A mammalian DNA repair enzyme that excises oxidatively damaged guanines maps to a locus frequently lost in lung cancer

Rongzhen Lu, Huw M. Nash and Gregory L. Verdine

Background: Guanine residues in the genome are vulnerable to attack by free radicals and reactive oxygen species. A major lesion thus produced, 8-oxoguanine (^{O}G), causes mutations by mis-pairing with adenine during replication. In bacteria and budding yeast, ^{O}G is removed from the genome through the action of base-excision DNA repair (BER) enzymes, which catalyze expulsion of the aberrant base and excision of its sugar moiety from the DNA backbone. Although ^{O}G is known to be produced in and cleansed from mammalian genomes, the enzymes responsible for ^{O}G repair in these cells have remained elusive.

Results: Here, we report the cloning and biochemical characterization of mammalian BER enzymes that specifically target ^{O}G residues in DNA. These 8-oxoguanine DNA glycosylases, hOgg1 (human) and mOgg1 (murine), are homologous to each other and to yeast Ogg1. They also contain an active site motif – the Helix–hairpin–Helix, Gly/Pro-rich–Asp motif – characteristic of a superfamily of BER proteins with a similar core fold and active site geometry. Both hOgg1 and mOgg1 exhibit exquisite selectivity for the base opposite ^{O}G in DNA, operating with high efficiency only on ^{O}G base-paired to cytosine. Furthermore, hOgg1 and mOgg1 are unable to process a panel of alternative lesions, including 8-oxoadenine, yet bind with high affinity to synthetic abasic site analogs. The proteins operate through a classical glycosylase/lyase catalytic mechanism; mutation of a catalytically essential lysine residue results in loss of catalytic potency but retention of binding to ^{O}G-containing oligonucleotides. The *hOGG1* gene is localized on the short arm of chromosome 3 (3p25/26) in a region commonly deleted in cancers.

Conclusions: These results conclusively establish the existence and identity of an 8-oxoguanine DNA glycosylase/lyase in human and murine cells, completing the triad of proteins that together protect mammals from the genotoxic effects of guanine oxidation. The observation that at least one allele of *hOGG1* is commonly deleted in cancer cells suggests that such cells may possess a reduced capacity to counter the mutagenic effects of reactive oxygen species, a deficiency that could increase their overall genomic instability. This speculation is fueled by recent observations that cells constitutively active for the Ras/Raf pathway constitutively produce high levels of superoxide, a known generator of ^{O}G.

Address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Correspondence: Gregory L. Verdine
E-mail: verdine@chemistry.harvard.edu

Received: **21 April 1997**
Accepted: **6 May 1997**

Published: **12 May 1997**

Current Biology 1997, 7:397–407
http://biomednet.com/elecref/0960982200700397

Background

The reduction of molecular oxygen to water is central to the process of energy utilization in cells. This four-electron process proceeds stepwise through a series of partially reduced intermediates, known collectively as reactive oxygen species (ROS), which include superoxide ion ($O_2^{\bullet-}$), hydrogen peroxide (H_2O_2) and hydroxyl radical ($HO^{\bullet}$). ROS are not only generated as by-products of metabolism, but have also been implicated as key mediators of intracellular signal transduction pathways [1–6]. Exogenous sources of ROS include ionizing radiation, free radicals, singlet-oxygen sensitizer dyes and redox-active organic molecules. Irrespective of their source, ROS are potent genotoxins that attack DNA and initiate the formation of oxidative lesions [7]. These ROS-induced alterations in genome structure have been implicated in a variety of pathophysiological processes, especially in cancer and aging [8].

A lesion produced in great amounts and probably having the most deleterious effects is 8-oxoguanine (^{O}G, also known as 8-hydroxyguanine) [9], which mis-pairs preferentially with adenine during replication and thereby gives rise to G:C to T:A transversion mutations [10]. Owing to their persistent generation, relative abundance and potent mutagenicity, ^{O}G residues are believed to represent a major source of spontaneous mutagenesis in all aerobic cells [7].

The mechanisms that confer resistance to guanine oxidation in *Escherichia coli* have been studied extensively [10–14]. This resistance pathway, known as the GO system, comprises three component enzymes, MutM, MutY and MutT. MutM (also known as Fpg) is a DNA glycosylase/lyase that recognizes primary ^{O}G lesions in DNA and catalyzes base excision and subsequent degradation of the sugar moiety so as to carve out the entire aberrant nucleoside from DNA. MutY is a monofunctional DNA glycosylase that recognizes the product of ^{O}G misreplication — ^{O}G:A base-pairs — and catalyzes hydrolytic excision of the inappropriate adenine base. The final component of the resistance triad is MutT, which is not a DNA repair enzyme, but rather a triphosphatase that acts to sanitize the nucleotide precursor pool by catalyzing the hydrolysis of 8-oxo-dGTP to 8-oxo-dGMP and inorganic pyrophosphate. Targeted disruption of the *mutM* and *mutY* genes alone or in combination results in a spontaneous G:C to T:A mutator phenotype under normal growth conditions [11,15,16], whereas disruption of *mutT* affords a spontaneous A:T to C:G mutator phenotype [15]. Interestingly, MutY actually increases the rate of A:T to C:G transversions in *mutT*$^{-}$ bacteria [17]. These data are fully consistent with the notion of a resistance system in which the individual components cooperate functionally.

Whereas the bacterial GO system is now well understood, the corresponding ^{O}G-resistance system in eukaryotic cells has only recently yielded to study. Mammalian MutT homologs have been cloned and functionally characterized [18,19], and a likely human MutY cDNA has been reported [20], but mammalian MutM homologs are unknown. Two ^{O}G-specific repair activities have been partially purified from human cells, one reportedly an endonuclease and the other a glycosylase [21]. Two ^{O}G-specific glycosylase/lyase activities have been detected in the budding yeast *Saccharomyces cerevisiae* [22,23]; one of these, yOgg1, has been cloned, overproduced and preliminarily characterized [23,24]. Interestingly, even though yOgg1 and MutM recognize the same substrates, ^{O}G and a related ring-opened formamidopyrimidine (FAPy) adduct, the yeast enzyme has no discernible sequence similarity to MutM, and the two enzymes can be inferred to use an altogether different catalytic machinery [23,25].

The amino acid sequence of yOgg1 contains a sequence element known as the Helix–hairpin–Helix, Gly/Pro-rich–Asp (HhH-G/PD) motif, which contains residues essential for substrate recognition and catalysis, and which defines yOgg1 as a member of an evolutionarily conserved base-excision DNA repair (BER) protein superfamily [23]. X-ray crystal structures of two members of the superfamily have been reported to date, *E. coli* endonuclease III [26] and AlkA [27]. Each of these represents a distinct sub-class of the BER superfamily, glycosylase/lyases (endo III) and monofunctional glycosylases (AlkA), which differ in the mechanism by which they process the substrate. Whereas monofunctional glycosylases simply displace the aberrant base using water as a nucleophile, glycosylase/lyases appear to use the ϵ-NH_2 group of a lysine side-chain as the nucleophile, thereby resulting in transient covalent attachment of the enzyme to its substrate. Importantly, the ordinarily fleeting Schiff base (imine) intermediate formed during the catalytic cycle can be intercepted by borohydride [28], resulting in conversion to an irreversibly linked glycosylase/lyase–DNA complex. This ability to 'tag' the protein with its DNA substrate allows whole cell extracts to be assayed for the presence of particular repair activities, and the activities to be isolated once identified. Borohydride trapping has been used to isolate yOgg1 [23] and a bovine endonuclease III homolog [29]. Here, we report the use of borohydride trapping to identify and characterize yOgg1 homologs in human and murine cells.

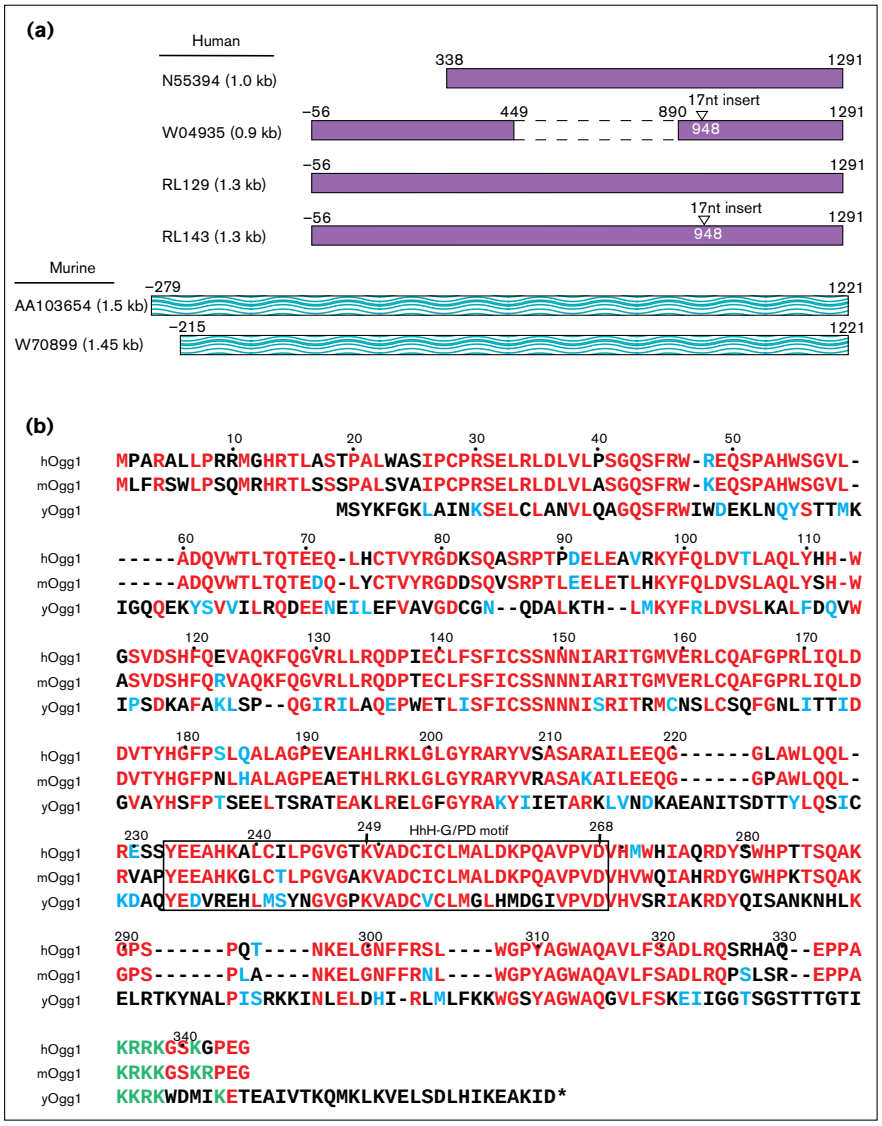
Results

Identification of cDNAs encoding mammalian Ogg1 homologs

In an attempt to identify mammalian 8-oxoguanine glycosylases, the amino acid sequence of yOgg1 [23,24] was used to query the database of expressed sequence tags (dbESTs) using the TBLASTN routine [30]. An EST from one human cDNA clone, N55934 (Fig. 1a), was identified as encoding a potential translated sequence similar to that of yOgg1. Another, overlapping human EST from clone W04935 was identified by BLASTN alignment [30] of the N55394 nucleotide sequence against the database. Complete DNA sequencing of these clones revealed them to be identical, with the exception that W04935 contained an extension at its 5′ end of ~400 bp, a large internal deletion (nucleotides 450–890), and a 17 bp insertion (after nucleotide 948). The region of N55394 deleted in W04935 encodes an open reading frame (ORF) with stretches of substantial sequence identity to yOgg1 (Fig. 1b, residues 143–279), including most or all of the enzyme active site (see below); furthermore, in the upstream direction, this yOgg1-homologous ORF continues uninterrupted to the 5′ end of N55394 and several hundred nucleotides further into the 5′ extension of W04935 before reaching an in-frame stop codon. Immediately downstream of the in-frame stop codon are two potential initiation codons, assigned as Met1 and Met11; initiation at the external site (Met1) is expected under ordinary circumstances, but the internal site (Met11) may also be used in some instances. This analysis established the coding sequence of most of the human Ogg1-homologous protein.

To determine whether the 17 nucleotide insertion in W04935 comprises part of the protein coding sequence, we generated predicted protein sequences that differed only in whether they possessed the 17 nucleotide segment, and found that inclusion led to a predicted 45.8 kDa protein

Figure 1



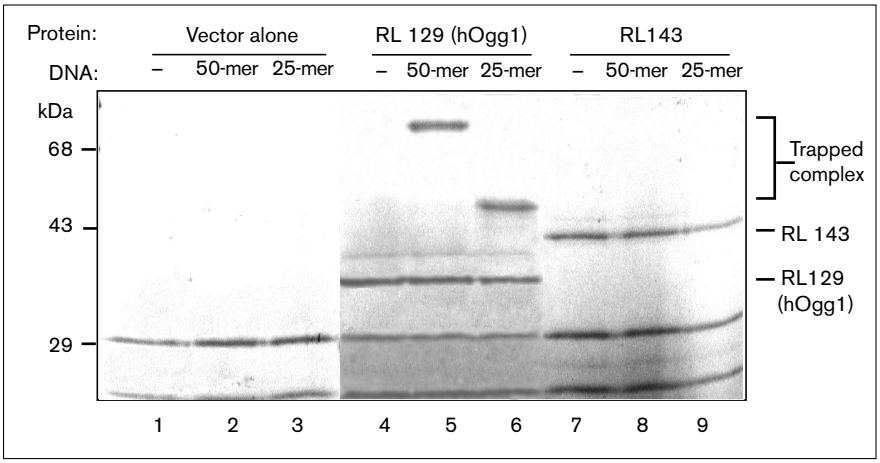
(a) Alignment of human and murine cDNA clones sequenced and characterized in this study. N55394 and W04935 are human EST cDNAs, and AA103654 and W70899 are murine EST cDNAs. RL129 (*hOGG1*) was constructed by the ligating the 5′ end of W04935 (nucleotides −56 to 366) to the 3′ end of N55394 (367–1291), and RL143 was constructed by substituting the 3′ end of RL129 by the 3′ end of W04935 (908–1291); see Materials and methods. Nucleotide numbering is according to the longest ORF for both *mOGG1* and *hOGG1*. **(b)** Amino acid alignment of the homologous human, murine and yeast Ogg1 proteins (hOgg1, mOgg1 and yOgg1, respectively). Both hOgg1 and mOgg1 are 345 amino acids in length, with predicted molecular weights of 38 780 Da and 38 820 Da, respectively. The amino acid sequences of the predicted hOgg1 and mOgg1 proteins share 84% identity, and both share 37% identity to yOgg1. All three proteins are members of the HhH-G/PD superfamily of BER enzymes [23], and the high degree of sequence conservation suggests that they share a common fold and active site configuration. According to the current mechanistic hypothesis [27], the carboxylate side chain of Asp268 acts as a general base to activate the ϵ-NH_2 group of Lys249 for nucleophilic displacement of ^{O}G by mOgg1 and hOgg1.

(not shown in Fig. 1b), whereas exclusion led to a predicted 38.8 kDa protein (shown in Fig. 1b). As shown in Figure 2 and discussed in further detail below, the shorter protein expressed from the cDNA lacking the 17 nucleotide insertion (clone RL129) exhibited 8-oxoguanine DNA glycosylase activity, but the longer protein expressed from the cDNA with the insertion (RL143) failed to exhibit such activity. Furthermore, the size of the shorter protein compared favorably with that of the 8-oxoguanine DNA glycosylase activity detected in whole-cell extracts (data not shown). We also note the presence of a nuclear localization signal (NLS) at the far carboxy-terminal end of the ORF lacking the 17 nucleotide segment (Fig. 1b, green), which is not present in the ORF containing the segment (not shown in Fig. 1b). Thus, although we cannot exclude the possibility that the longer ORF is expressed, we conclude that it does not encode a functional 8-oxoguanine DNA glycosylase. RL129 therefore contains the proper cDNA encoding the human Ogg1 homolog (hOgg1).

The hOgg1 cDNA sequence encodes a predicted 345 amino acid protein with a molecular mass of 38 780 Da and with 37% overall identity to the amino acid sequence of yOgg1. The level of identity is particularly high over a segment of the sequence known as the HhH-G/PD motif (see above) [23]. In particular, the residue corresponding to Asp268 is invariant amongst all BER proteins, whereas that corresponding to Lys249 is invariant amongst the glycosylase/lyase sub-class of BER enzymes.

Further database searching using the hOgg1 amino acid sequence led to the identification of two murine cDNA clones containing homologous ESTs, AA103654 and W70899 (Fig. 1a). Sequencing of these two clones

Figure 2



Borohydride trapping of the proteins encoded by human cDNAs RL129 and RL143 indicates that RL129 is the coding sequence of the ^{O}G DNA glycosylase/lyase hOgg1, and that the larger protein produced by RL143 lacks ^{O}G glycosylase/lyase activity. Both proteins were ^{35}S-labeled by transcription/translation *in vitro*, and were allowed to react with unlabeled 25-mer or 50-mer ^{O}G:C DNA in the presence of $NaBH_4$ so as to irreversibly trap active ^{O}G DNA glycosylase/lyase to the DNA; control reactions were performed with protein samples transcribed/translated *in vitro* using the empty vector (lanes 1–3). Reactions were analyzed by SDS-PAGE. The DNA size-dependent difference in mobility of the borohydride-trapped complex formed with the RL129-encoded protein (hOgg1) indicates that the complex contains irreversibly linked DNA (compare lanes 5,6).

revealed them to be identical over the 1.45 Kb region they share in common, and to contain a single ORF encoding a protein of 345 amino acids, with a predicted molecular mass of 38 820 Da (Fig. 1b). The amino acid sequence of the predicted mOgg1 protein has 84% identity to that of hOgg1 (Fig. 1b) and 37% identity to yOgg1.

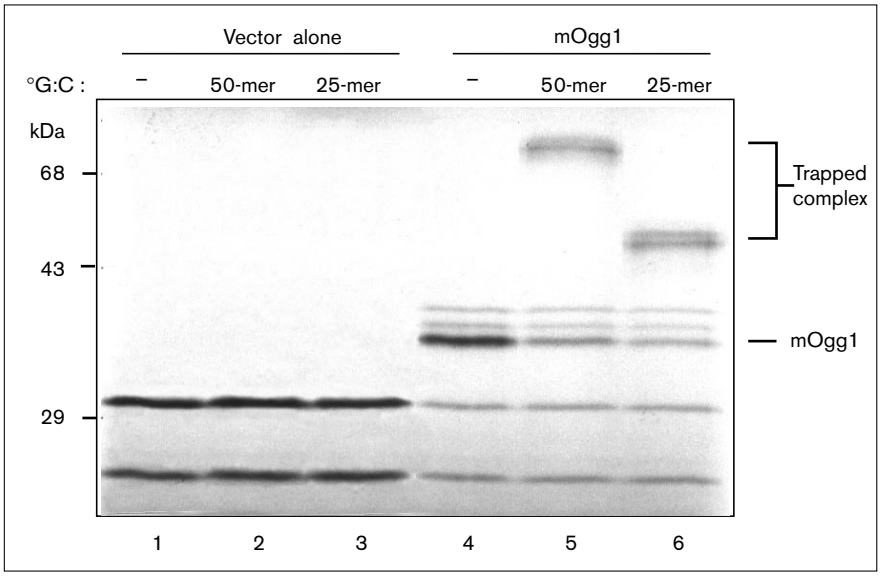
The mammalian Ogg1 homologs encode 8-oxoguanine DNA glycosylases

To characterize biochemically the mammalian Ogg1 homologs, we used borohydride trapping assays [23,28,31,32] on protein samples generated through *in vitro* transcription/translation. This assay detects the enzyme activity by intercepting an ordinarily transient Schiff base (imine) intermediate formed between the protein and its DNA substrate, and then reductively converting it to an irreversibly linked protein–DNA complex. As the formation of the Schiff base is absolutely dependent upon the catalytic activity of the enzyme, the trapping assay provides a stringent test for the function of the enzyme. Moreover, since Schiff base formation is a necessary step on the catalytic reaction pathway for both cleavage of the glycosidic bond and subsequent scission of the DNA backbone through conjugate elimination, the observation of borohydride trapping *per se* establishes that a BER enzyme operates *via* a glycosylase/lyase mechanism (the monofunctional glycosylase class of BER enzymes does not undergo borohydride trapping to any significant extent). In the case of yOgg1, the substrate-specificity in borohydride-trapping assays has been shown to correlate closely with the specificity of DNA strand cleavage [23].

In vitro transcription/translation of the human RL129 (*hOGG1*) and RL143 cDNAs yielded ^{35}S-labeled protein bands with apparent molecular weights of ~34 kDa (Fig. 2, lane 4) and ~40 kDa (lane 7), respectively; these bands were absent from control transcription/translation reactions using the empty vector (lane 1). To determine which of the human cDNA clones encodes a functional 8-oxoguanine DNA glycosylase/lyase, we carried out borohydride trapping assays using duplex 25-mer and 50-mer DNA substrates containing a single, centrally located ^{O}G:C base-pair (Fig. 2). In the case of the RL129-encoded protein (hOgg1), a single new band with retarded mobility was produced upon trapping, and only the ~34 kDa band was correspondingly diminished in intensity, consistent with the irreversible conversion of the RL129-encoded protein to a higher molecular weight species. The apparent size of the trapped species was increased by changing the DNA substrate from a 25-mer to a 50-mer (Fig. 2, lanes 6 and 5, respectively), thereby establishing conclusively that the increase in mobility of hOgg1 upon trapping is due to its being linked to DNA. As expected, trapping of unlabeled hOgg1 with ^{32}P-labeled ^{O}G:C 25-mer DNA gave a trapped band of identical mobility to that shown in Figure 2, lane 6 (data not shown). The RL143-encoded protein exhibited no detectable trapping activity with either ^{O}G substrate (Fig. 2, compare lanes 7–9). Results similar to that observed for RL129 (*hOGG1*) were obtained with the murine Ogg1 cDNA (*mOGG1*) encoded by the clones W70899 and AA103654 (Fig. 3).

We next examined the preference of mOgg1 for the base on the complementary strand opposite ^{O}G. As shown in Figure 4a, significant levels of borohydride-trapped protein–DNA complex were observed only with ^{O}G:C and not with ^{O}G:A, ^{O}G:T, or ^{O}G:G. Similar results were obtained for hOgg1, although the human protein exhibited a slightly greater ability to process ^{O}G:T than did the murine protein (data not shown). In contrast, although yOgg1 traps (and repairs) ^{O}G:A and ^{O}G:G very poorly, the yeast protein does act on ^{O}G:T with an efficiency approaching that of ^{O}G:C [23]. Thus, the mammalian Ogg1

Figure 3



Borohydride trapping of the protein encoded by murine cDNA W70899 indicates that it is the coding sequence of the ^{O}G DNA glycosylase/lyase mOgg1. The protein was ^{35}S-labeled by transcription/translation *in vitro*, and was allowed to react with unlabeled 25-mer or 50-mer ^{O}G:C DNA in the presence of $NaBH_4$ so as to irreversibly trap active ^{O}G DNA glycosylase/lyase to the DNA; control reactions were performed with protein samples transcribed/translated *in vitro* using the empty vector (lanes 1–3). Reactions were analyzed by SDS-PAGE. The DNA size-dependent difference in mobility of the borohydride-trapped complex formed with mOgg1 indicates that the complex contains irreversibly linked DNA (compare lanes 5,6). The origin of the two minor protein bands of greater size than mOgg1 in lanes 4–6 is unclear; these do not undergo borohydride trapping. The trapped mOgg1–DNA complexes appear as closely spaced doublets, perhaps reflecting some amount of DNA strand cleavage prior to trapping.

homologs exhibit exquisite selectivity for the base opposite ^{O}G, operating with great preference on ^{O}G:C only.

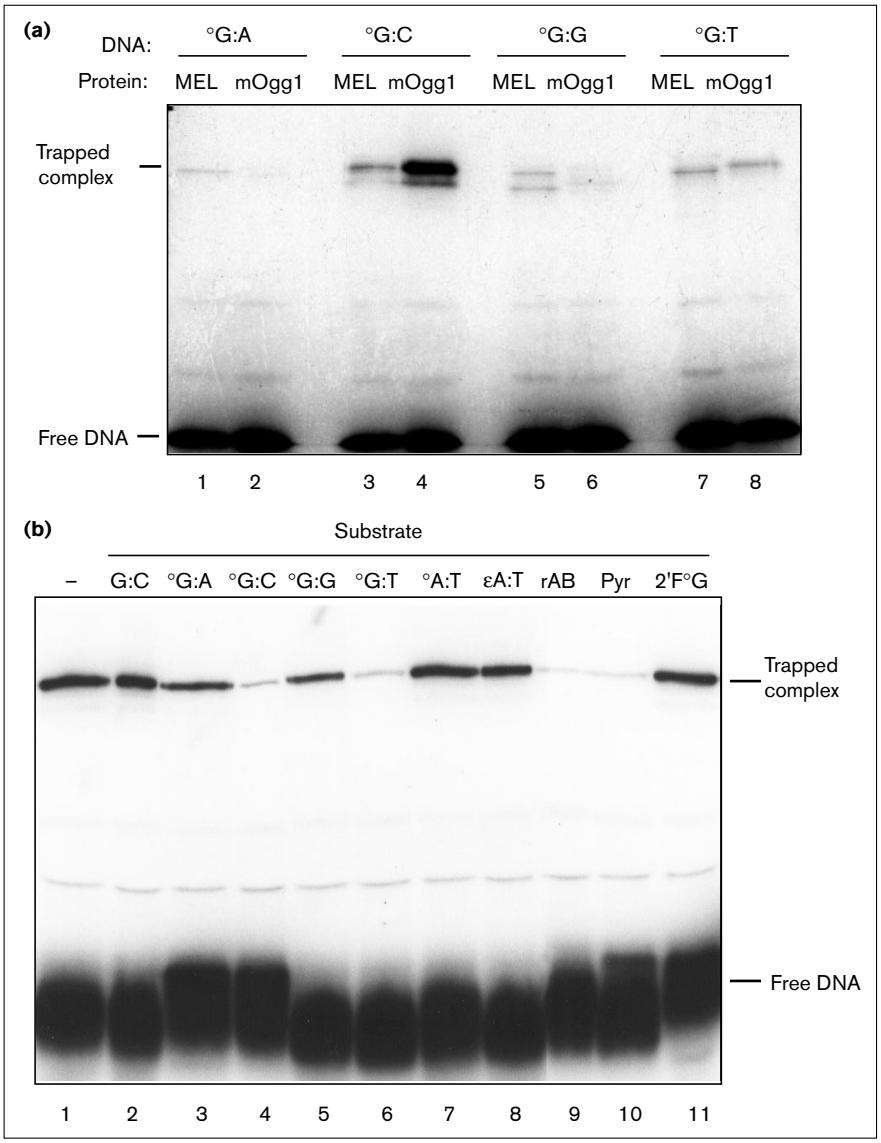
The ability of the mammalian Ogg1 homologs to recognize substrates other than ^{O}G was analyzed by competition trapping assays. A ^{32}P-radiolabeled ^{O}G:C oligonucleotide, mixed with a 100-fold excess of an unlabeled competitor oligonucleotide, was incubated with mOgg1 (unlabeled, synthesized *in vitro*), and then borohydride was added to effect the trapping reaction. Competitor oligonucleotides that interact strongly with mOgg1 reduce the intensity of the trapped protein–DNA complex. As shown in Figure 4b, ^{O}G:C was an effective competitor (lane 4), as was ^{O}G:T (lane 6). This interaction was specific, as neither the non-specific control G:C DNA (lane 2) nor oligonucleotides having ^{O}G paired to A or G (lanes 3,5) competed effectively. The reduced specificity of mOgg1 for ^{O}G:C *versus* ^{O}G:T in the competition assay, as compared with the direct trapping assay (Fig. 4b) may result from enzyme saturation by the relatively high concentrations of oligonucleotide used in the competition assay. Alternatively, it is possible that the enzyme catalyzes glycosidic bond cleavage more rapidly on the ^{O}G:C substrate than on ^{O}G:T; this kinetic substrate selection, which has been observed for other catalytic DNA-binding proteins [31], would be detected more readily in direct trapping assays than by competition. Importantly, mOgg1 appears incapable of binding the alternative oxidized purine substrate 8-oxoadenine (^{O}A:T, lane 7) or the alkylated purine N1,N^6-ethenoadenine (ϵA:T, lane 8).

Using the competition trapping assay, we also examined the ability of several synthetic inhibitors of DNA glycosylases to bind mOgg1 tightly. The reduced abasic site (rAB) binds with nanomolar affinity to yOgg1 and several other DNA glycosylases (H.M.N. and G.L.V., unpublished observations) [33], including the non-homologous bacterial 8-oxoguanine DNA glycosylase MutM [34]; it also competed very effectively for mOgg1 (lane 9). The pyrrolidine abasic site (Pyr), which binds DNA glycosylases with affinities extending to the picomolar range [33,35], also competed for mOgg1 efficiently (lane 10). The potential inhibitor 8-oxo-2′-flouro-2′-deoxyguanosine (2′F-^{O}G), which differs from the normal substrate only in having the 2′-βH replaced by F (lane 11), produced no significant level of competition; 2′F-^{O}G competes poorly against trapping by yOgg1 and MutM (H.M.N. and G.L.V., unpublished observations). Virtually identical trapping results were obtained with *in vitro* synthesized hOgg1 (data not shown). Overall, the trends observed in these competition trapping assays with the mammalian Ogg1 homologs are similar to those observed with yOgg1, fully consistent with the notion that mOgg1 and hOgg1 are both structural and functional homologs of yOgg1.

Probing the active site and mechanism of mammalian Ogg1 using site-directed mutagenesis

According to the current mechanistic understanding, the catalytic nucleophile of BER superfamily glycosylase/lyases is a conserved lysine residue within the HhH-G/PD motif, which in the case of hOgg1/mOgg1 is amino acid 249 (Fig. 1b). This residue becomes linked irreversibly to the DNA upon borohydride trapping. Consistent with this notion, mutation of the corresponding lysine residue in yOgg1 to glutamine (K241Q yOgg1) results in loss of glycosylase/lyase activity and failure to undergo borohydride trapping with ^{O}G-containing DNA, but retention of the ability to recognize DNA containing abasic site analogs

Figure 4



(a) The mOgg1 protein expressed *in vitro* and the predominant ^{O}G DNA glycosylase/lyase in murine (MEL) cells possess nearly identical specificities for the base paired opposite ^{O}G, and also possess similar apparent molecular weights, as determined by borohydride trapping with ^{32}P-radiolabeled ^{O}G:N 25-mer duplexes. The ^{O}G:N substrate preferences of mOgg1 and the MEL activity are: ^{O}G:C >> ^{O}G:T > ^{O}G:G ~ ^{O}G:A. Products of the trapping reactions were analyzed by SDS-PAGE. **(b)** mOgg1 possesses a high degree of specificity for ^{O}G:C as a substrate. Borohydride trapping of *in vitro* synthesized mOgg1 with radiolabeled ^{O}G:C 25-mer DNA in the presence of a 100-fold molar excess of the 10 indicated unlabeled DNA duplexes shows that ^{O}G:C (lane 4) and ^{O}G:T (lane 6) are the most specific substrate competitors, while DNA duplexes containing ^{O}G:A, ^{O}G:G, ^{O}A:T or 1,N^{6}εdA:T are poor specific competitors (lanes 3,5,7,8). Both the rAB site DNA and the pyrrolidine DNA are specific inhibitors (lanes 9,10), while 2'F^{O}G DNA fails to compete (lane 11).

[23]. To test this hypothesis with mOgg1, we constructed the corresponding K249Q mOgg1 protein and tested its ability to trap covalently and bind non-covalently to a ^{O}G:C oligonucleotide (Fig. 5). Unlike wild-type mOgg1 (Fig. 5a, lane 4), the K249Q mutant mOgg1 failed to undergo trapping (lane 6), but was fully capable of forming a non-covalent complex with a radiolabeled ^{O}G:C oligonucleotide (Fig. 5b, compare lanes 3,4).

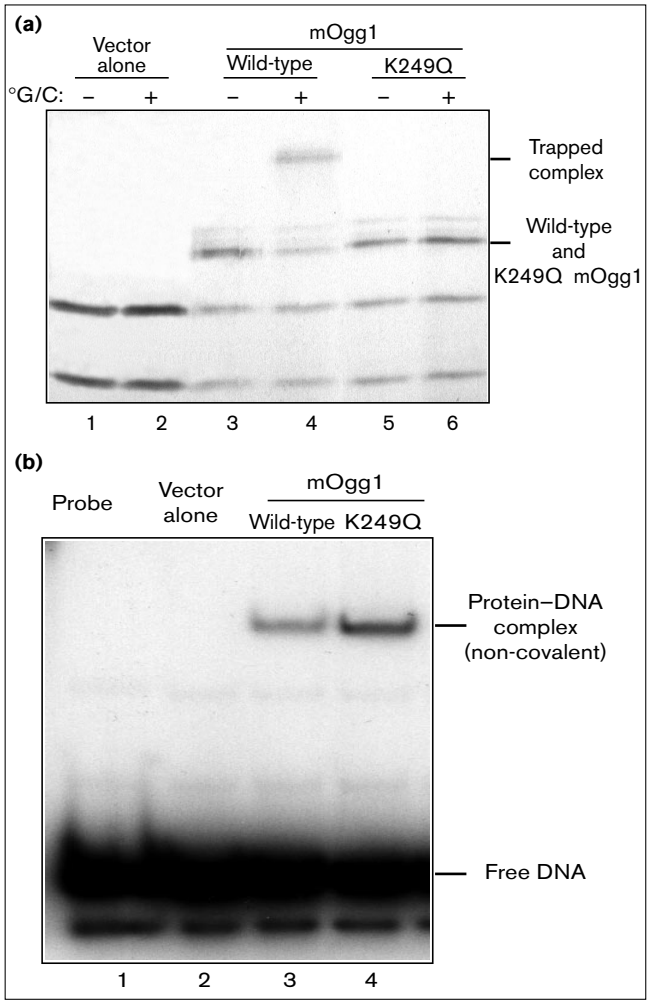
Chromosomal localization of *hOGG1*

The chromosomal localization of the *hOGG1* gene was determined by fluorescence *in situ* hybridization (FISH) analysis, using the *hOGG1* cDNA as a probe. Figure 6a shows paired sister chromatids of chromosome 3, in which a positive fluorescence signal can clearly be seen near the tip of both short arms. Figure 6c presents the statistical analysis for ten preparations analyzed by FISH, in which the dots represent probe fluorescence in a particular band. These data indicate that the *hOGG1* gene is located in the region near the junction of 3p25 and 3p26.

Discussion

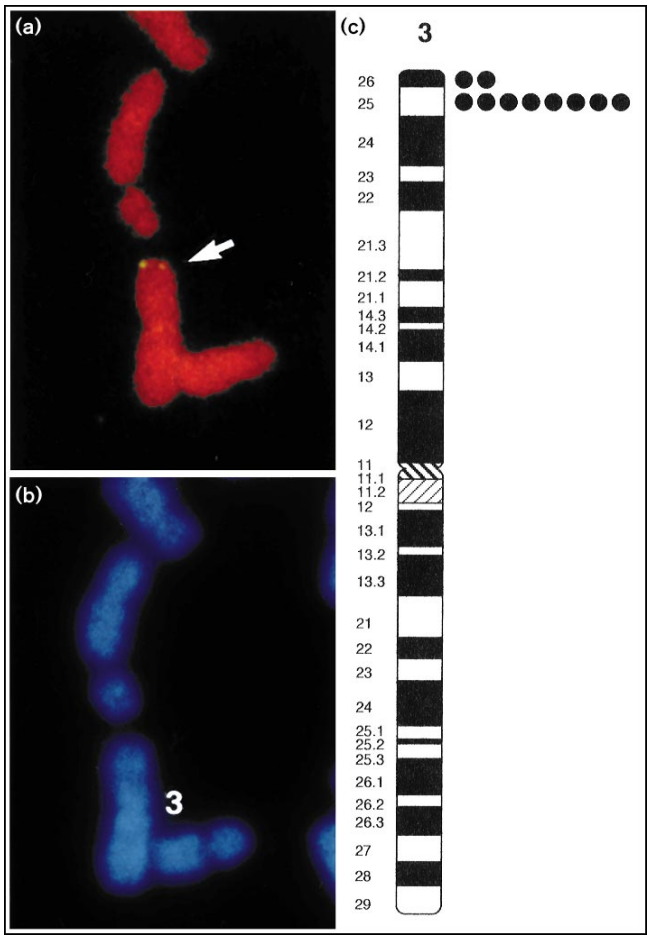
All cells that carry out aerobic metabolism suffer the oxidation of guanine residues in DNA and in the nucleotide precursor pool to 8-oxoguanine (^{O}G). The tendency of ^{O}G to mis-pair with A during replication creates a persistent threat to the integrity of the genome. Whereas the multi-enzyme (MutMYT) system responsible for purging ^{O}G from the genome and nucleotide precursor pool of *E. coli* has been extensively studied, the corresponding defense

Figure 5



(a) Borohydride trapping with the K249Q mutant of mOgg1 confirms that Lys249 is essential for ^{O}G DNA glycosylase/lyase activity. Both wild-type and K249Q mutant Ogg1 proteins (synthesized *in vitro* and labeled with ^{35}S) were allowed to react with unlabeled ^{O}G:C 25-mer DNA, and the products of the reaction were analyzed by SDS-PAGE; control reactions were performed with *in vitro* synthesized protein from the empty vector (lanes 1,2). The absence of a higher molecular weight species in lane 6 indicates that the K249Q mutant is unable to form the Schiff base intermediate required for ^{O}G DNA glycosylase/lyase activity. **(b)** Electrophoretic mobility shift assay (EMSA) of wild-type and K249Q mutant mOgg1 proteins (synthesized *in vitro*) confirms that mutation of the putative active site nucleophile, Lys249, does not affect the ability of the protein to bind ^{O}G:C DNA non-covalently. Wild-type and K249Q mutant mOgg1 proteins were incubated with ^{32}P-radiolabeled ^{O}G:C 25-mer DNA at room temperature for 10 min before analysis by 8% native acrylamide gel electrophoresis at 4°C (lanes 3 and 4, respectively). Control reactions included ^{32}P-radiolabeled ^{O}G:C 25-mer DNA alone (lane 1) or incubation of ^{32}P-radiolabeled ^{O}G:C 25-mer DNA with *in vitro* synthesized protein from the empty vector (lane 2).

Figure 6



FISH analysis of *hOGG1*. **(a)** The paired fluorescent signals obtained by specific hybridization of the *hOGG1* cDNA probe to the tips of the paired sister chromatids of chromosome 3. **(b)** The DAPI banding pattern of the chromosomes in (a). **(c)** The chromosomal localization of *hOGG1* is 3p25/26. Data from 10 photos, one of which is shown in (a), are summarized in this diagram, with each bullet representing an independent assignment.

systems in eukaryotic cells have remained poorly characterized. Here, we report the definitive identification and molecular cloning of hOgg1 and mOgg1, homologous 8-oxoguanine DNA glycosylases that presumably serve as the cornerstone of a defense mechanism against oxidative mutagenesis in human and murine cells.

The hOgg1 and mOgg1 ptoteins are highly homologous to each other, and also exhibit a significant level of sequence similarity to a recently identified 8-oxoguanine DNA glycosylase in budding yeast cells, yOgg1 [23,24]. Like yOgg1, mOgg1 and hOgg1 display a pronounced preference for the identity of the base opposite ^{O}G. All three of these eukaryotic 8-oxoguanine DNA glycosylases operate with high efficiency on ^{O}G:C, consistent with the notion that their primary function is to reverse the damage resulting from oxidation of normally base-paired G residues in DNA, prior to replication. The mammalian and yeast enzymes also are similar in that they act with poor efficiency on ^{O}G:A and ^{O}G:G. The aversion to these particular

substrates, especially ^{O}G:A, enables the enzyme to avoid catalyzing error-prone repair of misreplicated ^{O}G residues, which would thereby facilitate the overall mutagenic transversion of G:C to T:A or C:G, respectively. They differ, however, in that yOgg1 processes ^{O}G:T almost as well as ^{O}G:C, whereas hOgg1 and mOgg1 do not process ^{O}G:T efficiently. Like the yeast enzyme, recombinant hOgg1 expressed in *E. coli* preferentially cleaves ^{O}G opposite C by a β-elimination mechanism, and does not possess an efficient δ-lyase activity (H.M.N and G.L.V, unpublished observations). Also, like yOgg1, mOgg1 and hOgg1 fail to undergo borohydride-dependent trapping to any of the alternative substrates we have tested thus far, and are potently inhibited by the abasic site analog rAB [34] and the transition state mimetic Pyr [35], but are not inhibited by the fluorosugar-containing substrate 2′F-^{O}G (T. Kawate and G.L.V., unpublished observations). Taken together, these data conclusively establish that hOgg1 and mOgg1 are functional homologs of yOgg1 in mammalian cells.

The 37% sequence identity between yOgg1 and the mammalian proteins indicates that they are also structural homologs. Clearly discernible in the amino acid sequence of the Ogg1 proteins is a ~30 amino acid motif known as the HhH-G/PD motif [23,36], which comprises the active site region of a superfamily of BER enzymes conserved throughout evolution [23] (Fig. 1b). Although most members of this BER superfamily exhibit little or no sequence similarity outside the HhH-G/PD motif, two such seemingly divergent enzymes, *E. coli* endonuclease III [26] and AlkA [27], have been shown through crystallographic studies to contain a remarkably conserved core domain structure that includes the enzyme active site. A similar core domain structure has also been discovered in the 5′ deoxyribosephosphate DNA repair domain of human DNA polymerase β [37], and may also be present in a variety of non-repair-related proteins that bind DNA [36,38].

BER enzymes can be grouped into two major sub-classes: those that catalyze only the hydrolysis of the glycosidic bond (monofunctional glycosylases), and those that catalyze base-excision followed by degradation of the resulting abasic site (glycosylase/lyase). Both sub-classes are represented within the BER superfamily, and both are characterized by the presence of an invariant aspartate residue at the carboxy-terminal end of the HhH-G/PD motif (residue 268 in mOgg1/hOgg1). This aspartate residue plays a key role in catalysis, presumably serving as a general base to deprotonate the nucleophile that directly displaces the aberrant base from its sugar moiety [27]. The only other residue strongly implicated through biochemical studies as playing a key catalytic role is the lysine located at position 249 in mOgg1/hOgg1. This residue is invariant in glycosylase/lyases, but is lacking altogether in monofunctional glycosylases, suggesting that it represents a point of mechanistic bifurcation between the two classes of BER enzymes. Indeed, Lys249 and its counterpart in other enzymes is believed to function as the nucleophile that attacks the glycosidic bond and subsequently catalyzes DNA backbone cleavage through Schiff base chemistry. On the other hand, in monofunctional glycosylases, the attacking nucleophile is almost certainly an activated water molecule [27]. Here, we report that the K249Q mutant mOgg1 protein, which cannot form a Schiff base, fails to trap a ^{O}G:C-containing duplex, yet retains the ability to bind this DNA non-covalently (Fig. 5). These data establish that mOgg1 and hOgg1, like yOgg1, are 8-oxoguanine DNA glycosylase/lyases.

The identity of hOgg1 as a member of the BER superfamily provides a clue to the dysfunction we observed for the isoform containing the carboxy-terminal 17 nucleotide insertion (ORF encoded on RL143). The presence of this insert causes a shift in the reading frame, beginning at residue 316 and continuing on to the carboxyl terminus of the alternative protein. The region comprising residues ~310–320 of hOgg1/mOgg1, the carboxy-terminal end of which is directly affected by the insertion-directed frameshift, exhibits weak but notable sequence similarity to an α-helix in AlkA (α-M); amino acid side-chains at the far carboxy-terminal end of this helix project directly into the AlkA active site immediately alongside the key catalytic aspartate residue (W272 and Y273 in AlkA; refer to Figs 4 and 5c in [27]), where they are virtually certain to contact directly the substrate base. We propose that alteration of this key helix in the frameshifted RL143 cDNA causes a loss of 8-oxoguanine glycosylase activity. The present analysis furthermore suggests that residues in the 310–320 stretch of mammalian Ogg1s, all but one of which are invariant in yOgg1, contact the ^{O}G substrate base directly when bound in the enzyme active site. Whether the RL143-encoded protein has acquired a new function as the result of changes in its active site functionality, as seems possible if not probable, remains to be determined.

Curiously, the Ogg1 proteins share no detectable sequence similarity with MutM, the bacterial 8-oxoguanine DNA glycosylase protein. MutM contains no sequence that even vaguely matches the HhH-G/PD consensus motif, suggesting it contains an altogether different active site and fold from the Ogg1 proteins and all other members of the BER superfamily. This hypothesis is strongly supported by the fact that MutM uses its amino-terminal proline residue as the attacking nucleophile [25], whereas BER superfamily glycoylase/lyases use an internal lysine residue. Moreover, the substrate specificity of MutM is distinctly different from that of the Ogg1 proteins. On the basis of these considerations, we can rule out with reasonable certainty that MutM is structurally or evolutionarily related to yeast or mammalian Ogg1 proteins; it is most appropriate, therefore, to designate hOgg1 and mOgg1 according to their structural and functional similarity to yOgg1. Interestingly,

MutM may not have altogether evaded progression to the eukaryotic lineage: a homolog of MutM has been identified in the plant *Arabidopsis thaliana* (H.M.N., unpublished observations; Genbank z18192), but not biochemically characterized.

The cloning of a mammalian Ogg1 completes the triad of proteins that together protect mammalian cells from the genotoxic effects of ^{O}G. A human homolog of MutT (hMutT), a triphosphatase that cleanses the precursor pool of 8-oxo-dGTP by catalyzing its conversion to 8-oxo-dGMP and inorganic pyrophosphate, has been reported [18]. Database searching led to the identification of a cDNA having substantial sequence similarity to bacterial MutY, a monofunctional glycosylase of the BER superfamily, which catalyzes the hydrolysis of adenine residues mispaired to ^{O}G [20]. A human protein involved in a parallel BER pathway specific for oxidized pyrimidines has been cloned and characterized [39,40]. The importance of the Ogg1 proteins in preventing spontaneous oxidative mutagenesis has been demonstrated in yeast cells: targeted disruption of the *yOGG1* gene results in a 30-fold elevation of the rate of G:C to T:A transversions relative to *yOGG1*$^{+}$ cells (H.M.N and G.L.V, unpublished observations).

Chromosomal localization of *hOGG1*

Our FISH mapping results establish that *hOGG1* gene is located on the distal end of the short arm of chromosome 3 (3p25/26), a region that has aroused much attention within the cancer biology community. Nearly 100% of cells derived from small cell lung cancers — approximately one third of all lung cancers — possess allelic deletions of most of the short arm of chromosome 3 (3p) [41–45]. More than half of all non-small cell lung cancers are also characterized by the loss of heterozygosity at chromosome 3 [43–48]. Importantly, the majority of lung cancers have arisen in individuals with a history of cigarette smoking. The combustion of tobacco material results in the production of ROS, species that are particularly effective at generating ^{O}G residues in DNA; inhalation of tobacco smoke results in the delivery of high concentrations of ROS to the lung, thereby increasing the oxidative burden on these tissues. Lung cells that have lost one allele of *hOGG1* through a chromosomal deletion event may thus have a reduced capacity to counter the mutagenic burden imposed by massive influxes of exogenous ROS. This in turn may accelerate the acquisition of other mutations that contribute to the overall process of cell transformation.

Oxidative stress can also be imposed through endogenous processes in cells. Rapidly dividing cells have a high requirement for energy production through mitochondrial respiration, an aerobic process that liberates ROS into the cell [7]. Most intriguing in this regard is the finding that constitutive activation of the Ras/Raf signaling pathway causes a dramatic increase in the intracellular concentration of superoxide ($O_2^{\bullet-}$), an ^{O}G-generator that is proposed to act as a downstream mediator of this mitogenic signal in cells [1]. Deletions and other aberrations affecting 3p have been observed in a variety of tumors other than the lung, especially sporadic cancers of the kidney and breast [49], strongly suggesting the existence of one or more tumor suppressor genes on 3p. Two such candidate tumor suppressor genes have been identified, FHIT (3p14) [50] and VHL (3p25) [51]. In light of the recent discoveries that functional disruption of mismatched repair genes contributes to genetic instability and cancer [52–54], the present findings raise the prospect that *hOGG1* may function as a human tumor suppressor gene, and partial or total loss of the mammalian Ogg1 proteins may predispose cells toward oncogenic transformation.

Materials and methods

EST clones and DNA sequencing

Murine and human clones containing Ogg1 ESTs were products of the IMAGE consortium sequencing effort, and were purchased from Genome Systems. Murine and human *OGG1* sequences were sequenced on both strands with Sequenase version 2.0 (USB) as per the supplier's instructions using denatured plasmid DNA as the template.

In situ *hybridization and FISH detection*

FISH mapping was performed by See DNA Biotech. A 1.2 kb *hOGG1* cDNA probe was biotinylated with biotin-14-dATP using the BRL BioNick labeling kit (15°C, 1 h) [55]. The procedure for FISH detection was performed according to published procedures [55,56]. Briefly, slides were prepared by standard procedures with cultured human lymphocytes and baked at 55°C for 1 h. After RNase treatment, the slides were denatured in 70% formamide in 2× SSC for 2 min at 70°C, followed by dehydration with ethanol. Probes were denatured at 75 °C for 5 min in a hybridization mix consisting of 50% formamide and 10% dextran sulfate. Probes were loaded on the denatured chromosomal slides, and after overnight hybridization, the slides were washed, amplified and detected. FISH signals and the DAPI banding patterns were recorded with separate photographs, and the chromosomal localization was achieved by superimposing the FISH signals with the DAPI banded chromosomes [57]. Under the conditions used, the hybridization efficiency was approximately 68% (among 100 checked mitotic figures, 68 of them showed signals on one pair of chromosomes). The detailed chromosomal location, 3p25, was determined from the summary of 10 photos. No additional loci were observed.

In vitro *transcription/translation of mOgg1 and hOgg1 proteins*

The coding sequences of the human and murine OGG1 genes were subcloned into the pT7T3D expression vector (Pharmacia), and the T7 promoter was used to express the proteins *in vitro* using the T7 TNT Coupled Reticulocyte Lysate system (Promega) as per the supplier's instructions. The *mOGG1* coding sequence (W70899) was inserted into pT7T3D as an *Nhe* I/*Not* I fragment, with the *Nhe*I site occurring naturally in the 5′ UTR at position –8 and the *Not*I site originating from the EST carboxy-terminal primer. The *hOGG1* coding sequence (RL129) was assembled in pT7T3D by ligating the amino terminus of W04935 (an *Eco*RI/*Pfl*MI fragment) to the carboxyl terminus of N55394 (a *Pfl*MI/*Not*I fragment); the *Pfl*MI site occurs naturally at position 366 of the coding sequence, while the *Eco*RI and *Not*I sites originate from the EST primers. The larger, alternately-spliced, *hOGG1* coding sequence (RL143) was obtained by substituting the carboxy-terminal *Bsp*EI/*Not*I fragment of RL129 with the *Bsp*EI/*Not*I fragment of W04935; the *Bsp*EI site occurs naturally at position 908 of the coding sequence.

Oligonucleotide duplex substrates/competitors:

All oligonucleotides were synthesized by β-cyanoethyl solid-phase chemistry on a 1 μmole scale using an ABI model 392 DNA synthesizer, and

were purified by 20% denaturing PAGE. Full-length DNA was eluted from the acrylamide gel by crushing and soaking in 1 M triethylammonium bicarbonate, pH 7, followed by C_{18} Sep-Pak concentration/desalting. Concentrations were determined by A_{260} quantitation, and radiolabeling reactions were performed with T4 polynucleotide kinase (New England Biolabs) and γ-[^{32}P]ATP. Duplexes were prepared by annealing with a 50-fold molar excess of the complementary strand in 1× TE/100 mM NaCl. The radiolabeled strand of each duplex is the top strand as shown below.

1	8-oxodG:dA	5′ GGA TAG TGT CCA ^{o}G GTT ACT CGA AGC 3′
		3′ CCT ATC ACA GGT A CAA TGA GCT TCG 5′
2	8-oxodG:dC	5′ GGA TAG TGT CCA ^{o}G GTT ACT CGA AGC 3′
	25-mer	3′ CCT ATC ACA GGT C CAA TGA GCT TCG 5′
3	8-oxodG:dG	5′ GGA TAG TGT CCA ^{o}G GTT ACT CGA AGC 3′
		3′ CCT ATC ACA GGT G CAA TGA GCT TCG 5′
4	8-oxodG:T	5′ GGA TAG TGT CCA ^{o}G GTT ACT CGA AGC 3′
		3′ CCT ATC ACA GGT T CAA TGA GCT TCG 5′
5	8-oxodA:T	5′ GGA TAG TGT CCA ^{o}A GTT ACT CGA AGC 3′
		3′ CCT ATC ACA GGT T CAA TGA GCT TCG 5′
6	1,N^{6}ϵdA:T	5′ GGA TAG TGT CCA ^{ϵ}A GTT ACT CGA AGC 3′
		3′ CCT ATC ACA GGT T CAA TGA GCT TCG 5′
7	dG:dC	5′ GTG AAC CTG AGC G TAG CTC AGT AAC 3′
		3′ CAC TTG GAC TCG C ATC GAG TCA TTG 5′
8	rAB:dC	5′ GTG AAC CTG AGC rAB TAG CTC AGT AAC 3′
		3′ CAC TTG GAC TCG C ATC GAG TCA TTG 5′
9	pyrrolidine:	5′ GGA TAG TGT CCA Pyr GTT ACT CGA AGC 3′
		3′ CCT ATC ACA GGT T CAA TGA GCT TCG 5′
10	8-oxodG:dC	5′ GGC AAG TCT GAT GGA TAG TGT CCA ^{o}G 3′->
	50-mer	3′ CCG TTC AGA CTA CCT ATC ACA GGT C 5'->
		-> 5′ GTT ACT CGA AGC AGT TCG AAC TGG 3′
		-> 3′ CAA TGA GCT TCG TCA AGC TTG ACC 5′

Phosphoramidites of 8-oxo-dG [58], 8-oxo-dA [59], 1,N^{6}ϵdA [60], and Pyr [35] monomers were synthesized according to published procedures. The reduced abasic site (rAB) phosphoramidite was synthesized by an unpublished procedure (H.M.N., J-Y. Ortholand and G.L.V., unpublished results).

DNA trapping assays with $NaBH_4$

DNA trapping reactions were performed in a 25 μl volume in siliconized microcentrifuge tubes at 37°C. The final concentrations in all reactions were: 50 mM Tris, pH 7.5, 2 mM EDTA, 50 mM $NaBH_4$, and 1 nM ^{32}P-radiolabeled ^{O}G:N 25-mer oligonucleotide duplexes **1**–**4**. 2 μl of *in vitro* transcribed/translated mOGG1 or hOGG1 was added to the trapping reaction, and the reaction was incubated for 30 min (37°C) before termination with 1× SDS loading dye at 95°C for 5 min. Mammalian cell lysate reactions contained 3 μl of cell extract and were incubated for 30 min (37°C) before termination in 1× SDS loading buffer at 95°C for 5 min. Competition trapping reactions were performed by a 10 min, room temperature, preincubation of the *in vitro* transcribed/translated mOgg1 protein with a 100-fold molar excess of unlabeled duplex competitor, followed by the addition of the ^{32}P-radiolabeled ^{O}G:C duplex **2** and $NaBH_4$. The reactions were continued for 15 min at 37°C before termination in 1× SDS loading dye at 95°C for 5 min. 15 μl of each terminated DNA trapping reaction was loaded on 0.75 mm, 12% 37.5:1 acrylamide:bis-acrylamide, 0.1% SDS, SDS-PAGE gels. Gels were run at a constant 25 mA for 3 h, and were fixed for 15 min at room temperature in 10% methanol/10% acetic acid. Fixed gels were dried for 1.5 h and exposed to film (Kodak BioMax MR) and to a phosphorimaging plate (Fuji BAS1000). The gel figures presented herein were initially obtained as autoradiograms, scanned using a LACIE Silverscanner III digital scanner, then imported into Adobe Photoshop 3.0 for captioning. Relative trapping efficiencies were quantified by phosphorimager quantitation. ^{14}C molecular weight markers were from Life Technologies.

Electrophoretic mobility shift assays (EMSA)

Gel shift assays were performed with 3 μl of *in vitro* transcribed mOGG1 protein and 1 nM ^{32}P-radiolabeled ^{O}G:C oligonucleotide duplex **2** in a 20 μl binding reaction containing 50 mM tris, pH 8, 2 mM EDTA, 6% glycerol and 0.05 mg ml^{-1} plasmid DNA. Binding reactions were incubated at room temperature for 10 min before loading on a pre-run 8% 0.5× TBE acrylamide gel (37.5:1, Protogel, National Diagnostics). Gels were run at 250 volts for 2.5 h at 4°C, then dried and exposed to film (Kodak BioMax MR) and to a phosphorimaging plate (Fuji BAS1000). The gel figures presented herein were initially obtained as autoradiograms, scanned using a LACIE Silverscanner III digital scanner, then imported into Adobe Photoshop 3.0 for captioning.

Site-directed mutagenesis of mOGG1

Lysine 249 of the mOgg1 protein was mutated to glutamine 249 by replacing the wild-type *Nar*I/*Bsp*EI fragment of mOGG1 with a PCR-generated cassette bearing the AAG→CAG sequence change. The sequences of the two PCR primers are shown below, with the restriction sites in bold, and the K249Q mutation underlined:

*Nar*I: 5′-TG CCC GGG GTG **GGC GCC** CAG GTG GC-3′
*Bsp*EI: 5′-CCA CAG AT**T CCG GA**A AAA GTT TCC-3′

The mutagenesis cassette was PCR amplified with Deep Vent DNA polymerase (New England Biolabs), digested with *Nar*I and *Bsp*EI, and ligated into the *Nar*I/*Bsp*EI-digested mOGG1/pT7T3D *in vitro* expression construct. The presence of the K249Q mutation and the absence of any other undesired mutations was confirmed by DNA sequencing.

Acknowledgements

This research was supported by a grant from the National Institute of General Medical Sciences (GM51330). We thank John Minna and Eric Biesterveld for discussions and provision of human cell lines, Phil Familetti and Donna Stremlo of Hoffmann-LaRoche for MEL cells, Yong-Ping Wang for providing murine organs, and Pei Zhou, Steve Bruner and Li Jing Sun for assistance with graphics.

References

1. Irani K, Xia Y, Zweier JL, Solott SJ, Der CJ, Fearon ER, *et al.*: **Mitogenic signaling mediated by oxidants in Ras-transformed fibroblasts**. *Science* 1997, **275**:1649–1652.
2. Sundaresan M, Yu X-Z, Ferrans VJ, Irani K, Finkel T: **Requirement for generation of H_2O_2 for platelet-derived growth factor signal transduction**. *Science* 1996, **270**:296–299.
3. Ohba M, Shibanuma M, Kuroki T, Nose K: **Production of hydrogen peroxide by transforming growth factor-β1 and its involvement in induction of egr-1 in mouse osteoblast cells**. *J Cell Biol* 1994, **126:**1079–1088.
4. Lo YYC, Wong JMS, Cruz TF: **Reactive oxygen species mediate cytokine gene activation of c-Jun NH_2-terminal kinases**. *J Biol Chem* 1996, **271:**15703–15707.
5. Los M, Schenk H, Hexel K, Baeuerle PA, Droge W, Schultze-Osthoff K: **IL-2 gene expression and NF-κB activation through CD28 requires reactive oxygen production by 5-lipoxygenase**. *EMBO J* 1995, **14:**3731–3740.
6. Schmidt KN, Amstad P, Cerutti P, Baeuerle PA: **Identification of hydrogen peroxide as the relevant messenger inthe activation pathway of transcriptionfactos NF-κB**. *Adv Exp Med Biol* 1996, **387:**63–68.
7. Friedberg EC, Walker GC, Siede W: *DNA Repair and Mutagenesis*, Washington D.C.: ASM Press; 1995.
8. Ames BN, Shigenaga MK, Hagen TM: **Oxidants, antioxidants and the degenerative diseases of ageing**. *Proc Natl Acad Sci USA* 1993, **90:**7915–7922.
9. Kasai H, Nishimura S: **Formation of 8-hydroxydeoxyguanosine in DNA by oxygen radicals and its biological significance**. In *Oxidative Stress: Oxidants and Antioxidants*. Edited by Sies H. London: Academic Press; 1991:99–116.
10. Grollman AP, Moriya M: **Mutagenesis by 8-oxoguanine: An enemy within**. *Trends Genet.* 1993, **9:**246–249.
11. Michaels ML, Cruz C, Grollman AP, Miller JH: **Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA**. *Proc Natl Acad Sci USA* 1992, **89:**7022–7025.
12. Michaels ML, Tchou J, Grollman AP, Miller JH: **A repair system for 8-oxo-7,8-dihydrodeoxyguanine**. *Biochemistry* 1992, **31:**10964–10968.
13. Michaels ML, Miller JH: **The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine)**. *J Bacteriol* 1992, **174:**6321–6325.
14. Maki H, Sekiguchi M: **MutT protein specifically hydrolyzes a potent mutagenic substrate for DNA synthesis**. *Nature* 1992, **355:**273–275.
15. Tajiri T, Maki H, Sekiguchi M: **Functional cooperation of MutT, MutM and MutY proteins in preventing mutations caused by spontaneous oxidation of guanine nucleotide in *Escherichia coli***. *Mutat Res DNA Repair* 1995, **336:**257–267.

16. Duwat P, de Oliveira R, Ehrlich SD, Boiteux S: **Repair of oxidative DNA damage in gram-positive bacteria: the *Lactococcus lactis* Fpg protein**. *Microbiology* 1995, **141:**411–417.
17. Vidmar JJ, Cupples CG: **MutY repair is mutagenic in *mutT*$^+$ strains of *Escherichia coli***. *Can J Microbiol* 1993, **39:**892-894.
18. Sakumi K, Furuichi M, Tsuzuki T, Kakuma T, Kawabata S, Maki H, Sekiguchi M: **Cloning and expression of cDNA for a human enzyme that hydrolyzes 8-oxo-dGTP, a mutagenic substrate for DNA synthesis**. *J Biol Chem* 1993, **268:**23524–23530.
19. Cai JP, Kakuma T, Sekiguchi M: **cDNA and genomic sequences for rat 8-oxo-dGTPase that prevents occurrence of spontaneous mutations due to oxidation of guanine nucleotides**. *Carcinogenesis* 1995, **16:**2343–2350.
20. Slupska MM, Baikalov C, Luther WM, Chiang JH, Wei YF, Miller JH: **Cloning and sequencing a human homolog (hMYH) of the *Escherichia coli mutY* gene whose function is required for the repair of oxidative DNA damage**. *J Bacteriol* 1996, **178:**3885–3892.
21. Bessho T, Tano K, Kasai H, Ohtsuka E, Nishimura S: **Evidence for two DNA repair enzymes for 8-hydroxyguanine (7,8-dihydro-8-oxoguanine) in human cells**. *J Biol Chem* 1993, **268:**19416–19421.
22. de Oliveira R, van der Kamp PA, Thomas D, Geiger A, Nehls P, Boiteux S: **Formamidine DNA glycosylase in the yeast *Saccharomyces cerevisiae***. *Nucleic Acids Res* 1994, **22:**3760–3764.
23. Nash HM, Bruner SD, Schärer OD, Kawate T, Addona TA, Sponner E, Lane WS, Verdine GL: **Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily**. *Curr Biol* 1996, **6:**968–980.
24. van der Kemp PA, Thomas D, Barbey R, de Oliveira R, Boiteux S: **Cloning and expression in *Escherichia coli* of the *OGG1* gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro,8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyramidine**. *Proc Natl Acad Sci USA* 1996, **93:**5197–5202.
25. Zharkov DO, Rieger RA, Iden CR, Grollman AP: **NH_2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein**. *J Biol Chem* 1997, **272:**5335–5341.
26. Kuo C-F, McRee DE, Fisher CL, O'Handley SF, Cunningham RP, Tainer JA: **Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III**. *Science* 1992, **258:**434–440.
27. Labahn J, Scharer OD, Long A, Ezaz-Nikpay K, Verdine GL, Ellenberger TE: **Structural basis for the excision repair of alkylation-damaged DNA**. *Cell* 1996, **86:**321–329.
28. Dodson ML, Schrock RD,III, Lloyd RS: **Evidence for an imino intermediate in the T4 endonuclease V reaction**. *Biochemistry* 1993, **32:**8284–8290.
29. Hilbert TP, Boorstein RJ, Kung HC, Bolton PH, Xing D, Cunningham RP, Teebor GW: **Purification of a mammalian homolog of *Escherichia coli* endonuclease III: identification of a bovine pyrimidine hydrate-thymine glycol DNA-glycosylase/AP Lyase by irreversible cross linking to a thymine glycol-containing oligodeoxynucleotide**. *Biochemistry* 1996, **35:**2505–2511.
30. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *J Mol Biol* 1990, **215:**403–410.
31. Sun B, Latham KA, Dodson ML, Lloyd RS: **Studies on the catalytic mechanism of five DNA glycosylases. Probing for enzyme–DNA imino intermediates**. *J Biol Chem* 1995, **270:**19501–19508.
32. Tchou J, Grollman AP: **The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site**. *J Biol Chem* 1995, **270:**11671–11677.
33. McCullough AK, Schärer OD, Verdine GL, Lloyd RS: **Structural determinants for specific recognition by T4 endonuclease V**. *J Biol Chem* 1996, **271:**32147–32152.
34. Castaing B, Boiteux S, Zelwer C: **DNA containing a chemically reduced apurinic site is a high affinity ligand for the *E. coli* formamidopyrimidine-DNA glycosylase**. *Nucleic Acids Res* 1992, **20**:389–394.
35. Schärer OD, Ortholand J-Y, Ganesan A, Ezaz-Nikpay K, Verdine GL: **Specific binding of the DNA repair enzyme AlkA to a pyrrolidine-based inhibitor**. *J Am Chem Soc* 1995, **117:**6623–6624.
36. Thayer MM, Ahern H, Xing DX, Cunningham RP, Tainer JA: **Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure**. *EMBO J* 1995, **14:**4108–4120.
37. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J: **Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP**. *Science* 1994, **264:**1891–1903.
38. Doherty AJ, Serpell LC, Ponting CP: **The Helix-hairpin-Helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA**. *Nucleic Acids Res* 1996, **24:**2488–2497.
39. Aspinwall R, Rothwell DG, Roldan-Arjona T, Anselmino C, Ward C, Cheadle JP, *et al.*: **Cloning and characterization of a functional human homolog of *Escherichia coli* endonclease III**. *Proc Natl Acad Sci USA* 1997, **94:**109–114.
40. Hilbert TP, Chaung W, Boorstein RJ, Cunningham RP, Teebor GW: **Cloning and expression of the cDNA encoding the human homologue of the DNA repair enzyme, *Escherichia coli* endonuclease III**. *J Biol Chem* 1997, **272:**6733–6740.
41. Kok H, Osinga J, Carrit B, Davis MB, van der Veen AY, Landsvater RM, *et al.*: **Deletion of a DNA sequence at the chromosomal region 3p21 in all major types of lung cancer**. *Nature* 1987, **330:**578–581.
42. Naylor SL, Johnson BE, Minna JD, Sakaguchi AY: **Loss of heterozygosity of chromosome 3 markers in small-cell lung cancer**. *Nature* 1987, **329:**451–454.
43. Rabbitts P, Bergh J, Douglas J, Collins F, Waters J: **A submicroscopic homozygous deletion at the D3S3 locus in a cell line isolated from a small cell lung carcinoma**. *Genes Chromosomes Cancer* 1990, **2:**231–238.
44. Yokota A, Wada M, Shimosato Y, Terada M, Sugimura T: **Loss of heterozygosity on chromosomes 3, 13, 17 in small cell carcinoma and on chromosome 3 in adenocarcinoma of the lung**. *Proc Natl Acad Sci USA* 1987, **84:**9252–9256.
45. Brauch H, Johnson B, Hovis J, Yano T, Gadzar A, Pettengil O, *et al.*: **Molecular analysis of the short arm of chromosome 3 in small-cell and non-small cell carcinoma of the lung**. *N Engl J Med* 1987, **317:**1109–1113.
46. Hibi K, Takahashi T, Yamakawa K, Ueda R, Sekido Y, Ariyoshi Y, *et al.*: **Three distinct regions involved in 3p deletions in human lung cancer**. *Oncogene* 1992, **7:**445–449.
47. Horio Y, Takahashi T, Kuroishi T, Hibi K, Suyama M, Niimi T, *et al.*: **Prognostic significance of p53 mutations and 3p deletions in primary resected non-small cell lung cancer**. *Cancer Res* 1993, **53:**1–4.
48. Yokohama S, Yamakawa K, Tsuchiya E, Murata M, Sakiyama S, Nakamura Y: **Deletion mapping on the srort arm of chromosome 3 in squamous cell carcinoma and adenocarcinoma of the lung**. *Cancer Res* 1992, **52**:873–877.
49. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS: **Loss of heterozygosity in normal tissue adjacent to breast carcinomas**. *Science* 1997, **274:**2057–2059.
50. Sozzi G, Veronese ML, Negrini M, Baffa R, Cotticelli MG, Inoue H, *et al.*: **The *FHIT* gene is abnormal in lung cancer**. *Cell* 1996, **85:**17–26.
51. Latif F, Tory K, Gnarra J, Yao M, Duh FM, Orcutt ML, *et al.*: **Identification of the von Hippel-Lindau disease tumor suppressor gene**. *Science* 1993, **260:**1317–1320.
52. Papadopoulos N, Nicolaides NC, Wei Y-F, Ruben SM, Carter KC, Vogelstein B: **Mutation of a *mutL* homolog in hereditary colon cancer**. *Science* 1994, **363:**1625–1629.
53. Leach FS, Nicloaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, *et al.*: **Mutations of a *mutS* homolog in hereditary nonpolyposis colon cancer**. *Cell* 1993, **75:**1215–1225.
54. Papadopoulos N, Nicolaides NC, Liu B, Parsons R, Lengauer C, Palombo F, *et al.*: **Mutations of *GTBP* in genetically unstable cells**. *Science* 1995, **268:**1915–1917.
55. Heng HHQ, Squire J, Tsui L-C: **High resolution mapping of mammalian genes by *in situ* hybridization to free chromatin**. *Proc Natl Acad Sci USA* 1992, **89:**9509–9513.
56. Heng HHQ, Tsui L-C: **Modes of DAPI banding and simultaneous in situ hybridization**. *Chromosoma* 1993, **102:**325–332.
57. Heng HHQ, Tsui L-C: **FISH detection on DAPI banded chromosomes**. In *Methods in Molecular Biology: In situ Hybridization Protocols* Edited by Choo KHA. Clifton, NJ: Humana Press; 1994:35–49.
58. Roelen HCPF, Saris CP, Brugghe HF, van den Elst H, Westra JG, van der Marel GA, van Boom JH: **Solid-phase synthesis of DNA fragments containing the modified base 7-hydro-8-oxo-2′-deoxyguanosine**. *Nucleic Acids Res* 1991, **19:**4361–4369.
59. Guy A, Duplaa A-M, Harel P, Téoule R: **Synthesis and characterization of DNA fragments bearing an adenine radiation product: 7,8-dihydroadenin-8-one**. *Helvetica Chimica Acta* 1988, **71:**1566–1571.
60. Srivastava SC, Raza SK, Misra R: **1,N_6-etheno deoxy and ribo adenosine and 3,N_4-etheno deoxy and ribo cytidine phosphoramidites. Strongly fluorescent structures for selective introduction in defined sequence DNA and RNA molecules**. *Nucleic Acids Res* 1994, **22:**1296–1304.

Because *Current 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/cub – for further information, see the explanation on the contents page.