Comparative Mutagenesis of O^6-Methylguanine and O^4-Methylthymine in *Escherichia coli*†

Manjit K. Dosanjh,‡,§ B. Singer,§ and John M. Essigmann*,‡

Department of Chemistry and Division of Toxicology, Whitaker College of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Donner Laboratory, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720

Received January 9, 1991; Revised Manuscript Received April 17, 1991

ABSTRACT: The qualitative and quantitative features of mutagenesis by two DNA adducts of carcinogenic alkylating agents, O^6-methylguanine (m^6G) and O^4-methylthymine (m^4T), were examined in vivo. The deoxyhexanucleotides 5′-GCTAGC-3′ and 5′-GCTAGC-3′ were synthesized, where the underlined bases are the positions of m^4T or m^6G, respectively. By use of recombinant DNA techniques, the respective hexanucleotides or an unmodified control were inserted into a six-base gap in the otherwise duplex genome of the *Escherichia coli* virus M13mp19-*Nhe*I. The duplex adducted genome was converted to single-stranded form and introduced into an *E. coli* strain that was phenotypically normal with regard to m^6G/m^4T repair, a strain deficient in repair by virtue of an insertion in the gene encoding the Ada-m^6G/m^4T DNA methyltransferase, or the same two cell lines after challenge with *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine. Treatment with this alkylating agent chemically compromises alkyl-DNA repair functions. The mutation efficiency of m^6G was low or undetectable (0–1.7%) in all cell systems tested, owing, we believe, to rapid repair. In striking contrast, the mutagenicity of m^4T was high (12%) in cells fully competent to repair alkylation damage and was roughly doubled when those cells were pretreated with *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine to suppress repair. Taken together, these data suggest that m^4T is potentially more mutagenic than m^6G and, if formed by a DNA methylating agent, may pose a significant threat to the genetic integrity of an organism.

The mutations induced by chemical or physical agents most commonly arise from the misreplication or misrepair of chemically altered nucleotides within the genome. The specific biochemical rules underlying the mutagenic activities of such agents have been difficult to elucidate because the genome of a cell treated with a typical DNA-damaging agent forms not one, but usually a dozen or more discrete chemical–DNA adducts. Because each adduct has different rates of repair and different inherent base-pairing capabilities, it has been impossible, in all but exceptional cases, to specify which DNA adduct in a damaged genome gives rise to a given mutation. A simplifying approach that enables the relative mutagenicity of different DNA adducts to be gauged is provided by the tools of site-specific mutagenesis (Basu & Essigmann, 1988). This technology enables the placing of a single DNA adduct at a specific site in the genomes of viruses or plasmids, which are subsequently introduced into bacterial or mammalian hosts. Within the host, the adducted sequence is acted upon in a presumably normal way by the endogenous repair and replication systems. Finally, the progeny of the parental adducted sequence are recovered, and the area of the genome in the vicinity of the originally adducted site is analyzed. The type, amount, and genetic requirements for mutagenesis of the DNA adduct are determined, and these data provide a basis for comparison among the various adducts formed by a DNA damaging agent.

This work concerns a comparative examination of two well-studied DNA adducts involved in the mutagenicity of carcinogenic methylating agents. Among the 13 identified DNA lesions formed by these agents (Singer & Grunberger, 1983), only 2, O^6-methylguanine (m^6G)[1] and O^4-methylthymine (m^4T), have been assigned quantitatively and qualitatively important roles in mutagenesis and tumorigenesis (Goth & Rajewsky, 1974; Zarbl et al., 1985; Singer et al., 1981; Swenberg et al., 1984). In separate studies done by our respective laboratories and those of others, site-specific mutagenic analysis of both lesions has revealed that the adducts are mutagenic in vivo, with m^6G inducing G → A transitions (Loechler et al., 1984; Bhanot & Ray, 1986; Hill-Perkins et al., 1986; Ellison et al., 1989) and m^4T giving rise to T → C transitions (Preston et al., 1986; Klein et al., 1990). The literature also indicates that the specific host repair systems for alkyl-DNA adducts (Lindahl et al., 1988) can be strong modulators of the level of mutagenesis observed in vivo. For example, m^6G is repaired rapidly while m^4T has a much longer half-life (Singer, 1986). It is noteworthy that recent studies from our laboratories have revealed that both adducts are highly mutagenic in vitro and, at least with m^6G, its sequence context can significantly influence the kinetics of mispairing at the adduct site (Singer et al., 1989).

The goal of the present study was to compare the in vivo mutagenic properties of m^6G and m^4T under the same experimental conditions. The use of the same site-specific mutagenesis system for both DNA adducts circumvents problems

†This work was supported by Grants CA52127 (J.M.E.) and CA42736 (B.S.) from the National Institutes of Health.

* To whom correspondence should be addressed.

‡ Massachusetts Institute of Technology.

§ University of California, Berkeley.

[1] Abbreviations: DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene; EDTA, ethylenediaminetetraacetic acid; GHD, gapped-heteroduplex DNA; HPLC, high-pressure liquid chromatography; IPTG, isopropyl β-D-thiogalactopyranoside; M13mp19-*Nhe*I, insertion mutants of M13mp19 containing the recognition site *Nhe*I; m^4T, O^4-methylthymine; m^6G, O^6-methylguanine; MNNG, *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine; RF, replicative form; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside.

0006-2960/91/0430-7027$02.50/0

of interpretation that might arise owing to system-to-system technical differences. An additional feature of the experimental design reported here is the placement of the adducts in the same DNA sequence contexts used in our recent work on the changed base pairing of these lesions replicated in vitro (Singer et al., 1989; Dosanjh et al., 1990).

Experimental Procedures

Materials. Chemicals and supports required for oligonucleotide synthesis by the phosphoramidite method were purchased from Aldrich, Applied Biosystems, or Cruachem. Exonuclease III and the initial stock of bacteriophage M13mp19 DNA were from Pharmacia. *N*-Methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG), dithiothreitol, snake venom phosphodiesterase, bovine serum albumin, and Sepharose CL-4B were purchased from Sigma. 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal), isopropyl β-D-thiogalactopyranoside (IPTG), ATP, and calf intestinal phosphatase were supplied by Boehringer Mannheim. Bacteriophage T4 DNA ligase, T4 polynucleotide kinase, bacterial alkaline phosphatase, *Thermus aquaticus* DNA polymerase (*Taq* polymerase), and restriction enzymes were from New England Biolabs. [γ-^{32}P]ATP (6000 Ci/mol) was obtained from New England Nuclear. All other solvents and reagents were used at the highest purity commercially available. The *Escherichia coli* strains GW5100 (*lacZ*), AB1157 *ada*10::Tn*10* (GW5352), and MM294A (*lacZ*$^+$) were gifts from G. Walker (MIT, Cambridge, MA) and K. Backman (Bio-Technica International, Cambridge, MA). The *E. coli* m^{6}G/m^{4}T DNA methyltransferase (19-kDa fragment) was obtained from Applied Genetics. Oligonucleotide primers, 5'-AACAGCTATGACCATG-3' and 5'-TGACCGGCAGCAAAATG-3', for the polymerase chain reaction studies were synthesized as described below; these primers were complementary to bases 6209–6224 and 6290–6306 of the M13mp19 genome, respectively.

Methods for Oligonucleotide Synthesis. Oligonucleotides d(GCTAGC) and d[GCTA(m^{6}G)C] were synthesized on the 1 μmol scale on an Applied Biosystems Model 381A automated synthesizer by the cyanoethyl phosphoramidite method (Gait, 1984; Sinha et al., 1984). The deoxyoligonucleotide was prepared from the protected monomer 2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-*N*^{2}-isobutryl-*O*^{6}-methylguanosine 3'-[(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite], purchased from American Bionetics (Emeryville, CA), as previously reported (Singer et al., 1989). Deprotection in the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) prevented the possible conversion of m^{6}G to 2,6-diaminopurine (Kuzmich et al., 1983). Synthesis of the m^{4}T-containing deoxyhexanucleotide d[GC(m^{4}T)AGC] was carried out as previously reported (Dosanjh et al., 1990). Following treatment with snake venom phosphodiesterase and bacterial alkaline phosphatase (Fowler et al., 1982), the presence of m^{6}G and m^{4}T in the respective deoxyhexanucleotides was assessed by reversed-phase HPLC analysis on a Beckman system equipped with twin 114M pumps, a 421A gradient controller, and a Hewlett Packard 1040A UV diode-array detector. The nucleoside composition of the hexamers was found to be within experimental error of the theoretical values (data not shown). In parallel, the hexamers were 5'-phosphorylated with T4 polynucleotide kinase and [γ-^{32}P]ATP (vide intra). Following electrophoresis, autoradiography was used to evaluate oligonucleotide purity. Neither HPLC nor electrophoresis revealed detectable impurities.

Restriction enzyme digestions and 0.8% agarose gel electrophoresis were performed as described by Lasko et al. (1987). Denaturing 5% polyacrylamide [19:1 acrylamide:*N*,*N'*-methylenebis(acrylamide) ratio] gel electrophoresis was carried out at 350 V. DNA sequencing of single-stranded phage genomes was done by the method of Sanger et al. (1977) on an 8% denaturing polyacrylamide gel.

Construction of Bacteriophage M13 Genomes Containing m^{6}G or m^{4}T at a Unique Site. Construction of a six-base insertion mutant of M13mp19 containing a d(GCTAGC) sequence in the center of the unique *Sma*I site of the polylinker region was described by Basu et al. (1987). This genome, M13mp19-*Nhe*I, contains a unique recognition sequence for *Nhe*I in the center of the original *Sma*I site. This insertion clone produces colorless plaques on *Su*$^-$ strains of *E. coli* owing to an in-frame *amber* codon (underlined above) in the *lacZα* fragment. Base substitution mutations affecting the first or second base of the *amber* codon generate blue plaques in the presence of β-galactosidase indicator dyes.

Replicative-form (RF) DNA from both wild-type M13mp19 and M13mp19-*Nhe*I was used to prepare heteroduplexes containing a six-base gap (gapped heteroduplex, GHD) in the (+) or (−) strand within the *Nhe*I restriction site. The heating method of Green et al. (1984) was used to anneal 25 μg of *Sma*I-linearized M13mp19 DNA to an equal amount of M13mp19-*Nhe*I, which had previously been linearized with *Bgl*II and 5'-dephosphorylated with calf intestinal phosphatase (Maniatis et al., 1982). The hybrid molecules contain a nonligatable nick in the DNA strand opposite the six-base gap.

In separate reactions, 100 ng of d(GCTAGC), d[GCTA(m^{6}G)C], or d[GC(m^{4}T)AGC] was 5'-phosphorylated with [γ-^{32}P]ATP (20 μCi; 6000 Ci/mol) by T4 polynucleotide kinase at 37 °C for 30 min. Unlabeled ATP was added to a final concentration of 1 mM, and the incubation was continued for an additional 15 min. The reactions were terminated by heating at 65 °C for 15 min, and one-tenth of the reaction volume was removed for HPLC determination of the specific activity of the oligonucleotides. The respective hexamers were ligated into the GHD (1 μg) by using T4 DNA ligase as described by Lasko et al. (1987). After 16 h of incubation at 16 °C, any unincorporated hexamer, ATP, and salts was separated from the ligated DNA by chromatography on a Sepharose 4B column (15 × 0.75 cm) preequilibrated with 10 mM Tris-HCl buffer (pH 7.8), 0.1 M NaCl, and 1 mM EDTA. The site specifically modified genomes eluted in the excluded volume whereas unligated hexamers and other low molecular weight material were retained on the column. A portion (~10^3 cpm) of the [^{32}P]-labeled DNA fraction was electrophoresed through a 0.8% agarose gel. After autoradiography, the product was observed to have comigrated with a nicked circular (form II) DNA standard.

The ligation efficiency was calculated by using the specific activity of the labeled hexamers (determined by HPLC) and the amount of radioactivity incorporated into form II DNA. The ligation efficiency was 50–60% for each of the three genomes constructed (m^{6}G, m^{4}T, and control); because of the method of GHD preparation (Green et al., 1984), half of the ligation products had the adduct in the (+) strand (genome positions 6276 and 6274 for m^{6}G and m^{4}T, respectively), and half had the adduct in the (−) strand (genome positions 6273 and 6275 for m^{6}G and m^{4}T, respectively).

Site specifically modified genomes required for bacterial transformation experiments were prepared essentially as above except that the hexamers were phosphorylated with *unlabeled* ATP for 45 min. After ligation, the DNA was precipitated, washed, and resuspended in TE buffer (10 mM Tris-HCl/1 mM EDTA, pH 8). The genomes were incubated with *Sma*I

to cleave any circular M13mp19 DNA that had not hybridized during GHD preparation and then religated during the ligation step. Following phenol extraction, the DNA was precipitated and then dissolved in either TE buffer (100 μL) or water (5 μL), depending on which cell transformation procedure was to be used ($CaCl_2$ or electroporation, respectively; vide infra). Immediately before the transformation step, each solution was heated to 100 °C for 3 min to produce single-stranded genomes and then cooled rapidly on ice to prevent strand renaturation.

Transformation of E. coli with Site Specifically Modified M13 Genomes. M13 genomes were introduced into the respective *E. coli* host cells by either of two methods. By the first method, cells were made competent for DNA uptake by treatment with 30 mM $CaCl_2$ and transformed as described by Loechler et al. (1984). The second method of introducing DNA into the *E. coli* cells was by electroporation. In this procedure, a 100-mL culture of *E. coli* was grown in LB medium (Maniatis et al., 1982) to a density of ~1 × 10^8 cells/mL. The bacteria were harvested by centrifugation at 5000*g* at 4 °C for 10 min, resuspended in 100 mL of water, and recentrifuged at 5000*g* for 30 min. The procedure was repeated, with the sole exception that the cells were resuspended in 40 mL of H_2O prior to centrifugation. The bacterial pellet was finally resuspended in 200 μL of glycerol/H_2O (10%, v/v) and held on ice until required. For each transformation by electroporation, the DNA was dissolved in 5 μL of H_2O, combined with the 40-μL cell suspension, and transferred to the bottom of a cold Bio-Rad gene pulser cuvette. Cells were electroporated in a Bio-Rad gene pulser set at 25 μF and 2.5 kV with the pulse controller set at 200 Ω. Following electroporation, the cuvette was removed from the apparatus, and 1 mL of SOC medium [2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose (Hanahan, 1985)] was added immediately to the cell suspension. A 0.2-mL portion of the electroporated cell suspension was plated to determine transformation efficiency. The remainder of the sample was incubated for 1.5 h at 37 °C to allow for phage replication. Aliquots of the phage-containing supernatant were plated for determination of mutation frequency.

In some experiments, *E. coli* MM294A (*ada*$^+$) and AB1157 (*ada*10::Tn*10*) cells were challenged with MNNG in order to compromise the activity of the m^6G/m^4T DNA methyltransferase repair protein. Empirically, it was found that a dose of 25 μg of MNNG/mL for 2 min under the conditions as described in Loechler et al. (1984) gave a cell survival of approximately 50%. This dose was used in these studies.

Determination of the Mutation Frequency of m^6G. A portion of the progeny phage pool following transformation of *E. coli* MM294A or AB1157 (*ada*10::Tn*10*) with the m^6G-containing vector was plated on GW5100 in the presence of IPTG and the β-galactosidase indicator dye, X-gal (Loechler et al., 1984). A ratio (R_1) was determined of light blue and colorless plaques (wild type, adduct-induced G → A transitions, and small deletions likely to be introduced during genetic engineering manipulations) to dark blue plaques (primarily six-base deletions owing to regeneration of M13mp19 as a side product of the genome construction; see Figure 1). A mixed population of the phage (>10^6) from this first transformation (T_1) was allowed to infect GW5100, and RF DNA was prepared and treated with an excess of *Nhe*I followed by exonuclease III (Loechler et al., 1984). *Nhe*I linearizes duplex DNA that contains the wild-type sequence, whereas genomes carrying mutations within the six-base recognition sequence of this enzyme are refractory. The linear DNA was further digested by exonuclease III, rendering it biologically inactive. The selection procedure outlined above was repeated twice (transformations 2 and 3, T_2 and T_3) to eliminate *Nhe*I-sensitive DNA completely. After T_3, the ratio (R_3) of light blue and colorless to dark blue plaques was determined once again. Finally, 20–30 individual plaques after T_3 were picked, and the DNA in the vicinity of the originally adducted site was sequenced. From the genome containing a m^6G, two types of mutants were detected: the (+) strand of the genome of the first contained a G → A transition at position 6276, whereas the second had a C → T transition at position 6273. The relative proportion of these two mutations was approximately 5:1. We presume that the former mutation arose from the adduct in the (+) strand of the adducted genome while the latter came from the adduct in the (−) strand. Hence, in both cases, the m^6G would have induced a G·C → A·T transition, consistent with earlier observations (Loechler et al., 1984). The ratio, R_s, of G → A transitions to total plaques[2] picked can then be used to determine the G → A mutation frequency of m^6G as

$$MFm^6G = [(R_3/R_1)R_s]10^2$$

Determination of the Mutation Frequency of m^4T. The mutation frequency from m^4T was easier to determine because all targeted base substitutions reverted the *amber* mutation and thus generated an easily detectable dark blue plaque color phenotype (Figure 1). As with m^6G, the mutations appeared at two genome sites (positions 6274 and 6275), presumably reflecting the original positions of the m^4T residue in the (+) and (−) strands, respectively. Specifically, RF DNA prepared from progeny phage after T_1 was digested with *Sma*I to eliminate M13mp19 produced as a genetic engineering byproduct. After a total of three rounds of serial *Sma*I/exonuclease III selection, the ratio of dark blue to total plaques

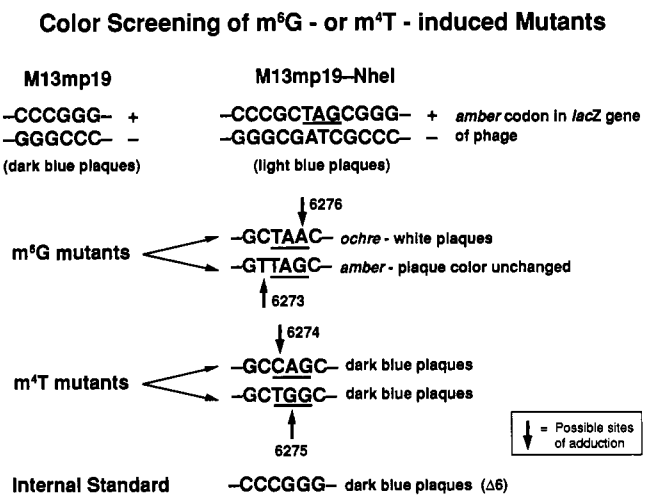

FIGURE 1: Color selection scheme for determining mutagenic efficiency. The strategy is illustrated for the detection of each of the expected mutants from m^6G and m^4T, as well as progeny derived from wild-type and contaminating M13mp19 genomes. The numbered arrows indicate the possible genome positions at which the alkylated bases were situated.

[2] The only targeted base substitution mutations observed for m^6G were G → A transitions. The only other mutations observed were small deletions which, in previous work, were found to be unrelated to m^6G mutagenesis (Loechler et al., 1984). They are likely to be the consequence of the accompanying exonuclease activity of enzymatic manipulations that were necessary to generate the site specifically modified genome.

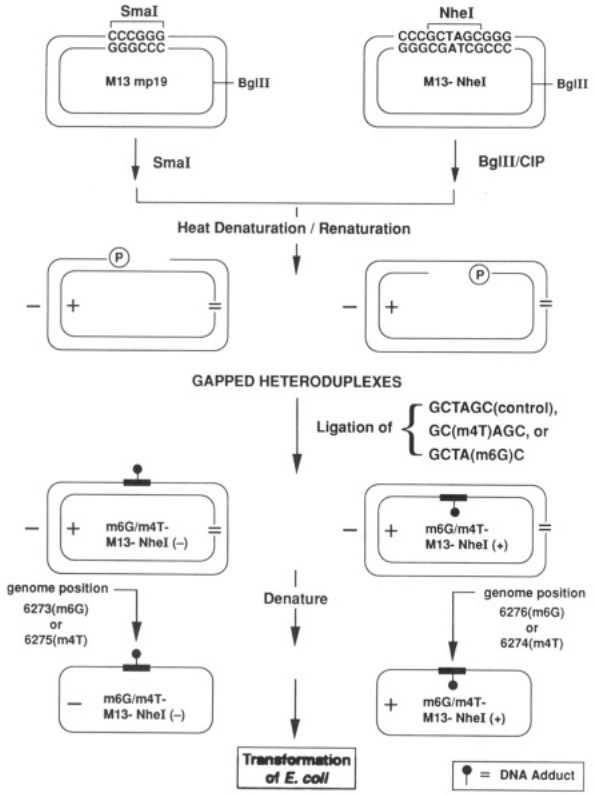

FIGURE 2: Scheme for preparation of a site specifically modified M13mp19-derived genome containing a single m^6G or m^4T at a defined site. The hexamers d(GCTAGC), d[GCTA(m^6G)C], or d[GC(m^4T)AGC] were ligated into the six-base gap in the *Nhe*I restriction site of the GHD. Following denaturation to remove the unadducted strand, the resulting single-stranded genomes were replicated in *E. coli*.

was determined as R_3, and the mutation frequency of m^4T was calculated:

$$MFm^4T = R_3R_{s'}10^2$$

where $R_{s'}$ was the fraction of all dark blue plaques after T_3 that were T → C transitions, as determined by DNA sequencing.[3]

Selected data points on both m^6G and m^4T mutagenesis were confirmed by an independent route. This approach involved preparing single- and double-stranded DNAs from approximately 100 randomly picked plaques from the progeny of T_1. Light blue plaques were selected from transfections with m^6G DNA, whereas dark blue plaques were analyzed from m^4T transfections. The double-stranded DNA was restricted with *Nhe*I in the case of m^6G and with *Sma*I in the case of m^4T. If the DNA were found to be resistant to cleavage, then the corresponding single-stranded genomes were sequenced.

Results

Construction of an M13 Genome with a Single m^6G or m^4T at a Preselected Site. Construction of these vectors followed an approach similar to previously published procedures [Figure 2 (Green et al., 1984; Basu et al., 1987)]. Each hexanucleotide (i.e., those containing m^6G, m^4T, or no modification) was inserted into a six-base gap in the otherwise double-stranded

[3] All sequenced dark blue plaques after T_3 were found to be T → C transitions ($R_{s'} = 1$). By comparison with the sequencing results on the m^6G mutants, it appears that all mutants introduced by the genetic engineering protocol, with the exception of the six-base deletion to M13mp19, give rise to light blue or colorless plaques.

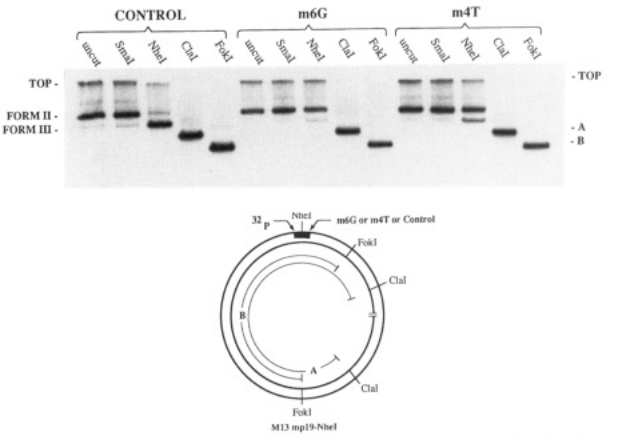

FIGURE 3: Characterization of M13mp19-*Nhe*I RF molecules into which the [γ-^{32}P]ATP-phosphorylated hexanucleotides d(GCTAGC), d[GCTA(m^6G)C], and d[GC(m^4T)AGC] were introduced. At the bottom is the physical map of a GHD ligation product, showing the *Nhe*I site into which the ^{32}P-labeled hexanucleotides were ligated. Each ligation product was subjected to electrophoresis through a 0.8% agarose gel, which subsequently was autoradiographed ("uncut" lanes, top). Also shown are the same DNAs following treatment with *Sma*I, *Nhe*I, *Cla*I, and *Fok*I. A and B denote the positions expected for radioactive *Cla*I and *Fok*I fragments, respectively. Forms II and III are relaxed circular and linear genomes, respectively.

genome of M13. The ligation efficiencies of the unmodified and modified bases were ~50–60%, values similar to those obtained in earlier related studies (Green et al., 1984).

Characterization of Site Specifically Modified Genomes. The ^{32}P label at the 5′ end of the ligated hexanucleotide was used as a marker to help establish the structure of each site specifically modified genome. The position of the ^{32}P label was physically mapped by digestion with the restriction endonucleases *Sma*I, *Nhe*I, *Cla*I, and *Fok*I. Following each digestion, samples were electrophoresed through an agarose gel. It was observed that the radiolabel from the inserted hexanucleotide always comigrated with the expected size fragment containing the *Nhe*I site (Figure 3). All three vectors were resistant to *Sma*I, as expected, since this site was disrupted during modified genome construction. The control genome, which contained only unmodified bases, was almost completely linearized by *Nhe*I, whereas the m^6G- or m^4T-containing genomes were largely resistant, presumably because the adducts within the restriction site inhibited the enzyme. Similar observations have been made with m^6G and other DNA adducts (Green et al., 1984; Lasko et al., 1987). We note, however, that a larger fraction of both site specifically modified genomes used in this study was sensitive to *Nhe*I digestion, as compared to those earlier studies in which the adducts were situated in a *Pst*I site. On the basis of other experience with DNA adducts in the *Nhe*I site (Basu et al., 1987; Basu and Essigmann, unpublished results), we suspect that the *Nhe*I enzyme is less sensitive to structural distortions than *Pst*I.

These experiments verified that d(GCTAGC), d[GCTA(m^6G)C], and d[GC(m^4T)AGC] can be ligated into a six-base gap within M13mp19-derived GHD. The extent of ligation at both 5′ and 3′ ends was determined by digesting the respective genomes with *Pvu*II to produce a 328-base fragment encompassing the hexamer [see Basu et al. (1987)]. This fragment was then electrophoresed through a 5% denaturing polyacrylamide gel, which was subjected to autoradiography. The presence of a single band corresponding to the 328-base single-stranded fragment would indicate that ligation occurred at both ends of the hexanucleotide; by contrast, fragments of 224 and 109 bases would show that ligation occurred at only

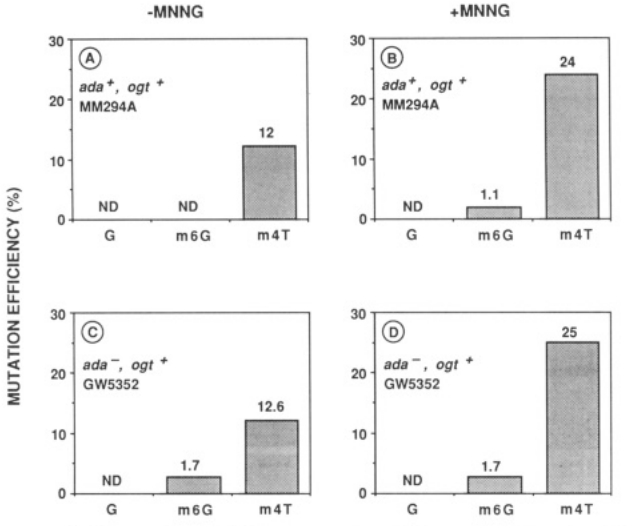

FIGURE 4: Comparison of the mutagenic efficiency of m^6G and m^4T in normal and repair-compromised *E. coli*.

the 3′ or 5′ end. For the genomes containing the unmodified oligonucleotide and both adduct-containing species, densitometric studies showed that 70–80% of each oligonucleotide was ligated at both ends (data not shown).

The presence of m^4T and m^6G within the M13-*Nhe*I genomes was assessed as in Ellison et al. (1989). The portion of each modified (and corresponding unmodified) genome containing the synthetic hexamers was subjected to the polymerase chain reaction. The samples were divided into two equal portions, and half was treated with purified *E. coli* m^6G/m^4T DNA methyltransferase. Both samples were amplified by using *Taq* polymerase in a polymerase chain reaction (25 cycles) from deoxyoligonucleotide primers annealed to sequences flanking the adduct site. The resultant amplified DNA, which was a fragment of 103 nucleotides, was then subjected to *Nhe*I digestion. It was found that the samples pretreated with the methyltransferase were completely cleaved by *Nhe*I, whereas the untreated m^6G and m^4T samples exhibited ~50% resistance, verifying the presence of the mutagenic alkyl group in one strand of the *Nhe*I site (data not shown; but as in Ellison et al. (1989)]. With or without methyltransferase treatment, the control genome, which contained only unmodified bases at the *Nhe*I site, was completely cleaved by *Nhe*I.

Comparative Mutagenicity of m^6G and m^4T. It was not a major goal of this work to establish the types of mutations induced by m^6G and m^4T, because the qualitative features of mutagenesis were established in earlier studies by us (Loechler et al., 1984; Preston et al., 1986) and others (Bhanot & Ray, 1986; Hill-Perkins et al., 1986). We note, however, that the DNA sequencing done in order to establish the mutation frequencies of these lesions (see Experimental Procedures) confirmed that m^6G and m^4T induce, nearly exclusively, G → A and T → C transitions, respectively. Some relatively infrequent nontargeted mutations were observed, primarily small deletions, but these occurred also in the controls at comparable frequencies.

Figure 4 compares the mutation frequencies of m^6G and m^4T in normal and repair-compromised *E. coli*. In each experiment, the mutation frequency of m^4T was severalfold higher than that of m^6G. When repair-proficient cells were used (panel A), no mutants could be detected with m^6G. By contrast, the m^4T-containing genome gave a mutation frequency of 12%. When the cells were challenged with MNNG (panel B) to reduce endogenous repair, the mutation frequency of m^6G was found to be 1.1% compared to 24% for m^4T. These results show that m^4T is significantly more mutagenic than m^6G in repair-proficient *E. coli*. Because both adducts show approximately equal mutagenicity in vitro, we suggest that the repair of m^6G is more efficient than that of m^4T in *E. coli*.

We next examined whether cells that lack the Ada protein (*E. coli* GW5352) showed increased mutagenesis by either of the two methylated bases. The data in Figure 4, panels C and D, show a slight increase in the mutagenicity of m^6G relative to that observed in the *ada*$^+$ cell line (panel A); we note, however, that GW5352 and MM294A are not isogenic, and hence this conclusion must be viewed with caution. With the same caveat, the data of panels A and C indicate that Ada plays a relatively minor role in m^4T repair under the conditions of this experiment. The mutagenicity of m^4T was increased substantially when the hosts for replication of the m^4T-containing genome were challenged with MNNG (panel D), but the magnitude of the increase was the same in *ada*$^+$ and *ada*$^-$ strains. We were initially surprised that the mutation frequency of m^6G following MNNG challenge was somewhat lower than shown in our previous study (Loechler et al., 1984). One plausible explanation for this difference is the method of transformation used to introduce the adducted genomes into *E. coli*. The present study employed electroporation, whereas in the study of Loechler et al. (1984) we used the $CaCl_2$ transformation technique. The latter method includes a 90-s "heat-shock" at 42 °C, and hence the cells experience a longer period of exposure to MNNG. In support of this explanation, when we repeated the m^6G experiment using the same vectors as used here and the $CaCl_2$ technique, the mutation frequency of m^6G increased to 4–4.2% (data not shown), which is within a factor of 2 of the value in our previous work. As a second, nonmutually exclusive explanation for the differences with the earlier study, we note that a large sequence context effect on mutagenesis has been observed for m^6G mutagenesis in vitro (Singer et al., 1989). Furthermore, DNA sequence context also appears to exert a modulating influence on m^6G repair [reviewed in Basu and Essigmann (1990)].

Our experiments were not designed to assess the possible lethal effects of m^6G and m^4T in a precisely quantitative way. Upon electroporation into each of the *E. coli* strains used, however, we noted a small, but consistent, reduction in the number of viable progeny produced. m^6G reduced survival by 50–70% and m^4T by 30–70%, as compared to the unmodified control. Typical transformation efficiencies were approximately 10^3–10^4 infective centers per nanogram of single-stranded input DNA.

Discussion

Previous site-specific mutagenesis studies from our laboratories and others have provided conclusive evidence that both m^6G and m^4T induce transition mutations (Loechler et al., 1984; Preston et al., 1986; Ellison et al., 1989); Chambers et al., 1985; Hill-Perkins et al., 1986; Rossi et al., 1989; Mitra et al., 1989). No comparative quantitative data are available, however, on the relative mutagenicities of the two lesions. Recently, such a comparison was made in vitro in the same DNA sequence context as was used in the present in vivo study (Singer et al., 1989; Dosanjh et al., 1990). The in vitro kinetics of replication by DNA polymerase in these studies demonstrated that m^4T and m^6G preferentially formed m^4T·G and m^6G·T pairs, respectively, and that the frequency of misincorporation was dependent on the DNA sequence context. These data suggested that the frequency of formation of m^6G·T and m^4T·G pairs is high and roughly equal in extent.

The present study of mutagenesis in vivo portrays a contrasting picture in that m^{4}T is strikingly more mutagenic than m^{6}G. The likely reason for this difference stems from the differential action of DNA repair enzymes on the two alkylated genomes prior to their replication in vivo. It appears from our data that a single m^{6}G was very efficiently repaired, whereas the removal of m^{4}T, even as a solitary modification, was so slow that a high level of mutation was induced even in cells with fully active repair functions (Figure 4). As expected, when the repair functions of these cells were compromised by pretreatment with MNNG, the mutation frequency of both m^{6}G and m^{4}T increased. Further evidence that the strikingly high mutation efficiency of m^{4}T in vivo is likely due to its evasion of repair comes from an examination of the known specificities and kinetic parameters of the *E. coli* enzymes that dealkylate m^{4}T and m^{6}G, as determined in vitro (Dolan et al., 1988; Graves et al., 1989; Wilkinson et al., 1989). The following argument assumes that the conclusions based on these parameters can be extrapolated from in vitro extracts to the in vivo situation. In a cell with fully active repair functions, e.g., Figure 4, panel A, the Ada and Ogt repair proteins are present at constitutive levels of approximately 20 (Mitra et al., 1989) and 1 (Rebeck et al., 1989) to 20 (Mitra et al., 1989) molecules per cell, respectively. The kinetic data show that the Ada protein repairs m^{6}G 10 000 times faster than m^{4}T. It has been reported that the Ogt protein repairs m^{4}T roughly 84 times faster than the Ada protein, but m^{6}G is still repaired faster than m^{4}T. Recently, a slightly contrasting picture was brought to light by Sassanfar et al. (1990), who, by using the same sequence used here, showed that Ogt repaired m^{4}T approximately 2–3-fold better than m^{6}G. This study confirmed, however, that Ada repaired m^{6}G much faster than m^{4}T (Graves et al., 1989; Wilkinson et al., 1989). These data, taken together, indicate that if a cell were to receive a small challenge with either alkylated base, a reasonable assumption in site-specific mutagenesis studies, the repair differential would be greatly biased favoring the repair of m^{6}G. Hence, in cells with fully functional Ogt and Ada repair systems (panel A), it is not surprising that the alkylated thymine shows such a high relative mutation efficiency. It is also expected that a genetic defect in *ada* would result in enhanced m^{6}G mutagenicity, whereas the effect on m^{4}T mutagenesis would be comparatively small. Comparison of panels A and C suggests that this is indeed true, although this result must be interpreted with caution since the strains being compared were not isogenic. Of interest would be the comparison of mutagenesis by the two adducts in isogenic wild-type, *ada*$^-$, *ogt*$^-$, and *ada*$^-$/*ogt*$^-$ strains; such strains have recently become available (Rebeck et al., 1989), and these studies are planned.

The data of Figure 4 also suggest, albeit not definitively, that the Ogt protein repairs m^{4}T in vivo. The *ada* strain used in panel C has an insertion in the *ada* gene (Lemotte & Walker, 1985), and thus the only known alkyltransferase repair function is supplied by Ogt. The data of panel D were obtained by using the same strain after challenge with MNNG, and hence any increase in the mutagenicity of either m^{6}G or m^{4}T would be due, in principle, to compromised Ogt activity. Such an effect was observed for m^{4}T mutagenicity. One caveat, however, is that repair systems *other* than those encoded by the *ada* and *ogt* genes may act on m^{4}T. It is known, for example, that the Uvr system of *E. coli* can repair m^{6}G in vitro (Voigt et al., 1989; Schendel et al., 1978; Samson et al., 1988), and it is possible that this or another system might act on m^{4}T. Arguing against this possible modulating role of the Uvr system on m^{4}T mutagenesis is the lack of an effect of MNNG challenge on m^{6}G mutagenesis (panel D).

These studies illustrate the use of site-specific mutagenesis for gauging the relative contributions of two different DNA adducts to the overall mutagenic activity of a DNA damaging agent. Our earlier work showed that both m^{4}T and m^{6}G can be highly mutagenic in vitro. The present study, by contrast, demonstrates that the latter lesion is considerably less mutagenic than the former in vivo. In vivo, the higher mutation frequency of m^{4}T is probably owed to its low repair efficiency (Singer, 1986). We note, however, that m^{4}T typically forms to less than $^1/_{10}$th the level of m^{6}G and hence that the alkylated thymine may pose a lesser overall threat to the genetic integrity of an organism. These studies underscore the quantitatively important role of cellular repair systems in defense against these specific forms of DNA damage.

ACKNOWLEDGMENTS

We are grateful to Ashis Basu and Edward Loechler for their many helpful suggestions during the course of this work. We also thank Laura Niedernhofer for valued technical assistance.

REFERENCES

Basu, A. K., & Essigmann, J. M. (1988) *Chem. Res. Toxicol. 1*, 1–18.

Basu, A. K., & Essigmann, J. M. (1990) *Mutat. Res. 233*, 189–201.

Basu, A. K., Niedernhofer, L. J., & Essigmann, J. M. (1987) *Biochemistry 26*, 5626–5635.

Bhanot, O. S., & Ray, A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7348–7352.

Chambers, R. W., Sledziewska-Gojska, E., Hirani-Hojati, S., & Borowy-Borowski, H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7173–7177.

Dolan, M. E., Oplinger, M., & Pegg, A. E. (1988) *Mutat. Res. 193*, 131–137.

Dosanjh, M. K., Essigmann, J. M., Goodman, M. F., & Singer, B. (1990) *Biochemistry 29*, 4698–4703.

Ellison, K. S., Dogliotti, E., Connors, T. D., Basu, A. K., & Essigmann, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8620–8624.

Fowler, K. W., Buchi, G., & Essigmann, J. M. (1982) *J. Am. Chem. Soc. 104*, 1050–1054.

Gait, M. J. (1984) in *Oligonucleotide Synthesis: A. Practical Approach*, IRL Press, Washington, D.C.

Goth, R., & Rajewsky, M. F. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 639–643.

Graves, R. J., Li, B. F. L., & Swann, P. F. (1989) *Carcinogenesis 10*, 661–666.

Green C. L., Loechler, E. L., Fowler, K. W., & Essigmann, J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 13–17.

Hanahan, D. (1985) in *DNA Cloning: A Practical Approach* (Glover, D. M., Ed.) Vol. 1, pp 109–135, IRL Press, Oxford.

Hill-Perkins, M., Jones, M. D., & Karran, P. (1986) *Mutat. Res. 162*, 153–163.

Klein, J. C., Bleeker, M. J., Lutgerink, J. T., van Dijk, W. J., Brugghe, H. F, van der Elst, H., van der Marel, G. A., van Boom, J. H., Westra, J. G., Berns, A. J. M., & Kriek, E. (1990) *Nucleic Acids Res. 18*, 4131–4138.

Kuzmich, S., Marky, L. A., & Jones, R. A. (1983) *Nucleic Acids Res. 11*, 3393–3407.

Lasko, D. D., Basu, A. K., Kadlubar, F. F., Evans, F. E., Lay, J. O., Jr., & Essigmann, J. M. (1987) *Biochemistry 26*, 3072–3081.

Lemotte, P. K., & Walker, G. C. (1985) *J. Bacteriol. 161*, 888–895.

Lindahl, T., Sedgwick, B., Sekiguchi, M., & Nakabeppu (1988) *Annu. Rev. Biochem. 57*, 133-157.
Loechler, E. L., Green, C. L., & Essigmann, J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6271-6275.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Mitra, G., Pauly, G. T., Kumar, R., Pei, G. K., Hughes, S. H., Moschel, R. C., & Barbacid, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8650-8654.
Mitra, S., Pal, B. C., & Foote, R. S. (1982) *J. Bacteriol. 152*, 534-537.
Preston, B. D., Singer, B., & Loeb, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8501-8505.
Rebeck, G. W., Smith, C. M., Goad, D. L., & Samson, L. (1989) *J. Bacteriol. 171*, 4563-4568.
Rossi, S. C., Conrad, M., Voigt, J. M., & Topal, M. D. (1989) *Carcinogenesis 10*, 373-377.
Samson, L., Thomale, J., & Rajewsky, M. F. (1988) *EMBO J. 7*, 2261-2267.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Sassanfar, M., Dosanjh, M. K., Essigmann, J. M., & Samson, L. (1991) *J. Biol. Chem. 266*, 2767-2771.
Schendel, P. F., Defais, M., Jeggo, P., Samson, L., & Cairns, J. (1978) *J. Bacteriol. 135*, 466-475.
Singer, B. (1986) *Cancer Res. 46*, 4879-4886.
Singer, B., & Grunberger, D. (1983) *Molecular Biology of Carcinogens and Mutagens*, 347 pp, Plenum Press, New York.
Singer, B., Spengler, S., & Bodel, W. J. (1981) *Carcinogenesis 2*, 1069-1073.
Singer, B., Chavez, F., Goodman, M. F., Essigmann, J. M., & Dosanjh, M. K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8271-8274.
Sinha, N. D., Biernat, J., McManus, J., & Koster, H. (1984) *Nucleic Acids Res. 12*, 4539-4551.
Swenberg, J. A., Dyroff, M. C., Bedell, M. A., Popp, J. A., Huh, N., Kirstein, U., & Rajewsky, M. F. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1692-1695.
Voigt, J. M., Van Houten, B., Sancar, A., & Topal, M. D. (1989) *J. Biol. Chem. 264*, 5172-5176.
Wilkinson, M. C., Potter, P. M., Cawkwell, L., Georgiadis, P., Patel, D., Swann, P. F., & Margison, G. P. (1989) *Nucleic Acids Res. 17*, 8475-8484.
Zarbl, H., Sukumar, S., Arthur, A. V., Martin-Zanca, D., & Barbacid, M. (1985) *Nature (London) 315*, 382-385.

Proton Nuclear Magnetic Resonance as a Probe of Differences in Structure between the C102T and F82S,C102T Variants of Iso-1-cytochrome *c* from the Yeast *Saccharomyces cerevisiae*†

Yuan Gao,‡ Jonathan Boyd,§ Gary J. Pielak,*,∥ and Robert J. P. Williams‡

Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, U.K., Department of Biochemistry, University of Oxford, Oxford OX1 3QU, U.K., and Department of Chemistry and Program in Molecular Biology and Biotechnology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3290

Received February 20, 1991; Revised Manuscript Received April 23, 1991

ABSTRACT: Differences in chemical shifts and in nuclear Overhauser effects between the C102T and F82S,C102T variants of *Saccharomyces cerevisiae* iso-1-cytochrome *c* in both the reduced and oxidized forms are reported and analyzed. There is evidence for small conformational differences in both oxidation states of the double variant near position 82. Differences in structure are more evident in the oxidized forms of the variants. These differences extend to distant parts of the protein. It is concluded that the oxidized double variant has undergone a small rearrangement of several regions of the protein that are linked by a hydrogen-bond network. It is shown that the rearrangement involves hydrogen bonds associated with the two heme propionates and associated water molecules. The deductions from nuclear magnetic resonance data are compared with the differences in the crystal structures of the reduced forms of wild-type protein and the F82S variant [Louie, G. V., Pielak, G. J., Smith, M., & Brayer, G. D. (1988) *Biochemistry 27*, 7870-7876].

Studies of unnatural variants of cytochromes *c* have allowed existing ideas about the relationships between the sequence, structure, and function of this important class of electron transfer proteins to be tested and new ideas about the relationships to be proposed [for a review see Mauk (1991)]. Data from functional studies of variants can only be fully interpreted, however, with prior knowledge of the way changes in primary sequence affect tertiary structure and dynamics. One of the most well-studied variants of cytochrome *c* from both functional and structural standpoints is the F82S variant of *Saccharomyces cerevisiae* iso-1-cytochrome *c*. It has been found that this substitution reduces the reduction potential (Pielak et al., 1985; Rafferty et al., 1990) and stability of the protein (Pielak et al., 1987; Y. Gao, G. J. Pielak, J. Boyd, and R. J. P. Williams, manuscript in preparation) and alters the electron transfer properties [Liang et al., 1987, 1988; Rafferty

† This work was supported by grants from the National Institutes of Health (G.J.P.), the Medical Research Council, U.K., and the Royal Society (R.J.P.W.) and by the donors of the Petroleum Research Fund, administered by the American Chemical Society (G.J.P.).
* Corresponding author.
‡ Inorganic Chemistry Laboratory, University of Oxford.
§ Department of Biochemistry, University of Oxford.
∥ University of North Carolina at Chapel Hill.

0006-2960/91/0430-7033$02.50/0