Environmental Health Perspectives
Vol. 76, pp. 29-32, 1987

Nearest Neighbor Affects G:C to A:T Transitions Induced by Alkylating Agents

by Barry W. Glickman,* Michael J. Horsfall,* Alasdair J. E. Gordon,* and Philip A. Burns*

The influence of local DNA sequence on the distribution of G:C to A:T transitions induced in the *lacI* gene of *E. coli* by a series of alkylating agents has been analyzed. In the case of nitrosoguanidine, two nitrosoureas and a nitrosamine, a strong preference for mutation at sites proceeded 5′ by a purine base was noted. This preference was observed with both methyl and ethyl donors where the predicted common ultimate alkylating species is the alkyl diazonium ion. In contrast, this preference was not seen following treatment with ethylmethanesulfonate. The observed preference for 5′PuG-3′ site over 5′-PyG-3′ sites corresponds well with alterations observed in the Ha-*ras* oncogene recovered after treatment with NMU. This indicates that the mutations recovered in the oncogenes are likely the direct consequence of the alkylation treatment and that the local sequence effects seen in *E. coli* also appear to occur in mammalian cells.

Introduction

Alkylating agents represent a major class of potential human mutagens and carcinogens. These agents give rise to a large number of DNA adducts, both *in vivo* and *in vitro* (*1,2*). The major adducts suspected of being responsible for the mutational and carcinogenic effects of these agents are the O^6-alkylguanine and the O^4-alkylthymine adduct. Evidence indicating the potency of these adducts is derived from a number of sources. These include the observation that persistence and/or existence of adducts generally correlates with carcinogenesis in the target organs (*3–5*). Furthermore, these adducts have been shown to mispair in *in vitro* systems (*6–8*). In addition, O^6-methylguanine and O^4-methylthymine have been shown to mispair with thymine and guanine, respectively, in semi- *in vivo* site-specific mutational systems (*9–11*). Hence, there would appear to be more than adequate evidence indicating a role for these adducts in alkylation-induced mutagenesis.

Our approach to these questions has concentrated on the analysis of mutational specificity. These studies also indicate major roles in mutagenesis for the O^6-guanine and O^4-thymine adducts. For example, agents producing almost exclusively the O^6 adduct such as ethylmethanesulphonate (EMS), *N*-methyl-*N*-nitro-*N*-nitrosoguanidine (MNNG), and *N*-nitroso-*N*-methylurea (NMU) yield almost exclusively G:C to A:T transitions (*12–14*). In contrast, N-nitroso-ethylurea (ENU), which produces both the O^6 and O^4 lesions, induces both G:C to A:T and A:T to G:C transitions in a ratio similar to that of the adducts (*15,16*). Finally, the analysis of mutation in strains defective in DNA repair are also consistent with these models (*12,15*).

While these mutagenic lesions thus appear capable of accounting for the specificity of the base substitutions encountered following treatment with a wide variety of alkylating agents, there are multiple steps between the exposure to the chemical and the final observation of the mutation. Among the major factors requiring consideration are the deposition of damage and the influence of DNA repair. Others include the nature and quantities of other molecules in the cell that can act as traps and thereby protect the DNA by essentially altering the effective dose. Further considerations need to include the nature of the mutation being selected, as a large fraction of DNA sequence changes are unlikely to lead to selectable phenotypic alterations.

The analysis of mutation at the DNA sequence level is also capable of shedding light on some of these questions as well. For example, a knowledge of mutational specificty at the DNA sequence level can also provide information on the influence of local DNA sequence. We have recently reported a nearest neighbor effect on G:C to A:T transition mutations induced by MNNG (*13*). The nature of this effect was that mutations were found to be much more likely to occur at guanine residues proceeded 5′ by a purine base; thus developed the "PuG rule" for mutation induced by MNNG. In this communication we examine the validity and possible cause of the PuG rule following mutagenesis with other alkylating agents.

*Biology Department, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3.

Materials and Methods

The approach used in this study involves the analysis of mutation in the *lacI* gene of *Escherichia coli*. Mutations in the first 180 base pairs of the gene were genetically identified and then cloned onto a specially constructed M13 vector, mRS81, by *in vivo* recombination. The methodologies in obtaining mutants are identical to those described elsewhere (*13*). Those methods for the cloning and sequencing of mutations are as described by Schaaper et al. (*17*). All the *lacI* mutations reported here were induced in a repair proficient strain of *E. coli* known as NR3835.

Results and Discussion

Mutation Induction by Alkylating Agents

The survival and LacI$^-$ mutation induction data for a number of alkylating agents is given in Table 1. In addition to agents already mentioned, the table includes data on the nitrosamine *N*-nitroso-*N*-methyl-*N*-α-acetoxybenzylamine (NMAB). This is a direct-acting analog of the esophogeal carcinogen *N*-nitroso-*N*-methyl-*N*-benzylamine (NMB). As can be seen in Table 1, the treatments cover a range of doses and some of the agents are clearly more toxic and/or mutagenic than others. From the table, however, it can be concluded that only a very small fraction of mutations sequenced following treatment are likely to be spontaneous in origin.

Preference for Mutation at 5′-PuG-3′Sites

The reported experiments with MNNG (*13*) indicate an influence of the 5′ base on mutation induction. This effect is documented in Table 2, where it can be seen that 125 MNNG-induced mutations occurred at the 8 available 5′-PuG-3′ sites, compared with 39 mutations occurring at the 17 potential 5′-PyG-3′ sites. From these data it appears that mutation is almost seven times more likely to occur at a G residue proceeded by a purine than a pyrimidine. A similar analysis of the data for MNU, ENU, and MNAB also indicate a strong preference for mutation occurring at 5′-PuG-3′ over that seen for 5′-PyG-3′ sites.

In the case of the EMS treatment, the preference for 5′-PuG-3′ sites is much less obvious. 5′-PuG-3′ sites are preferred over 5′-PyG-3′ sites by a factor of only 2.0. Moreover, mutations are most commonly recovered at sites proceeded by a C. This observation, however, is further complicated by the fact that DNA repair influences the distribution of mutation (*12*). As we previously demonstrated, the excision of the O^6-ethylguanine residues by the *E. coli uvrABC* repair pathway did not occur with equal efficiency at all sites; sites flanked by A:T base pairs are more efficiently repaired than sites flanked by G:C base pairs. In the absence of excision repair, i.e., in a UvrB$^-$ strain, the slight bias favoring 5′-PuG-3′ sites disappears (*12*), and we concluded that the PuG rule does not apply to EMS-induced mutation.

The mutational spectra for the above treatments are summarized in Figure 1.

Table 1. Survival and mutagenesis data.

Treatment	Survival, %	Mutation frequency[a]	Fold increase[b]
MNNG, 0.16 mM	88	2400	800
NMAB, 1 mM	57	1900	630
MNU, 10 mM	32	270	90
ENU, 30 mM	4	630	210
EMS, 280 mM	51	80	27

[a] Mutation frequency × 10^6.
[b] Ratio of induced over spontaneous mutation frequency.
Abbreviations: MNNG, *N*-methyl-*N*-nitrosoguanidine; NMAB, *N*-nitroso-*N*-methyl-*N*-α-acetoxybenzylamine; MNU, methylnitrosourea; ENU, *N*-nitroso-*N*-ethylurea; EMS, ethylmethanesulfonate.

Table 2. Influence of 5′ flanking base.

Treatment	Mutations per available site[a] 5′ Flanking base: A	G	C	T	Ratio Pu/Py
MNNG	7.7 (23)	20.4 (102)	2.0 (16)	2.6 (23)	6.8
NMAB	7.0 (21)	12.8 (64)	1.8 (14)	1.3 (12)	7.6
MNU	9.0 (27)	8.6 (43)	1.5 (12)	0.4 (4)	9.8
ENU	3.7 (11)	7.2 (36)	1.0 (8)	1.4 (13)	4.9
EMS	2.0 (6)	4.2 (21)	2.8 (22)	0.8 (7)	2.0

[a] Number of known available sites in this 180 bp target sequence are A, 3; G, 5; C, 8; T, 9. The numbers in parentheses give the actual number of mutations recovered. See footnotes of Table 1 for abbreviations.

Origin of the Nearest Neighbor Effect

The observation that the mutability of a site might be influenced by local DNA sequence was first proposed by Benzer (*18*) as an explanation for the occurrence of mutational hotspots in the bacteriophage T4 rII gene. Since then, a number of examples of sequence effects have been observed (*16,19*). These influences have been shown to reflect reactivity (*20*), secondary structure of DNA (*21,22*), and repairability (*12,23*).

If we assume that the primary source of G:C to A:T transitions induced by these agents is the O^6-alkylguanine adduct, then the 5′ flanking base pair may influence mutation distribution by affecting the initial distribution of these lesions; the efficiency of repair of these lesions; and the fidelity of DNA polymerase at those positions in the sequence.

The 5′ flanking sequence may exert its effect by affecting the distribution of O^6-methylguanine lesions by modifying the reactivity of the O^6 position of guanine. Pullman and Pullman (*24*) have shown that the molecular electrostatic potential, an important component of a site's reactivity, can be strongly influenced by neighboring base pairs. Another important component of reactivity is steric accessibility, clearly a factor also subject to influence by local DNA sequence.

Similarly, any influence of local DNA base sequence on the efficiency of repair of the O^6 adduct will affect mutation distribution. Topal et al. (*23*) demonstrated

```
MNNG               3          0  51                 9    0  11    16
MNAB               3          0  62                 9    0  15    10
MNU                3          0  40                 7    1   9     0
ENU                2          0  40                 4    1   3     4
EMS                4          0  91                 2    1   3     2
      GTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTC
                 :                        :                        :
                40                       65                       90

MNNG     71 0        0               9        0          5     0
MNAB     23 0        0               6        0          6     0
MNU      23 0        0               7        0         13     0
ENU      11 0        1               3        0          4     0
EMS      54 0        1               0        0          4     0
        CCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGC
                 :                        :                        :
               100                      125                      150

MNNG                          14    0    114 1  2      0 22    0
MNAB                          11    0    156 1  3      0 11    0
MNU                            6    0    191 0  1      0  9    1
ENU                            7    0    141 0  4      0  7    4
EMS                            6    0     91 0  3      0  1    0
        GATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTC
                 :                        :                        :
               160                      185                      210
```

FIGURE 1. Distribution of G:C to A:T transitions induced by alkylating agents in the *lacI* gene of *E. coli*. The numbering is as given by Farabaugh (*27*). Available sites where mutants were not recovered in this study are indicated.

that the efficiency of repair of O^6-methylguanine could vary several-fold with respect to the position of the lesion. However, in this case the influence of the 5′ flanking base was modest, less that twofold. We note that differential repairability is unlikely to play a major role in the observed preference, since under the treatment conditions employed here, the repair capacity of the methyltransferase system is most likely fully saturated (*25*). Finally, the distribution of mutations may reflect the influence of neighboring base sequence on the fidelity of DNA synthesis past miscoding lesions.

On the whole we tend to favor the hypothesis that it is the initial DNA damage distribution that is responsible for the observed 5′-PuG-3′ preference. The argument is that this preference is observed under a number of conditions and at different doses (results not all shown). It is seen when the alkyltransferase repair pathway is quite probably saturated and when ethyl adducts, which are very much less efficiently repaired, are involved. We prefer the suggestion that the reactivity of the O^6 position of guanine is substantially enhanced when proceeded by a purine.

Relevance of This Effect to Carcinogenesis

The 5′ flanking base influence seen in these studies has an important counterpart in the results of studies on the activation of Ha-*ras* oncogenes by methylnitrosourea (MNU) (*26*). MNU was found to preferentially induce the activation of the protooncogene by a G:C to A:T transition at the second guanine in codon 12 (GGC), even though a transition at the first guanine can activate this gene. The finding that all the alkylating agents predicted to have a common intermediate, the alkyl diazonium ion, show this preference can be taken as a strong indication that the oncogene alteration was a direct consequence of the mutagenic treatment. Hence, this reactive intermediate may demonstrate the same mutational specificity in bacteria as in mammalian cells.

We are grateful to F. Allen for her technical assistance in carrying out so many cloning and sequencing reactions. This research was supported by the Natural Science and Engineering Council of Canada through grants A2814 and G1598.

REFERENCES

1. Pegg, A. E. Alkylation and subsequent repair of DNA after exposure to dimethylnitrosamine and related carcinogens. In: Reviews in Biochemical Toxicology (E. Hodgson, J. R. Bend, and R. M. Philpot, Eds.), Elsevier, New York, 1983, pp. 83–133.
2. Singer, B., and Kusmierek, J. T. Chemical mutagenesis. Annu. Rev. Biochem. 52: 655–693 (1982).
3. Goth, R., and Rajewsky, M. F. Persistence of O^6-ethylguanine in brain DNA: Correlation with nervous system-specific carcinogenesis by ethyl nitrosourea. Proc. Natl. Acad. Sci. (U.S.) 71: 639–643 (1974).
4. Pegg, A. E. Methylation of the O^6 position of guanine in DNA is the most likely initiating event in carcinogenesis by methylating agents. Cancer Invest. 2: 223–231 (1984).
5. Swenberg, J. A., Dyroff, M. C., Bedell, M. A., Popp, N., Huh, N., Kirstein, U., and Rajewsky, M. F. O^4-ethylthymidine, but not O^6-ethyl-deoxyguanosine, accumulates in DNA of hepatocytes of rats exposed continuously to diethylnitrosamine. Proc. Natl. Acad. Sci. (U.S.) 81: 1692–1695 (1984).
6. Mehta, J. R., and Ludlum, D. B. Synthesis and properties of O^6-methyldeoxyguanylic acid and is copolymers with deoxycytidylic acid. Biochim. Biophys. Acta 521: 770–778 (1978).
7. Abbott, P. J., and Saffhill, R. DNA synthesis with methylated poly(dC-dG) templates. Evidence for a competitive nature to miscoding by O^6-methylguanine. Biochim. Biophys. Acta 562: 51–61 (1979).
8. Singer, B., and Grunberger, D. Molecular Biology of Mutagens and Carcinogens. Plenum Press, New York, 1983, pp. 31–44.

9. Loechler, E. L., Green, C. L., and Essigmann, J. M. *In vivo* mutagenesis by O^6-methyloxyguanine built into a unique site in a viral genome. Proc. Natl. Acad. Sci. (U.S.) 81: 6271–6275 (1984).
10. Bhanot, O. S., and Ray, A. The *in vivo* mutagenic frequency and specificity of O^6-methylguanine in OX174 RF DNA. Proc. Natl. Acad. Sci. (U.S.) 83: 7348–7352 (1986).
11. Preston, B. D., Singer, B., and Loeb, L. A. Mutagenic potential of O^4-methylthymine *in vivo* determined by an enzymatic approach to site-specific mutagenesis. Proc. Natl. Acad. Sci. (U.S.) 83: 8501–8505 (1986).
12. Burns, P. A., Allen, F. L., and Glickman, B. W. DNA sequence analysis of mutagenicity and site specificity of ethyl methanesulfonate in Uvr^+ $UvrB^-$ strains of *Escherichia coli*. Genetics 113: 811–819 (1986).
13. Burns, P. A., Gordon, A. J. E., and Glickman, B. W. Influence of neighbouring base sequence of N-methyl-N′-nitro-itrosoguanidine mutagenesis in the *lacI* gene of *Esherichia coli*. J. Mol. Biol. 194: 385–390 (1987).
14. Burns, P. A., and Glickman, B. W. Mutational specificity of MNU in the *lacI* gene of *Escherichia coli*. Submitted to Carcinogenesis.
15. Burns, P. A., Gordon, A. J. E., Kunzman, K., and Glickman, B. W. Neighbouring base sequence influences both the distribution and repair of ENU-induced lesions. Submitted to Proc. Natl. Acad. Sci. (U.S.).
16. Richardson, K. K., Richardson, F. C., Crosby, R. M., Swenberg, J. A., and Skopek, T. R. DNA base changes and alkylation following *in vivo* exposure of *Escherichia coli* to N-methyl-N-nitrosourea or N-ethyl-N-nitrosourea. Proc. Natl. Acad. Sci. (U.S.) 84: 344–348 (1987).
17. Schaaper, R. M., Danforth, B. N., and Glickman, B. W. Rapid repeated cloning of mutant *lac* repressor genes. Gene 39: 181–189 (1985).
18. Benzer, S. On the topography of the genetic fine structure. Proc. Natl. Acad. Sci. (U.S.) 47: 403–415 (1961).
19. Conkling, M. A., Koch, R. E., and Drake, J. W. Determination of mutation rates in bacteriophage T4 by unneighbourly base pairs. J. Mol. Biol. 143: 303–315 (1980).
20. Mattes, W. B., Hartley, J. A., and Kohn, K. W. DNA sequence selectivity of guanine-N7 alkylation by nitrogen mustards. Nucl. Acids Res. 14: 2971–2987 (1986).
21. Glickman, B. W., and Ripley, L. S. Structural intermediates of deletion mutagenesis: A role for palindromic DNA. Proc. Natl. Acad. Sci. (U.S.) 81: 512–516 (1984).
22. Fix, D. F., and Glickman, B. W. Differential enhancement of spontaneous transition mutations in the *lac*I gene of and Ung^- strain of *Escherichia coli*. Mutat. Res. 175: 41–45 (1986).
23. Topal, M. D., Eadie, J. S., and Conrad, M. O^6-methylguanine mutation and repair is nonuniform. J. Biol. Chem. 261: 9879–9885 (1986).
24. Pullman, A., and Pullman, B. Molecular electrostatic potential of the nucleic acids. Q. Rev. Biophys. 14: 289–380 (1981).
25. Schendel, P. F., and Robins, P. E. Repair of O^6-methylguanine in adapted *Escherichia coli*. Proc. Natl. Acad. Sci. (U.S.) 75: 6017–6020 (1978).
26. Zarbl, H., Sukumar, S., Arthur, A. V., Martin-Zanca, D., and Barbacid, M. Direct mutagenesis of Ha-*ras*-1 oncogenes by N-nitroso-N-methylurea during inititation of mammary carcinogenesis in rats. Nature 315: 382–385 (1985).
27. Farabaugh, P. J. The sequence of the *lacI* gene. Nature 274: 765–769 (1978).