Context-dependent mutagenesis by DNA lesions

James C Delaney and John M Essigmann

Background: Detailed analyses of mutational hotspots following DNA damage provide an understanding of oncogene activation and tumor suppressor gene inactivation, and hence provide an insight into the earliest steps in the induction of cancer. A mutational hotspot might be created by preferential lesion formation, decreased lesion repair, or increased misinsertion past the lesion during DNA replication. The respective contribution of these factors might be influenced by the DNA sequence context of the hotspot.

Results: As a prelude to addressing the contribution of all possible nearest-neighbor contexts on the replication past O^6-methylguanine (m^6G) and repair of m^6G *in vivo*, we have devised a mutation frequency (MF) detection strategy on the basis of the properties of type IIs restriction enzymes. We also report a method for constructing site-specific single-stranded viral DNA genomes that should yield identical ligation efficiencies regardless of the lesion or its surrounding sequence context. Using repair-deficient *Escherichia coli*, we discovered that m^6G in three sequence contexts was nearly 100% mutagenic *in vivo*, showing that the DNA polymerase holoenzyme almost always placed a thymine base opposite m^6G during replication. In partially repair-proficient cells, the Ada O^6-methylguanine-DNA methyltransferase repair protein was twice as efficient on m^6G when a guanine base rather than an adenine base was 5′ to the lesion.

Conclusions: The system allows the mutagenic potential of, theoretically, any DNA lesion that exhibits point mutations, in any varied local sequence context, to be rapidly determined. The assay demonstrates low background, high throughput, and does not require phenotypic selection, making it possible to discern the effects of sequence context on the processing of m^6G.

Address: Department of Chemistry and Division of Bioengineering and Environmental Health, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Correspondence: John M Essigmann
E-mail: jessig@mit.edu

Key words: DNA context, DNA repair, DNA replication, O^6-methylguanine, mutagenesis

Received: 17 May 1999
Revisions requested: 18 June 1999
Revisions received: 5 July 1999
Accepted: 13 July 1999

Published: 14 September 1999

Chemistry & Biology October 1999, 6:743–753
http://biomednet.com/elecref/1074552100600643

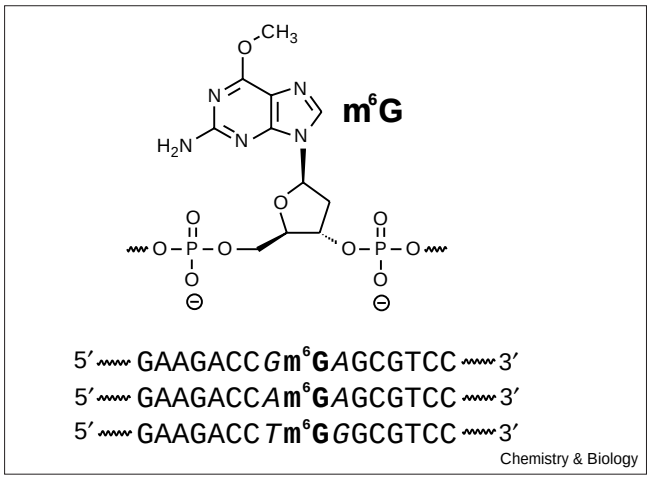
1074-5521/99/$ – see front matter

Introduction

The analysis of the mutational properties of specific types of DNA damage has advanced through the past decade to the point where it is possible to determine the mutational specificity of nearly all types of DNA base damage [1,2]. By chemically synthesizing and placing a modified base at a specific site in a genome, and propagating that genome within cells, one can obtain feedback on how the lesion is processed during its replication and repair. The degree to which a DNA lesion is repaired by a particular system prior to replication can be determined by comparing the mutant to total progeny ratio (the mutation frequency; MF) between cells that possess or lack that repair system. Many methods have been developed for determining the MF of a DNA lesion. For some local sequence contexts, the lesion can be flanked by base combinations whereby the mutation-dependent alteration of a start [3] or stop [4] codon is easily visualized from plated progeny. To analyze the genetic effects of a DNA adduct in any sequence context, one could use differential DNA hybridization [5] or simply sequence individual progeny one at a time [6]. These systems work well but, because they rely on the screening of a subset from the progeny, large numbers of individual clones must be counted or manipulated per transformation to determine the MF for a single sequence context scenario. Thus it would be an arduous task to study all nearest-neighbor contributions to DNA lesion processing within multiple cell lines. Here, we describe and validate a method whereby the easy detection of mutagenesis in all sequence contexts is accomplished. The method is demonstrated using the base-substitution lesion O^6-methylguanine (m^6G).

We chose m^6G (Figure 1) as the inaugural lesion for our assay because the persistence of O^6-alkylguanine DNA adducts correlates well with the mutagenic and carcinogenic effects of alkylating agents [7,8], which can be formed endogenously by nitrosation of amides and amines [9,10]. The mutagenic consequence of unrepaired m^6G is the creation of exclusively G→A transition mutations, as demonstrated in *Escherichia coli* [11]. *E. coli* has proteins that remove m^6G, thus preventing mutations. The inducible Ada and constitutive Ogt methyltransferases have active site cysteine residues that receive the alkyl group directly, converting O^6-alkylguanine to guanine, but the active site is not regenerated [12]. Both the UvrABC nucleotide excision repair [13,14] and mismatch repair [15] pathways have also been implicated in m^6G repair. We used two isogenic

Figure 1



*O*6-Methylguanine (m^6G) in the three nearest-neighbor sequence contexts studied.

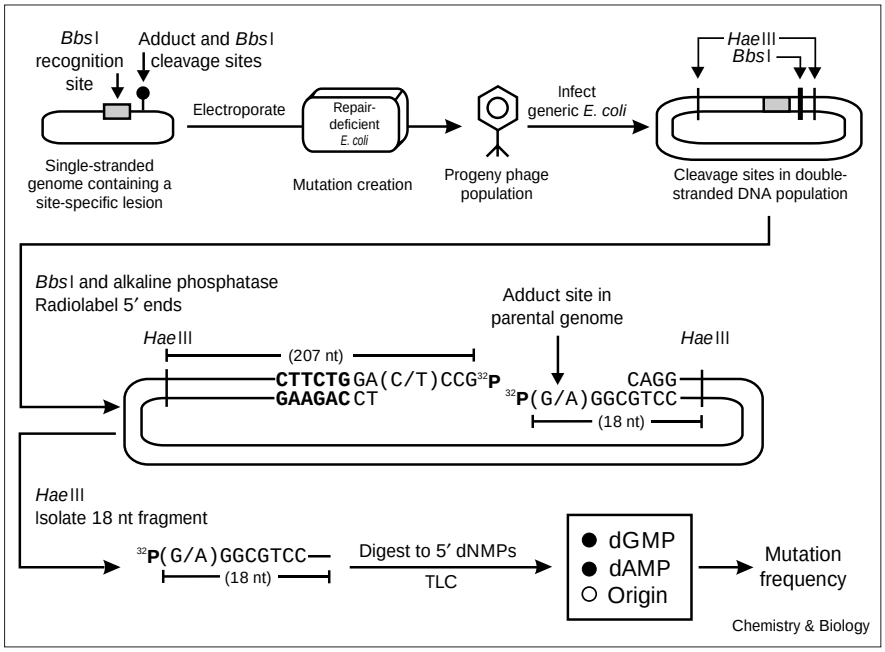
strains of *E. coli* that lacked both the UvrB and Ogt proteins, and one additionally lacked the Ada repair protein. We validated our MF detection assay and used it to show that genomes containing m^6G were nearly 100% mutagenic at the lesion site in repair-deficient *E. coli*, in several different sequence contexts. In an *E. coli* cell that possessed Ada, by contrast, mutation was suppressed in a context-dependent manner. We found that the Ada protein best repaired m^6G when a guanine base was 5′ to the lesion.

Results

Determination of the MF of DNA lesions in any sequence context

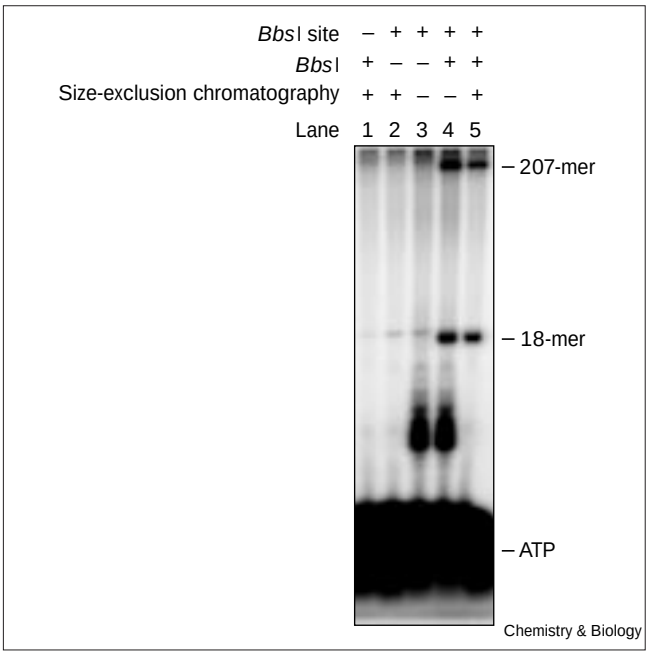
The newly developed assay (Figure 2) was named REAP (restriction endonuclease and postlabeling determination of MF) and is based upon the properties of type IIs restriction endonucleases, which cleave a specified phosphodiester bond a defined number of bases away from the enzyme recognition sequence. As shown in Figure 2, an M13 derivative bacteriophage genome was constructed in which a unique site for a type IIs enzyme, *Bbs*I, was positioned such that the phosphodiester bond to be cleaved was immediately 5′ to the interrogation site (the site that contained the DNA lesion, m^6G, before it was biologically processed). Following replication *in vivo*, double-stranded DNA progeny was cleaved with *Bbs*I, dephosphorylated with shrimp alkaline phosphatase, and treated with RNase A. The linearized, dephosphorylated vector was purified by size-exclusion chromatography and 5′ ^{32}P-labeled using T4 polynucleotide kinase, resulting in the incorporation of radiolabel into both DNA strands. Complete digestion of this substrate with *Hae*III liberated a radiolabeled 18-mer from the plus strand and an undesired 207-mer from the minus strand, which were resolved by 20% denaturing polyacrylamide gel electrophoresis (Figure 3). The base composition at the 5′ end of the 18-mer provided the MF. After excision, elution, and desalting, the 5′ ^{32}P-labeled 18-mer was digested to 5′ deoxynucleotide monophosphates (5′ dNMPs) with snake venom phosphodiesterase. At this stage, the sample was applied to a thin layer chromatography (TLC) plate and the

Figure 2



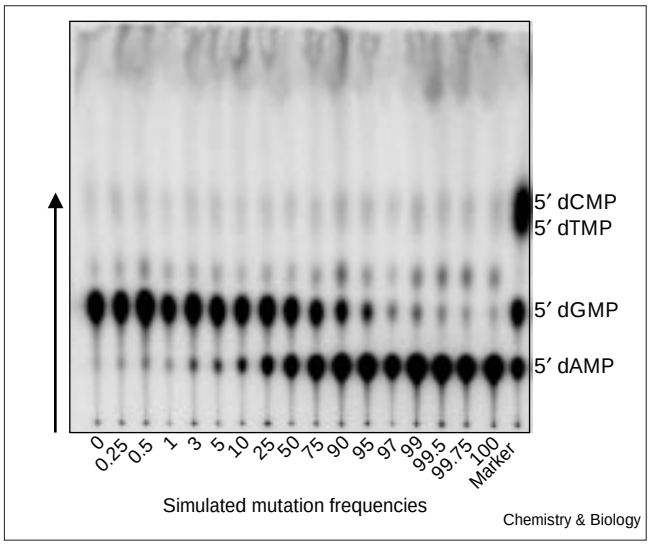
Detection of mutations in any sequence context using the REAP assay. A site-specifically adducted genome was introduced into a repair-deficient (or proficient) cell line, and the progeny phage from the biological processing of the lesion was used to infect a generic strain of *E. coli*, thus creating a double-stranded DNA population. DNA was cleaved at the position that had contained the lesion in the parental vector using the type IIs restriction endonuclease *Bbs*I (recognition sequence in bold). The 5′ ends of the linearized duplex were radiolabeled at the lesion site which, in the case of m^6G, contained a mixture of G and A after passage through the cell. Treatment with *Hae*III liberated an 18-mer, of which the base composition at the 5′ end provided the MF. To determine the MF quantitatively, the purified 18-mer was digested to 5′ deoxynucleotide monophosphates (5′ dNMPs) with snake venom phosphodiesterase. Partitioning of the radioactive 5′ dNMPs on a thin layer chromatography (TLC) plate, followed by PhosphorImager analysis, provided the fractional base composition at the lesion site, from which the mutational specificity and MF were determined.

Figure 3



Gel purification of 18-mer containing a 5′ ^{32}P phosphate at the interrogation site. Lane 5 is the result of the REAP assay as described, whereas the omission of certain elements in other lanes emphasizes the importance of size-exclusion chromatography prior to radiolabeling. The genome in lane 1 was from double-stranded M13mp7L2 wild-type.

Figure 4



TLC showing the MF analysis by the REAP assay for a range of simulated MFs at the 'lesion' site. Model mutant (A at lesion site) and wild-type (G at lesion site) phage were mixed at known input ratios and taken through the REAP assay. Following isolation of ^{32}P-labeled 18-mer and enzymatic hydrolysis, 5′ dNMPs were resolved by TLC and visualized by PhosphorImager analysis. The numbers from 0 to 100 are simulated target input MFs. The marker consisted of a hydrolysate of a 5′ ^{32}P-labeled oligonucleotide of which the 5′ end was degenerate.

5′ dNMPs were resolved (Figure 4). The relative amounts of radioactivity in the partitioned spots gave the fractional base composition at the lesion site after mutation creation, which ultimately provided the MF.

The mutation analysis system was validated by performing a reconstruction experiment in which various ratios of a model wild-type and mutant phage were passaged through cells, and the output 'MFs' were determined by two independent methods. The mutational specificity (G→A mutations) [11] and various MFs for m^{6}G in the sequence 5′-Tm^{6}GG-3′ were simulated by mixing varying amounts of phage containing genomes of sequence 5′-T*G*G-3′ and 5′-T*A*G-3′ (lesion site italicized) from two pure M13 pools. For this particular nearest-neighbor sequence context, the simulated MFs could be analyzed using both the new assay described and an α-complementation assay in which plaques of different shades of blue were counted. In this assay, lawns of amber suppressing *E. coli*, which contain an F′ episome for M13 infection and express the β-galactosidase omega peptide, were infected with the mixed populations of M13 phage containing 5′-T*G*G-3′ and 5′-T*A*G-3′ sequences within the M13 *lacZ* coding region. Phage with the 5′-T*G*G-3′ sequence formed dark blue plaques on IPTG/X-gal indicator plates because β-galactosidase gene activity was restored through α-complementation [16], thus allowing hydrolysis of the chromogenic substrate X-gal; however, phage that harbored the 5′-T*A*G-3′ sequence inside the suppressing host formed light blue plaques because a substantial amount of *lacZ* mRNA was not fully translated. The phage with the 5′-T*A*G-3′ sequence modeled m^{6}G induced mutants.

Success of the REAP assay relied on the mutant phage to wild-type phage ratio remaining the same for both phage (generated from an electroporation of genome construct or a contrived mixture) and the replicative form DNA generated from said phage (Figure 2). The contrived phage mixtures, designated as input (Table 1), were used to infect, in triplicate, cultures of *dam*, *dcm E. coli* containing an F′ episome so that double-stranded DNA could be harvested for MF analysis (Figure 2). The composition of the phage and replicative form DNA populations after infection and growth are designated as output (Table 1). The input and output (for one set) MFs were obtained by plating dilutions of the appropriate phage on indicator plates and counting more than 1500 total dark and light blue plaques per mixture (Table 1). The difference between each output and input MF revealed that no bias occurred during growth, with 95% confidence using the difference of proportions statistical test. Furthermore, each output MF determined by the REAP assay fell within the respective 95% confidence interval obtained from plaque counting (as

Table 1

Determination of simulated mutation frequencies for m^{6}G.

	Input (plaque counting)			Output set 1 (plaque counting)			Output MF (REAP)			
Target input MF (%)	Dark	Light	MF	Dark	Light	MF	Set 1	Set 2	Set 3	Average*
0	1740	2	0.1	5409	9	0.2	0.2	0.2	0.2	0.2 ± 0.0
0.25	ND	ND	ND	ND	ND	ND	0.4	0.5	0.2	0.4 ± 0.1
0.50	ND	ND	ND	ND	ND	ND	0.6	0.5	0.5	0.5 ± 0.1
1	1747	18	1.0	1696	16	0.9	1.0	0.7	0.6	0.8 ± 0.2
3	1564	44	2.7	2162	81	3.6	3.5	2.4	2.6	2.8 ± 0.6
5	1801	86	4.6	1690	87	4.9	5.6	4.5	4.8	5.0 ± 0.6
10	1502	154	9.3	1445	156	9.7	10.7	10.0	9.7	10.1 ± 0.5
25	1303	391	23.1	1353	437	24.4	26.0	25.6	27.3	26.3 ± 0.9
50	850	816	49.0	1389	1425	50.6	52.3	52.8	51.1	52.1 ± 0.9
75	531	1528	74.2	569	1678	74.7	75.7	76.8	76.0	76.1 ± 0.6
90	173	1531	89.8	163	1365	89.3	90.2	90.6	90.9	90.6 ± 0.4
95	120	1984	94.3	100	1668	94.3	95.2	95.3	96.0	95.5 ± 0.4
97	47	1588	97.1	36	1494	97.6	98.4	97.8	98.1	98.1 ± 0.3
99	37	3647	99.0	25	1797	98.6	99.1	99.0	99.3	99.1 ± 0.1
99.5	ND	ND	ND	ND	ND	ND	99.5	99.6	99.7	99.6 ± 0.1
99.75	ND	ND	ND	ND	ND	ND	99.7	99.7	99.8	99.7 ± 0.1
100	0	1613	100	1	1621	99.9	99.9	99.9	99.9	99.9 ± 0.0

Simulated MFs were determined from contrived mixtures of phage from two pure phage pools (5′-T*G*G-3′ and 5′-T*A*G-3′) in proportions to yield the target input MFs for m^{6}G at the italicized site. Dilutions of these mixtures were plated on indicator plates and wild-type (dark) and mutant (light) blue plaques were counted to provide the true input MFs. The mixtures were propagated in a bacterial cell line and progeny phage from the supernatant was used to determine the output MFs, again by plaque counting, whereas progeny double-stranded DNA inside the cell pellets was used to determine the output MFs by the REAP assay. *The average output MF determined by the REAP assay was based on MFs obtained using the double-stranded DNA output from three cultures of *E. coli* that were independently infected with a common phage input mixture, and is reported as ± one standard deviation. ND, not determined.

analyzed using the double-stranded DNA and phage progeny from the same culture tube), suggesting that the REAP and plaque counting assays provided the same results. As shown in Tables 1 and 2, the MFs analyzed in triplicate using the REAP assay were reproducible.

A TLC plate used to generate the data in Table 1 (Set 1) by the REAP assay is shown in Figure 4. For all samples, the amount of radioactivity in the C and T region constituted less than one percent of the total radiation from all 5′ dNMP spots. For the simulated 0% and 100% MF determinations, the amount of radioactivity in the expected spot represented greater than 99% of the total radiation emanating from all four 5′ dNMPs. The low background and small standard deviation in the MFs obtained by the REAP assay instils confidence that a subtle difference in the MF of a biologically processed DNA lesion in different sequence contexts can be deemed statistically significant.

Construction of site-specific genomes

After the REAP assay was validated, three viral single-stranded genomes were constructed that contained m^{6}G in the contexts 5′-Gm^{6}GA-3′, 5′-Am^{6}GA-3′, and 5′-Tm^{6}GG-3′. A procedure was developed whereby a DNA lesion would easily be placed in all possible contexts (Figure 5). Initially, an M13 circular single-stranded genome was linearized at a unique *Eco*RI site within a hairpin. A previous study [6] has used one long scaffold to bridge the resulting vector termini, thus creating a gap into which the insert containing the DNA lesion could be placed prior to ligation. Instead of using a different long scaffold to achieve a high ligation efficiency for each sequence to be tested (16 scaffolds would be needed to test all Nm^{6}GN contexts), we used two shorter scaffolds from which the variant genomes were made. It is noteworthy that the same scaffolds and M13 genome preparations are used in all samples; only the adduct-containing 16-mer varies from sample to sample. An

Table 2

Influence of sequence context on m^{6}G replication and repair.

		Cellular repair status			Output MF (REAP)			
	Sequence (5′→3′)	Ada	Ogt	UvrB	Set 1	Set 2	Set 3	Average*
Experiment 1	Gm^{6}GA	No	No	No	99.1	99.2	98.8	99.1 ± 0.2
	Am^{6}GA	No	No	No	99.8	99.6	99.6	99.6 ± 0.1
	Tm^{6}GG	No	No	No	99.2	99.3	99.3	99.3 ± 0.1
	Gm^{6}GA	Yes	No	No	ND	53.5	53.1	53.3 ± 0.3
	Am^{6}GA	Yes	No	No	90.5	90.6	90.2	90.4 ± 0.2
	Tm^{6}GG	Yes	No	No	82.1	82.1	81.9	82.0 ± 0.1
Experiment 2	Gm^{6}GA	No	No	No	97.9	97.5	97.9	97.8 ± 0.2
	Am^{6}GA	No	No	No	99.8	99.7	99.7	99.7 ± 0.1
	Tm^{6}GG	No	No	No	98.2	97.8	98.1	98.0 ± 0.2
	Gm^{6}GA	Yes	No	No	44.8	45.0	44.4	44.7 ± 0.3
	Am^{6}GA	Yes	No	No	90.6	90.3	88.1	89.7 ± 1.4
	Tm^{6}GG	Yes	No	No	81.3	81.7	80.8	81.2 ± 0.5

Single-stranded genomes containing m^{6}G in three different sequence contexts were constructed, electroporated into two strains of repair-deficient *E. coli*, and their subsequent mutation frequencies were determined using the REAP assay. Each experiment refers to an independent genome construction and subsequent electroporation, whereas genomes containing identical sequence contexts within the same experiment were from the same genome construction. *The average output MF determined by the REAP assay was based on MFs obtained using the double-stranded DNA output from three cultures of *E. coli* that were independently infected with the same progeny phage pool obtained after electroporation, and is reported as ± one standard deviation. ND, not determined.

equimolar mixture of components was created, with each scaffold annealing to 20 bases of the vector and six or seven bases of the 5′ phosphorylated insert. The lesion-containing oligonucleotide was ligated into the vector and the scaffolds were degraded with exonuclease III until they no longer annealed to the vector (data not shown). Because the technique described uses two scaffolds that do not span the variable positions that flank the lesion, although not essential for this project, identical ligation efficiencies should be achieved for the construction of all nearest-neighbor lesion-containing genomes. Furthermore, this method should improve the ligation efficiency of bulky DNA adducts because there is no destabilizing base placed opposite the lesion. One could adapt this method for constructing genomes containing thermally unstable adducts by pre-annealing the scaffolds to the vector and adding the adduct-containing insert at low temperature. We also found that passing the M13mp7L2 vector source through a hydroxylapatite column allowed M13 genomes to remain viable after being exposed to high temperature prior to electroporation.

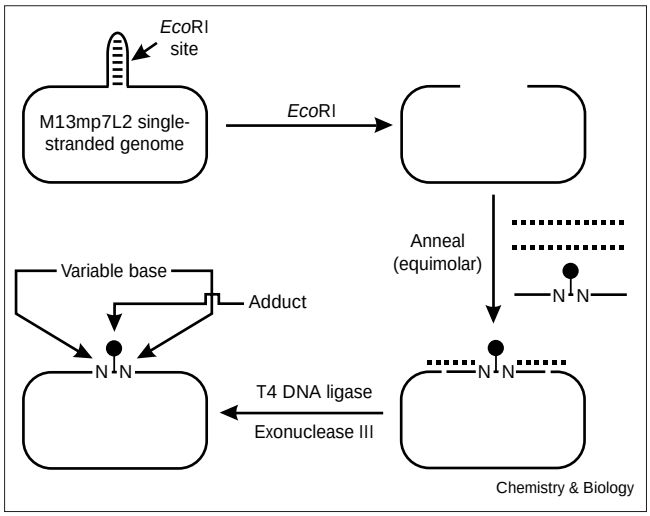
Influence of sequence context on replication past m^{6}G by DNA polymerase III and repair of m^{6}G by Ada *in vivo*

Viral genomes containing m^{6}G in the 5′-Gm^{6}GA-3′, 5′-Am^{6}GA-3′, and 5′-Tm^{6}GG-3′ contexts were made in duplicate and each construct was electroporated into repair-deficient *E. coli* as described in the Materials and methods section. The MF from each transformation was analyzed in triplicate using harvested progeny phage to infect *E. coli* for isolation of double-stranded DNA employed in the REAP assay (Figure 2). As shown in Table 2, the MF in an *ada*, *ogt*, *uvrB* cell line was nearly 100% for all three sequence contexts. This result demonstrates the nearly complete miscoding property of m^{6}G *in vivo* in the absence of repair. The isogenic cell line that expressed uninduced Ada demonstrated a sequence context dependence by this protein in the repair of m^{6}G, as reflected in the difference in the MFs. For the 5′-Gm^{6}GA-3′, 5′-Am^{6}GA-3′, and 5′-Tm^{6}GG-3′ contexts, the respective MFs obtained from the REAP assay were 53%, 90%, and 82% for the first independent trial and 45%, 90%, and 81% for the second. The 5′-Tm^{6}GG-3′ context was additionally analyzed by counting dark and light blue plaques, providing MFs of 81% and 79% for the first and second trials, respectively (data not shown). The MFs obtained by the REAP and plaque counting assays were statistically identical.

Discussion

There has been no generally applicable system by which the genetic effects of lesions in all sequence contexts could be conveniently determined. The system described here was developed in order to address that need. A key component of the methodology for mutation detection is

Figure 5



Scheme for constructing single-stranded genomes containing an adduct in any variable surrounding sequence context. The M13 single-stranded circular genome contains an *Eco*RI site that is in a hairpin. Cleavage with *Eco*RI yielded a linear product. Scaffolds (indicated by broken lines) of 27 and 26 nucleotides were annealed to the 3′ and 5′ ends, respectively, of the linearized M13 genome. The same mixture included a 16-mer in which m^6G was present in a defined sequence context (indicated by -N(lollipop)N-). Ligation covalently inserted the adduct-containing oligonucleotide into the M13 genome, and exonuclease III removed the scaffolds, affording a single-stranded, adduct-containing genome. It is noteworthy that the scaffold did not overlap with the -N(lollipop)N- trinucleotide; this feature is believed to promote a high ligation efficiency.

the use of a type IIs restriction endonuclease. Many parameters were taken into account in choosing the *Bbs*I recognition sequence to provide scission at the interrogation site. The site was introduced as a unique site within the oligonucleotide used to construct the genome, so only genomes that had contained the DNA lesion incorporated radioactivity during postlabeling. Cleavage by *Bbs*I generated a 5′ overhang of four bases, which gave unhindered access to alkaline phosphatase and polynucleotide kinase. We assumed that there was low exonuclease contamination because after 10-fold overdigestion of lambda DNA, more than 95% of the fragments generated from a *Bbs*I digest could be ligated and recut [17]. The recognition sequence was not far away from the cleavage site, which might be important if a large distance between the two promotes DNA slippage with respect to the type IIs catalytic site, resulting in some cleavage deviating from the canonical distance. Furthermore, because the genomes were constructed so that two intervening bases lay between the type IIs recognition sequence and the lesion, next-to-nearest neighbor base influences on lesion processing could have also been analyzed using the same assay (i.e., NNm^6GNN contexts). An advantage of the experimental design was that double-stranded DNA could be generated on demand by simply mixing the progeny phage, which we found to be stable for several years at 4°C, with an *E. coli* strain that possessed an F′ episome allowing for bacteriophage infection. The SCS110 *E. coli* strain was chosen for this purpose because it lacked the sequence-specific methylases Dam and Dcm, thus eliminating potential cutting bias of *Bbs*I from host methylation of certain vector sequence contexts surrounding the lesion site.

The REAP assay would have provided inaccurate MFs if the rates of enzymatic activity for *Bbs*I, alkaline phosphatase, and polynucleotide kinase varied with respect to the base at the interrogation site, and the reactions had not gone to completion. To address this possibility, separate experiments were performed, under the conditions described in the Materials and methods section, revealing that the *Bbs*I and alkaline phosphatase reactions went to completion (data not shown). The 4 hour *Bbs*I and alkaline phosphatase reactions were performed simultaneously, because the *Bbs*I digestion was complete within 2 hours, and the alkaline phosphatase reaction was complete within 1 hour. It is also noteworthy that excess RNase A was used during the double-stranded vector isolation and *Bbs*I cleavage steps prior to size-exclusion chromatography, which ensured efficient degradation of large molecular weight RNA to small fragments that were later trapped by the resin as shown in Figure 3. Given that small phosphorylatable oligonucleotide contaminants were removed by size-exclusion chromatography prior to labeling, a greater than 10-fold molar excess of ATP to 5′ phosphorylatable vector ends should have been maintained throughout the entire reaction. This was necessary to ensure that different substrates (e.g., 5′-GG...-3′ and 5′-AG...-3′) would be labeled equally, regardless of differences in the rate constants for substrate phosphorylation. Separate experiments were also performed revealing that the shrimp alkaline phosphatase and polynucleotide kinase were completely inactivated before the next step in the protocol (data not shown). This result ensured that the ATP was not depleted by futile cycling during labeling (see above) and also eliminated the possibility of ^{32}P incorporation into the 5′ ends of the newly exposed *Hae*III fragments via the minor 5′ phosphate exchange reaction of polynucleotide kinase. It is reassuring that the intensities of the 18-mer and 207-mer bands were equal to each other, and that these bands dominated the labeled species (Figure 3, lane 5), again indicating that excess ATP was maintained.

Size-exclusion chromatography effectively trapped all small phosphorylatable oligonucleotide contaminants generated during double-stranded M13 DNA isolation (less than 3% of a radiolabeled double-stranded 24-mer co-eluted with the M13 genome (data not shown)); however, there appeared to be a specific degradation band generated after size-exclusion chromatography that

ran slightly above the desired 18-mer (Figure 3, visible in lanes 1–3). Because of the close proximity this contaminant was probably excised with the 18-mer band and carried through the assay; however, the intensity of this band was always less than 3% of the desired 18-mer. It has yet to be determined where partitioning of the radiation from the contaminant would occur on the TLC plate. We also noticed a well-resolved contaminant that migrated slightly above the 5′ dGMP spot on the TLC plate (Figure 4) that was absent when a 'marker' oligonucleotide with a degenerate 5′ terminus was treated exactly as described for the MF samples but omitting the size-exclusion chromatography step. Under the conditions of the REAP assay, therefore, the material within this spot was not created by an enzymatic impurity (e.g., adenosine deaminase) or thermal degradation, and was tentatively a 5′ ribonucleotide monophosphate, which would not impact MF assessment.

As can be seen in Table 2, the MFs for m^{6}G in the *ada, ogt, uvrB* cell line were nearly 100% for all three contexts tested. Our values were therefore very similar to the results obtained by Pauly *et al.* [18,19] who, by using a single context opposite a gap, show a 94% MF for m^{6}G in an *ada, ogt, mutS E. coli* cell line. Our results showed this high MF holds for multiple contexts. An *in vitro* primer extension study using a template containing m^{6}G has shown that the Klenow fragment of *E. coli* DNA polymerase I will place a T opposite the lesion ~50% or ~85% of the time depending on the flanking base [20]; our data suggest that the DNA polymerase holoenzyme almost always places a T opposite m^{6}G during replication *in vivo*, at least for the three contexts tested. Our data further suggest that differential DNA polymerase fidelity with respect to bases flanking m^{6}G does not play an important role in the creation of mutational hotspots.

When genomes containing m^{6}G were passaged through the isogenic *ogt, uvrB* cell line (i.e., cells with a constitutive level of Ada), the MFs dropped for all three sequence contexts tested (Table 2). Although the MFs attributed to Ada were similar for identical genomes containing m^{6}G in the context 5′-Am^{6}GA-3′ (~90%) or 5′-Tm^{6}GG-3′ (~82%), they differed significantly for the 5′-Gm^{6}GA-3′ context (53% compared with 45%). This difference is probably attributed to the high repair efficiency of Ada in the 5′-Gm^{6}GA-3′ context and cannot be explained by experimental error, which was a maximum of approximately 2% (with 95% confidence) for a 50% MF (Table 1). It has been shown that the concentration of uninduced Ada in *E. coli* can vary depending on when the culture is harvested [21]. Given that m^{6}G was relatively well repaired in the 5′-Gm^{6}GA-3′ context, a slight variation in methyltransferase concentration during different batch preparations of electrocompetent cells might have had a greater impact on the absolute MF; however, within each batch preparation, the sought-after rank order of MF with respect to sequence context will probably remain the same.

The Ada protein is one of the few DNA repair proteins that can work on single-stranded DNA. Although single-stranded viral m^{6}G genomes were introduced into the repair-deficient cell lines, the observed effects might well have arisen from the reaction of Ada with m^{6}G in duplex form. The reaction with m^{6}G in single-stranded DNA occurs at only 0.1% of the rate for duplex DNA [22] and it has been estimated that there are only one or two molecules of Ada in uninduced *E. coli* [23]. Indeed, if the protein was acting on m^{6}G in duplex form, the opposing base would be T because the work presented here demonstrates that the DNA polymerase placed a T opposite the lesion. The fact that the MF dropped by half when a single base 5′ to m^{6}G was changed from an A to a G suggested that the activity of the Ada repair protein, either directly or indirectly, is strongly influenced by the 5′ flanking base. One would expect, therefore, that genomes of alkyltransferase-proficient cells treated with alkylating agents would show an excess of mutations in 5′-A*G*-3′ sites as compared with 5′-G*G*-3′ sites (mutation site italicized). There are indirect data supporting this prediction. Specifically, alkylation treatment of alkyltransferase-proficient *E. coli*, with respect to their isogenic alkyltransferase-deficient counterpart, results in increased guanine mutagenesis in the 5′-A*G*-3′ sequence [24,25], and decreased mutagenesis in the 5′-G*G*-3′ sequence [26]. One caveat is that in most of these published studies on alkyltransferase-deficient cells, Ogt was the variable modulating the site of mutation, whereas our conclusion was on the basis of cells expressing uninduced Ada. Although a nearest-neighbor analysis seems like a good method by which to start for discovering rules regarding the catalytic activity of DNA repair proteins, one should keep in mind that the topology of chromatin, the amount of gene transcription, and the local sequence context surrounding a DNA lesion (outside of nearest-neighbors) might also influence repair of DNA lesions.

We are cautious in attributing the lower MF for the 5′-Gm^{6}GA-3′ context directly to Ada alone because our cells had mismatch repair capability. Three formal scenarios can explain the differences in the MF in the Ada-proficient cell line without Ada being affected by sequence context. First, sequence context may affect binding of the mismatch repair recognition protein MutS to an m^{6}G–T pair, making some sequences better shielded by MutS from repair by Ada. Second, the m^{6}G–T mispairs may be converted equally to G–T mispairs by Ada, but initiation of the methyl-directed mismatch repair process by MutS, that would ultimately lead to a G–C pair, might be context dependent. Third, m^{6}G might be a stronger block to the DNA polymerase in some sequence contexts, thus giving

Ada more time to convert m^6G to G in single-stranded DNA before the mutation is created.

Another method does exist for locating lesion-induced hotspots in a site-specific fashion [27]. Although this method might work well for obtaining consensus sequences flanking a lesion that are simultaneously most favorable for miscoding and least favorable for base excision repair, the *in vitro* system might not accurately reflect the realities of lesion processing *in vivo*. Furthermore, many sequencing reactions would have to be performed to see a subtle difference in the contribution of flanking bases to mutations, as would any *in vivo* system whereby a lesion-containing oligonucleotide with flanking degenerate nearest-neighbors is placed in a vector and replicated in cells. In addition to being able to detect the MF of a DNA lesion placed in any varied local sequence context in single-stranded or double-stranded vectors, the obvious utility of the REAP assay is that it is fast and accurate with a little background. We note that the MF is generated from the entire population of isolated progeny, which might account for the small standard deviations shown in Tables 1 and 2. The high throughput afforded by one-dimensional TLC made possible the processing of 100 samples by one person in about a week, starting with the progeny phage populations (Figure 2).

Significance

We have described and validated an assay whereby the mutation frequencies of DNA lesions can be determined in a high-throughput fashion. Because there are no restrictions on the bases flanking the lesion, the same experimental system can be used for determining in which sequence contexts a lesion is more likely to miscode or be poorly repaired, as reflected by their mutation frequencies when placed in the appropriate repair-deficient or repair-proficient cell lines. The *in vivo* nearest-neighbor rules generated from such an experiment might provide an insight into the important contacts made between the region surrounding a DNA adduct and either a DNA polymerase for effecting nucleotide triphosphate misincorporation, or a DNA repair protein for providing poor binding and/or catalytic activity. We have demonstrated that O^6-methylguanine (m^6G) almost completely miscodes in the three contexts tested, suggesting that the replicative DNA polymerase in *Escherichia coli* does not contribute to the non-uniform distribution of mutational hotspots when cells are treated with alkylating agents. We have also shown that m^6G is better repaired when flanked by a 5′ guanine base; an exhaustive analysis of all contexts will be presented later. The restriction endonuclease and postlabeling (REAP) assay can be used to study lesions other than m^6G that exhibit point mutations, in single-stranded or double-stranded vectors; and, with the use of the appropriate shuttle-vector, the assay should be able to determine the effect of sequence context on DNA lesion processing in mammalian cells.

Materials and methods

Simulation of MF

MFs were simulated by mixing varying amounts of phage from two homogeneous pools, of which DNA was identical except at the interrogation site. The pools were derived from the single-stranded parent vector M13mp7L2 (C. Lawrence, University of Rochester), in which the hairpin at the *Eco*RI site was replaced with the sequence 5′-GAAGAC-CTXGGCGTCC-3′, with X being a G or A base representing pure wild-type or mutant progeny, respectively. Double-stranded DNA was obtained by infecting a 1:100 dilution of an overnight culture of SCS110 (JM110, *end A1*; Stratagene) cells grown in unsupplemented Luria–Bertani (LB) media [28] with the 500 µl of the phage mixtures. Approximately 10^{10} phage were mixed with 10^8 cells in 10 ml LB and grew on a culture wheel at 37°C for 12 h. The entire culture was transferred to a 15 ml polypropylene tube, pelleted at 15000 × g for 10 min and drained. Some supernatant was saved to assess the output MF by plaque color screening as described later.

Analysis of MF using the REAP assay

Double-stranded DNA was prepared as follows, using a modification of the technique described by Lee and Rasheed [29]. In sets of 10, the cell pellets obtained as described above were resuspended in 500 µl R1: 50 mM glucose, 25 mM Tris-HCl buffer (pH 8.0), 10 mM EDTA, supplemented with 5 mg/ml lysozyme powder (grade I from chicken egg white, Sigma no. L6876) and 400 µg/ml of boiled RNase A (type II-A from bovine pancreas, Sigma no. R5000) solution prepared as described [16]. To 2 ml polypropylene tubes containing the above suspension was added 250 µl R2: 2% sodium dodecyl sulfate, 0.4 N NaOH obtained from 10% and 2 N stock solutions, respectively. The tubes were agitated gently at room temperature on an Orbitron (Boekel) for 5 min, after which was added 400 µl ice cold R3: 7.5 M NH_4OAc (pH 7.6). Tubes were inverted gently 15 times and placed on ice for 5 min, after which they were centrifuged at 4°C for 10 min. The supernatant was transferred to 1.5 ml tubes, placed on ice for 5 min, and again centrifuged at 4°C for 10 min. The supernatant was then decanted into 2 ml tubes containing 650 µl isopropanol, and the contents were mixed and kept at room temperature for 10 min. Tubes were centrifuged at room temperature for 10 min and the supernatant was discarded. A 750 µl solution of 2 M NH_4OAc (pH 7.4) was added and the tubes were agitated for 5 min at high speed on a Vortex Genie 2™ (VWR) equipped to hold thirty 1.5 ml tubes. The contents were pulsed at 1500 × g for 5 sec, transferred to 1.5 ml tubes and placed at 4°C overnight. Tubes were centrifuged at 4°C for 10 min and the supernatant was decanted into 1.5 ml tubes containing 750 µl isopropanol. All following manipulations were performed at room temperature. The contents were mixed, and after 10 min, centrifuged for 10 min. The supernatant was discarded and 1 ml 70% ethanol was added, after which the tubes were shaken, centrifuged for 10 min, then drained. After another 70% ethanol rinsing, the samples were dried under vacuum in a Speed Vac® (Savant) for 20 min and stored at –20°C. The yield was ~15 µg (1 A_{260} = 50 µg/ml).

Double-stranded DNA was linearized at the lesion site and radiolabeled as follows. The samples obtained above were resuspended in 50 µl 1 × buffer 2 (New England Biolabs: 10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol (DTT) (pH 7.9)) containing 10 U *Bbs*I (New England Biolabs), 3 U shrimp alkaline phosphatase (Boehringer-Mannheim), and 30 µg RNase A prepared as previously described. After incubation at 37°C for 4 h, the samples were heated at 80°C for 5 min to inactivate the alkaline phosphatase and stored at –20°C until further use.

The linearized vector was purified away from small contaminating oligonucleotides by size-exclusion chromatography as follows. Disposable home-made stationary columns were made from polyethylene

transfer pipets (Corning Samco no. 202) and empty spin columns equipped with a 30 micron frit (American Bioanalytical). The tapered and closed ends of the pipet were snipped off and the smaller end was forced ~9 mm into the top of the spin column. To the assembly that was primed with water, was added 4.6 ml of well-suspended Sephracryl S-400 resin (Pharmacia). The open end of the pipet bulb acted as a solvent reservoir as the drained column (91 × 7 mm) was washed with 3 ml 10 mM Tris-HCl buffer (pH 8.5). For ease of processing multiple samples, the columns were inserted into holes made in a styrofoam rack (Sarstedt no. 95.064.249) designed to hold microfuge tubes (16 holes were made on one side of each rack), and an identical rack was aligned underneath for fraction collection. The sample volume was adjusted to 100 μl with 10 mM Tris-HCl buffer (pH 8.5) and applied to the drained column. Adsorption was allowed to occur and after 10 min, 0.18 g sand, dried and washed (Mallinckrodt), was sprinkled on top. Test runs in which 100 μl Tris-HCl buffer were added, collected, and spotted with SYBR Green I® dye (Molecular Probes, Inc.) on Saran-Wrap™ revealed that M13 DNA elution consistently started after fraction 13 and ended before fraction 20. Therefore, 1.3 ml 10 mM Tris-HCl buffer (pH 8.5) was applied to the sample columns that were allowed to drain, after which 600 μl Tris-HCl buffer was added and collected into tubes containing 150 μl 10 M NH_4OAc (pH 7.6) and 750 μl isopropanol. All following manipulations were also performed at room temperature. Samples were mixed, kept overnight, centrifuged in sets of 10 for 10 min, and the supernatant discarded. Approximately 1 ml 70% ethanol was added, after which the tubes were shaken, centrifuged for 10 min, then drained. After another 70% ethanol wash, the samples were dried in a Speed Vac® for 20 min and stored at –20°C. The yield of M13 was approximately 1 pmol (1 pmol = 4.72 μg).

The vector ends were ^{32}P-labeled as follows. Samples were resuspended in 50 μl 10 mM Tris-HCl buffer (pH 8.5) and 20 μl were dried down (45 min in a Speed Vac® at room temperature) and used for labeling. The DNA was then resuspended in 7.5 μl 1 × buffer 2 containing the following: supplemental DTT for a final concentration of 10 mM, [γ-^{32}P]ATP (New England Nuclear, 6000 Ci/mmol) diluted with ATP (Pharmacia) to provide 10 pmol ATP at specific activity of 1000 Ci/mmol, and 5 U T4 polynucleotide kinase (New England Biolabs). This provided a greater than 10-fold molar excess of ATP to 5′ phosphorylatable ends. After incubation at 37°C for 1 h, the samples were heated at 65°C for 20 min to inactivate the T4 polynucleotide kinase.

The 18-mer with variable 5′ base composition (Figure 2) was created and isolated as follows. To the above solution was added 2.5 μl 1 × buffer 2 containing 30 U *Hae*III (New England Biolabs, 50 U/μl stock). After this solution was incubated at 37°C for 2 h, 10 μl denaturing dye (98% formamide, 10 mM EDTA (pH 8.0), 0.025% bromophenol blue, 0.025% xylene cyanol FF) was added and the samples were stored at –20°C until further use. Small, thin 17 cm × 13.5 cm × 0.8 mm 20% denaturing polyacrylamide gels (7 M urea, 19:1 acrylamide:*N*-*N*′-methylene-bis-acrylamide, 1 × TBE (0.09 M Tris-HCl, 0.09 M boric acid, 0.002 M EDTA (pH 8.3))) were prerun at 550 volts for 15 min. Every other lane was loaded with sample, and 10 samples per gel were run at 550 volts for 70 min. The gel was bloted with Kimwipes®, placed between SaranWrap™, and set under a storage phosphor screen for 30 min; the output was generated on a transparency using a PhosphorImager (Molecular Dynamics). After the free ATP was cut out of the gel and discarded, the gel was sliced vertically down the unloaded lanes to prevent cross-contamination and for each sample, the well and 18-mer from the transparency placed underneath the gel strip was aligned with the well and the bromophenol blue and xylene cyanol dye markers flanking the 18-mer from the gel. The 18-mers were excised in 1 cm × 0.5 cm slices that were crushed and had 200 μl water added to them in 0.5 ml flip-top tubes (Sarstedt no. 72.699), of which caps and lips had been removed with a utility knife. After sitting at room temperature overnight, the gel bits were agitated by drawing and expelling 100 μl of solution, and 20 min passed before the liquid was separated from the gel. Alternatively, the crushed gel can be soaked for 1 h, agitated, and soaked for an extra hour to allow for complete equilibration between the gel and the liquid. A home-made device was made that quickly forced the liquid out of the gel solution, leaving most of the gel bits behind. Two 1.5 ml flip-top tubes (Sarstedt no. 72.690) were stacked; the top one had the cap pulled off, a small hole at the bottom made by a 26 gauge needle, and contained 0.09 g sand. The 0.5 ml tube containing the gel solution was placed into the assembly inverted and spun at high speed in a microcentrifuge for 1 min (with every other chamber loaded). Greater than 60% of the radioactivity came out in the eluant. Home-made desalting spin columns were made using Sephadex G-25 Fine resin (Pharmacia) that had been swelled and washed with 3 volumes of water. Resin was added to primed empty columns (American Bioanalytical) and the liquid was allowed to drain. Each column was placed into two 1.5 ml Sarstedt flip-top tubes that were stacked, had both lids pulled off, and a hole made in the top supporting tube. The assembly was placed in a swinging-bucket centrifuge (RC-3B Sorvall) and spun at 700 × g (1500 rpm) for 2 min. The packed bed height was 42 mm. All of the gel eluant (approximately 160 μl) was layered onto the desalting column, which was placed inside the supporting tube and stacked on a new 1.5 ml Sarstedt tube, which was spun at 700 × g for 4 min. Greater than 75% of the radioactivity loaded on the column came out in the eluant, with a volume change of less than 10%. The sample to sample recovery from both gel elution and Sephadex desalting were consistent. The eluant was dried at room temperature under vacuum in a Speed Vac® overnight.

The 18-mers were digested to monomers and analyzed by TLC as follows. The 5′ ^{32}P-labeled 18-mers were digested to 5′ deoxynucleotide monophosphates using snake venom phosphodiesterase I (ICN Biomedical no. 100978, 40 U/mg solid). A 20% glycerol stock at 50 mU/μl had been made from the lyophilized enzyme and stored at –20°C. A 5 μl solution of 100 mM Tris-HCl buffer (pH 8.8), 15 mM $MgCl_2$, and 50 mU snake venom phosphodiesterase I was used to resuspend the 18-mers and digestion was carried out at 37°C for 1 h. Approximately 0.5 μl sample was drawn into a 1 μl glass micro capillary tube (Drummond) attached to an adjustable positive pressure dispenser and spotted on a 20 × 20 cm polyethyleneimine cellulose TLC plate (J.T. Baker). The spots were spaced 2.5 cm from the bottom and 1 cm apart. After the last spot was dry (5 min) the plates were placed in TLC tanks (27 L × 7.5 W × 26 H cm) containing 200 ml saturated $(NH_4)_2SO_4$. The plates were developed until the solvent almost reached the top, dried at room temperature for 1 h, and placed under storage phosphor screens for 28–36 h (below pixel saturation). MFs were calculated using ImageQuant 5.0 software (Molecular Dynamics) by drawing ellipses (local average background correction) around the A (5′ dAMP) and G (5′ dGMP) spots and reporting the percentage value of 100 × [A/(A + G)].

Synthesis of m^{6}G oligonucleotides

All oligonucleotides were made on an Applied Biosystems 391 DNA Synthesizer using the 0.2 μmol scale. The three 16-mers, 5′-GAAGACC- (Gm^{6}GA) or (Am^{6}GA) or (Tm^{6}GG) -GCGTCC-3′ and a control oligonucleotide lacking G residues, 5′-TAATACCTm^{6}GTA-CATCC-3′ were made using the *N*-isobutyryl m^{6}G protected phosphoramidite (Glen Research). Deprotection was carried out by placing the controlled pore glass resin in dry glass vials equipped with septa to which were added, under argon, 1 ml 10% 1,8-diazabicyclo[5.4.0]undec-7-ene in anhydrous methanol (Aldrich). The solution stirred at room temperature for three weeks in a dark desiccator, after which the contents were transferred to microfuge tubes and lyophilized to an oil. After resuspension in 1 ml 10 mM NaOH, the DNA was ethanol precipitated, mixed with 50% formamide, and purified by migration through ~25 cm of a 20% denaturing polyacrylamide gel containing 7 M urea and 1 × TBE. Bands were excised, crushed and soaked with agitation in 10 ml water, and desalted on Sep-Pak® C-18 cartridges (Waters). Oligonucleotides were characterized as follows. In a 40 μl solution of 100 mM Tris-HCl buffer (pH 8.8) and 15 mM $MgCl_2$ was digested approximately 2 nmol of each oligonucleotide with 300 mU snake venom phosphodiesterase I (ICN Biomedical) and 20 U alkaline phosphatase (type VII-T from bovine intestinal mucosa, Sigma no. P6774) at 37°C for 1 h. All protected and deprotected nucleosides were separated by gradient

reverse phase HPLC using an analytical Beckman Ultrasphere™ 5 μ C-18 column (250 × 4.6 mm) and a flow of 1 ml/min. The linear gradient was 0–8.2% B over 35 min, then 8.2–30% B over 25 min (solvent A, 0.1 M NH_4OAc; solvent B, CH_3CN (neat)). Detection and quantitation were performed with an in-line diode array detector (Hewlett Packard 1040A). A comparison of the retention times and UV peak profiles between a mixture of authentic standards and the nucleoside digests from the synthesized m^6G oligonucleotides revealed the following. The isobutyryl group had been completely removed from m^6G, no guanine was formed at the lesion site (as judged by the identically deprotected control oligonucleotide), and no O^6-ethylguanine was created during the ethanol precipitation step (data not shown). The limit of detection was less than 0.5% for each event monitored.

Construction of site-specific, single-stranded viral genomes

M13mp7L2 was prepared as follows. To a 2 l baffled flask was added 1 l 2 × YT media [28] and 2 ml of a saturated culture of GW5100 *E. coli* (JM103, P1⁻; G.C. Walker, Massachusetts Institute of Technology). The flask was shaken (275 rpm) at 37°C for 2.5 h, after which was added 2 ml of saturated M13mp7L2 progeny phage supernatant (~6×10^9 plaque forming units (pfu)) and growth continued for 9 h. Cells were pelleted and a solution that was 20% polyethylene glycol (MW 8000, Sigma no. P2139) and 2.5 M NaCl was added to the supernatant (1:4 v/v). The phage in the supernatant was precipitated at 4°C overnight, pelleted, combined and resuspended in 19 ml TE (10 mM Tris-HCl, 1 mM EDTA (pH 8.0)), split into thirds, and extracted with 5 × 3 ml phenol:chloroform:isoamyl alcohol (25:24:1 v/v/v) until the aqueous-organic interface remained clear. The aqueous phase (~15 ml) was applied to a drained column (35 mm × 10 mm) containing 1 g DNA-Grade Bio-Gel® HTP hydroxylapatite resin (BioRad), which had been conditioned with 5 ml TE (an acceptable flow-rate was achieved by applying moderate air pressure over the solvent bed). The column was then washed with 5 ml TE and the DNA was eluted with a solution of phosphate buffer (0.07 M KH_2PO_4, 0.07 M K_2HPO_4) at the molarity suggested by the manufacturer that would elute single-stranded DNA. Fractions containing the most concentrated DNA were pooled (~5 ml), split in half, and concentrated/dialyzed through two Centricon-100 spin-dialysis devices (Amicon no. 4212) using three, 2 ml TE rinses. The final yield of single-stranded M13mp7L2 was ~4.5 mg.

Genomes containing m^6G were constructed as follows. *Eco*RI, T4 polynucleotide kinase, and T4 DNA ligase were from New England Biolabs. ATP and exonuclease III were from Pharmacia. Oligonucleotides containing m^6G (10 pmol) were 5′-phosphorylated in a 30 μl volume with T4 polynucleotide kinase (0.5 U/μl) in 70 mM Tris-HCl buffer (pH 7.5), 10 mM $MgCl_2$, 5 mM DTT, and 1 mM ATP at 37°C for 1 h. Single-stranded M13mp7L2 (0.15 pmol/μl, (2.36 μg/pmol)) was linearized with *Eco*RI (0.37 U/μl) in 100 mM Tris-HCl buffer (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, and 100 μg/ml bovine serum albumin at 23°C for 4 h, after which, an equimolar amount of the scaffolds 5′-GGTCTTCCACTGAATCATGGTCATAGC-3′ and 5′-AAAACGACGGCCAGTGAATTGGACGC-3′ were added, with negligible volume increase. Approximately 100 μl of this solution (10 pmol in each component) was added to the tube containing the phosphorylated oligonucleotide. The components were annealed in a PCR machine using the following program: 80°C for 5 min, 80°C–50°C at 1°C/min, 50°C–0°C at 0.33°C/min. The solution was made 10 mM DTT, 1 mM ATP, and 5.4 U/μl T4 DNA ligase, with negligible volume increase, after which it was incubated at 16°C for 2 h. Scaffolds were degraded to approximately 4-mers (data not shown) by incubation with exonuclease III (1 U/μl) at 37°C for 2 h. The constructed genomes were desalted by spin-dialysis through Centricon-100 devices using two, 2 ml rinses with TE, after which each sample was divided equally into eight tubes (~8 μl/tube) and stored at –20°C.

Electroporation of genomes into repair-deficient isogenic E. coli

The drug-resistant alleles Δ*ada*-25::Cam^r [30], *ogt*-1::Kan^r [31], and *uvr*B5::Tn*10*Tet^r [32] had been previously used to construct the triple mutant *E. coli* strain CJM2 (*ada, ogt, uvrB*) [33] by P1 *vir* transduction into the wild-type strain FC215 [34]. We also used P1 *vir* transduction [35] to transfer the *ogt*-1::Kan^r and *uvr*B5::Tn*10*Tet^r alleles from CJM2 into FC215 to create the isogenic repair-deficient *E. coli* strain C216/uvrB⁻ (*ogt, uvrB*). Repair-deficient *E. coli* cells were made electrocompetent as follows. A 5 ml saturated culture of cells was used to innoculate 500 ml LB (supplemented with 2.5 g/l $MgSO_4$, 0.2% (w/v) maltose, and the appropriate drugs at concentrations of 10 μg/ml for chloramphenicol, 50 μg/ml for kanamycin, and 15 μg/ml for tetracycline). After growing with aeration in a baffled flask at 37°C, the cells were harvested at mid-log phase (OD_{600} = 0.5). At 0°C, cells from 350 ml of culture were washed thoroughly with 350 ml, then 175 ml, of ice-cold sterile water. Cells were finally resuspended in 4 ml 10% glycerol, providing ~2×10^9 cells per 100 μl, as judged by counting plated dilutions of cells. Approximately 100 μl of electrocompetent cells were mixed with a thawed aliquot of desalted genome construct (see above). To a chilled electroporation cuvette (0.2 cm gap) containing the above solution was applied 2.5 kV at 129 ohms from an Electro Cell Manipulator 600® electroporation system (BTX). The voltage delivered (2.36 kV) and time constant (5.60 msec) were consistent between electroporations for all genomes and cell lines, and rendered a reduction in cell survival (electroporation of cells alone) of no more than 60%. Immediately, 1 ml SOC media [16] was added to the cuvette and the contents transferred to 9 ml supplemented LB, lacking drugs. The number of initial independent events was determined by immediately plating a dilution of the mixture onto a lawn of NR9050 cells (see below). One million electroporated cells secreted phage and were thus initially transformed with the M13 construct. The 10 ml culture was immediately placed on a roller drum at 37°C for 2.5 h, after which the supernatant containing progeny phage at ~10^{10} pfu/ml to be used for MF analysis was stored at 4°C.

Determination of MF by counting individual plaques

MFs were obtained by counting dark and light blue plaques for the simulated MF experiment and for the *in vivo* m^6G processing experiment in which a 5′-Tm^{6}GG-3′ sequence context was used. An overnight culture of NR9050 *E. coli* (*suB, F′ prolacIZΔM15* (relevant genotype); R.M. Schaaper, NIEHS) was diluted 1:5 in 2 × YT and grew on a culture wheel at 37°C for 90 min before being used as plating bacteria. In triplicate or greater, 300 μl plating bacteria, 25 μl 1% thiamine, 10 μl isopropyl β-D-thiogalactopyranoside (IPTG, 24 mg/ml stock in water), 40 μl 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal, 40 mg/ml stock in *N,N*-dimethylformamide), and phage to yield about 500 plaques per 100 × 15 mm plate were added to 2 ml B-broth soft agar (10 g/l tryptone, 8 g/l NaCl, 10 mg/l thiamine, and 0.6% (w/v) Bacto agar) kept at 47°C and spread immediately on 25 ml B-broth plates (2% (w/v) Bacto agar). After 10 min, the plates were inverted and incubated at 37°C for 12–16 h. This technique gave excellent discrimination between dark and light blue plaques.

Acknowledgements

We thank John Marquis and Leona Samson (Harvard School of Public Health) for the triple mutant *E. coli* strain, CJM2. This work was supported by grants CA09112 (J.C.D.) and CA80024 (J.M.E.) from the National Institutes of Health.

References

1. Basu, A.K. & Essigmann, J.M. (1988). Site-specifically modified oligodeoxynucleotides as probes for the structural and biological effects of DNA-damaging agents. *Chem. Res. Toxicol.* **1**, 1-18.
2. Loechler, E.L. (1996). The role of adduct site-specific mutagenesis in understanding how carcinogen–DNA adducts cause mutations: perspective, prospects and problems. *Carcinogenesis* **17**, 895-902.
3. Pauly, G.T., Hughes, S.H. & Moschel, R.C. (1994). Response of repair-competent and repair-deficient *Escherichia coli* to three O^6-substituted guanines and involvement of methyl-directed mismatch repair in the processing of O^6-methylguanine residues. *Biochemistry* **33**, 9169-9177.
4. Basu, A.K., Wood, M.L., Niedernhofer, L.J., Ramos, L.A. & Essigmann, J.M. (1993). Mutagenic and genotoxic effects of three vinyl chloride-induced DNA lesions: 1,N^6-ethenoadenine, 3,N^4-ethenocytosine, and 4-amino-5-(imidazol-2-yl)imidazole. *Biochemistry* **32**, 12793-12801.

5. Chary, P., *et al*., & Lloyd, R.S. (1995). *In vivo* and *in vitro* replication consequences of stereoisomeric benzo[*a*]pyrene-7,8-dihydrodiol 9,10-epoxide adducts on adenine N^6 at the second position of N-*ras* codon 61. *J. Biol. Chem.* 270, 4990-5000.
6. Lawrence, C.W., Borden, A., Banerjee, S.K. & LeClerc, J.E. (1990). Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Res*. 18, 2153-2157.
7. Goth, R. & Rajewsky, M.F. (1974). Persistence of O^6-ethylguanine in rat-brain DNA: correlation with nervous system-specific carcinogenesis by ethylnitrosourea. *Proc. Natl Acad. Sci. USA* 71, 639-643.
8. Kleihues, P. & Margison, G.P. (1974). Carcinogenicity of *N*-methyl-*N*-nitrosourea: possible role of excision repair of O^6-methylguanine from DNA. *J. Natl Cancer Inst*. 53, 1839-1841.
9. Taverna, P. & Sedgwick, B. (1996). Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli. J. Bacteriol.* 178, 5105-5111.
10. Sedgwick, B. (1997). Nitrosated peptides and polyamines as endogenous mutagens in O^6-alkylguanine-DNA alkyltransferase deficient cells. *Carcinogenesis* 18, 1561-1567.
11. Loechler, E.L., Green, C.L. & Essigmann, J.M. (1984). *In vivo* mutagenesis by O^6-methylguanine built into a unique site in a viral genome. *Proc. Natl Acad. Sci. USA* 81, 6271-6275.
12. Samson, L. (1992). The suicidal DNA repair methyltransferases of microbes. *Mol. Microbiol.* 6, 825-831.
13. Samson, L., Thomale, J. & Rajewsky, M.F. (1988). Alternative pathways for the *in vivo* repair of O^6-alkylguanine and O^4-alkylthymine in *Escherichia coli*: the adaptive response and nucleotide excision repair. *EMBO J.* 7, 2261-2267.
14. Voigt, J.M., Van Houten, B., Sancar, A. & Topal, M.D. (1989). Repair of O^6-methylguanine by ABC excinuclease of *Escherichia coli in vitro*. *J. Biol. Chem.* 264, 5172-5176.
15. Rasmussen, L.J. & Samson, L. (1996). The *Escherichia coli* MutS DNA mismatch binding protein specifically binds O^6-methylguanine DNA lesions. *Carcinogenesis* 17, 2085-2088.
16. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* (2nd edn), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY.
17. *New England Biolabs catalog* (1998), New England Biolabs, Beverly, MA.
18. Pauly, G.T., Hughes, S.H. & Moschel, R.C. (1998). Comparison of mutagenesis by O^6-methyl- and O^6-ethylguanine and O^4-methylthymine in *Escherichia coli* using double-stranded and gapped plasmids. *Carcinogenesis* 19, 457-461.
19. Pauly, G.T., Hughes, S.H. & Moschel, R.C. (1995). Mutagenesis in *Escherichia coli* by three O^6-substituted guanines in double-stranded or gapped plasmids. *Biochemistry* 34, 8924-8930.
20. Singer, B., Chavez, F., Goodman, M.F., Essigmann, J.M. & Dosanjh, M.K. (1989). Effect of 3′ flanking neighbors on kinetics of pairing of dCTP or dTTP opposite O^6-methylguanine in a defined primed oligonucleotide when *Escherichia coli* DNA polymerase I is used. *Proc. Natl Acad. Sci. USA* 86, 8271-8274.
21. Taverna, P. & Sedgwick, B. (1996). Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli. J. Bacteriol.* 178, 5105-5111.
22. Lindahl, T., Demple, B. & Robins, P. (1982). Suicide inactivation of the *E. coli* O^6-methylguanine-DNA methyltransferase. *EMBO J.* 1, 1359-1363.
23. Rebeck, G.W., Smith, C.M., Goad, D.L. & Samson, L. (1989). Characterization of the major DNA repair methyltransferase activity in unadapted *Escherichia coli* and identification of a similar activity in *Salmonella typhimurium. J. Bacteriol.* 171, 4563-4568.
24. Vidal, A., Abril, N. & Pueyo, C. (1995). DNA repair by Ogt alkyltransferase influences EMS mutational specificity. *Carcinogenesis* 16, 817-821.
25. Vidal, A., Abril, N. & Pueyo, C. (1997). The influence of DNA repair by Ogt alkyltransferase on the distribution of alkylnitrosourea-induced mutations in *Escherichia coli. Environ. Mol. Mutagen.* 29, 180-188.
26. Jurado, J., Ferrezuelo, F. & Pueyo, C. (1995). Mutational specificity of 1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea in the *Escherichia coli lacI* gene of O^6-alkylguanine-DNA alkyltransferase-proficient and -deficient strains. *Mol. Carcinog.* 14, 233-239.
27. Hatahet, Z., Zhou, M., Reha-Krantz, L.J., Morrical, S.W. & Wallace, S.S. (1998). In search of a mutational hotspot. *Proc. Natl Acad. Sci. USA* 95, 8556-8561.
28. Ausubel, F.M., *et al.,* & Struhl, K. (1993). *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., New York, NY.
29. Lee, S.Y. & Rasheed, S. (1990). A simple procedure for maximum yield of high-quality plasmid DNA. *Biotechniques* 9, 676-679.
30. Shevell, D.E., Abou-Zamzam, A.M., Demple, B. & Walker, G.C. (1988). Construction of an *Escherichia coli* K-12 *ada* deletion by gene replacement in a *recD* strain reveals a second methyltransferase that repairs alkylated DNA. *J. Bacteriol.* 170, 3294-3296.
31. Rebeck, G.W. & Samson, L. (1991). Increased spontaneous mutation and alkylation sensitivity of *Escherichia coli* strains lacking the *ogt* O^6-methylguanine DNA repair methyltransferase. *J. Bacteriol.* 173, 2068-2076.
32. Backendorf, C., Spaink, H., Barbeiro, A.P. & van de Putte, P. (1986). Structure of the *uvrB* gene of *Escherichia coli.* Homology with other DNA repair enzymes and characterization of the uvrB5 mutation. *Nucleic Acids Res*. 14, 2877-2890.
33. Samson, L., Han, S., Marquis, J.C. & Rasmussen, L.J. (1997). Mammalian DNA repair methyltransferases shield O^4MeT from nucleotide excision repair. *Carcinogenesis* 18, 919-924.
34. Mackay, W.J., Han, S. & Samson, L.D. (1994). DNA alkylation repair limits spontaneous base substitution mutations in *Escherichia coli. J. Bacteriol.* 176, 3224-3230.
35. Miller, J.H. (1992). *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for* Escherichia coli *and Related Bacteria.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.