Construction and characterization of extrachromosomal probes for mutagenesis by carcinogens: Site-specific incorporation of *O*⁶-methylguanine into viral and plasmid genomes

(chemical carcinogenesis/mutagenesis/*O*⁶-alkylguanine/carcinogen-DNA adducts)

CALVERT L. GREEN*, EDWARD L. LOECHLER*, KERRY W. FOWLER^{†‡}, AND JOHN M. ESSIGMANN*§

*Laboratory of Toxicology, Department of Nutrition and Food Science, and [†]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by George H. Büchi, August 22, 1983

ABSTRACT Organic synthesis and recombinant DNA technology were used to situate a putatively premutagenic DNA lesion, *O*⁶-methylguanine (*O*⁶MeGua), at a specific location in the genomes of two bacterial viruses, M13mp8 and ϕ X174, and of the bacterial plasmid pBR322. In each genome the first guanine residue in the unique recognition sequence for restriction endonuclease *Pst* I (5'-C-T-G-C-A-G-3') was replaced with *O*⁶MeGua. This was accomplished by ligating a chemically synthesized tetranucleotide, 5'-pTpm⁶GpCpA-3', into a circular, genome-length heteroduplex in which the four internal nucleotides of the *Pst* I recognition site had been removed from one strand of the DNA double helix (ligation yield, ~50%). It was established that the tetranucleotide was located specifically at the *Pst* I site and that the presence of *O*⁶MeGua rendered the ligation product resistant to cleavage by *Pst* I. Sensitivity of the genome to *Pst* I was restored upon treatment with purified *Escherichia coli* *O*⁶MeGua DNA-methyltransferase, a repair protein that removes the methyl group from DNA-bound *O*⁶MeGua. This result, in combination with other data, showed unambiguously that *O*⁶MeGua was incorporated with high yield into the *Pst* I recognition sequence.

Most chemical carcinogens induce mutations, and a quantitative relationship often has been observed between their carcinogenic and mutagenic potencies (1). Taken together with other evidence, this has added support to the view that DNA is the critical cellular target for many carcinogens and that certain of the genetic changes induced by these agents may contribute to the multistep process resulting in malignant transformation. Mechanistically, the first step in the program of events leading to carcinogenesis is believed to involve entry of the carcinogen into the cell, where it is converted by either metabolism or spontaneous reactions into chemically reactive derivatives. These usually are electrophilic and react readily with cellular nucleophiles, including DNA, to form covalent adducts (2). In recent years a detailed picture has emerged as to the structures of the adducts that form in the DNA of cells treated with members of virtually every class of chemical carcinogen (2-6). Some of these adducts are thought to be more biologically significant than others, although in no case has the specific involvement of any individual adduct in the induction of genetic change *in vivo* been unambiguously demonstrated.

The *O*⁶-alkylguanines have been suggested to play a major role in the mutagenic and carcinogenic effects of certain alkylating agents (7-10). It was originally proposed that these adducts were likely to cause replication errors because of their theoretical ability to form base pairs (bp) with thymine as well as cytosine during DNA synthesis (7). To investigate this possibility, others have randomly incorporated *O*⁶-alkyl-

guanines into DNA or RNA polymers, copied them with prokaryotic and eukaryotic polymerases, and analyzed the products of nucleic acid synthesis for the presence of non-complementary nucleotides (10-12). In related work, the accuracy of DNA synthesis has been assayed by using the deoxynucleoside triphosphate derivative of *O*⁶-methylguanine (*O*⁶MeGua) as a synthetic precursor during *in vitro* replication of bacteriophage T7 (13). Both of these approaches have confirmed the original prediction that *O*⁶-alkylguanines can mispair with thymine *in vitro*. *In vivo*, however, few data are available with which to construct a detailed model of *O*⁶-alkylguanine mutagenesis, a model that would include both the quantitative and qualitative features of mutagenesis as well as information on whether the adduct induced targeted or nontargeted mutations. Clearly, studies on *in vivo* processing of adducts would be facilitated by the availability of defined DNA substrates with known carcinogen lesions at specific genome sites.

Recent developments in several areas have provided the experimental tools with which these questions on the mechanism of mutagenesis by carcinogens can be addressed. In the work reported here, we have used a combination of recombinant DNA technology, oligonucleotide synthesis, and carcinogen adduct synthesis to construct viral and plasmid genomes that contain *O*⁶MeGua at a specific site. In the future, these and similarly constructed genomes containing other adducts will be used to probe exactly what happens *in vivo* when DNA polymerases and repair enzymes encounter these carcinogen adducts as they attempt to carry out their roles of replicating or maintaining the structural and functional integrity of the genome.

MATERIALS AND METHODS

Restriction endonucleases were purchased from either New England BioLabs or Bethesda Research Laboratories. Polynucleotide kinase and DNA ligase from bacteriophage T4 were obtained from New England BioLabs. Bacterial alkaline phosphatase and T4 DNA polymerase were obtained from Bethesda Research Laboratories. Calf intestinal alkaline phosphatase was the product of Miles. The *Escherichia coli* *O*⁶MeGua DNA-methyltransferase (*O*⁶MeGua MeTase) was the gift of B. Demple and T. Lindahl (Imperial Cancer Research Fund, London). Electrophoresis of DNA was carried out in horizontal agarose gels prepared and run in Tris borate buffer (14) containing 1 μ g of ethidium bromide per ml.

Abbreviations: bp, base pairs; *O*⁶MeGua², *O*⁶-methylguanine; *O*⁶MeGua MeTase, *O*⁶MeGua DNA-methyltransferase; *O*⁶MeGua-M13, the double-stranded genome of M13mp8 in which the first guanine of the recognition site for *Pst* I has been replaced with *O*⁶MeGua.

[‡]Present address: G. D. Searle and Co., Chicago, IL 60680.

[§]To whom reprint requests should be addressed.

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

Bacteriophage, Plasmid, and Bacterial Strains and DNA Isolation. *E. coli* HF4704 (Su⁻) with ϕ X174 *am3* were obtained from D. Dressler (Harvard University); *E. coli* HB101 and pBR322, from C. Tibbetts (Vanderbilt University); *E. coli* JM103 and M13mp8, from H. N. Munro (Massachusetts Institute of Technology). Viral or plasmid form I DNA was isolated by a modification of the cleared lysate procedure of Godson and Vapnek (15) from *E. coli* strain HF4704 (infected with ϕ X174), strain HB101 (which harbored pBR322), and JM103 (infected with M13mp8); DNAs of ϕ X174 and pBR322 were amplified in the presence of chloramphenicol (30 μ g/ml and 170 μ g/ml, respectively).

Synthesis of a Tetranucleotide Containing O^6 MeGua. The tetranucleotide 5'-HO-Tpm⁶GpCpA_{OH}-3' was prepared as described (16). Phosphorylation of the 5'-hydroxyl group was accomplished by incubating 3 nmol ($\approx$ 4 μ g) of 5'-HO-Tpm⁶GpCpA_{OH}-3' at 37°C for 45 min with 24 units of T4 polynucleotide kinase and 4 nmol of [γ -³²P]ATP (250 μ Ci; 1 Ci = 37 GBq) in 50 mM Tris-HCl, pH 7.8/10 mM MgCl₂/20 mM dithiothreitol/50 μ g of bovine serum albumin per ml (final volume, 80 μ l). The reaction was terminated by heating at 65°C for 20 min, and the phosphorylated tetranucleotide was analyzed by HPLC (ref. 16; the product was eluted 2 min earlier than the precursor).

Construction of Viral and Plasmid Genomes Modified to Contain O^6 MeGua in Their *Pst* I Recognition Sites. The following protocol details the procedure used to build O^6 MeGua into M13mp8 DNA (see Fig. 1); the aspects of the procedure that are specific for pBR322 and ϕ X174 are indicated at the end of the section.

Initially, 200 μ g of replicative form M13mp8 DNA was cleaved with 450 units of *Pst* I. The product (50 μ g/ml) in 50 mM Tris-HCl, pH 8/50 mM NaCl/10 mM MgCl₂ was incubated at 37°C with 100 units of T4 DNA polymerase. After 30

min an additional 100 units of T4 polymerase were added along with dATP, dGTP, and dTTP, each at 120 μ M, and dCTP at 14.4 μ M. Coincubation of DNA with polymerase and DNA synthesis precursors continued at 37°C for 30 min, after which the solution was adjusted to 12 mM disodium EDTA and cooled on ice, and the DNA was precipitated with EtOH.

In a separate reaction, 200 μ g of M13mp8 was linearized with 400 units of *Bgl* II. The product was treated with 25 units of calf intestinal alkaline phosphatase for 30 min at 37°C in 50 mM Tris-HCl, pH 9/1 mM MgCl₂/0.1 mM ZnCl₂/1 mM spermidine. After a second identical treatment with calf intestinal alkaline phosphatase, the product was dephosphorylated with phenol and precipitated with EtOH.

Equal amounts (200 μ g) of *Bgl* II-cleaved, 5'-dephosphorylated linear DNA and *Pst* I-cleaved, blunt-ended linear DNA were mixed in 0.01 M Tris-HCl/1 mM EDTA, pH 8/0.1 M NaCl to give a final DNA concentration of 25 μ g/ml. The mixture was heated at 100°C for 6 min, after which it was placed in an insulated H₂O bath at 80°C and then cooled to 65°C over 40 min [the half-time of strand reassociation was calculated (17) to be 3 min]; after 30 min at 65°C, the DNA was precipitated with EtOH. Electrophoretic analysis of this material revealed that $\approx$ 40% had formed circular structures, presumably the gapped heteroduplexes in Fig. 1.

Incorporation of pTpm⁶GpCpA_{OH} into the gapped heteroduplex M13 genome was accomplished by using DNA ligase. The ligation mixture was prepared in 40 mM Tris-HCl, pH 7.8/8 mM MgCl₂/16 mM dithiothreitol/20 μ g of bovine serum albumin per ml and included, in a total volume of 0.17 ml, 4 μ g (3 nmol) of pTpm⁶GpCpA_{OH} (185 μ Ci), $\approx$ 7.5 μ g (1.5 pmol) of gapped duplex DNA, 1 mM ATP, and 200 units of T4 DNA ligase. The ligation mixture was incubated at 16°C for 6 hr, after which an additional 200 units of ligase were added. After a further 22 hr, the M13 genome was separated from unincorporated tetranucleotide by chromatography on a Bio-Gel P-6DG column (25 $\times$ 1.5 cm; BioRad) preequilibrated with 0.01 M Tris-HCl/1 mM EDTA, pH 8. DNA that was eluted in the void volume was collected and concentrated by evaporation, and a portion was analyzed by electrophoresis through an agarose gel. The ligation yield was determined by relating the amount of ³²P incorporated into high molecular weight DNA to the number of potential ligation sites, as determined by electrophoretic analysis of the ligation mixture and estimation of the amount of DNA present as gapped duplexes (i.e., form II DNA structures). The ligation product will be referred to as O^6 MeGua-M13.

Procedures for construction of pBR322 and ϕ X174 genomes containing O^6 MeGua were identical with a few exceptions. In place of *Bgl* II, *Pvu* II was used to cleave pBR322 prior to the heteroduplexing step; in addition, calf intestinal alkaline phosphatase treatment was omitted. For ϕ X174, *Bgl* II was replaced with *Xho* I, and dephosphorylation of the resultant linear product was accomplished with bacterial alkaline phosphatase. The products obtained from these procedures are defined O^6 MeGua-pBR322 and O^6 MeGua- ϕ X174, respectively.

Characterization of Genomes Containing O^6 MeGua. Each adduct-containing genome was digested with a multi-cutting restriction endonuclease to determine whether the radioactive oligonucleotide had been incorporated into the correct restriction fragment. Approximately 1 fmol (100 cpm) of each modified genome was combined with 0.4 pmol (2 μ g) of the respective unmodified viral or plasmid DNA. The mixtures containing O^6 MeGua-M13, O^6 MeGua-pBR322, and O^6 MeGua- ϕ X174 were cleaved, respectively, with *Pvu* II, *Hinf* I, and *Hae* III. The restriction products were resolved by agarose gel electrophoresis and autoradiographed.

Treatment of O^6 MeGua-M13 with *Pst* I was performed in

FIG. 1. Modification of the unique *Pst* I recognition site of M13mp8 to contain O^6 MeGua. Procedural details are given in *Materials and Methods*.

order to determine the effect of the adduct on the ability of this enzyme to cleave DNA. O^6 MeGua-M13 (1 fmol; ≈ 100 cpm) was treated at 37°C with 10 units of *Pst* I for 60 min. In a separate reaction, O^6 MeGua-M13 (1.5 fmol; ≈ 150 cpm) was treated with 1.5 units of *E. coli* O^6 MeGua MeTase at 37°C for 30 min in 50 mM Hepes, pH 7.8/10 mM dithiothreitol, 1 mM Na_2EDTA /50 μM spermidine/5% glycerol. Subsequently, carrier DNA was added (0.5 μg of M13mp8 form I), the mixture was made 50 mM in NaCl and 10 mM in MgCl_2 , and half of the mixture was treated with *Pst* I (10 units under the conditions described above), while half was not. Both *Pst* I-treated and *Pst* I-untreated samples were analyzed by electrophoresis and autoradiography.

RESULTS

The key feature of the strategy used for situating a carcinogen residue at a specific genome location was to create a small gap of defined size at an exactly known site in duplex DNA. Into the gap was ligated a synthetic oligonucleotide, which had exactly the sequence of nucleotides removed from DNA in forming the gap with one exception: a guanine in the natural genome was replaced with O^6 MeGua, a putatively mutagenic DNA adduct. The specifics of this scheme are presented in Fig. 1.

Placement of O^6 MeGua in the *Pst* I Recognition Site of Viral and Plasmid Genomes. Cleavage of M13mp8 DNA with *Pst* I produced a linear duplex with 3' overhanging termini (5'-T-G-C-AOH-3'). These were removed by using T4 DNA polymerase in the blunt-ending reaction, which was conducted in two steps. Initially, T4 polymerase was used in the absence of deoxynucleoside triphosphates; under these conditions, the 3'-exonuclease activity of the enzyme removed the four nucleotides of the 3' overhang and then continued inward into the duplex genome. It was determined that ≈ 30 nucleotides were removed from the 3' end of each strand of the linear duplex (18). At this point, addition of DNA synthesis precursors and more DNA polymerase resulted in replacement of DNA removed during the exonuclease step, producing ultimately a linear, blunt-ended duplex, which presumably was 7,225 bp in length (i.e., the length of the genome minus four bp). The missing four nucleotides were originally present in the center of the *Pst* I recognition site (5'-C-T-G-C-A-G-3') of the M13mp8 genome. Because ϕ X174 and pBR322 also have a unique recognition site for *Pst* I, the same technique removed the analogous 5'-T-G-C-A-3' se-

quence from their linearized genomes.

A circular DNA structure containing a four-nucleotide gap was created in the heteroduplexing step. The linear product of the blunt-ending reaction was combined with M13mp8 DNA that previously had been linearized with *Bgl* II. Heat denaturation followed by controlled cooling of the mixture resulted in approximately equal amounts of the original linear structures (not shown as products in Fig. 1A) and circular heteroduplexes. In these heteroduplexes, there was an exactly positioned four-nucleotide gap in the strand originally treated with *Pst* I and then blunt-ended. The intact DNA strand opposite the gap had the sequence 3'-ApCpGpTp-5' and, thus, displayed partial complementarity to 5'- ^3H Tpm 6 GpCpA-3', which we had synthesized previously (16).

The 5'-terminal deoxythymidine of the synthetic tetranucleotide was phosphorylated by using T4 polynucleotide kinase and [γ - ^{32}P]ATP (Fig. 1B) prior to the ligation step. As judged by HPLC analysis of the reaction mixture, 97% phosphorylation was achieved. The high specific activity of 5'-[^{32}P]pTpm 6 GpCpA-3' ($\approx 10^9$ dpm/ μg) provided a means for measuring the yield of the ligation step. In the presence of T4 DNA ligase and a 2,000-fold molar excess of 5'-pTpm 6 GpCpA-3', the yield of ligation into gapped duplex DNA was $\approx 50\%$. This high yield indicated that the tetramer, when present in large excess, was sufficiently complementary to the ligation target for DNA ligase to effectively form phosphodiester bonds. Electrophoresis followed by autoradiography of the ligation mixture revealed that incorporation of ^{32}P occurred principally in the circular heteroduplexes (Fig. 2, lane b). However, there was some tetranucleotide incorporation into DNA that remained near the origin after electrophoresis. The exact structure of this material is unknown, although preliminary characterization suggests it may consist of DNA networks or concatenated circular structures.

The heteroduplex DNA structures that received the modified oligonucleotide had gaps that formed with equal probability in either of the two complementary strands (Fig. 1A). Thus, ligation occurred into these gaps in either of two possible orientations, resulting in placement of O^6 MeGua at position 6,255 when it was ligated into the (+) strand and at position 6,256 when it was introduced into the (-) strand (Fig. 1B).

Restriction Analysis of O^6 MeGua Genomes. If site-specific ligation were accomplished successfully, digestion of an

FIG. 2. Characterization of O^6 MeGua-M13 and O^6 MeGua- ϕ X174 by agarose gel electrophoresis. Lanes: a, M13mp8 standard containing circular, supercoiled duplex DNA [replicative form I (RF-I)], linear duplex DNA (form II), and open circular duplex DNA (form III); b, autoradiogram of ^{32}P -labeled O^6 MeGua-M13 (the ligation product from Fig. 1); c, M13mp8 DNA (from lane a) after *Pst* I treatment; d, autoradiogram of ^{32}P -labeled O^6 MeGua-M13 (lane b) after *Pst* I treatment; e, M13mp8 DNA (lane a) after *Pvu* II treatment (A, B, and C denote positions of restriction fragments of *Pvu* II-digested M13mp8); f, autoradiogram of ^{32}P -labeled O^6 MeGua-M13 after *Pvu* II treatment (numbers between lanes f and g are fragment sizes in bp); g, ϕ X174 after treatment with *Hae* III; h, autoradiogram of ^{32}P -labeled O^6 MeGua-M13 after *Hae* III treatment. Letters to the right indicate positions of restriction fragments.

O^6 MeGua-containing genome with a restriction endonuclease with multiple recognition sites would result in localization of radiolabel from the 5'-end of the tetranucleotide within the restriction fragment that originally contained the *Pst* I site. Digestion of M13mp8 with *Pvu* II produced three fragments, which were resolved by electrophoresis (Fig. 2, lane e). The *Pst* I site (the ligation target) is located in the 301-bp B fragment. Radiolabel from the tetranucleotide in O^6 MeGua-M13 was present exclusively within this fragment (Fig. 2, lane f), which represents 4% of the genome. Similar analyses of O^6 MeGua- ϕ X174 (Fig. 4, lanes g and h) and O^6 MeGua-pBR322 (5) cleaved with *Hae* III and *Hinf* I, respectively, also accurately mapped the radioactivity within the restriction fragment that contained the ligation target. Taken together, these data indicate that the synthetic tetranucleotide is associated uniquely with the area of each genome that contains the *Pst* I recognition site.

O^6 MeGua-M13 Was Insensitive to *Pst* I. Many bacteria methylate specific nucleotides in their chromosomes to protect against endogenous restriction endonucleases (19). Although methylation of the O^6 atom of guanine is not known to be involved in bacterial restriction/modification systems, it seemed plausible that the presence of this alkylated base in the *Pst* I recognition site might confer immunity to cleavage by this enzyme. To test this possibility, 32 P-labeled O^6 MeGua-M13 was mixed with nonradioactive carrier M13mp8 and treated with *Pst* I. Carrier DNA was completely cleaved, producing a linear, double-stranded product (Fig. 2, lane c). In marked contrast, O^6 MeGua-M13 was almost completely resistant to cleavage by *Pst* I (Fig. 2, lane d). The inability of *Pst* I to cleave at its six-base recognition site was consistent with the presence within this sequence of a structural alteration, presumably O^6 MeGua.

A careful examination of the autoradiogram (Fig. 2, lane d) reveals that a small amount of linearization of O^6 MeGua-M13 occurred after *Pst* I treatment. This could have resulted either from incorporation of a small amount of *O*-dealkylated tetranucleotide into the gapped duplex or from the slow cleavage of the O^6 MeGua-containing sequence by *Pst* I. If the former possibility is true, it will be necessary to purify the O^6 MeGua-M13 by *Pst* I treatment in order to study O^6 MeGua mutagenesis *in vivo*.

***In Vitro* Repair of O^6 MeGua-M13 Restored Sensitivity to *Pst* I.** Repair of O^6 MeGua *in vivo* is thought to occur principally by the action of the O^6 MeGua MeTase. Lindahl and co-workers (20) have purified this protein from *E. coli*, and several of its properties have been characterized, including the ability to restore guanine at the site of an O^6 MeGua adduct (21). The reaction mechanism involves transferring the methyl-group from the adduct to one of the O^6 MeGua MeTase cysteine residues.

In the previous section it was shown that *Pst* I failed to cleave O^6 MeGua-M13. If this inhibition were solely due to the presence of the adduct, then treatment of O^6 MeGua-M13 with the MeTase should convert O^6 MeGua to guanine in the *Pst* I site and, thereby, restore sensitivity to this restriction endonuclease. This was tested by digesting O^6 MeGua-M13 with *Pst* I before and after incubation with the O^6 MeGua MeTase repair enzyme. As expected, before treatment with the MeTase, most of the DNA was resistant to cleavage by *Pst* I (Fig. 3, lanes 1 and 2), whereas after MeTase treatment (Fig. 3, lanes 4), sensitivity to *Pst* I was acquired by $\approx 95\%$ of the previously insensitive DNA (the possibility that the MeTase possessed a nuclease activity that would provide an alternative explanation of the results was ruled out; Fig. 3, lanes 3). Thus, on the basis of the known stringent requirements of two DNA-acting enzymes, *Pst* I and O^6 MeGua MeTase, it is concluded that the O^6 MeGua-M13 genome has O^6 MeGua specifically in the reconstructed recognition site for *Pst* I.

FIG. 3. Evidence that O^6 MeGua was located at the *Pst* I recognition site of O^6 MeGua-M13. O^6 MeGua-M13 (1.5 fmol, 150 cpm) was either treated with 1.5 units of O^6 MeGua MeTase (+MT; lanes 3 and 4) or not (-MT; lanes 1 and 2). Subsequently, 0.5 μ g of marker M13mp8 was added to each sample, and one-half was treated with *Pst* I (+*Pst* I; lanes 2 and 4), while the other half was not (-*Pst* I; lanes 1 and 3). After agarose gel electrophoresis, the ethidium-stained DNA was photographed (lanes a, c, e, and g). The same material was autoradiographed to visualize radioactive DNA (lanes b, d, f, and h). Indicated to the left are the positions of replicative form I DNA (RF-I), open circular DNA (form II), and genome-length linear duplex DNA (form III).

Both Ends of the Tetranucleotide Ligated into the Gapped Duplex Genome. Although the experiments described above collectively demonstrate that 5'-pTpm⁶GpCpA-3' was located at the *Pst* I site, they do not prove that the tetranucleotide was ligated into the gapped DNA structure at both ends. This was of concern, because failure to ligate at one of the ends of the tetranucleotide would leave a nick near the O^6 MeGua lesion, and this could provide a mechanism for its efficient repair upon entry of the adducted genome into cells. To test whether complete ligation had occurred, 32 P-labeled O^6 MeGua-M13 was digested with *Bgl* II, the complementary strands were separated with alkali, and the length(s) of the single-stranded fragment(s) generated were determined by electrophoresis (Fig. 4). *Bgl* II cleaves M13mp8 once, 661 bp away from the *Pst* I site (see Fig. 1A).

FIG. 4. Evidence that the 5'- 32 P-labeled tetranucleotide was ligated into the gapped duplex at both 5' and 3' ends. O^6 MeGua-M13 (1 fmol, 100 cpm) was cleaved with 10 units of *Bgl* II, which linearized the genome. The complementary strands were denatured for 1 min at 25°C with 0.1 M NaOH in 25 mM Na₂ EDTA, and the length of the single-stranded fragment produced was determined by agarose gel electrophoresis followed by autoradiography (lane 2). Single-stranded (ss) DNA of specific lengths was produced by treating 2 μ g of M13mp8, prior to strand denaturation, with *Bgl* II alone (10 units; lane 1) or with *Bgl* II and *Pst* I (10 units each; lane 3); DNA in these lanes was visualized by ethidium bromide staining. In addition to ss DNAs of 7,229, 6,568, and 661 bases, there was some complementary strand reassociation (electrophoresis was under nondenaturing conditions), resulting in material migrating at the positions of double-stranded (ds) DNA of 6,568 and 661 bp.

If both ends of the tetranucleotide were ligated into the four-base gap, there would be no nick at the *Pst* I site, and the radioactivity associated with this strand would appear in genome-length, single-stranded material. However, incomplete ligation would leave a nick at the *Pst* I site, and the strand containing *O*⁶MeGua (and the radioactivity) would separate into two shorter single-stranded fragments [on average 661 and 6,568 bases in length; an analysis of all of the possibilities—i.e., a nick on the 3' or 5' end of the tetranucleotide in the (+) or (−) strand—reveals that equal amounts of radioactivity would appear in each fragment]. The results of this experiment are shown in Fig. 4. The flanking, ethidium-stained lanes are size markers for single-stranded DNA of 7,229 bases (lane 1), 6,568 (average) bases (lane 3), and 661 (average) bases (lane 3). The center lane in Fig. 4 is an autoradiograph of *O*⁶MeGua-M13, which was cleaved with *Bgl* II and then strand-denatured. This lane shows a single prominent band of genome length (7,229 bases). The fact that we observed only labeled DNA of genome length and no smaller fragments demonstrates conclusively that the ³²P-labeled tetranucleotide was ligated into the genome on both 3' and 5' ends.

Additional data in support of successful ligation of the 5' end of the tetranucleotide came from treatment of the ³²P-labeled ligation product, *O*⁶MeGua-pBR322, with bacterial alkaline phosphatase. It was found that >98% of the radioactivity incorporated by ligation into high molecular weight DNA was resistant to the 5' phosphatase activity of this enzyme (5).

DISCUSSION

Based upon several lines of evidence, this work resulted in the assembly of three intact genomes in which a single natural base has been replaced with a suspected carcinogen adduct. The specific approach taken was to ligate pTpm⁶GpCpA into a gap created by genetic engineering techniques in the genomes of M13mp8, ϕ X174, and pBR322. It was established that both ends of the tetranucleotide ligated with high yield into duplex DNA, and, by restriction analysis of each modified genome, it was shown that incorporation occurred exclusively within the restriction fragment that originally contained the ligation target. Additional experiments more precisely located the adduct within the unique recognition site for restriction endonuclease *Pst* I. It was found that the presence of the adduct within this DNA sequence protected it from endonucleolytic cleavage by *Pst* I. As expected, removal of the carcinogen residue from the adducted genome by *in vitro* repair almost completely restored susceptibility to *Pst* I hydrolysis. Taken together, these data are most consistent with the conclusion that *O*⁶MeGua has been incorporated successfully into the *Pst* I recognition site of these viral and plasmid genomes.

The *O*⁶MeGua-M13 vector described in this work can be used in either double- or single-stranded form to infect *E. coli* because its method of construction situated a nonligatable nick at the *Bgl* II site (Fig. 1B). This nick provides the ability to remove the wild-type sequence opposite the adduct (with alkali, heat, or an exonuclease) and produce a single-stranded, adduct-containing DNA vector for infection of *E. coli*. Such a vector would be less subject to accurate repair by repair pathways that either require double-stranded DNA (e.g., excision repair) or prefer double-stranded DNA, as is the case for repair of *O*⁶MeGua by the *O*⁶MeGua MeTase (22). In recent work to be described elsewhere, the *O*⁶MeGua-M13 vector was made single-stranded and introduced into *E. coli*, where the mutation frequency of *O*⁶MeGua was determined to be $\approx 0.5\%$ in cells with normal DNA repair capability and at least 25% in cells in which the

repair activity of the *O*⁶MeGua MeTase had been compromised (unpublished observations). Mutations were G-to-A transitions located specifically at position 6,255, which was the target position in the present study for introduction of *O*⁶MeGua into the (+) strand of M13mp8. The high efficiency at which mutations were observed at this site adds support to the conclusion that the adduct has been located accurately and with good yield at a specific position within the genome.

The system described here is the prototype for future investigations on the genetic effects of a number of structurally different chemical carcinogens. We anticipate that the ability to produce DNA substrates containing uniquely situated carcinogen adduct will greatly facilitate investigation of the genetic fate of these DNA lesions in both prokaryotic and eukaryotic cells and eventually will provide insights into the chemistry and enzymology of mutagenesis. Analysis of these events is central to the development of accurate concepts on the initiation of chemical carcinogenesis.

For their contributions to this work, we express our gratitude to K. Berkner, M. Chow, L. Couto, B. Demple, J. King, D. Lasko, U. RajBhandary, D. Russell, P. Sharp, and D. Shortle. We are indebted especially to Prof. G. Büchi for his interest and support of this work. Financial support for this research was provided by Grants GM09869 (to G. H. Büchi), 5 PO1 ES 00597, and T 32 ES07020 from the National Institutes of Health and by a fellowship to C.L.G. from the Monsanto Fund.

1. Meselson, M. & Russell, K. (1977) in *Origins of Human Cancer*, Cold Spring Harbor Conferences on Cell Proliferation, eds. Hiatt, H. H., Watson, J. D. & Winsten, J. A. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 4, Book C, pp. 1473–1481.
2. Miller, E. C. (1978) *Cancer Res.* **38**, 1479–1496.
3. Singer, B. & Kusmierek, J. T. (1982) *Annu. Rev. Biochem.* **52**, 655–693.
4. Lawley, P. D. (1979) in *Chemical Carcinogens and DNA*, ed. Grover, P. L. (CRC, Boca Raton, FL), Vol. 1, pp. 1–36.
5. Essigmann, J. M., Green, C. L., Croy, R. G., Fowler, K. W., Büchi, G. H. & Wogan, G. N. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 327–337.
6. Conney, A. H. (1982) *Cancer Res.* **42**, 4875–4917.
7. Loveless, A. (1969) *Nature (London)* **223**, 206–207.
8. Goth, R. & Rajewsky, M. F. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 639–643.
9. Margison, G. P. & Kleihues, P. (1974) *J. Natl. Cancer Inst.* **53**, 1839–1841.
10. Gerchman, L. L. & Ludlum, D. B. (1973) *Biochim. Biophys. Acta* **308**, 310–316.
11. Mehta, J. R. & Ludlum, D. B. (1978) *Biochim. Biophys. Acta* **512**, 770–778.
12. Abbott, P. J. & Saffhill, R. (1979) *Biochim. Biophys. Acta* **562**, 51–61.
13. Dodson, L. A., Foote, R. S., Mitra, S. & Masker, W. E. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7440–7444.
14. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 148–152.
15. Godson, G. N. & Vapnek, D. (1973) *Biochim. Biophys. Acta* **299**, 516–520.
16. Fowler, K. W., Büchi, G. & Essigmann, J. M. (1982) *J. Am. Chem. Soc.* **104**, 1050–1054.
17. Wetmur, J. G. & Davidson, N. (1968) *J. Mol. Biol.* **31**, 349–370.
18. O'Farrell, P. H., Kutter, E. & Nakanishi, M. (1980) *Mol. Gen. Genet.* **179**, 421–435.
19. Modrich, P. (1979) *Q. Rev. Biophys.* **12**, 315–369.
20. Demple, B., Jacobsson, A., Olsson, M., Robins, P. & Lindahl, T. (1982) *J. Biol. Chem.* **257**, 13776–13780.
21. Olsson, M. & Lindahl, T. (1980) *J. Biol. Chem.* **255**, 10569–10571.
22. Lindahl, T., Demple, B. & Robins, P. (1982) *EMBO J.* **1**, 1359–1363.