify the resonances representing the spin systems of the imidazole rings of the three histidyl residues in the protein, of six of the nine tyrosyl rings, and of four of the five phenylalanyl rings. The spin systems of the phenyl rings of $Phe^{28}$, $Phe^{78}$, and $Phe^{82}$ could be assigned by using mutant proteins, since no severe structure-induced spectral changes in the aromatic part of the spectra of the mutant proteins were detected. Sequence-specific assignments of the histidine imidazole resonances could be obtained by comparison of the distance information obtained by nuclear Overhauser enhancement spectroscopy (NOESY) experiments with the crystal structure. The change in the chemical shift values of the H1′ proton and the $\alpha$-phosphate of the bound GDP in the NMR spectra of the p21(F28L) mutant and the 28-fold increase in the GDP dissociation rate constants of this mutant suggest a strong interaction between $Phe^{28}$ and the p21-bound nucleotide. In solution, the p21-bound GDP·$Mg^{2+}$ has an anti conformation, and the phenyl ring of $Phe^{28}$ is close to the ribose of the bound GDP·$Mg^{2+}$.

The products of the *ras* gene family are highly related proteins of molecular weight 21 000 termed p21. They have chain lengths of 189 amino acids; their sequences are identical for the N-terminal 80 amino acids and over 85% identical up to amino acid 164/165. The C-terminal 25 amino acids are very divergent, except for the Cys-A-A-X-OH motif at the end of the chain (A being an aliphatic residue). Normal (cellular) *ras* genes acquire transforming properties by single point mutations within their coding sequences. *ras* genes carrying these mutations have been detected in a significant fraction of human cancers as well as in experimentally induced animal tumors [for a recent review, see, e.g., Barbacid (1987)]. p21 proteins are thus believed to play an essential role in cellular growth and/or development (Bishop, 1983). They bind guanine nucleotides with high affinity and specificity and exhibit low GTPase activity (Feuerstein et al., 1987; Gibbs et al., 1984; McGrath et al., 1984; Sweet et al., 1984; Manne et al., 1985). Due to their significant sequence homology and biochemical similarities to guanine nucleotide (G) binding proteins such as transducin and the bacterial elongation factor Tu (EF-Tu),[1]

[‡] Present address: INSERM U-248 Faculté de Medecine Lariboisiere—Saint-Louis, 10, avenue de Verdun, 75010 Paris, France.
[§] Present address: Molecular Crystal Group, Max-Planck-Institut for Medical Research, Molecular Crystal Group, Jahnstrasse 29, D-6900 Heidelberg, West Germany.

[1] Abbreviations: 1D, one dimensional; 2D, two dimensional; COSY, correlated spectroscopy; DQF, double quantum filtered; DSS, sodium 2,2-dimethyl-2-silapentanesulfonate; EDTA, ethylenediaminetetraacetic acid; EF-Tu, bacterial elongation factor EF-Tu; GMP, guanosine 5′-monophosphate; GDP, guanosine 5′-diphosphate; GTP, guanosine 5′-triphosphate; H-ras, Harvey *ras*; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; NMR, nuclear magnetic resonance; $p21_c$, cellular H-*ras* p21; $p21_v$, viral H-*ras* p21 (G12R, A59T); $pK_a$, apparent p*K* value; TPPI, time-proportional incrementation method.

0006-2960/90/0429-504$02.50/0 