# Base-excision repair of oxidative DNA damage

Sheila S. David[1], Valerie L. O'Shea[2] & Sucharita Kundu[2]

**Maintaining the chemical integrity of DNA in the face of assault by oxidizing agents is a constant challenge for living organisms. Base-excision repair has an important role in preventing mutations associated with a common product of oxidative damage to DNA, 8-oxoguanine. Recent structural studies have shown that 8-oxoguanine DNA glycosylases use an intricate series of steps to locate and excise 8-oxoguanine lesions efficiently against a high background of undamaged bases. The importance of preventing mutations associated with 8-oxoguanine is shown by a direct association between defects in the DNA glycosylase MUTYH and colorectal cancer. The properties of other guanine oxidation products and the associated DNA glycosylases that remove them are now also being revealed.**

The specific pairing of DNA bases — A with T and G with C — is crucial for preserving the information content of the genome. However, the structural properties of DNA bases, and therefore their pairing properties, are often modified by reactions with environmental toxins (for example, ultraviolet radiation and cigarette smoke) and endogenous metabolic products or by-products (for example, *S*-adenosylmethionine and hydroxyl radicals)[1–3]. Over three decades ago, armed with the knowledge that cytosine was readily deaminated to yield uracil, T. Lindahl searched for bacterial enzymes that mediated the repair of uracil and discovered uracil DNA glycosylase (as discussed in ref. 4). This enzyme catalyses the removal of uracil from DNA, the first step of the base-excision repair (BER) pathway[5–7]. We now know that there are many DNA glycosylases that specifically recognize and excise abnormal DNA bases. After the base has been removed, other enzymes are recruited to excise the remaining sugar fragments and to reinstall an undamaged nucleotide, thereby repairing the damaged base site[8].

Recent studies have propelled the DNA-damage-specific glycosylases into the limelight. Structural, biophysical and biochemical approaches have provided exquisite insight into the remarkable process by which these enzymes locate damaged bases in the context of a large excess of normal DNA. Genetic studies have revealed direct correlations between defects in the repair of damaged bases and human disease. And sophisticated analytical and chemical approaches have identified new base lesions and their associated glycosylases. Here, we highlight the BER of guanine oxidation products (Fig. 1) and illustrate notable features of the chemistry, structural biology and clinical aspects of these repair pathways.

## The special problem of 8-oxoguanine

The production of reactive oxygen species (ROS), such as hydrogen peroxide, superoxide and hydroxyl radicals, has been linked to the initiation and progression of cancer[9]. ROS are by-products of cellular respiration and components of inflammatory responses. In addition, ROS can be produced as a consequence of ionizing radiation or environmental exposure to transition metals, chemical oxidants and free radicals. Normal cells have enzymatic and non-enzymatic mechanisms to counter the production of ROS. But aberrantly functioning cells are often in a state of 'oxidative stress', in which the balance between oxidants and antioxidants has been disrupted, resulting in increased levels of cellular damage.

DNA bases are particularly susceptible to oxidation mediated by ROS[10]. The low redox potential of guanine makes this base particularly vulnerable and leads to a plethora of oxidized guanine products[10]. The most thoroughly examined guanine oxidation product is 7,8-dihydro-8-oxoguanine (also known as 8-oxoguanine; 8-oxoG)[10,11]; the presence of 8-oxoG is often used as a cellular biomarker to indicate the extent of oxidative stress[9]. The 8-oxoG lesion is particularly deleterious because of its two-atom change from G (the introduction of an oxo group on the carbon at position 8 (C8) and an hydrogen atom on the nitrogen at position 7 (N7)) and because of its ability to mimic T functionally in the *syn* conformation, forming a stable 8-oxoG(*syn*)•A(*anti*) base pair (Fig. 2a). In contrast to many other types of DNA damage, these structural features allow efficient, although inaccurate, bypass of 8-oxoG by replicative DNA polymerases[12]. A series of X-ray crystal structures of a DNA polymerase I fragment (BF) from *Bacillus stearothermophilus* (also known as *Geobacillus stearothermophilus*) illustrated the consequences of 8-oxoG presence in the template before insertion of dAMP or dCMP, after insertion, and after extension beyond 8-oxoG•A and 8-oxoG•C base pairs[13]. A comparison of these structures with those containing DNA mismatches showed that inaccurate bypass of 8-oxoG results from the ability of the 8-oxoG(*syn*)•A(*anti*) base pair to mimic a normal base pair when processed by BF. By contrast, the formation of an 8-oxoG(*anti*)•C(*anti*) base pair during replication induces template and polymerase distortions similar to those seen when the active site of DNA polymerase encounters mismatches. The structural data provide a molecular rationale for both the more facile insertion of A than C opposite 8-oxoG and the lack of error detection by replicative DNA polymerases after this insertion.

Failure to remove 8-oxoG before replication results in G-to-T transversion mutations[10]. Repair pathways in all organisms are tailored to deal with the special features of 8-oxoG and to mitigate its high mutagenic potential. The MutT, MutM (also known as Fpg) and MutY enzymes in bacteria[14] and the corresponding human enzymes[7], MTH1, OGG1 and MUTYH (formerly hMYH), constitute the 8-oxoG repair pathway (also known as the GO repair pathway) (Fig. 1). MutT (or MTH1) hydrolyses 8-oxo-dGTP, removing it from the nucleotide

[1]Department of Chemistry, University of California at Davis, 1 Shields Avenue, Davis, California 95616, USA. [2]Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, USA.

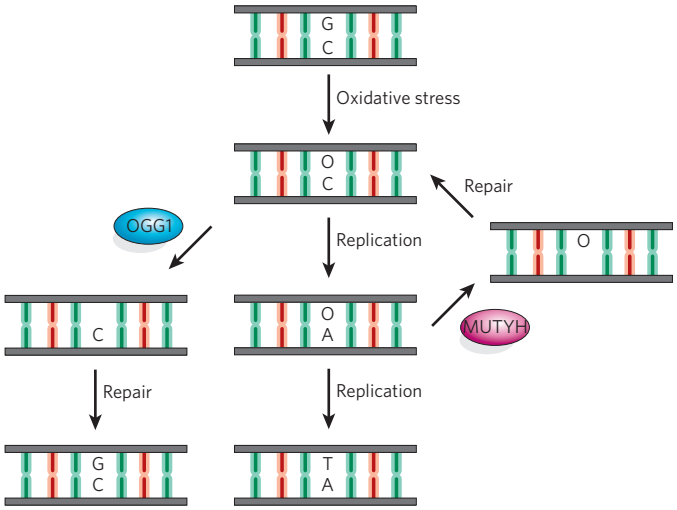


**Figure 1 | Short-patch BER pathway for 8-oxoG.** The presence of 8-oxoG (O) in DNA causes G-to-T transversions, as illustrated in the central pathway. The human DNA glycosylases OGG1 and MUTYH are involved in excision of bases from the DNA. OGG1 removes 8-oxoG from 8-oxoG•C base pairs, and MUTYH removes A from 8-oxoG•A base pairs, both generating AP sites in the DNA. The corresponding bacterial enzymes are MutM and MutY. The steps labelled 'repair' summarize the actions of AP endonuclease, deoxyribophosphate lyase, DNA polymerase and DNA ligase. Notably, OGG1 also has AP lyase activity. MutT and its human homologue MTH1 (not shown) have an important role in preventing the incorporation of 8-oxoG, through hydrolysis of free 8-oxo-dGTP.

pool so that it cannot be incorporated into DNA by polymerases. The BER glycosylase MutM (OGG1) excises 8-oxoG from the 8-oxoG•C base pair so that subsequent processing by other enzymes in the BER pathway can restore the G•C base pair. However, if this does not occur, and replication takes place, then the DNA glycosylase MutY (MUTYH) intercepts the resultant 8-oxoG•A base pair and removes the inappropriate A. Subsequent processing of the apurinic/apyrimidinic (AP) site, followed by replication by a repair polymerase, provides an opportunity to create an 8-oxoG•C substrate for the 8-oxoG DNA glycosylase, MutM (OGG1). Notably, structural studies of the eukaryotic repair polymerase-β with 8-oxoG-containing templates have shown that the active site of this polymerase more easily accommodates 8-oxoG•C base pairs than 8-oxoG•A mismatches[15]. This specificity is exactly the opposite of the results obtained with the bacterial polymerase BF. Moreover, this structural information explains the higher propensity of polymerase-β to insert dCTP over dATP opposite 8-oxoG, a property that is required of this repair polymerase to prevent futile cycles of repair and replication[15].

## Seeking and removing 8-oxoG lesions

### Recognition of 8-oxoG lesions by OGG1

In eukaryotic nuclear DNA, it has been estimated that there are several 8-oxoG bases per $10^6$ guanine bases[16]. This low frequency and the similarity in structure of 8-oxoG and G make it a formidable challenge for 8-oxoG DNA glycosylases to find 8-oxoG lesions among a huge number of undamaged bases. Other BER glycosylases, such as uracil DNA glycosylase and MutY (MUTYH), also need to distinguish relatively minor perturbations from the normal DNA bases or base pairs. Structural and biophysical studies of many DNA glycosylases bound to their relevant substrate duplexes have revealed common themes in lesion recognition, including enzyme-initiated DNA bending, disruption of the base pair, extrusion of the damaged nucleotide from the interior of the DNA helix, and placement of the target base into a base-specific enzyme pocket[17–21]. This enzyme-driven process is often referred to as 'base flipping' but might be described more accurately as 'nucleotide flipping', because the entire nucleotide is rotated out of the helix to accommodate the base in the base-specific pocket[20]. The molecular 'gymnastics' mediated by the DNA-damage-specific glycosylases have been described as a routine involving specific 'pinch–push–plug–pull' steps[17,18]. On encountering a specific DNA lesion, the DNA-damage-specific glycosylase induces bending and distortion of the DNA double helix, constituting the 'pinch'. This step, as well as a 'push' of the target base out of the helix, is facilitated by intercalation of a glycosylase amino-acid side chain into the DNA helix. The same or another intercalating amino acid functions as a 'plug' to fill the position of the extruded base and stabilize the contortion of the DNA duplex. A 'pull' by active-site residues specific for the relevant target base secures it into the recognition and clipping pocket.

These general features used for lesion recognition are well illustrated by the X-ray crystal structure of a catalytically inactive variant of OGG1 (in which the lysine residue at position 249 is replaced by a glutamine residue; Lys249Gln) bound to a substrate duplex containing an 8-oxoG•C base pair[22]. This variant of OGG1 retains high-affinity binding to an 8-oxoG•C-containing duplex but cannot excise 8-oxoG because of the mutation of Lys 249, which participates in catalysis of the base-displacement reaction. Such enzyme complexes with a lesion-containing DNA substrate are referred to as lesion-recognition complexes (LRCs). In the X-ray crystal structure of the LRC described above, the 8-oxoG•C base pair has been disrupted, and the 8-oxoG nucleotide has been rotated

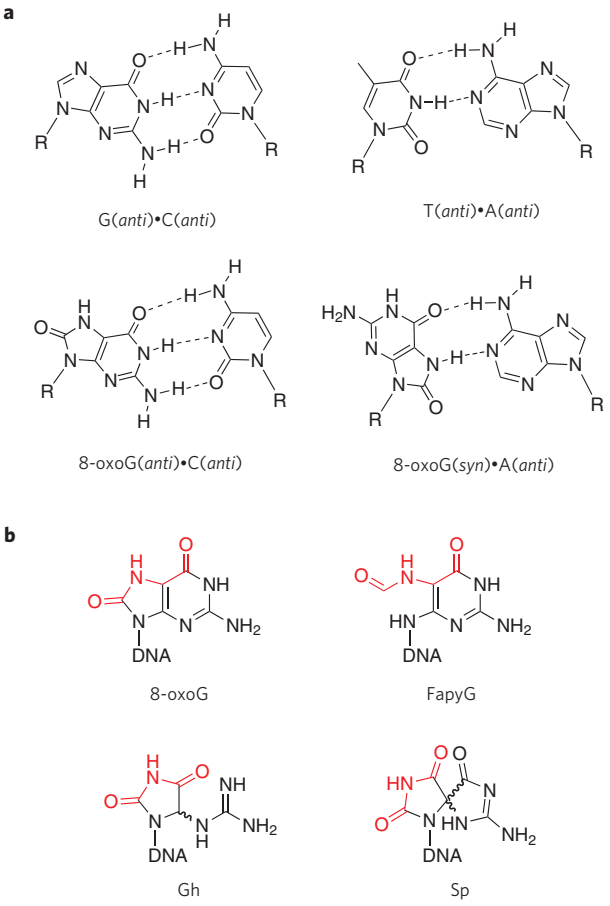


**Figure 2 | Structures of 8-oxoG-containing base pairs and of several nucleosides of guanine oxidation products. a**, The structures of the base pairs G•C and T•A are compared with those of 8-oxoG•C and 8-oxoG•A. 8-oxoG differs from G by an oxo group at C8 and an NH at N7. This subtle change allows 8-oxoG to base-pair easily with either C or A. **b**, A large number of guanine oxidation products have been observed[10]. Here, we focus on the hydantoin lesions: Gh and Sp. The structures of these lesions resemble 8-oxoG and FapyG in retaining hydrogen-bonding functionality (red) that can mimic T.

to place the attached base within an 8-oxoG-specific pocket (Figs 3 and 4a). Although OGG1 makes many specific contacts with both the extrahelical 8-oxoG and the orphaned C, only one hydrogen-bonding contact would not be possible with G. This crucial hydrogen bond is between the carbonyl oxygen of Gly 42 of OGG1 and the hydrogen at N7 of 8-oxoG. With G instead of 8-oxoG, not only would such a hydrogen bond be absent, but there would also be unfavourable repulsion between the carbonyl oxygen and the lone pair at N7. It seems unlikely that specificity for 8-oxoG over G would be due to a single hydrogen bond, so the mechanism for discrimination between these two bases is probably much more complex.

### Revealing encounters of OGG1 with G•C base pairs

Understanding the strategy used by OGG1 to discriminate between 8-oxoG•C and G•C required a structural glimpse of OGG1 encountering a G•C base pair. Because these enzymes do not recognize undamaged G•C base pairs, such an encounter is necessarily transient. For this reason, Verdine and co-workers used an innovative covalent trapping method to secure OGG1 in proximity to a G•C base pair[23]. Using the X-ray crystal structure of the OGG1 LRC as a guide[22], a cysteine residue was installed at position 149 of the enzyme (in place of asparagine), and a single thiol-modified-base-containing nucleotide was positioned within the DNA duplex such that a disulphide crosslink would be formed when OGG1 encountered the targeted G•C base pair. The remarkable structure[23] revealed that even though G is forcibly presented to the enzyme, it does not gain access to the 8-oxoG site but is lodged within an alternative, peripheral, 'exo' site (Fig. 4b). Calculations of differences in free energy indicate that favourable 8-oxoG interactions and unfavourable G interactions in both the active and exo sites result in the $10^5$-fold preference for 8-oxoG over G. Quantum calculations also show that the 8-oxoG and G bases have local dipoles that have opposite orientations as a result of the charge inversion at C8 and N7. The active site of OGG1 has a complementary dipole for 8-oxoG created by a Lys-249-$NH_3^+$ and Cys-253-$S^-$ amino-acid pair within the active site (Fig. 3). Indeed, the contribution of this dipole–dipole interaction to recognition is calculated to be greater than the contribution of the single hydrogen bond to NH7 of 8-oxoG. However, the importance of the specific hydrogen bond to 8-oxoG was shown by replacing Gly 42 with alanine and carrying out structural analysis of the resultant complexes using the disulphide-crosslink approach with 8-oxoG or G presented to the enzyme[24]. Surprisingly, in all of the structures examined, the backbone conformation at position 42 remains identical to that in the LRC (Fig. 3), even though the alanine substitution is energetically unfavourable as a result of the steric demands of the side-chain methyl group. These structures show the role of the surrounding protein in 'hard-wiring' the conformation of Gly 42 to preserve the ability of OGG1 to recognize 8-oxoG and effectively reject G from the active site[24].

The presence of the exo base-binding site, and the calculations showing preferences for binding to 8-oxoG at both base-binding sites, suggest that the damaged base is extracted from the duplex by OGG1 in discrete steps: an initial step involves placement into the exo site, and a final step involves engagement in the active-site pocket. The use of a multistep base-excision process allows many opportunities to check the identity of the base before its excision. In the case of OGG1, only 8-oxoG has the proper 'credentials' to make it through all of the checkpoints.

### Additional intermediates in the 8-oxoG repair pathway

A later intermediate in the 8-oxoG recognition and excision process was revealed recently by structural analysis of several constructs containing a mutation (Gln315Phe) in OGG1 that enlarges a side chain to occlude the active-site pocket[24] (Fig. 3). Notably, in one structure, the 8-oxoG lesion is almost completely inserted into the active-site pocket; however, the 8-oxoG base is not clipped out of the DNA by this mutated enzyme. This demonstrates another level of quality control, requiring perfect geometric alignment of the base within the specificity pocket to attain the proper transition state for base cleavage.

The disulphide-crosslink approach also provided a fortuitous structure of OGG1 'interrogating' a G•C base pair located adjacent to an 8-oxoG•C base pair in a DNA duplex[25] (Fig. 4c). This structure reveals how 8-oxoG might first be detected by OGG1 during the evaluation of the neighbouring base pair. The overall structure is similar to the LRCs and the interrogation complex with the G•C base pair discussed earlier. The target G, however, is destacked from the helix, but it does not reside at the exo site. Instead, the target G is folded back into the major groove and forms an unusual triple hydrogen-bonding interaction with the major groove face of the neighbouring 8-oxoG. The altered mode of examination of G in the vicinity of 8-oxoG seems to be associated with severely repulsive steric and electrostatic 'clashes' with the phosphate backbone as a result of the presence of the 8-oxo group of 8-oxoG. The altered backbone conformations do not allow complete presentation of the G to the exo site, thus favouring the folded-in conformation that is further stabilized by hydrogen bonding with 8-oxoG. The authors of this study, Banerjee and Verdine, noted that such base–phosphate clashes only occur as a result of the remodelling of the duplex by OGG1 (ref. 25). Thus, the active role of the DNA glycosylase in probing and altering the duplex is a key feature of the mechanism for detecting subtle alterations in the base structure.

### Intrahelical interrogation of normal base pairs

The structures described here indicate that G•C base pairs might be paid special attention by OGG1. But does OGG1 scrutinize normal base pairs by removing one of the bases from the helix? Such a process would seem to be energetically costly and inefficient. To explore further the interrogation process used by 8-oxoG DNA glycosylases, the disulphide-crosslink approach was used to generate several structures of *B. stearothermophilus* MutM (which is functionally similar to OGG1) tethered to normal DNA in proximity to both an A•T and a G•C base pair[26] (Fig. 4d–f). In these structures, the normal bases are not extruded from the helix, suggesting that MutM initially locates 8-oxoG lesions by intrahelical interrogation. Indeed, the side chain of a phenylalanine residue (Phe 114) was found to be wedged into the helix from the minor groove above the A•T or G•C base pair. This invasion by Phe 114 results in severe bending of the DNA and buckling of the base pair. In addition, two other amino-acid residues (Arg 112 and Met 77) that take the place

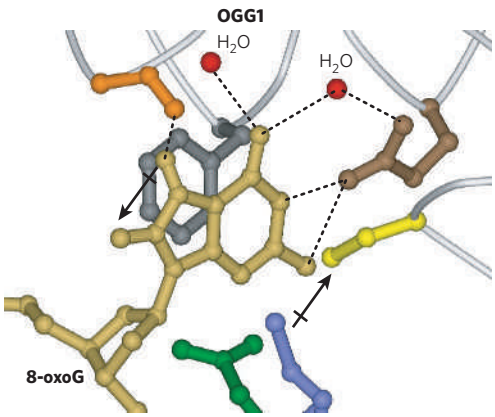


**Figure 3 | Recognition of 8-oxoG by OGG1 observed in the LRC of OGG1 with 8-oxoG•C-containing duplexes.** This is a view of the base-specific pocket of OGG1, showing residues involved in the recognition of 8-oxoG. DNA is shown in gold, with the OGG1 backbone in light grey and coordinating water molecules in red. Residues that are important for recognition are shown as follows: orange (Gly 42), purple (Lys 249), yellow (Cys 253), green (Asp 268), brown (Gln 315) and dark grey (Phe 319). Hydrogen bonds are shown as dashed lines. An 8-oxoG-specific hydrogen bond between NH7 and the carbonyl of Gly 42 provides specificity for 8-oxoG over G. In addition, the dipole (crossed arrow) that is associated with the Lys 249–Cys 253 pair makes favourable dipole–dipole interactions with 8-oxoG but not with G. Image generated from pdb file 1EBM from the Worldwide Protein Data Bank, based on data from ref. 22.

of the flipped-out 8-oxoG base in the corresponding LRC[27] are drawn back, as though they are waiting to lunge in when space has been created. The overall structure of these non-specific complexes is similar to that of the LRC, except that many contacts between the enzyme and the DNA are water-mediated rather than direct. This might allow a looser association with the DNA, facilitating movement along the helix. Interestingly, despite using the disulphide-crosslink strategy, it was difficult to trap a complex in which MutM was interrogating a G•C base pair; indeed, the protein seemed to avoid interrogating G•C base pairs. To obtain the desired structure, the DNA sequence was altered so that a string solely consisting of G•C base pairs was presented to the enzyme. Taken together, the various structures indicate that Phe 114 is a sensor of the stability and/or deformability of the target base pair. The authors, Banerjee *et al.*, suggest that although an 8-oxoG•C base pair is only slightly less stable than an A•T base pair, the non-standard properties of this base pair might render it unable to withstand the insertion of Phe 114 of MutM[26]. After the base pair has been disrupted, the liberated 8-oxoG would then be captured by the active site of MutM.

### A redundant search of the genome for 8-oxoG lesions

It might be expected that such a search process would not be extremely precise; therefore, Banerjee *et al.* suggest that DNA glycosylases might compensate for this imprecision by being extremely fast, allowing repeated opportunities to find a damaged base[26]. This fast search process was visualized recently using single-molecule detection to track OGG1 movements along a normal DNA duplex[28]. In these studies, OGG1 was found to move along the DNA with a diffusion constant approaching the theoretical upper limit for one-dimensional diffusion, indicating that OGG1 samples millions of base pairs per second. On the basis of these measurements, the estimated barrier to sliding is extremely small (0.5 kcal $mol^{-1}$). The smaller barrier and the observed unbiased random movement of OGG1 on DNA suggest that OGG1 rapidly searches along DNA as a consequence of brownian motion. Effects on the measured diffusion constants as a result of mutation of His 270, which is located at the DNA-binding interface[22–24], and pH suggest that the DNA–protein interface has been optimized to allow fast enzyme sliding.

Taken together, a scheme of the events involved in the search for 8-oxoG lesions by OGG1 or MutM can be proposed from the results of these studies (Fig. 5). First, the 8-oxoG DNA glycosylase moves rapidly along the helix, inserting the probe ligand (for example, Phe 114 in MutM) into the helix to look for vulnerable sites. Intercalation of an amino-acid residue of the enzyme at a normal base pair merely buckles the base pair; however, it can be envisaged that such a probing event could disrupt an abnormal base pair such as 8-oxoG•C. Such a search process would be extremely fast, so an 8-oxoG base might be missed

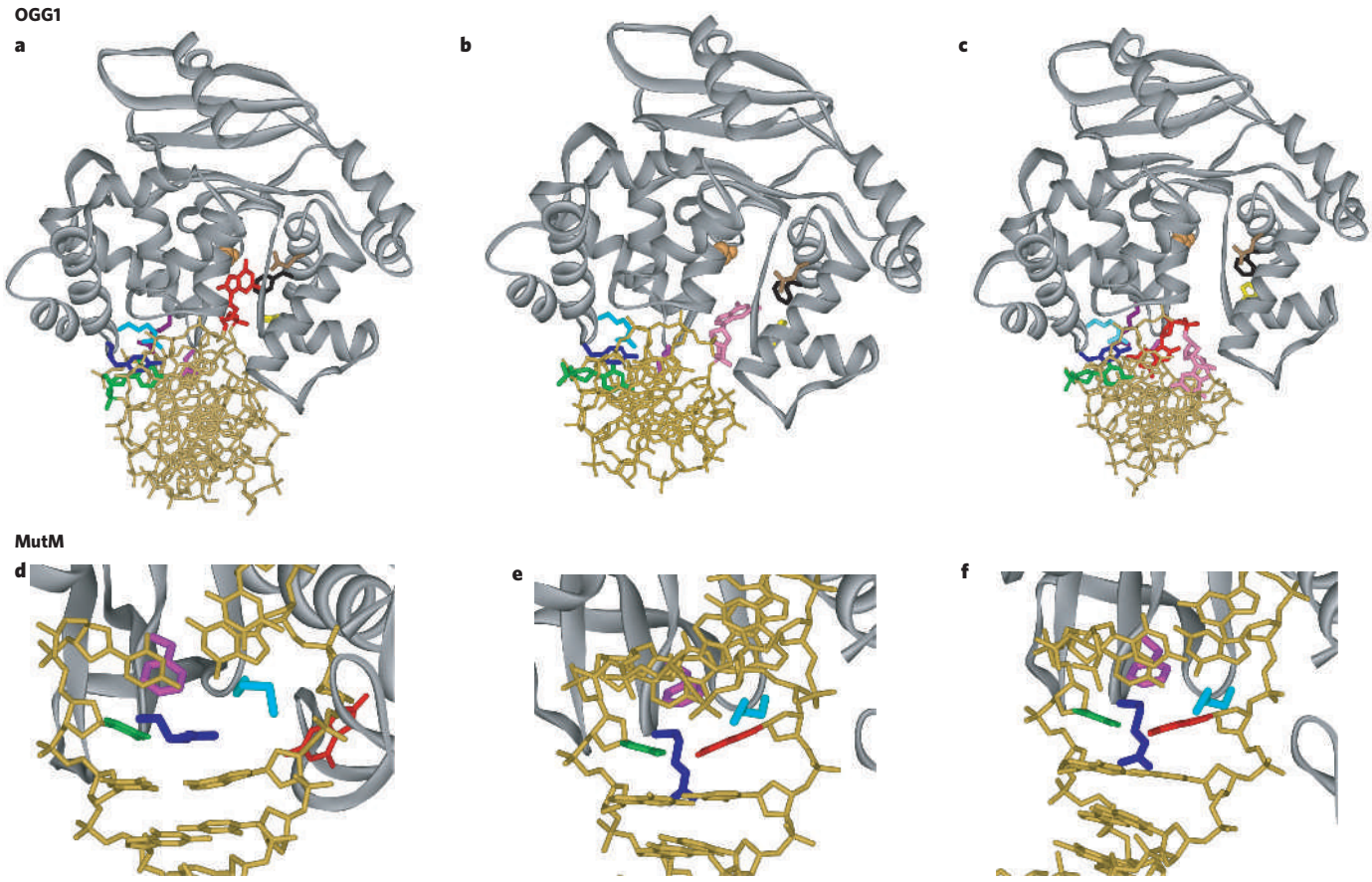


**Figure 4 | The LRCs of OGG1 and MutM with non-specific complexes (normal base pairs).** In all structures, the protein backbone is shown as a dark grey ribbon, and the DNA is in gold. In **a**–**c**, important amino-acid side chains are shown in orange (Gly 42), dark pink (Asn 149 or Cys 149), light purple (Arg 154), dark purple (Tyr 203), light blue (Arg 204), yellow (His 270), brown (Gln 315) and black (Phe 319). **a**, OGG1 LRC with 8-oxoG•C-containing DNA. 8-oxoG is shown in red, and the estranged C in green. **b**, OGG1 interrogating a G•C base pair. The target G in shown in light pink, and the target C in green. (Arg 154 is not shown.) **c**, OGG1 interrogating a G•C base pair adjacent to an 8-oxoG lesion. 8-oxoG is shown in red, the target G in light pink, and the target C in green. In **d**–**f**, the intercalating residue, Phe 114, is shown in dark pink, and Met 77 and Arg 112 are in light blue and dark purple, respectively. **d**, MutM LRC with 8-oxoG•C-containing DNA. 8-oxoG is shown in red, and the estranged C in green. **e**, MutM interrogating a G•C base pair. The target G is shown in red, and the target C in green. **f**, MutM interrogating an A•T base pair. The target A is shown in red, and the target T in green. Images generated from pdb files from the Worldwide Protein Data Bank, based on data from the following: ref. 22, file 1EBM (**a**), ref. 23, file 1YQK (**b**), ref. 25, file 2I5W (**c**), ref. 27, file 1R2Y (**d**), ref. 26, file 2F5O (**e**) and ref. 26, file 2F5N (**f**).

sometimes. But this might be preferable to wasting time interrogating undamaged DNA. This fast, random and redundant process would provide many opportunities to find 8-oxoG. After the 8-oxoG has been partly expelled from the helix, it is captured by the exo site and passed quickly to the 8-oxoG-specific active site. Occasionally, a G might be displaced from the helix; however, it would be intercepted in the exo base-binding site and would not proceed to the 8-oxoG-specific pocket; instead, it would be placed back in the helix. Moreover, if a G adjacent to 8-oxoG is interrogated, it does not fully dislodge from the helix, thus allowing the 8-oxoG DNA glycosylase to home in more rapidly on the adjacent 8-oxoG lesion. Structural and biophysical studies suggest other DNA glycosylases probably operate by similar mechanisms[18,29]. Indeed, some damaged bases might be expelled more readily from the helix or be distinguished more readily from their normal counterparts, making such a search process more streamlined.

### Recognition of 8-oxoG•A mismatches by MutY and MUTYH

The partners for MutM and OGG1 in the 8-oxoG repair pathway, MutY and MUTYH, also need to recognize 8-oxoG precisely, for selective removal of only the A bases present in 8-oxoG•A mismatches. MutY has a special carboxy-terminal domain that is not found in other BER glycosylases, and this domain has an important role in the recognition of 8-oxoG[30–32]. In the X-ray crystal structure of an inactive variant (Asp144Asn) of *B. stearothermophilus* MutY[32], there are a plethora of contacts with 8-oxoG but minimal contacts with A. Despite the implied importance of 8-oxoG, MutY also mediates the removal of A opposite other bases, including G, FapyG (a formamidopyrimidine) and C[5,33]. Time-resolved fluorescence experiments of the MutY A-excision reaction using 8-oxoG•A substrates indicated a multiphase reaction profile, with a fast process being associated with changes at 8-oxoG and a slower process associated with altering the environment of the A[34]. The fluorescence data were originally interpreted as indicating sequential extrusion of 8-oxoG then A from the helix. However, in the *B. stearothermophilus* MutY LRC structure[32], 8-oxoG is within the helix but has an altered conformation about the N-glycosidic bond. Although the molecular mechanism of the 8-oxoG-associated fluorescence changes is unclear, these experiments, together with the structure, suggest that MutY relies heavily on initial recognition of 8-oxoG to locate A bases for excision. Additional structures of MutY bound to other base pairs (such as T•A) would be of considerable interest to elaborate how this enzyme prevents inadvertent excision of A opposite T.

## BER and colorectal cancer

The importance of preventing mutations associated with 8-oxoG was emphasized when a direct link was uncovered between colorectal cancer and mutations in the gene encoding the human MutY homologue (MUTYH)[35]. This work also established the first link between inherited defects in BER and cancer. In a British family (denoted family N), several siblings presented clinical symptoms characteristic of familial adenomatous polyposis, a common form of familial colon cancer. In familial adenomatous polyposis, the colon of afflicted individuals is littered with adenomatous polyps as a result of mutations in the gene adenomatous polyposis coli (*APC*)[36]. The APC protein has many important roles in controlling the proliferation of colon cells and is mutated in most colorectal tumours. Although the familial nature of familial adenomatous polyposis is usually a consequence of inherited mutations in *APC*, this was not the case in family N. Consistent with the polyposis phenotype, DNA from tumours of afflicted members of family N had inactivating mutations in *APC*. The types of mutation, G-to-T transversions, was particularly revealing, because this type of mutation is commonly associated with 8-oxoG and therefore prompted sequencing of the genes encoding the enzymes of the human 8-oxoG repair pathway. This revealed germline biallelic missense mutations in *MUTYH* that would result in two variants of MUTYH, each containing a single amino-acid substitution, Tyr165Cys and Gly382Asp. Owing to the high similarity of MUTYH and *Escherichia coli* MutY,

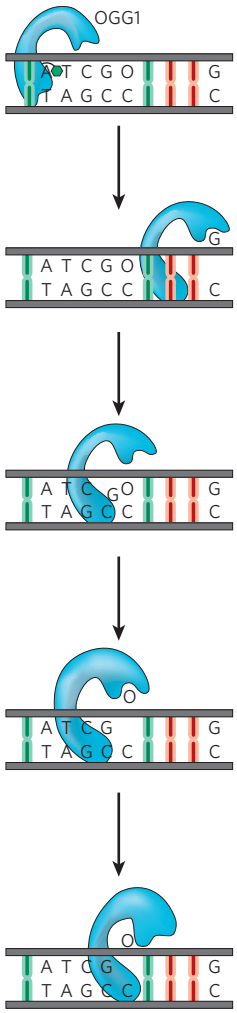


**Figure 5 | The 8-oxoG lesion search process.** The pathway is based on structures of OGG1 and MutM 'trapped' with DNA. The enzyme tracks rapidly along DNA, inserting a 'probe' amino-acid residue (green hexagon) at various base pairs to test the stability and/or deformability of the duplex. This results in preferential expulsion of 8-oxoG (O) from 8-oxoG•C base pairs. However, if a G from a G•C base pair is extruded, it is captured by the exo site and then replaced within the helix. Encountering a G•C base pair adjacent to an 8-oxoG•C base pair enables 8-oxoG to be detected easily, because this G cannot be extruded in the same manner, promoting movement of the enzyme to the 8-oxoG•C base pair. At this point, the 8-oxoG might first be extruded to the exo site and then be quickly captured in the 8-oxoG-specific pocket, where it is excised from the DNA.

the activities of the two corresponding variants in *E. coli* MutY were analysed[35]. Both were found to have compromised A removal activity relative to wild-type MutY[35,37]. This finding supported the hypothesis that the MUTYH variants have a reduced capacity to initiate the repair of 8-oxoG•A mismatches, leading to increased numbers of G-to-T transversions in *APC* and eventually resulting in inactivation of the APC protein. This is a novel mechanism by which inherited defects in a gene encoding a BER enzyme (*MUTYH*) lead to mutations in a gene associated with predisposition to cancer (*APC*). This new mechanism of predisposition to colorectal cancer is now referred to as MUTYH-associated polyposis[38,39].

To evaluate the consequences of the two amino-acid variants on DNA-damage and -mismatch recognition independently of base cleavage, dissociation constants of the two variant MutY enzymes were measured with DNA duplexes containing a non-cleavable A analogue, 2′-deoxy-2′-fluoroadenosine (FA), opposite 8-oxoG and G[37]. In these experiments, the two MutY enzymes (with the substitutions Tyr82Cys and Gly253Asp) showed a lower affinity for both 8-oxoG•FA-containing duplexes and G•FA-containing duplexes than did the wild-type enzyme. Surprisingly, the affinity of the variants for the duplex was not increased by the presence of 8-oxoG, in contrast to the case for the wild-type enzyme, indicating that the mutations abrogate the ability to recognize 8-oxoG over G. Insight into the role of these two amino acids in 8-oxoG recognition was also provided by determining the structure of a *B. stearothermophilus* MutY variant that has an Asp144Asn substitution bound to an 8-oxoG•A-containing duplex[32] (Fig. 6a). The corresponding tyrosine residue in *B. stearothermophilus* MutY (Tyr 88) was found to be intercalated 5′ to 8-oxoG and to participate in a hydrogen-bond

network with NH7 of 8-oxoG, whereas the glycine residue (Gly 260) is located in a turn region of the protein where backbone amides are involved in close hydrogen bonds with the phosphodiester backbone adjacent to 8-oxoG.

Interestingly, the location of Tyr 82 in MutY corresponds to the Phe 114 'probe' amino acid revealed in the structure of MutM trapped with non-specific DNA[26]. This suggests that the ability of MutY to seek and find 8-oxoG•A mismatches is hampered by mutation of the large tyrosine probe ligand to a smaller residue, cysteine. The ability to replace the tyrosine residue in MutY with a leucine residue without sacrificing catalytic activity established the importance of intercalation of a bulky residue for indirect recognition of 8-oxoG•A mismatches[40], because a leucine side chain would be unable to participate in the hydrogen-bond network with NH7 of 8-oxoG. Many DNA glycosylases have a phenylalanine, leucine or tyrosine residue at the analogous position[26]. In a more subtle manner, the turn region of *B. stearothermophilus* MutY that contains Gly 260 seems to be stabilizing a pinched and distorted conformation of the phosphodiester backbone adjacent to 8-oxoG, and this might be required for promoting A extrusion. This glycine residue is located in the C-terminal domain of MutY, which has been shown to be important for 8-oxoG recognition[30–32]. Interestingly, the single amino-acid change in the Gly253Asp MutY variant resulted in a loss in affinity for duplexes containing 8-oxoG over G that was similar to that observed with a truncated form of MutY completely lacking the C-terminal domain[37]. Study of a structure of the *B. stearothermophilus* Asp144Asn MutY variant bound to 8-oxoG•A-mismatch-containing DNA[32] revealed that the peptide-backbone torsion angles of Gly 260 are highly unusual, and glycine is the only amino acid that would easily accommodate such a conformation. Thus, replacement of this glycine with aspartic acid or alanine destabilizes the turn region and alters the ability to recognize 8-oxoG at this site specifically[40].

Although both variant MutY enzymes have a DNA-binding defect, the binding defect of the Tyr82Cys variant *in vitro* translates into a more deleterious effect on A excision[37,40]. The *in vitro* adenine glycosylase activity of Tyr165Cys and Gly382Asp MUTYH variants (A. L. Livingston and S.S.D., unpublished observations) and the corresponding variants of mouse MUTYH[41] (also known as MYH) were analogous to those observed for the *E. coli* variants. However, in contrast to the *E. coli* variants, the activity of the mouse MUTYH variants is further reduced in the presence of other DNA-binding proteins[41,42]. For example, the presence of the human AP endonuclease (APE1; also known as APEX1) increases the A-excision activity of wild-type mouse MUTYH by stimulating release of MUTYH from the AP-site DNA product. By contrast, for the mouse MUTYH variants, the presence of APE1 reduces the efficiency of A-excision activity[41]. This is a consequence of competition between MUTYH and APE1 for the 8-oxoG•A-containing DNA substrate, which is not a problem for the wild-type enzyme, owing to its greater affinity for 8-oxoG•A. These competition effects might also explain why, in

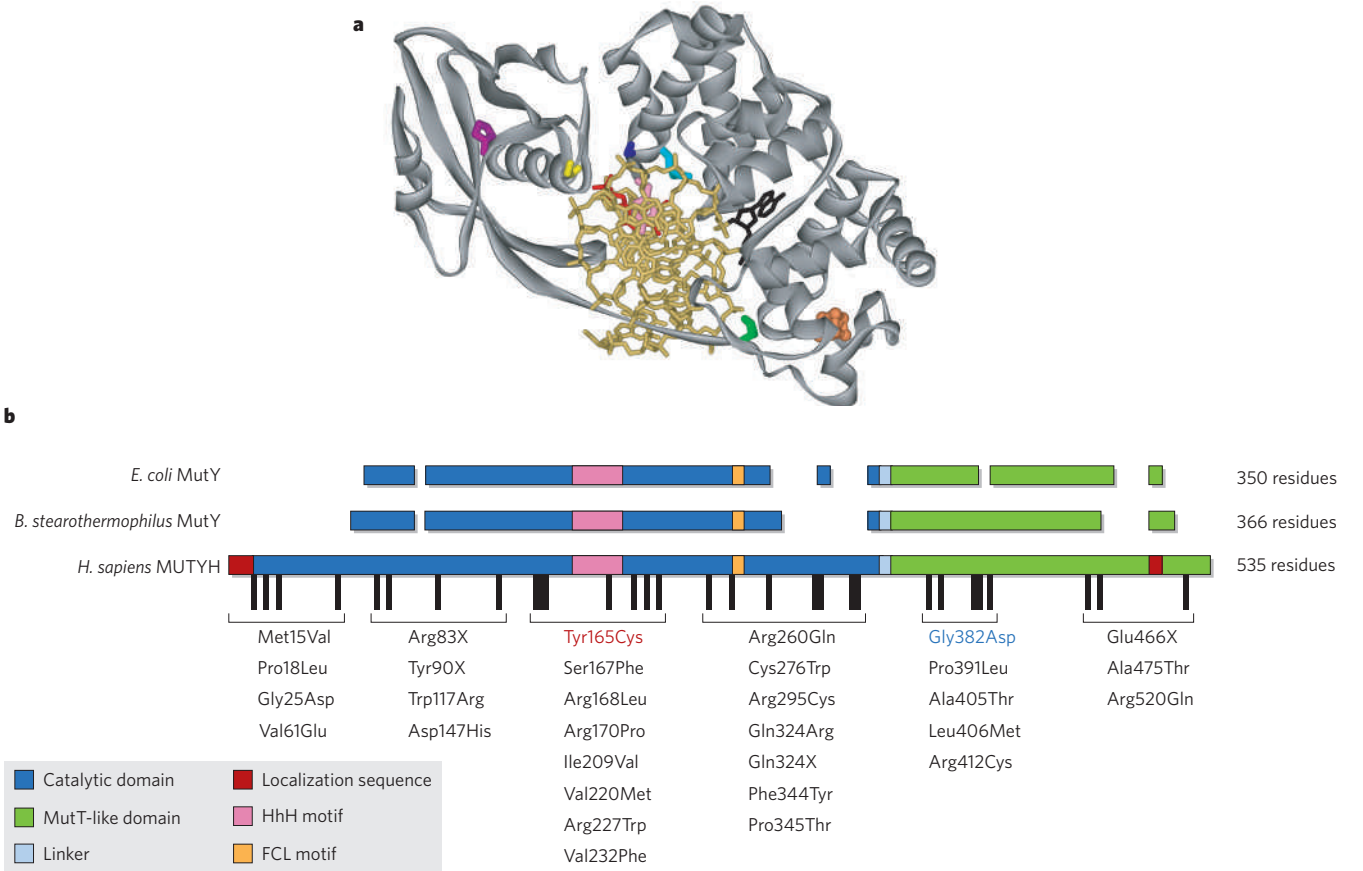


**Figure 6 | Germline mutations observed in *MUTYH* in individuals with MUTYH-associated polyposis. a**, Structure of *B. stearothermophilus* MutY with an Asp144Asn substitution bound to an 8-oxoG•A-mismatch-containing duplex. The view is down the helical axis, with the MutY backbone (grey ribbon) encircling the DNA (gold). The A is shown in black, bound in its recognition pocket, and 8-oxoG is shown in red. The 4Fe–4S cluster is shown in orange. Amino-acid residues in MutY that correspond to some of the residues observed to vary in individuals with MUTYH-associated polyposis are indicated on the MutY ribbon: MutY Tyr 88 (light pink), MUTYH Tyr165Cys; Ser 90 (dark purple), Ser167Phe; Arg 91 (light blue), Arg168Leu; Pro 226 (green), Pro345Thr; Gly 260 (yellow), Gly382Asp; and Pro 269 (dark pink), Pro391Leu. Image generated from pdb file 1RRQ from the Worldwide Protein Data Bank, based on data from ref. 32. **b**, Alignment of *E. coli* MutY, *B. stearothermophilus* MutY and human MUTYH. Representative mutations that have been observed in individuals with MUTYH-associated polyposis are indicated. The original MUTYH variants, Tyr165Cys and Gly382Asp, are highlighted in red and blue, respectively. Crucial DNA-binding motifs are the helix–hairpin–helix (HhH) motif and the Fe–S cluster loop (FCL) motif[99,100]. X, stop codon results in truncation at this amino acid.

cell-extract preparations, both human MUTYH variants seemed to be completely inactive[43]. Taken together, these observations suggest that defects in finding and recognizing damaged DNA might be magnified in a cellular context compared with *in vitro* experiments as a result of the presence of many other DNA-binding enzymes. In addition, the search process is much more demanding in a cellular environment because of the much larger size of chromosomal DNA than the oligonucleotide substrates used *in vitro*. These ideas could explain why MUTYH variants seem to be less active in cellular complementation assays[37,44].

Consistent with a global defect in 8-oxoG•A repair, a high proportion of tumours from patients with biallelic mutations in *MUTYH* have been observed to contain G-to-T transversions in the first G of codon 12 of the oncogene *Ki-Ras*[45]. This finding suggests that other genes, in addition to *APC*, might be susceptible to mutation as a result of dysfunctional MUTYH; however, so far, no further compelling connections with other types of cancer have been reported. The observed mutation of *Ki-Ras* in MUTYH-associated polyposis tumours was also interesting, because the same type of inactivating mutation is observed in mice that are deficient in both MUTYH and OGG1 (ref. 46). These double-knockout mice show high susceptibility to tumour formation, mainly lung and ovarian tumours and lymphomas. In addition, these mice also have increased amounts of 8-oxoG in the lungs and small intestine[47]. Surprisingly, mice that are deficient only in MUTYH do not show any atypical properties[46]. However, crossing MUTYH-deficient mice with multiple intestinal neoplasia ($Apc^{Min/+}$) mice, which carry a nonsense mutation in *Apc*, resulted in greater intestinal tumorigenesis than in $Apc^{Min/+}/Mutyh^{+/+}$ mice[48]. These mouse model studies showed that MUTYH deficiency does result in an increased intrinsic mutation rate; however, this factor alone does not lead to tumour formation in mice. Deficiencies in other proteins, such as APC or OGG1, help to fuel the progression towards cancer. In humans, the extent of oxidative stress and the level of mutations mediated by 8-oxoG, especially in the gastrointestinal tract, might be greater than in mice. It should also be noted that in these mouse models, MUTYH protein is absent, rather than present in a variant form. Studies of *Mutyh*-variant knock-in mice would be interesting, because they would shed light on whether the presence of the variant MUTYH proteins increases mutagenesis and is thereby more deleterious than the complete absence of the protein.

## Beyond family N

Since the original discovery of family N, considerable work has established the relationship between mutations in *MUTYH* and colorectal adenomas and carcinomas, and this disorder is now referred to as MUTYH-associated polyposis[38,49–54]. MUTYH-associated polyposis seems to be an autosomal recessive disorder (requiring mutations in both alleles). Present estimates of the frequency indicate that MUTYH-associated polyposis accounts for ~1% of all colorectal cancer, although this might increase as more patients are tested for mutations in *MUTYH*[55]. This contribution to colorectal cancer is lower than the present estimate of 1–6% for hereditary non-polyposis colorectal cancer (HNPCC)[56,57]. HNPCC, like MUTYH-associated polyposis, stems from an inherited DNA-repair defect, specifically in genes involved in mismatch repair (MMR)[57]. Of the MMR gene mutations, those in the genes encoding the human MutS homologue, MSH2, and the human MutL homologue, MLH1, predominate[57]. However, HNPCC differs from MUTYH-associated polyposis in being an autosomal dominant disorder and in resulting from a different type of DNA-repair defect. Thus, the evolutionary mechanisms that led to cancer originating from these two DNA-repair defects are not likely to be the same. Although familial adenomatous polyposis and HNPCC are relatively rare, understanding the genetic basis of these disorders has provided important insight into pathways that result in cancer[58,59]. For example, a feature of the MMR defect is the expansion or contraction of short nucleotide repeats in the DNA of tumours, a process commonly referred to as microsatellite instability[56,57]. Interestingly, microsatellite instability is also characteristic of many sporadic colorectal cancers[60]. The presence of microsatellite instability, family history, the presence and type of MMR gene mutation, and clinical manifestations are all important pieces of the puzzle that are used to diagnose HNPCC and to advise and treat patients and family members appropriately[61]. Although much is known about HNPCC, many questions still remain about the molecular mechanism that leads from an MMR defect to cancer.

Because MUTYH-associated polyposis was discovered only recently, more information about the clinical and molecular properties of MUTYH-associated polyposis is needed to aid in the diagnosis and treatment of affected patients and family members. Such knowledge might also provide insight into how *MUTYH* mutations contribute more globally to cancer. At present, testing for mutations in *MUTYH* is recommended for patients who have clinical features of familial adenomatous polyposis but either do not have inherited mutations in *APC* or have a family history consistent with recessive inheritance, as is the case for MUTYH-associated polyposis[50]. In classical familial adenomatous polyposis, the sheer number of adenomatous polyps leads to a lifetime risk of eventually developing carcinoma of almost 100% (ref. 36). On the basis of preliminary studies[53,54,62,63], a similar lifetime risk for colorectal cancer has been suggested for MUTYH-associated polyposis[57]. The potential increased colorectal cancer risk for heterozygous individuals (who have only one mutated *MUTYH* allele) is a subject of debate, and more clinical testing is required to resolve this issue[39].

In addition to the two variants in MUTYH that were initially identified (namely Tyr165Cys and Gly382Asp), ~82 germline mutations have been found in the *MUTYH* alleles of patients with colorectal adenomas and carcinomas[39] (Fig. 6b). The original variants seem to be the most common mutations found in Caucasians of northern European origin[38,39], and different mutations are more prevalent in other ethnic populations[39]. Although many of these mutations might result in lack of MUTYH, a large number of the 82 mutations (~52) are missense mutations, presenting a diagnostic challenge because their significance is unknown. On the basis of the type of mutation, the location in the sequence of MUTYH, and the corresponding position in the structure of the *B. stearothermophilus* MutY–DNA complex, it is possible to make a prediction about the consequences in some cases (Fig. 6a). For example, early truncating mutations (such as Tyr90X, where X denotes that a stop codon results in truncation at this amino acid) in the core of the C-terminal domain or late truncating mutations (such as Glu466X) are likely to be destabilizing and to result in the absence of MUTYH. Some of the missense variants that are conserved in *B. stearothermophilus* MutY might affect DNA binding (for example, Arg168Leu and Ser167Phe) or folding of one of the domains (for example, Trp117Arg). However, the consequences of other conserved missense variants (for example, Pro325Thr and Pro391Leu) are not obvious. Moreover, some missense variants are not conserved in the bacterial MutY enzymes, so it is difficult to make predictions. At this early stage in our understanding of MUTYH-associated polyposis, the possibility that some mutations are non-pathogenic polymorphisms that are found coincidentally in patients with colorectal adenomas and carcinomas cannot be ruled out. It is also possible that some variants are only mildly pathogenic and result in less severe disease symptoms. This idea is supported by *in vitro* results indicating that the Gly382Asp MUTYH variant is less catalytically compromised than the Tyr165Cys variant; however, it is unclear whether these findings correlate with clinical data. Reduced A-removal activity has also been reported for four other variants (Arg227Trp, Arg231Leu, Val232Phe and Ala459Asp)[64–66]. Importantly, mutations in *MUTYH* might also alter MUTYH expression levels and stability, protein–protein interactions and post-translational modifications. These and other factors, together with reduced 8-oxoG•A mismatch affinity and A-removal activity, could cooperate to reduce the efficiency of 8-oxoG•A repair in the *APC* gene. Additional clinical and functional data will be particularly important for revealing any associations between colorectal cancer and specific mutations in *MUTYH*.

Why are inherited mutations in *MUTYH* and not other BER enzymes associated with cancer? One possibility is that dysfunction of MUTYH is especially problematic because there are no other mechanisms for

repairing 8-oxoG•A mismatches. By contrast, nucleotide excision repair seems to function as a 'back-up' repair mechanism to mediate removal of 8-oxoG in mouse cells lacking OGG1 (refs 67–70). It is also important to note that the clinical feature of having colonic polyps provides a key piece of information that prompts testing for *APC* or *MUTYH* mutations, thereby increasing the probability of detecting a relatively rare defective gene. We predict that dysfunction of other BER glycosylases will also be found to modify cancer susceptibility, and these correlations might be uncovered as screening for genetic mutations progresses. Moreover, as information on the molecular basis of cancer becomes available, additional links between BER defects and cancer will be more readily revealed.

## Beyond 8-oxoG

Although 8-oxoG has garnered much attention, a plethora of oxidized guanine lesions have been identified[10] (Fig. 2b). Numerous studies indicate that the most common mutations observed in cells in conditions of oxidative stress are G-to-T and G-to-C transversions[10]. Whereas G-to-T transversions can be readily explained by the presence of 8-oxoG or FapyG, other oxidized lesions are probably responsible for mediating G-to-C transversions. Of these other lesions, two hydantoin products — spiroiminodihydantoin (Sp) and 5-guanidinohydantoin (Gh) — are particularly interesting and are emerging as important lesions to be dealt with by repair enzymes.

For many years, the product of singlet-oxygen damage to G was thought to be 4-OH-8-oxoG[71,72], and the presence of this product was used routinely as an indicator of singlet-oxygen damage[73]. However, when Burrows and co-workers determined the structure of Sp, as a major product arising from further oxidation of 8-oxoG-containing nucleosides and oligonucleotides, suspicion was raised about the structure of 4-OH-8-oxoG, because the molecular weights are the same[74]. It was later confirmed that the structure of 4-OH-8-oxoG had been misassigned and that 4-OH-8-oxoG was Sp[75,76]. When the identity of Sp was unmasked, it was observed to be generated by oxidation of G or 8-oxoG by a large number of oxidants, including singlet oxygen, high-valent metal ions, peroxynitrite and ionizing radiation[10,77]. Interestingly, the studies of oxidation of 8-oxoG revealed that, depending on the conditions, Gh is formed in addition to Sp[78]. When the substrate is a nucleoside or single-stranded DNA, the main product is Sp, whereas Gh predominates in double-stranded DNA[77].

Hydantoin lesions have been demonstrated to be highly mutagenic both *in vitro* and *in vivo*. Single-nucleotide insertion and primer extension experiments using an *E. coli* Klenow fragment of DNA polymerase without exonuclease activity indicate that dAMP and dGMP are inserted opposite these oxidized lesions[79,80]. In *E. coli*-based mutagenesis assays with single-stranded lesion-containing viral DNA, Sp was found to be a strong block to replication, whereas Gh was more readily bypassed[81,82]. Moreover, both lesions are 100% mutagenic, mediating both G-to-C and G-to-T transversions, whereas 8-oxoG is only mildly mutagenic (3%) in these assays[81,82].

Important features that influence the mutagenic potential of a given lesion are the efficiency of its repair and the sensitivity of the repair enzymes to the correct base-pairing context. The relatively low mutation frequency of 8-oxoG in both *E. coli* and mammalian cells results from the efficient repair of 8-oxoG. *In vitro* assays have shown that Gh and Sp are substrates for *E. coli* Fpg (MutM)[83]. Efficient repair in the correct base-pairing context would be expected to mitigate the mutagenic potential of lesions; however, removal in the wrong context might increase mutagenesis. This is shown by studies of Gh, which was found to be excised by Fpg from Gh•G and Gh•C base pairs with similar efficiencies. The Gh•G base pair is particularly interesting because it might mediate the G-to-C transversions observed *in vivo*. Removal of Gh or Sp opposite A by Fpg was less than that opposite C; however, the extent of removal of these oxidized lesions is greater than removal of 8-oxoG opposite A[83]. The activity in base-pairing contexts involving A might also be more problematic, because MutY was found to be unable to remove A from Gh•A or Sp•A base pairs[83].

Surprisingly, Gh and Sp are not substrates for OGG1, although they are removed by the *Saccharomyces cerevisiae* homologues Ogg1 and Ntg1 (ref. 84). Another bacterial glycosylase, endonuclease VIII (Nei), can also remove Gh and Sp[85]. Interestingly, this enzyme usually targets oxidized pyrimidines, such as thymine glycol and 5-hydroxycytosine[86]. In fact, Gh and Sp, like 8-oxoG, have a thymine-like Watson–Crick hydrogen-bonding face (Fig. 2b). On the basis of the idea that repair might thwart the ability to detect lesions that are present at low levels in cells, Sugden and co-workers used various repair-deficient *E. coli* strains to identify the products formed by chromate oxidation[87]; Sp was detected by mass spectrometric techniques in chromate-treated cells deficient in the BER glycosylase Nei.

Recently, mammalian orthologues of *E. coli* Fpg and Nei were identified and designated as the Nei-like (NEIL) family of enzymes (which consists of NEIL1, NEIL2 and NEIL3)[88–91]. The overall structure of NEIL1 is similar to that of *E. coli* Nei and Fpg, although it contains an unusual 'zincless'-finger motif, which is required for its glycosylase activity[92]. The substrate specificity of NEIL1 is more similar to that of Nei than that of Fpg[88]. Reported substrates include 5-hydroxyuracil, Fapy nucleotides and thymine glycol, whereas the activity towards 8-oxoG is minimal. In contrast to OGG1, NEIL1 also operates on single-stranded DNA and 'bubble DNA' (that is, single-stranded DNA flanked by duplex regions), thus prompting the suggestion that it is involved in replication and/or transcription-coupled repair[93]. Qualitative reports also showed that Gh and Sp lesions are removed by the mouse counterparts of Nei, NEIL1 and NEIL2 (ref. 94). Interestingly, quantitative examination of the substrate specificity of human NEIL1 revealed Gh and Sp lesions as the best substrates identified so far (N. Krishnamurthy and S.S.D., unpublished observations). Notably, mouse and human NEIL1 have been found to remove Gh and Sp from all base-pairing contexts that would be mutagenic, except when paired with C. The detection of Sp lesions in cells, together with the high mutagenic potential of both Gh and Sp and the potentially muddled processes of repair, suggests that these lesions will begin to capture more and more of the attention that is presently directed toward 8-oxoG.

The importance of human NEIL1 in response to oxidative stress has been illustrated by the observation of increased amounts of *NEIL1* mRNA in response to ROS[95]. Moreover, knockdown of NEIL1 using RNA interference produces cells that are extremely sensitive to ionizing radiation[96]. In addition, an intriguing report correlated inactivating mutations in *NEIL1* with human gastric cancer[97]. These varied biological consequences are not surprising for an enzyme involved in the repair of oxidative damage; however, some peculiarities in the biological role of human NEIL1 were revealed in studies of the recently generated *Neil1*-knockout mouse[98]. In the absence of oxidative stress, both *Neil1*-knockout mice ($Neil1^{-/-}$) and mice that are heterozygous for *Neil1* ($Neil1^{+/-}$) show symptoms consistent with metabolic syndrome in humans[98]. These symptoms include severe obesity, dyslipidaemia and fatty liver disease. It seems somewhat odd for a defect in a BER glycosylase to be correlated with a metabolic disorder, but these types of disorder have been linked to oxidative stress. The authors of that study, Vartanian and co-workers, suggest that disruption of energy homeostasis consistent with metabolic disorders might result from the extensive mitochondrial DNA damage observed in NEIL1-deficient mice[98]. Alternatively, damage to nuclear DNA in specific cell types, such as liver and pancreatic cells, might be extensive in the absence of NEIL1, leading to symptoms of metabolic syndrome. Clearly, these interesting results warrant further study, and they illustrate the important role of BER in the overall well-being of organisms.

## Concluding remarks

Since the discovery of BER more than three decades ago, the field has clearly taken its place in the mainstream of DNA repair. A plethora of structural information has provided amazing insight into the processes of DNA-damage recognition and excision by these enzymes. This structural work will probably be complemented in the future by new biophysical and biochemical approaches allowing direct visualization of

the search processes by a variety of DNA glycosylases. This could allow additional intermediates to be observed and provide a more generalized model for damage recognition by these enzymes.

Uncovering the molecular details of events involved in the processing of damaged DNA bases by DNA glycosylases will be important to understand further how the aberrant function of these enzymes causes disease. Here, this relationship was illustrated by discussing MUTYH variants, for which the biochemical and structural data are consistent with a reduced efficiency to recognize and repair 8-oxoG•A mismatches. This provides a molecular basis for the observed G-to-T transversion mutations in *APC*, leading to MUTYH-associated polyposis. Understanding of MUTYH-associated polyposis is still at an early stage, however, and many unanswered questions about the role of the variant enzymes in initiating the process of carcinogenesis remain. For example, does severity of the disease correlate with the type of *MUTYH* mutation? Why are the colon and the *APC* gene particularly sensitive to a BER defect? The intriguing metabolic disorder phenotype of the *Neil1*-knockout mouse also suggests that interesting relationships between BER and human disease will continue to be discovered. Our sophisticated understanding of the intricate steps used to recognize and excise aberrant DNA bases could now be exploited to selectively target steps in the 'search and rescue' process mediated by DNA glycosylases. For example, small-molecule modulators of glycosylase activity could be new tools for chemical biology studies and might lead to new therapeutic approaches. ■


1. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362,** 709–715 (1993).
2. Friedberg, E. C. DNA damage and repair. *Nature* **421,** 436–440 (2003).
3. Pfeifer, G. P. *et al.* Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* **21,** 7435–7451 (2002).
4. Friedberg, E. C. Inroads into base excision repair II. The discovery of the DNA glycosylases. *DNA Repair (Amst.)* **3,** 1531–1536 (2004).
5. David, S. S. & Williams, S. D. Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev.* **98,** 1221–1261 (1998).
6. Fromme, J. C. & Verdine, G. L. Base excision repair. *Adv. Protein Chem.* **69,** 1–41 (2004).
7. Barnes, D. E. & Lindahl, T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* **38,** 445–476 (2004).
8. Sung, J.-S. & Demple, B. Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. *FEBS J.* **273,** 1620–1629 (2006).
9. Klaunig, J. E. & Kamendulis, L. M. The role of oxidative stress in carcinogenesis. *Annu. Rev. Pharmacol. Toxicol.* **44,** 239–267 (2004).
10. Neeley, W. L. & Essigmann, J. M. Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem. Res. Toxicol.* **19,** 491–505 (2006).
11. Burrows, C. M. & Muller, J. Oxidative nucleobase modifications leading to strand scission. *Chem. Rev.* **98,** 1109–1152 (1998).
12. Shibutani, S., Takeshita, M. & Grollman, A. P. Insertion of specific bases during DNA synthesis past the oxidation damaged base 8-oxodG. *Nature* **349,** 431–434 (1991).
13. Hsu, G. W., Ober, M., Carell, T. & Beese, L. S. Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature* **431,** 217–221 (2004).
14. Michaels, M. L. & Miller, J. H. The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J. Bacteriol.* **174,** 6321–6325 (1992).
15. Krahn, J. M., Beard, W. A., Miller, H., Grollman, A. P. & Wilson, S. H. Structure of DNA polymerase β with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure* **11,** 121–127 (2003).
16. Gedik, C. M. & Collins, A. Establishing the background level of base oxidation in human lymphocyte DNA: results on an interlaboratory validation study. *FASEB J.* **19,** 82–84 (2005).
17. Parikh, S. S., Putnam, C. D. & Tainer, J. A. Lessons learned from structural results on uracil-DNA glycosylase. *Mutat. Res.* **460,** 183–199 (2000).
18. Stivers, J. T. Site-specific DNA damage recognition by enzyme-induced base flipping. *Prog. Nucleic Acid Res. Mol. Biol.* **77,** 37–65 (2004).
19. Fromme, J. C., Banerjee, A. & Verdine, G. L. DNA glycosylase recognition and catalysis. *Curr. Opin. Struct. Biol.* **14,** 43–49 (2004).
20. Huffman, J. L., Sundheim, O. & Tainer, J. A. DNA base damage recognition and removal: new twists and grooves. *Mutat. Res.* **577,** 55–76 (2005).
21. Hitomi, K., Iwai, S. & Tainer, J. A. The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal and repair. *DNA Repair (Amst.)* **6,** 410–428 (2007).
22. Bruner, S. D., Norman, D. P. & Verdine, G. L. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* **403,** 859–866 (2000).
23. Banerjee, A., Yang, W., Karplus, M. & Verdine, G. L. Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* **434,** 612–618 (2005).
24. Radom, C. T., Banerjee, A. & Verdine, G. L. Structural characterization of human 8-oxoguanine DNA glycosylase variants bearing active site mutations. *J. Biol. Chem.* **282,** 9182–9194 (2007).
25. Banerjee, A. & Verdine, G. L. A nucleobase lesion remodels the interaction of its normal neighbor in a DNA glycosylase complex. *Proc. Natl Acad. Sci. USA* **103,** 15020–15025 (2006).
26. Banerjee, A., Santos, W. L. & Verdine, G. L. Structure of a DNA glycosylase searching for DNA lesions. *Science* **311,** 1153–1157 (2006).
27. Fromme, J. C. & Verdine, G. L. DNA lesion recognition by the bacterial repair enzyme MutM. *J. Biol. Chem.* **278,** 51543–51548 (2003).
28. Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L. & Xie, X. S. A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc. Natl Acad. Sci. USA* **103,** 5752–5757 (2006).
29. Jiang, Y. L. *et al.* Recognition of an unnatural difluorophenyl nucleotide by uracil DNA glycosylase. *Biochemistry* **43,** 15429–15438 (2004).
30. Noll, D. M., Gogos, A., Granek, J. A. & Clarke, N. D. The C-terminal domain of the adenine-DNA glycosylase MutY confers specificity of 8-oxoguanine–adenine mispairs and may have evolved from MutT, an 8-oxo-dGTPase. *Biochemistry* **38,** 6374–6379 (1999).
31. Chmiel, N. H., Golinelli, M.-P., Francis, A. W. & David, S. S. Efficient recognition of substrates and substrate analogs by the adenine glycosylase MutY requires the C-terminal domain. *Nucleic Acids Res.* **29,** 553–564 (2001).
32. Fromme, J. C., Banerjee, A., Huang, S. J. & Verdine, G. L. Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature* **427,** 652–656 (2004).
33. Wiederholdt, C. J., Delaney, M. O., Pope, M. A., David, S. S. & Greenberg, M. M. Repair of DNA containing FapydG and its C-nucleoside analogue by formamidopyrimidine DNA glycosylase and MutY. *Biochemistry* **42,** 9755–9760 (2003).
34. Bernards, A. S., Miller, J. K., Bao, K. K. & Wong, I. Flipping duplex DNA inside out: a double base-flipping reaction mechanism by *Escherichia coli* MutY adenine glycosylase. *J. Biol. Chem.* **277,** 20960–20964 (2002).
35. Al-Tassan, N. *et al.* Inherited variants of MYH associated with somatic G:C to T:A mutations in colorectal tumors. *Nature Genet.* **30,** 227–232 (2002).
36. Fearnhead, N. S., Britton, M. P. & Bodmer, W. F. The ABC of APC. *Hum. Mol. Genet.* **10,** 721–733 (2001).
37. Chmiel, N. H., Livingston, A. L. & David, S. S. Insight into the functional consequences of inherited variants of the hMYH adenine glycosylase associated with colorectal cancer: complementation assays with hMYH variants and pre-steady-state kinetics of the corresponding mutated *E. coli* enzymes. *J. Mol. Biol.* **327,** 431–443 (2003).
38. Sampson, J. R., Jones, S., Dolwani, S. & Cheadle, J. P. MutYH (MYH) and colorectal cancer. *Biochem. Soc. Trans.* **33,** 679–683 (2005).
39. Cheadle, J. P. & Sampson, J. R. MUTYH-associated polyposis — from defect in base excision repair to clinical genetic testing. *DNA Repair (Amst.)* **6,** 274–279 (2007).
40. Livingston, A. L., Kundu, S., Henderson-Pozzi, M., Anderson, D. W. & David, S. S. Insight into the roles of tyrosine 82 and glycine 253 in the *Escherichia coli* adenine glycosylase MutY. *Biochemistry* **44,** 14179–14190 (2005).
41. Pope, M. A., Chmiel, N. H. & David, S. S. Insight into the functional consequences of hMYH variants associated with colorectal cancer: distinct differences in the adenine glycosylase activity and the response to AP endonuclease of Y150C and G365D murine MYH. *DNA Repair (Amst.)* **4,** 315–325 (2005).
42. Tominaga, Y. *et al.* MUTYH prevents OGG1 or APEX1 from inappropriately processing its substrate or reaction product with its C-terminal domain. *Nucleic Acids Res.* **32,** 3198–3211 (2004).
43. Wooden, S. H., Bassett, H. M., Wood, T. G. & McCullough, A. K. Identification of critical residues required for the mutation avoidance function of human MutY (hMYH) and implications in colorectal cancer. *Cancer Lett.* **205,** 89–95 (2004).
44. Hirano, S. *et al.* Mutator phenotype of MutYH-null mouse embryonic stem cells. *J. Biol. Chem.* **278,** 38121–38124 (2003).
45 Lipton, L. *et al.* Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway. *Cancer Res.* **63,** 7595–7599 (2003).
46. Xie, Y. *et al.* Deficiencies in mouse *Myh* and *Ogg1* results in tumor predisposition and G to T mutations in codon 12 of the *K-ras* oncogene in lung tumors. *Cancer Res.* **64,** 3096–3102 (2004).
47. Russo, M. T. *et al.* Accumulation of the oxidative base lesion 8-hydroxyguanine in DNA of tumor-prone mice defective in both the Myh and Ogg1 DNA glycosylase. *Cancer Res.* **64,** 4411–4414 (2004).
48. Sieber, O. M. *et al.* Myh deficiency enhances intestinal tumorigenesis in multiple intestinal neoplasia ($Apc^{Min/+}$) mice. *Cancer Res.* **64,** 8876–8881 (2004).
49. Sampson, J. R. *et al.* MYH polyposis: a new autosomal recessive form of familial adenomatous polyposis demanding reappraisal of genetic risk and family management. *Lancet* **362,** 39–41 (2003).
50. Chow, E., Thirlwell, C., Macrae, F. & Lipton, L. Colorectal cancer and inherited mutations in base-excision repair. *Lancet Oncol.* **5,** 600–606 (2004).
51. Lipton, L. & Tomlinson, I. The multiple colorectal adenoma phenotype and *MYH*, a excision repair gene. *Clin. Gastroenterol. Hepatol.* **2,** 633–638 (2004).
52. Tenesa, A. *et al.* Association of MutYH and colorectal cancer. *Br. J. Cancer* **95,** 239–242 (2006).
53. Lipton, L. & Tomlinson, I. The genetics of FAP and FAP-like syndromes. *Fam. Cancer* **5,** 221–226 (2006).
54. Farrington, S. M. *et al.* Germline susceptibility to colorectal cancer due to base-excision repair gene defects. *Am. J. Hum. Genet.* **77,** 112–119 (2005).
55. Fleischmann, C. *et al.* Comprehensive analysis of the contribution of germline MYH variation of early-onset colorectal cancer. *Int. J. Cancer* **109,** 554–558 (2004).
56. Strate, L. L. & Syngal, S. Hereditary colorectal cancer syndromes. *Cancer Causes Control* **16,** 201–213 (2005).
57. Jo, W. S. & Chung, D. C. Genetics of hereditary colorectal cancer. *Semin. Oncol.* **32,** 11–23 (2005).
58. Bodmer, W. F. Cancer genetics: colorectal cancer as a model. *J. Hum. Genet.* **51,** 391–396 (2006).
59. Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87,** 159–170 (1996).
60. Soreide, K., Janssen, E. A. M., Soiland, H., Korner, H. & Baak, J. P. Microsatellite instability in colorectal cancer. *Br. J. Surg.* **93,** 395–406 (2006).
61. Lindor, N. M. *et al.* Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome. *JAMA* **296,** 1507–1517 (2006).

62. Venesio, T. *et al.* High frequency of *MYH* gene mutations in a subset of patients with familial adenomatous polyposis. *Gastroenterology* **126,** 1681–1685 (2004).
63. Leite, J. S. *et al.* Is prophylactic colectomy indicated in patients with MYH-associated polyposis? *Colorectal Dis.* **7,** 327–331 (2005).
64. Bai, H. *et al.* Functional characterization of two human MutY homolog (hMYH) missense mutations (R227W and V232F) that lie within the putative hMSH6 binding domain and are associated with hMYH polyposis. *Nucleic Acids Res.* **33,** 597–604 (2005).
65. Bai, H. *et al.* Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis. *Cancer Lett.* **250,** 74–81 (2007).
66. Alhopuro, P. *et al.* A novel functionally deficient MYH variant in individuals with colorectal adenomatous polyposis. *Hum. Mutat.* **26,** 393 (2005).
67. Klunglund, A. *et al.* Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage. *Proc. Natl Acad. Sci. USA* **96,** 13300–13305 (1999).
68. Osterod, M. *et al.* A global DNA repair mechanism involving Cockayne syndrome B (*CSB*) gene product can prevent the *in vivo* accumulation of endogenous oxidative DNA base damage. *Oncogene* **21,** 8232–8239 (2002).
69. Osterod, M. *et al.* Age-related and tissue-specific accumulation of oxidative DNA base damage in 7,8-dihydro-8-oxoguanine-DNA glycosylase (*Ogg1*) deficient mice. *Carcinogenesis* **22,** 1459–1463 (2001).
70. Sunesen, M., Stevnsner, T., Brosh, R. M., Dianov, G. L. & Bohr, V. A. Global genome repair of 8-oxoG in hamster cells requires a functional *CSB* gene product. *Oncogene* **21,** 3571–3578 (2002).
71. Cadet, J., Decarroz, C., Wang, S. Y. & Midden, W. R. Mechanisms and products of photosensitized degradation of nucleic acids and related model compounds. *Isr. J. Chem.* 1**983,** 420–429 (1983).
72. Ravanat, J. L. & Cadet, J. Reaction of singlet oxygen with 2′-deoxyguanosine and DNA. Isolation and characterization of the main oxidation products. *Chem. Res. Toxicol.* **8,** 379–388 (1995).
73. Ravanat, J. L., Berger, M., Bernard, F., Langlois, R. & Ouellet, R. Phthalocyanine and naphthalocyanine photosensitized oxidation of 2′-deoxyguanosine: distinct type I and type II products. *Photochem. Photobiol.* **55,** 809–814 (1992).
74. Luo, W., Muller, J. G., Rachlin, E. M. & Burrows, C. J. Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7,8-dihydroguanosine. *Org. Lett.* **2,** 613–616 (2000).
75. Niles, J. C., Wishnok, J. S. & Tannenbaum, S. R. Spiroiminodihydantoin is the major product of the 8-oxo-7,8-dihydroguanosine reaction with peroxynitrite in the presence of thiols and guanosine photooxidation by methylene blue. *Org. Lett.* **3,** 963–966 (2001).
76. Adam, W. *et al*. Spiroiminodihydantoin is a major product in the photooxidation of 2′-deoxyguanosine by the triplet states and oxyl radicals generated from hydroxyacetophenone photolysis and dioxetane thermolysis. *Org. Lett.* **4,** 537–540 (2002).
77. Burrows, C. J. *et al*. Structure and potential mutagenicity of new hydantoin products from guanosine and 8-oxo-7,8-dihydroguanosine oxidation by transition metals. *Environ. Health Perspect.* **110,** 713–717 (2002).
78. Luo, W., Muller, J. G., Rachlin, E. M. & Burrows, C. J. Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. *Chem. Res. Toxicol.* **14,** 927–938 (2001).
79. Kornyushyna, O., Berges, A. M., Muller, J. G. & Burrows, C. J. *In vitro* nucleotide misinsertion opposite the oxidized guanosine lesions spiroiminodihydantoin and guanidinohydantoin and DNA synthesis past the lesions using *Escherichia coli* DNA polymerase I (Klenow fragment). *Biochemistry* **41,** 15304–15314 (2002).
80. Kornyushyna, O. & Burrows, C. J. Effect of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin on proofreading by *Escherichia coli* DNA polymerase I (Klenow fragment) in different sequence contexts. *Biochemistry* **42,** 13008–13018 (2003).
81. Henderson, P. T. *et al.* The hydantoin lesions from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors *in vivo*. *Biochemistry* **42,** 9257–9262 (2003).
82. Delaney, S., Neeley, W. L., Delaney, J. C. & Essigmann, J. M. The substrate specificity of MutY for hyperoxidized guanine lesions *in vivo*. *Biochemistry*, **46,** 1448–1455 (2007).
83. Leipold, M. D., Muller, J. G., Burrows, C. J. & David, S. S. Removal of hydantoin products of 8-oxoguanine oxidation by the *Escherichia coli* DNA repair enzyme, Fpg. *Biochemistry* **39,** 14984–14992 (2000).
84. Leipold, M. D., Workman, H., Muller, J. G., Burrows, C. J. & David, S. S. Recognition and removal of oxidized guanines in duplex DNA by the base excision repair enzymes hOGG1, yOGG1 and yOGG2. *Biochemistry* **42,** 11373–11381 (2003).
85. Hazra, T. K. *et al.* Repair of hydantoins, one electron oxidation product of 8-oxoguanine, by DNA glycosylases of *Escherichia coli*. *Nucleic Acids Res.* **29,** 1967–1974 (2001).
86. Wallace, S. S., Bandaru, V., Kathe, S. D. & Bond, J. P. The enigma of endonuclease VIII. *DNA Repair (Amst.)* **2,** 441–453 (2003).
87. Hailer, M. K., Slade, P. G., Martin, B. D. & Sugden, K. D. Nei-deficient *Escherichia coli* are sensitive to chromate and accumulate the oxidized guanine lesion spiroiminodihydantoin. *Chem. Res. Toxicol.* **18,** 1378–1383 (2005).
88. Bandaru, V., Sunkara, S., Wallace, S. S. & Bond, J. P. A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair (Amst.)* **1,** 517–529 (2002).
89. Hazra, T. K. *et al.* Identification and characterization of a human DNA glycosylase for repair of modified oxidatively damaged DNA. *Proc. Natl Acad. Sci. USA* **99,** 3523–3528 (2002).
90. Hazra, T. K. *et al.* Identification of a novel human DNA glycosylase for repair of cytosine-derived lesions. *J. Biol. Chem.* **277,** 30417–30420 (2002).
91. Morland, I. *et al.* Human DNA glycosylases of the bacterial Fpg/MutM superfamily: an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA. *Nucleic Acids Res.* **30,** 4926–4936 (2002).
92. Doublie, S., Bandaru, V., Bond, J. P. & Wallace, S. S. The crystal structure of human endonuclease VIII-like 1 (NEIL1) reveals a zincless finger motif required for glycosylase activity. *Proc. Natl Acad. Sci. USA* **101,** 10284–10289 (2004).
93. Dou, H., Mitra, S. & Hazra, T. K. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J. Biol. Chem.* **278,** 49679–49684 (2003).
94. Hailer, K. M., Slade, P. G., Martin, B. D., Rosenquist, T. A. & Sugden, K. D. Recognition of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin in DNA by the base excision repair glycosylases NEIL1 and NEIL2. *DNA Repair (Amst.)* **4,** 41–50 (2005).
95. Das, A., Hazra, T. K., Boldogh, I., Mitra, S. & Bhakat, K. K. Induction of the human oxidized base-specific DNA glycosylase NEIL1 by reactive oxygen species. *J. Biol. Chem.* **280,** 35272–35280 (2005).
96. Rosenquist, T. A. *et al.* The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair (Amst.)* **2,** 581–591 (2003).
97. Shinmura, K. *et al.* Inactivating mutations of the human base excision repair gene *NEIL1* in gastric cancer. *Carcinogenesis* **25,** 2311–2317 (2004).
98. Vartanian, V. *et al.* The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase. *Proc. Natl Acad. Sci. USA* **103,** 1864–1869 (2006).
99. Guan, Y. *et al.* MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nature Struct. Biol.* **5,** 1058–1064 (1998).
100. Lukianova, O. L. & David, S. S. A role for iron–sulfur clusters in DNA repair. *Curr. Opin. Chem. Biol.* **9,** 145–151 (2005).


**Acknowledgements** We thank members of the David laboratory for reading the manuscript. We also apologize to all scientists whose original studies and reviews were not included because of space limitations. Research in the laboratory of S.S.D. is funded by the National Institutes of Health, and V.L.O. has been supported by pre-doctoral fellowships from the National Institutes of Health.

**Author Information** Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to S.S.D. (david@chem.ucdavis.edu).