REVIEW ARTICLE

DNA glycosylases in the base excision repair of DNA

Hans E. KROKAN*, Rune STANDAL and Geir SLUPPHAUG

UNIGEN Center for Molecular Biology, The Medical Faculty, Norwegian University of Science and Technology, N-7005 Trondheim, Norway

A wide range of cytotoxic and mutagenic DNA bases are removed by different DNA glycosylases, which initiate the base excision repair pathway. DNA glycosylases cleave the N-glycosylic bond between the target base and deoxyribose, thus releasing a free base and leaving an apurinic/apyrimidinic (AP) site. In addition, several DNA glycosylases are bifunctional, since they also display a lyase activity that cleaves the phosphodiester backbone 3′ to the AP site generated by the glycosylase activity. Structural data and sequence comparisons have identified common features among many of the DNA glycosylases. Their active sites have a structure that can only bind extrahelical target bases, as observed in the crystal structure of human uracil-DNA glycosylase in a complex with double-stranded DNA. Nucleotide flipping is apparently actively facilitated by the enzyme. With bacteriophage T4 endonuclease V, a pyrimidine-dimer glycosylase, the enzyme gains access to the target base by flipping out an adenine opposite to the dimer. A conserved helix–hairpin–helix motif and an invariant Asp residue are found in the active sites of more than 20 monofunctional and bifunctional DNA glycosylases. In bifunctional DNA glycosylases, the conserved Asp is thought to deprotonate a conserved Lys, forming an amine nucleophile. The nucleophile forms a covalent intermediate (Schiff base) with the deoxyribose anomeric carbon and expels the base. Deoxyribose subsequently undergoes several transformations, resulting in strand cleavage and regeneration of the free enzyme. The catalytic mechanism of monofunctional glycosylases does not involve covalent intermediates. Instead the conserved Asp residue may activate a water molecule which acts as the attacking nucleophile.

INTRODUCTION

The structural integrity of DNA is continuously challenged by a number of exogenous and endogenous agents, as well as by some cellular processes [1]. To counteract these threats, cells have several defence mechanisms that act at different levels to prevent or repair damage or to eliminate damaged cells. DNA damage causes a temporary arrest of cell-cycle progression, allowing DNA repair to take place prior to replication. DNA repair is thus only one of several processes that co-operate to maintain the integrity of the genome (Figure 1). A large number of factors co-ordinate the complex interplay between different cellular processes.

Damaged bases may be cytotoxic, miscoding or both. Mutations resulting from miscoding are thought to be a major mechanism by which DNA-reactive agents cause diseases such as cancer. In the simplest type of DNA repair, the damaged base is repaired directly, e.g. dealkylated, by a one-step mechanism, rather than being excised and replaced by the correct one. However, most lesions in DNA are repaired by the much more complex recombination repair or excision repair systems. The latter include nucleotide excision repair (NER), mismatch repair and base excision repair (BER). NER is the most complicated of the excision repair systems, and involves the products of about 30 genes. The system acts upon a wide range of alterations that result in large local distortions in DNA. The damage is removed as part of an oligonucleotide, and new DNA is synthesized using the intact strand as a template. Mismatch repair refers to the repair of mispaired bases in DNA, and may occur by several biochemical pathways, including NER and BER.

As the name implies, the initial step in BER is the removal of a base rather than a nucleotide. This step is carried out by DNA glycosylases, which are the focus of the present review. In general, the damaged or mismatched base recognized by a DNA glycosylase does not cause major helix distortions, although the damage may be caused by a variety of agents and processes, such as spontaneous deamination of bases, radiation, oxidative stress, alkylating agents or replication errors. BER is quantitatively probably the most important mechanism of DNA repair, yet no disease has so far been clearly related to deficiencies in this pathway. Most DNA glycosylases remove several structurally different damaged bases, while a few have very narrow substrate specificities. The information on enzymes involved in BER has flourished in the last few years. This is particularly true for some of the DNA glycosylases, which are now among the best understood enzymes involved in nucleic acid metabolism. These studies have also contributed new concepts about the interactions between proteins and DNA in general.

DNA N-GLYCOSYLASES AND THE BER PATHWAY

DNA N-glycosylases hydrolyse the N-glycosylic bond between the target base and deoxyribose, thus releasing a free base and leaving an apurinic/apyrimidinic (AP) site in DNA [2]. Such AP sites are cytotoxic and mutagenic, and must be further processed [3]. Some DNA glycosylases also have an associated AP lyase activity that cleaves the phosphodiester bond 3′ to the AP site (β-lyase). DNA glycosylases are relatively small monomeric proteins that do not require cofactors for their activity, and they are

Abbreviations used: AP site, apurinic/apyrimidinic site; BER, base excision repair; dRpase, deoxyribophosphodiesterase; dsDNA, double-stranded DNA; EndoIII (etc.), endonuclease III (etc.); fapy, 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine; Fpg, formamidopyrimidine-DNA glycosylase; GPD motif, Pro/Gly-rich stretch with a conserved Asp residue C-terminal to it; HhH, helix–hairpin–helix; hmUra, 5-hydroxymethyluracil; 3-meA (etc.), 3-methyladenine (etc.); MGMT, O^6-alkylguanine-DNA alkyltransferase; MPG, 3-meA-DNA glycosylase (in plants and mammals); NER, nucleotide excision repair; 8-oxoG, 7,8-dihydro-8-oxoguanine; RPA, replication protein A; T4endoV, bacteriophage T4 endonuclease V; TDG, thymine-DNA glycosylase; UDG, uracil-DNA glycosylase; XRCC1, X-ray cross-complementation protein 1.

* To whom correspondence should be addressed.

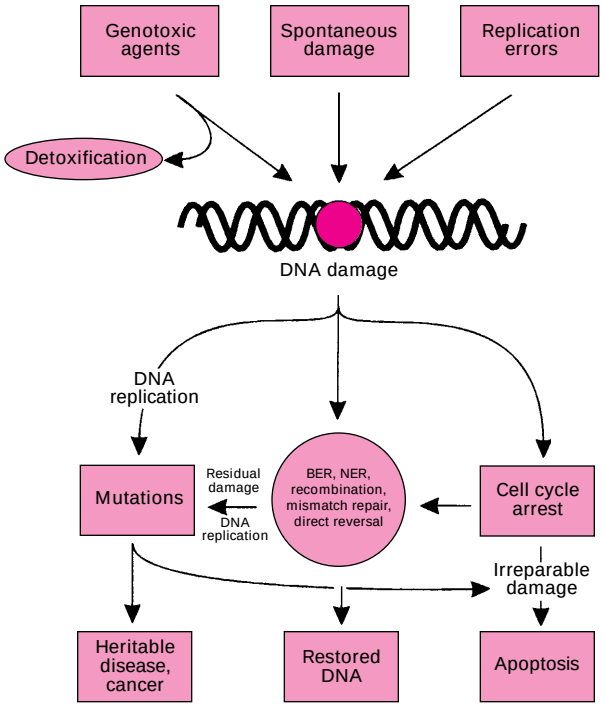


Figure 1 Cellular mechanisms contributing to the maintenance of the genome

therefore excellent models for studying interactions between damaged DNA and proteins. In Figure 2, typical base lesions and DNA glycosylases involved in their removal are shown.

In vitro reconstitution of the BER pathway with cell-free extracts or purified components has been carried out using uracil-DNA glycosylase (UDG) [4–6] or mismatch-specific thymine-DNA glycosylase (TDG) [7] as the initiating enzyme. These studies have established the minimal enzymic requirements for the repair of the resulting AP sites, and indicate that the repair may proceed via two alternative pathways (Figure 3). In the first of these, a 5′-AP endonuclease [8] and a deoxyribophosphodiesterase (dRpase) create a single nucleotide gap, which in eukaryotic cells is filled in by DNA polymerase β and DNA ligase III [9]. This route may also be initiated by bifunctional DNA glycosylases, in which the gap may be created by successive β- and δ-eliminations [10,11]. A role for the presumably non-catalytic XRCC1 (X-ray cross-complementation protein 1) in BER was recently also established. This protein interacts through its N-terminal half with DNA polymerase β, whereas the C-terminal region interacts with DNA ligase III [9,12]. DNA strand displacement and excessive gap-filling were observed in cell-free extracts lacking XRCC1, indicating that XRCC1 might serve as a scaffold protein during BER. Moreover, human UDG was recently found to interact with both the 34 kDa subunit and the trimeric form of replication protein A (RPA) [13]. In the alternative BER pathway, a short patch containing the abasic site is excised and replaced by normal nucleotides [14,15]. In eukaryotes this nucleotide patch is apparently displaced by polymerase δ or ϵ, and the resulting overhang is removed by the flap-endonuclease FEN-1 in a process involving PCNA (proliferating cell nuclear antigen) [16].

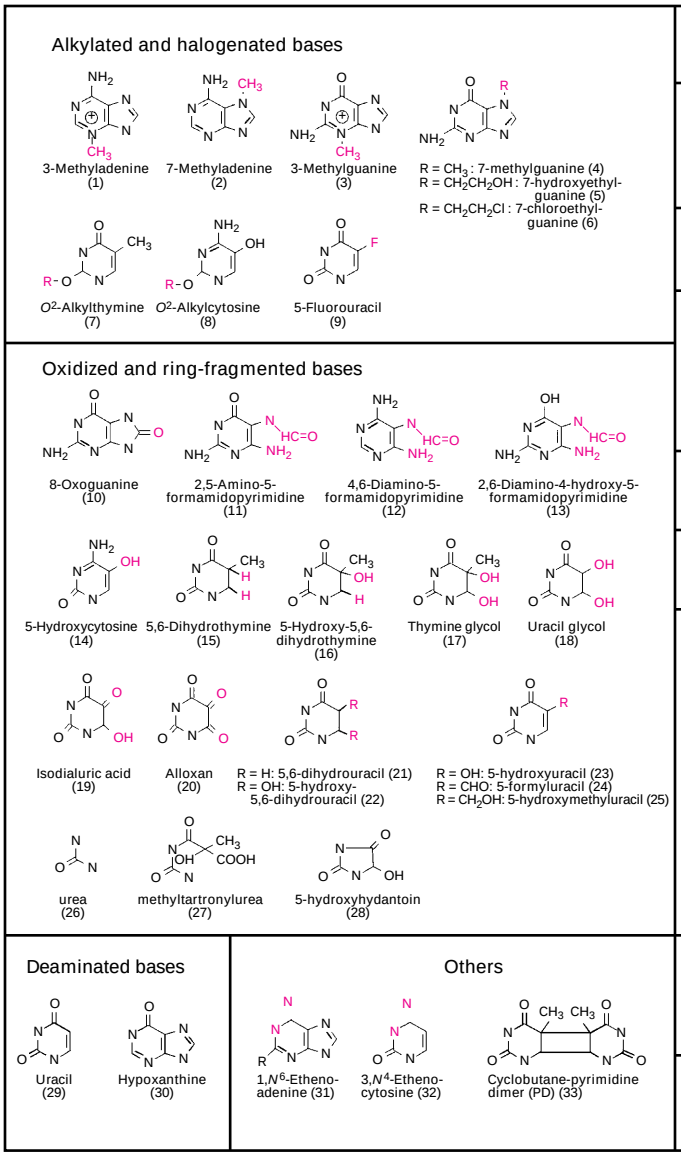
In *Escherichia coli*, the AP-endonucleolytic step is catalysed by endonuclease IV (EndoIV), which is damage-inducible, or by exonuclease III, which is constitutively expressed [1]. Exonuclease III was originally characterized as an exonuclease, but was later shown to be a multifunctional enzyme which probably functions primarily as a 5′ AP-endonuclease [17]. The dRpase activity is a property of the *recJ* gene product [11]. At least one *E. coli* DNA glycosylase, the formamidopyrimidine-DNA glycosylase (Fpg) protein, may also create a DNA gap alone. Fpg releases both some purine bases with damaged imidazole rings and a deoxyribose derivative, leaving a gap bordered by 5′- and 3′-phosphoryl groups [10]. Further processing requires a 3′-phosphatase, probably contributed by EndoIV [18,19], to produce a primer for DNA polymerase. If an AP site has already been incised by an AP-endonuclease 5′-terminal to the deoxyribose, Fpg also promotes release of deoxyribophosphate [10,11]. A reported dRpase activity associated with *E. coli* exonuclease I [20] has not been reproduced by others [11]. Studies with double mutants deficient in RecJ and Fpg have indicated that the dRpase activities of these enzymes are not essential for BER, indicating the presence of a back-up function so far not identified [11]. The gaps resulting from dRpase activities are filled by DNA polymerase I and the single DNA ligase present in *E. coli* [6].

UDGs

UDG from *E. coli* was the first DNA glycosylase to be discovered. It was discovered as a consequence of a search for an enzymic activity that would recognize uracil resulting from the deamination of cytosine, a process that introduces pre-mutagenic U/G mispairs [21]. Subsequently similar enzyme activities were demonstrated in other bacteria, yeast, plants, mammalian cells and mitochondria [1]. UDGs are highly conserved in evolution and, except for UDGs from pox viruses, the active site is completely conserved. Uracil is removed both from U:A pairs resulting from misincorporation of dUMP during replication and from mutagenic U/G mispairs resulting from the deamination of cytosine [1]. The latter is estimated to result in some 100–500 uracil residues per mammalian genome per day [22]. Another possible source of uracil in DNA is the enzymic deamination of cytosine by (cytosine-5)-methyltransferase under certain conditions [23]. *E. coli* and yeast mutants in UDG have, depending on the sequence context, 4–30-fold increases in G:C → A:T transition mutations [24,25]. It is believed that the primary function of UDG is to remove uracil from U/G mispairs resulting from the deamination of cytosine [1,25].

Although UDG is highly selective for uracil in DNA, it does remove certain closely related bases at rates some three orders of magnitude lower than that with uracil. These include 5-fluorouracil in DNA found after treatment with 5-fluorouracil [26], as well as isodialuric acid, 5-hydroxyuracil and alloxan, which are all formed from cytosine in DNA after exposure to γ-irradiation or oxidative stress [27–29]. Uracil at the 3′-end of a DNA chain is not removed by human [30] or *E. coli* [31] UDGs, whereas uracil at the 5′-end is removed provided that it is phosphorylated. Consistent with this, the minimal substrate for UDG was found to be pd(UN)p [31]. *In vitro* experiments have indicated a processive mechanism for substrate recognition. In this mechanism UDG slides along the DNA and scans the macromolecule for uracil residues. This applies to both bacterial and mammalian UDGs [32,33], although the mechanism shifts from a processive to a distributive mode when the salt concentration is increased [32].

Variations in both damage induction and DNA repair in different sequence contexts may contribute to a non-random distribution of mutations. Interestingly, UDGs from human, bovine and bacterial sources remove uracil at different rates from different double-stranded DNA (dsDNA) sequence contexts, but



Enzyme	Source/gene	Reported DNA substrates	β-Lyase activity
Uracil-DNA glycosylases (UDGs)	Viral	29	No
	Bacterial (*ung*)	9, 19, 23, 29	No
	S. cerevisiae (*UNG1*)	29	No
	Plants	29	No
	Human (*UNG*)	19, 20, 23, 29	No
G/T(U)mismatch-DNA glycosylases	*M. thermoautotropicum*	T, 29, G/G, A/G, T/C, U/C	No
	Insects	29	No
	Human	T, 29	No
Alkylbase-DNA glycosylases	*E. coli* (*tag*)	1, 3	No
	E. coli (*alkA*)	1–8, 24, 25, 30–32 (G, A)	No
	S. cerevisiae (*MAG*)	1–6, 30, 31	No
	S. pombe (*mag1*)	1	?
	A. thaliana (*MPG*)	1	?
	Rodent/human (*MPG*)	1–6, 10, 30, 31	No
5-Methylcytosine-DNA glycosylase	Chick	5-MeC, T in G/T	No
	Human	5-MeC	No
Adenine-specific mismatch-DNA glycosylases	*E. coli* (*mutY*)	A in G/A, 10/A, C/A	Yes/no
	Bovine, human (*MYH*)	A in G/A, 10/A, C/A	Yes?
DNA glycosylases removing oxidized pyrimidines (EndoIII-like)	*E. coli* EndoIII (*nth*)	14–18, 20–23, 26–28	Yes
	S. cerevisiae (*NTG1*)	11, 12, 13, 17	Yes
	S. pombe (*nth*)	17, 26	Yes
	Bovine/human EndoIII homologue	17, 26	Yes
EndoVIII	*E. coli*	15, 17	Yes
EndoIX	*E. coli*	26	?
Hydroxymethyl-DNA glycosylase	Mouse	25, 29	No
	Bovine	25	Yes?
Formyluracil-DNA glycosylase	Human	24	?
DNA glycosylases removing oxidized purines	*E. coli* (*fpg*)	10–13	Yes
	S. cerevisiae (*OGG1*)	10 (opposite C or T)	Yes
	S. cerevisiae (*OGG2*)	10 (opposite G or A)	Yes
	D. melanogaster S3	10	Yes
Pyrimidine-dimer-DNA glycosylases	T4	33, 12	Yes
	M. luteus (*pdg*)	33	Yes
	N. mucosa	33	Yes

Figure 2 Typical damaged DNA bases and DNA glycosylases acting upon them

Chemical groups not found in the normal DNA bases are shown in red. The reported substrates for the various DNA glycosylases are numbered in accordance with the left panel. For DNA glycosylases recognizing multiple substrates, the number of the proposed preferred substrate is shown in red. Organisms: *S. cerevisiae*, *Saccharomyces cerevisiae*; *M. thermoautotrophicum*, *Methanobacterium thermoautotrophicum*; *S. pombe*, *Schizosaccharomyces pombe*; *A. thaliana*, *Arabidopsis thaliana*; *D. melanogaster*, *Drosophila melanogaster*; *M. luteus*, *Micrococcus luteus*; *N. mucosa*, *Neisseria mucosa*.

at essentially similar rates from single-stranded DNA. The rate differs as much as 20-fold between the 'best' (A/T/UAA) and 'worst' (G/CUG/C/T) sequences. Usually, but not always [34], the rate of removal is greater from U/G mispairs than from U:A pairs [35–37]. A low removal rate tends to be correlated with the occurrence of 'hot spots' for mutation [36], similar to the correlation between slow spots for the repair of cyclobutyl dimers and hot spots for mutations [38,39].

UDGs belong to a highly conserved ancient family

The cloning of genes or cDNAs for UDGs from bacteria [40], yeast [41], the fish *Xiphophorus* [42], mouse [43] and human cells [43,44] and several herpes viruses [45–52] has demonstrated a striking similarity between these enzymes, ranging from 40.3% (yeast) to 90% (mouse) amino acid identity relative to human UDG. The similarity is confined to several discrete boxes. The exon–intron boundaries in UDGs from human, mouse and *Xiphophorus* are completely conserved, indicating a conservation of exon–intron organization for more than 450 million years [42]. Phylogenetic analysis of UDG and other protein sequences from mammalian members of the family *Herpesviridae* distinguished the three recognized subfamilies, and it was estimated that the three subfamilies arose approx. 180–200 million years ago [53]. UDGs from pox viruses [54–58] are more distantly related, but the active-site region is also highly conserved in these viruses.

The gene for human UDG (*UNG*) has been isolated from P1 clones that contain the complete gene. *UNG* spans approx. 13.5 kb, comprises seven exons, and was assigned to chromosome 12q23-q24.1 by radiation hybrid mapping [43,59]. The promoter

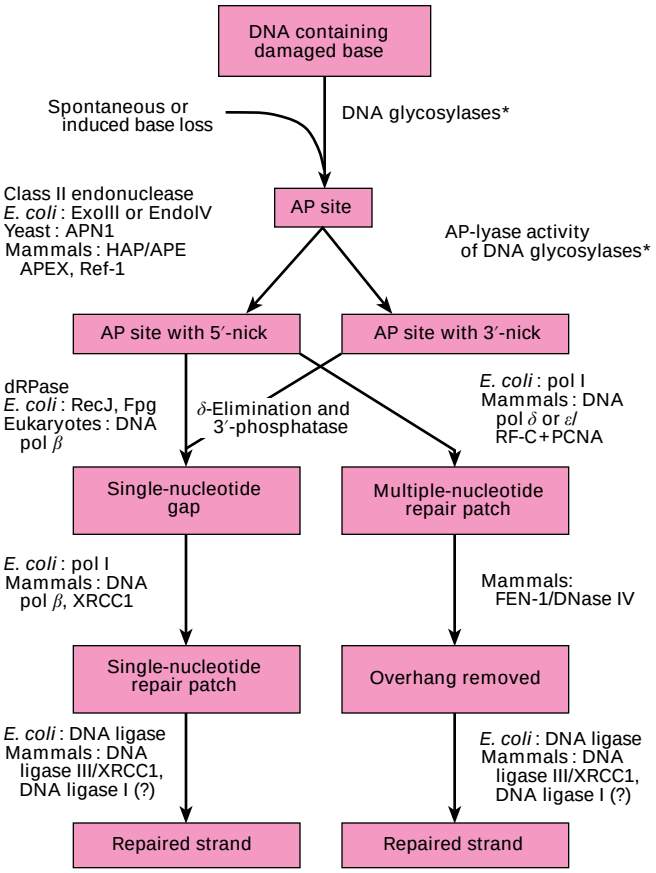


Figure 3 Alternative pathways in BER

Enzymes that are reported to be involved in the individual steps are shown. * See Figure 2. Abbreviations: Exo, exonuclease; Endo, endonuclease; pol, polymerase; APE/APEX/Ref-1, alternative names for the major mammalian AP-endonuclease; HAP, human AP-endonuclease; APN1, yeast AP-endonuclease; PCNA, proliferating cell nuclear antigen; FEN, flap-endonuclease.

region and exons 1A and 1 constitute a CpG island. A core region rich in putative transcription factor binding elements is unmethylated. The *UNG* gene contains a number of different repetitive elements of unknown significance. Alternative splicing and transcription from two different GC-rich and TATA-less promoters (P_A and P_B) in the *UNG* gene result in distinct mitochondrial and nuclear forms of UDG, as described below.

Similarities between the nuclear and mitochondrial forms of UDG

The presence of UDG activity [60–64] and of an AP-endonuclease activity [65] in mitochondria indicates that mitochondria are proficient in the repair of uracil-containing DNA. Mutational inactivation of the nuclear yeast UDG without affecting mitochondrial UDG activity [66], as well as different biochemical properties [61], have indicated that these enzymes are encoded by different genes. However, early studies demonstrated that the two forms are inhibited to the same extent by the very selective protein inhibitor Ugi. This was, in fact, the first indication that UDGs from different sources are structurally related [67]. Ugi forms an essentially irreversible complex with UDG [68,69], and the specificity of the interaction [70,71] indicates that only structurally related UDGs will bind to Ugi. Furthermore, antibodies raised against homogeneous human UDG detected both mitochondrial and nuclear forms, indicating that these forms are closely related [37,64,72].

Recently it was demonstrated that mRNAs for the nuclear and mitochondrial forms of mammalian UDG are generated from the *UNG* gene by alternative transcription start points and alternative splicing. The nuclear form (UNG2; 313 amino acids) and the mitochondrial form (UNG1; 304 amino acids) differ in their N-terminal sequences that direct nuclear and mitochondrial import respectively, but they have identical catalytic domains [43]. To our knowledge, this is the only known example of this mechanism for generating mitochondrial and nuclear forms of enzymes from one gene.

Regulation of UDG expression

In general, UDG activity is higher in proliferating cells than in non-cycling cells [73–75]. Furthermore, the induction of DNA synthesis in resting lymphocytes increases UDG activity several-fold [76,77]. Using synchronized cell cultures, it was demonstrated that UDG in human fibroblasts is cell-cycle-regulated [78,79]. The mRNA for UDG increases 8–12-fold late in G_1-phase, whereas enzyme activity increases just prior to S-phase and reaches a maximum early in S-phase. This induction is mainly regulated at the transcriptional level [79]. Several putative transcription factor binding elements often involved in cell cycle regulation, such as E2F, c-Myc and Yi, are present in the P_B promoter of human *UNG*. Mutational studies have indicated that the c-Myc element is a positive regulator of UDG expression, whereas the E2F element (and overexpression of E2F) regulates UDG expression negatively. In addition, Sp1 elements are present in both the P_A and the P_B promoters, and are required for effective expression ([80]; T. Haug, P. A. Aas, F. Skorpen, V. Malm, C. Skjeldbred and H. E. Krokan, unpublished work). The nuclear and mitochondrial forms of UDG are probably not equally regulated, since their promoters (P_A and P_B) are structurally different.

Structure–function analyses of UDG

The elucidation of the crystal structures and mutational studies of human [70,81–83] and herpes-viral [84] UDGs have demonstrated the mechanism for the selective binding of uracil over the structurally closely related normal pyrimidine in DNA, and have revealed the catalytic mechanism. DNA binds along a positively charged groove in the enzyme, but the tight-fitting uracil-binding pocket located at the base of this groove is too deep and narrow to allow binding of DNA-uracil unless it is 'flipped out' of the DNA helix. The complex of human UDG and uracil-containing DNA has demonstrated the basis for the enzyme-assisted nucleotide flipping [83]. The three-dimensional structure of human UDG is very similar to the structures of the Herpes simplex virus type-1 (HSV-1) enzyme [84] and UDG from *E. coli* (C. D. Mol, D. W. Mosbaugh and J. A. Tainer, personal communication), and consists of a single α/β domain containing eight α-helices and a central four-stranded parallel and twisted β-sheet. The UDG–DNA structure has been studied using a double mutant ($Asp^{145} \rightarrow Asn/Leu^{272} \rightarrow Arg$) of human UDG. In this UDG–DNA complex [83] the positively charged groove traversing the UDG surface orients the enzyme active site along the DNA. A conserved leucine (Leu^{272} in the human wild-type enzyme, but Arg^{272} in the mutant studied) located directly above the buried uracil-binding pocket aids in minor-groove scanning and expulsion ('push') of the dUMP residue from the dsDNA base stack via the major groove. A concomitant compression of the DNA backbone phosphates flanking the uracil and specific

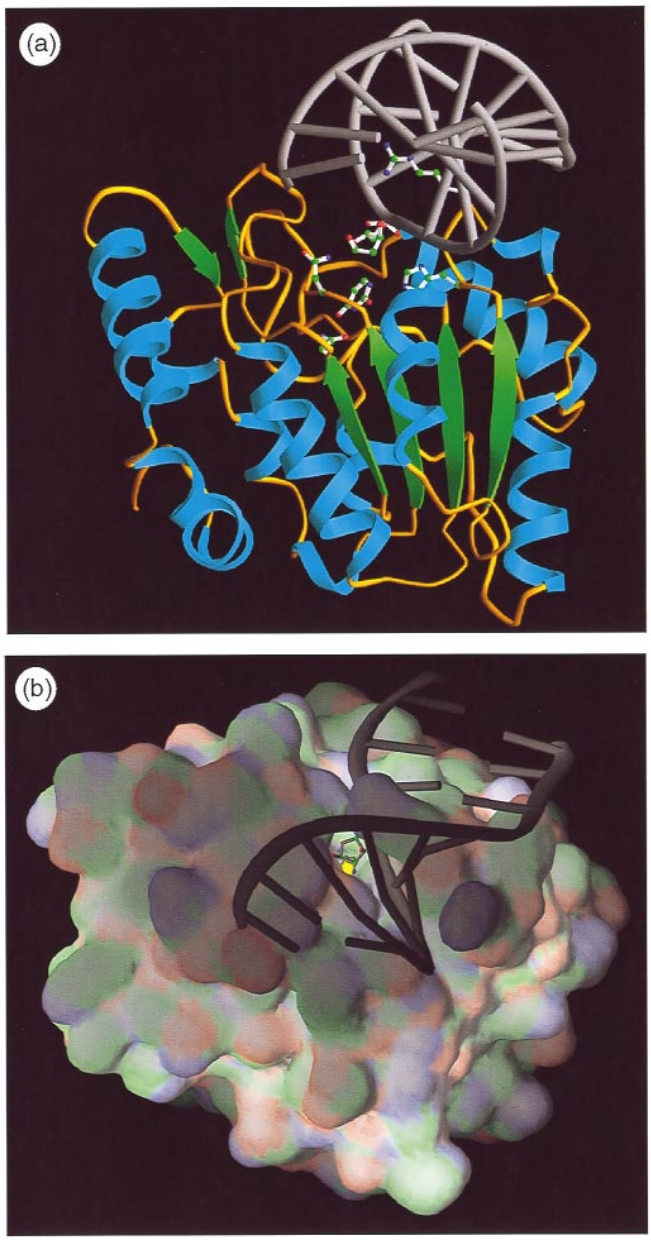


Figure 4 Uracil is actively flipped out from the dsDNA helix by UDG

(**a**) UDG–DNA complex viewed from the 5′-end of the uracil-containing strand (grey tubes) with the DNA helical axis nearly perpendicular to the UDG central β-sheet (green arrows). Catalytically important residues (His^{268}; mutants Asn^{145}, Asn^{204}) and the extrahelical deoxyribose and uracil are in white. The mutant Arg^{272} side chain penetrates the helix from the minor groove, thus stabilizing the extrahelical conformation of the uracil nucleotide. (**b**) UDG–DNA complex viewed from above the DNA-binding groove and down into the uracil-binding pocket occupied by the flipped-out uracil (yellow) and abasic sugar. The UDG surface is coloured according to atom type (N, blue; O, red; C, green).

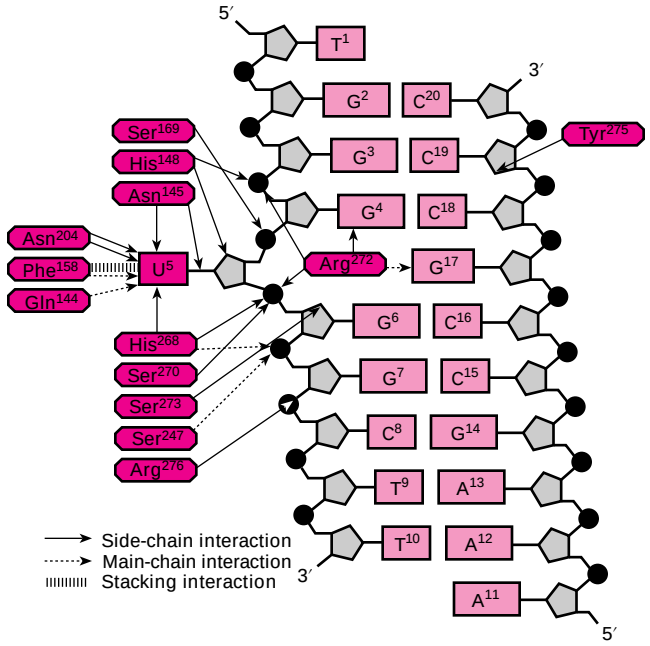


Figure 5 UDG–DNA interactions

Nucleotide numbering follows the 5′–3′direction, starting from the uracil-containing strand. DNA bases are shown in light pink, deoxyriboses in grey and phosphates in black. Amino acids (dark pink) interact with the DNA primarily along the sugar–phosphate backbone surrounding uracil.

recognition of the 5′-phosphate, deoxyribose and uracil by UDG active-site residues ('pull') stabilizes the extrahelical nucleotide conformation, and promotes concerted condensation of the surrounding catalytic residues to form a productive complex specific for uracil cleavage [83]. In the UDG–DNA complex both uracil and the deoxyribose phosphate are rotated by nearly 180 ° from their normal positions, and the mechanism should thus be described as 'nucleotide flipping' (Figures 4a and 4b). The conserved and buried uracil-binding pocket is characterized by extensive shape and electrostatic complementarity to uracil. Several hydrogen bonds are observed between conserved amino acid residues and the 2-, 3- and 4-positions in uracil (Figure 5). A conserved Phe stacks with uracil, while a conserved Tyr (Tyr^{147} in the human enzyme) sterically hinders the entry of bases containing bulky substituents at the 5-position (such as the 5-methyl group of thymine) and excludes the entry of water. There is, however, sufficient space to allow small substituents such as hydroxy or carbonyl groups at the 5- or 6-positions. Thus certain uracil analogues generated from cytosine by γ-irradiation are recognized by UDG, but are excised at considerably lower rates than uracil [27,29]. This steric shielding has been verified by site-specific mutagenesis of Tyr^{147} of human UDG to Ala [82], which results in a mutant that excises thymine as well as uracil from DNA (Figure 6). Furthermore, the replacement of Asn^{204} of human UDG by Asp allows the binding and excision of cytosine in addition to uracil [82]. The specificity against uracil in RNA resides in the steric hindrance of hydrogen-bond formation between the catalytic His^{268} and uracil O-2 resulting from the ribose 2-hydroxy and 3′-*endo* puckering which blocks His^{268} movement.

Site-directed mutagenesis of human UDG has demonstrated that effective catalysis is critically dependent upon Gln^{144}, Asp^{145}, Tyr^{147}, Asn^{204} and His^{268}. In the buried active site, Asp^{145} may rotate towards bound uracil and may activate a water molecule, with the resultant hydroxy nucleophile making a direct in-line attack on deoxyribose C-1′, as suggested in [84]. Gln^{144} and Tyr^{147} shield the active site from the bulk solvent, and the severely reduced k_{cat} of a $Tyr^{147} \rightarrow Ala$ mutant suggests that bulk-water exclusion is important for catalysis. His^{268} interacts with the uracil 3′-phosphate, thus delivering a charged hydrogen

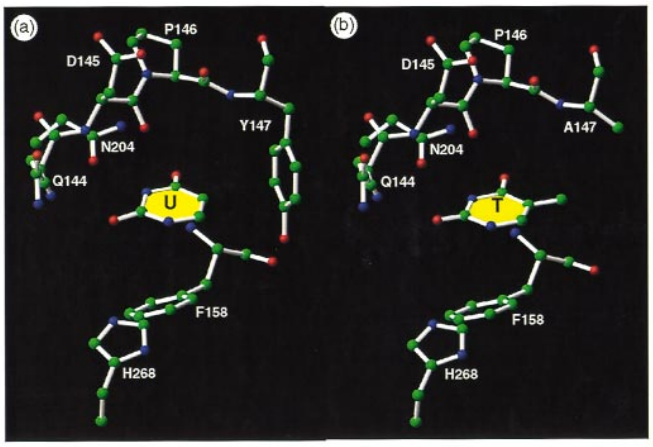


Figure 6 Design of the UDG active-site pocket

(**a**) Extensive steric and electrostatic complementarity between uracil and the uracil-binding pocket explains the narrow substrate specificity of UDG. Tyr^{147} (Y147) sterically hinders the entry of bases containing bulky substituents at their 5-positions, and Asn^{204} (N204) confers selectivity for uracil over cytosine. (**b**) Substituting Ala for Tyr^{147} opens the pocket and allows the entry of thymine, and this TDG confers a mutator phenotype when expressed in *E. coli*.

bond to uracil O-2 that can facilitate bond cleavage by stabilizing the developing oxyanion.

The similar conformations of the flipped-out uracil nucleotide [83] and dCMP in the structures of two bacterial DNA deoxycytidine methyltransferases bound to DNA [85,86] suggest that flipping of the target nucleotide might be actively facilitated by several classes of DNA-modifying enzymes (reviewed in [87]). The DNA deoxycytidine methyltransferases, however, interact with DNA primarily through the major groove, and flip the nucleotide out via the minor groove. Which groove is employed might thus be specific for each enzyme depending on the structural determinants characteristic of each type of DNA damage.

DNA GLYCOSYLASES SPECIFIC FOR MISMATCHES

In addition to the well characterized MutH/MutL/MutS nucleotide excision system in *E. coli* and an analogous system in mammalian cells [88], as well as a very-short-patch system for the repair of G/T(U) in *E. coli* [88], bacteria and eukaryotic cells have DNA glycosylases for the repair of different types of single-base mismatches.

G/T(U)-mismatch DNA glycosylases

A mismatch-specific DNA glycosylase (TDG) removing T from G/T mispairs was originally detected in simian cell extracts [89], and the corresponding human activity was subsequently demonstrated to be a DNA glycosylase [7]. TDG removes thymine from G/T mispairs in a CpG context, although G/T mispairs in other sequence contexts and thymine opposite O^6-methylguanine, cytosine and thymine are also substrates [90]. Interestingly, the enzyme excises uracil from G/U mispairs more efficiently than it excises thymine from G/T mispairs, whereas neither U nor T in single-stranded DNA nor U:A are substrates. Compared with the specialized and efficient UDGs, the mismatch-specific UDG has a very low turnover number [90]. cDNA for the human mismatch-TDG gene has been cloned and has no significant identity with the coding sequences of other known DNA-metabolizing enzymes. Deletion of amino acids from the C-terminal and N-terminal ends of the human protein resulted in a core enzyme of 248 amino acids that had lost TDG activity but retained double-strand-specific UDG activity. Interestingly, homologues of this core enzyme are present in bacteria as well as in insect cells, and it thus appears to be an ancient enzyme [91]. It may be subservient to the catalytically more efficient form of UDG present in most organisms, but may constitute a first-line defence against the effects of cytosine deamination in insects. Recently, a G/T(U) mismatch DNA glycosylase was also cloned from the thermophile *Methanobacterium thermoautotrophicum*, but this enzyme is not significantly related to the human mismatch glycosylase [92].

A/G[7,8-dihydro-8-oxoguanine (8-oxoG)]-mismatch DNA glycosylase (MutY)

An *E. coli* DNA mismatch glycosylase (MutY) that removes adenine from A/G mismatches was originally identified as a gene that prevented C:G → A:T transversion mutations [93,94]. The mutator locus was designated *mutY*, and the repair activity was found to be independent of the methylation state of the DNA [94,95]. *mutY* encodes a DNA glycosylase of 36 kDa which, in addition to excision of A opposite G and C, also removes A opposite the oxidized purines 8-oxoG and 7,8-dihydro-8-oxoadenine (8-oxoA) [96]. If 8-oxoG is not removed prior to replication, a C or an A is inserted opposite it. Thus MutY apparently serves an important function in protection against oxidative damage by removing adenine mispaired to 8-oxoG after replication. The substrate specificity of MutY has been extensively studied, but it is not clear whether A/G or A/8-oxoG is the preferred substrate [97,98]. As described below, several glycosylases take part in the repair of oxidized bases.

The amino acid sequence (350 residues) of MutY shares 66.3 % similarity and 23.8 % identity with that of *E. coli* endonuclease III (EndoIII), a DNA glycosylase/AP lyase that recognizes oxidized and ring-fragmented pyrimidines. The similarity spans a 181-amino-acid region [99], suggesting that these proteins have a common evolutionary origin. MutY also shows sequence similarity to the *Micrococcus luteus* pyrimidine-dimer DNA glycosylase [100]. Proteolytic cleavage of MutY yields 13 and 26 kDa fragments, the latter of almost equivalent size to EndoIII. Interestingly, the 26 kDa fragment retains normal DNA binding and adenine-DNA glycosylase/β-lyase activity against G/A, whereas the activity is dramatically decreased against 8-oxoG/A [101]. Homology modelling has indicated that the 26 kDa proteolytic fragment of MutY has the same overall conformation as EndoIII [102], and structural similarity is consistent with preliminary crystallographic studies (Y. Guan and J. A. Tainer, personal communication). EndoIII and MutY are likely to have related catalytic mechanisms, although their substrate specificities are very different.

Functional homologues of MutY have been detected in calf thymus [103] and HeLa cell extracts [104], and designated MYH. Bovine MYH is a 65 kDa protein that is apparently degraded to a functional 36 kDa species [103]. The enzyme removes mispaired A from G/A, C/A and 8-oxoG/A mismatches, with 8-oxoG/A being the best substrate. In addition, an associated or co-purified endonuclease nicks the phosphodiester bond 3′ to the AP site generated by the N-glycosylase activity. Structural homology to MutY is suggested by recognition of MYH and inhibition of the AP-nicking activity by anti-MutY antibodies, and inhibition by potassium ferricyanide, which oxidizes Fe–S clusters [104]. Recently a human homologue (hMYH) with 41 % identity to the *mutY* gene was cloned and sequenced. The gene maps on the short arm of chromosome 1 between p32.1 and p34.3, contains 16 exons encoding 535 amino acids and is 7.1 kb long [105].

DNA GLYCOSYLASES FOR ALKYLATED BASES

Historically, DNA glycosylases removing alkylated bases were called 3-methyladenine (3-meA)-DNA glycosylases, because this was the first substrate identified [106]. The substrate specificities of these enzymes are, however, usually much wider (Figure 2). 3-MeA has been demonstrated to be a major cytotoxic and mutagenic DNA lesion [107,108], although it is estimated to be some 40-fold less mutagenic than O^6-methylguanine (O^6-meG) [108], which is directly dealkylated by an alkyltransferase.

Alkylating agents are widely present in the environment and are also formed endogenously. A number of alkylating anti-cancer drugs are used routinely, and the efficacy of these may be modified by DNA repair processes. Methyl chloride, used in industrial processes and produced even more abundantly in Nature, has been shown to alkylate DNA [109] and to induce repair responses [110,111]. *N*-Nitroso compounds also exert their mutagenic effects through alkylation of DNA. Such compounds are formed endogenously [112], although the most significant human exposure may be from tobacco-specific nitrosamines [113,114]. In addition, the cellular methyl donor *S*-adenosylmethionine has been shown to alkylate DNA directly, indicating a potential intracellular source of alkylating agents [115]. More indirect evidence also suggests endogenous sources of alkylating agents [116].

N-glycosylases that excise alkylated bases have been identified in *E. coli* (Tag and AlkA), [117–119], other bacteria [120,121], *Saccharomyces cerevisiae* (MAG) [122–124], *Schizosaccharomyces pombe* (Mag1) [125], plants [3-meA-DNA glycosylase (MPG)] [126] and mammalian cells (MPG) [127–133]. These studies have demonstrated that AlkA and yeast glycosylases are related, whereas plant and mammalian alkyl-DNA glycosylases constitute a different family.

Substrate specificities of DNA glycosylases for alkylated bases

The Tag protein in *E. coli* is fairly specific for 3-meA, although it also removes 3-meG with much lower efficiency [134,135] (Figure 2). No other known alkylbase-DNA glycosylase has a similarly narrow substrate specificity. In contrast, AlkA, the other *E. coli* enzyme, has the broadest substrate specificity of all known DNA glycosylases. It is the only alkylbase-DNA glycosylase known to remove both damaged purines and pyrimidines [91]. AlkA is also 10–20-fold more efficient than Tag in the removal of 3-meA from single-stranded DNA [136]. Many of the damaged bases recognized by AlkA carry a positive charge or a weakened glycosylic bond, and this represents the only obvious common characteristic of the substrates for AlkA. AlkA removes all adducts caused by simple monofunctional alkylating agents, except those repaired by the alkyltransferase Ada. AlkA substrates include 3-meA, 3-meG, 7-meG, 7-meA, O^2-alkylcytosine and O^2-alkylthymine [91]. Not only are these substrates different in the sense that they represent pyrimidines and purines but, in addition, the methyl group of 7-meG protrudes into the major groove, whereas the other alkyl groups protrude into the minor groove. AlkA also removes alkylation products of the bifunctional alkylating agent chloroethylnitrosourea, such as 7-hydroxyethylguanine, 7-chloroethylguanine and some minor alkylation products [137], as well as the cyclic etheno adducts 1,N^6-ethenoadenine, 1,N^2-ethenoguanine and 3,N^4-ethenocytosine induced by vinyl chloride or chloroacetaldehyde [138,139]. Adducts formed by sulphur mustard [140] and nitrogen mustards [141] are also recognized by AlkA. Finally, hypoxanthine [142], 5-formyluracil and 5-hydroxymethyluracil (hmUra) [143] are all removed by AlkA. Apparently, AlkA does not remove 8-oxoG. This lesion is instead removed by the DNA glycosylase Fpg (see below).

Although the substrate specificities of the yeast protein MAG and the mammalian MPGs do not include damaged pyrimidines, these enzymes seem to recognize most, if not all, of the other products that are substrates for AlkA. Apparently different cyclic etheno adducts of adenine and cytosine are removed by different glycosylases in human cells [144]. In addition, MPG [130,145,146], but apparently not MAG [91], removes 8-oxoG. However, even the closely related MPGs from human and mouse sources differ somewhat in their preferences. Thus mouse MPG removes 7-meG and 3-meG some 2–3-fold faster than does human MPG [147].

Genes for alkylbase-DNA glycosylases

Although AlkA and Tag both remove 3-meA efficiently, their genes (*alkA* and *tag*) are not related [91]. AlkA is, however, clearly related to the *S. cerevisiae* protein MAG [122–124] and to the *S. pombe* protein Mag1 [125]. In contrast, the single MPG identified in human, rat and murine cells is unrelated to bacterial and yeast glycosylases. MPGs from different mammalian sources are, however, closely related to each other [127–133] and to MPG from the higher plant *Arabidopsis thaliana* [126]. The human MPG gene is located on chromosome 16p close to the telomere [129,133], while the mouse MPG gene has a similar location in chromosome 11 [131]. Both genes are localized close to the α-globin gene cluster in a G+C-rich isochore. The human gene comprises five exons, the representation of which differs in the isolated cDNA clones mainly due to differences in the 5′-regions. The mouse MPG gene apparently has only four exons, and their sizes are clearly different from those of the human gene, indicating that the exon–intron boundaries are not conserved [130]. The mouse MPG promoter is TATA-less, but has a CAAT element and is GC-rich, with putative AP-2 elements and Sp1-complementary sequences [132]. The rat MPG gene promoter has also been characterized. Like the mouse promoter it is TATA-less, and it has a CAAT box as well as putative binding sites for the transcription factors Sp1, AP-2, Ets-1, PEA3, NF-1, p53, c-Myc, NF-κB and the glucocorticoid receptor [148].

Since bacteria have (at least) two genes for alkylbase-DNA glycosylases, eukaryotic cells would also be expected to have more than one gene. This was also indicated by biochemical heterogeneity observed in early studies [149]. This question is not yet settled, but the discovery of different, but closely related, cDNAs for human MPG [128–130], and their generation by alternative splicing and transcription from two promoters [133], might explain the observed biochemical heterogeneity. The alternative transcripts were found simultaneously in different cell lines and tissues, and a tissue-specific mode of expression of the two forms would therefore seem to be ruled out [150]. The functional implications of these findings are not known, but one could speculate that the different N-terminal sequences could serve a function in subcellular targeting, since such a mechanism has recently been demonstrated for human UDGs UNG1 and UNG2 [43].

Regulation of alkylbase-DNA glycosylases

Resistance to alkylating agents in *E. coli* is strongly and specifically induced by small amounts of alkylating agents due to the induction of AlkA, AlkB (unknown function) and Ada [O^6-alkylguanine-DNA alkyltransferase (MGMT)], which are

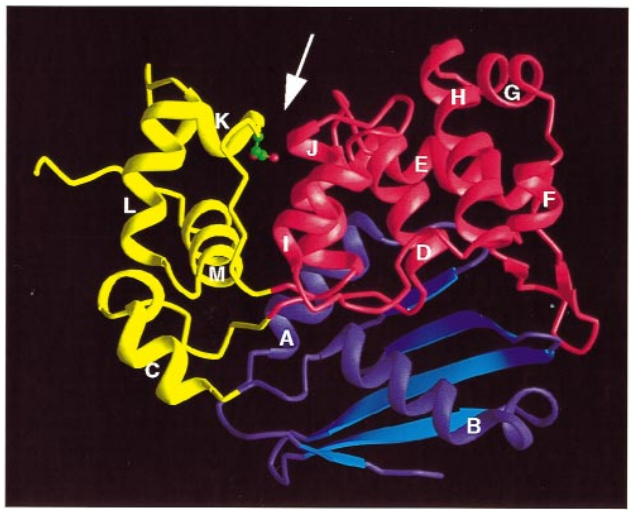


Figure 7 Structural domains of AlkA

Domains 1–3 are coloured blue, red and yellow respectively. The 13 α-helices are labelled A–M. The HhH motif comprises helices I and J. The conserved and catalytic Asp^{238} (green) resides in the region connecting helices J and K and protrudes into the putative DNA-binding cleft (white arrow).

encoded by genes in the same operon. In contrast, Tag is constitutively expressed [1]. The MAG protein of yeast is also inducible but, unlike the bacterial homologue, this induction is not specific for alkylating agents [151]. Mammalian MPGs seem to be essentially constitutively expressed [152,153], except in some rodent cells where MPG, as well as MGMT, are induced by a number of agents, including alkylating agents and X-rays [148,152,154,155].

MPG is cell-cycle-regulated, with the highest levels found just prior to and early in S-phase [77,152,156]. This is similar to the expression of UDGs [76,78,79], but very different from the expression of MGMT, which actually decreases significantly late in G_1-phase [152]. Although MPG is ubiquitously expressed, expression varies significantly in different tissues and cells. In mouse the highest levels were found in the stomach, and high levels were also found in the brain. This was somewhat surprising, because MGMT expression is highest in liver and very low in the brain. Thus the two different classes of enzymes responsible for repair of alkylation damage are differently regulated [152]. Possibly it can be argued that, in non-proliferating brain cells, cytotoxic lesions such as 3-meA that may block transcription are more harmful than the miscoding O^6-meG residues. Since the removal of alkylation lesions from DNA prior to replication is necessary to avoid miscoding and cytotoxicity, one would expect that both MGMT and MPG would be present at higher levels in fetal tissues and tissues from sucklings than in adult tissues. However, the opposite appears to be the case for both MGMT [157] and MPG [152]. One possible explanation for these unexpected findings could be that fetal tissues are efficiently protected from alkylating agents by the placenta.

Three-dimensional structure of AlkA

The recently described structure of the putative active site of AlkA is consistent with the broad substrate specificity of the enzyme. Its surface has a prominent cleft lined with electron-rich aromatic residues that may guide an extrahelical, positively charged, alkylated base into a position where it may be subject to nucleophilic attack by water deprotonated by an Asp residue. AlkA is a compact globular protein consisting of 13 α-helices and a five-stranded anti-parallel β-sheet [158,159]. The structure consists of three roughly equal-sized domains (Figure 7). The N-terminal 88 residues form domain 1, consisting of the five-stranded anti-parallel β-sheet at the surface flanked by two α-helices (A and B). Domain 1 is similar in shape and topology to the conserved tandem repeat of the TATA-binding protein, but is probably not functionally similar to this domain. A long peptide segment connecting domains 1 and 2 contains a small α-helix (C) that packs against domain 3. Domain 2 is a globular bundle of seven α-helices and contains a hydrophobic core. A pendulous loop connecting the second (E) and third (F) α-helices of domain 2 contains two short β-strands which run at almost 90 ° to one edge of the twisted β-sheet of domain 1, exposing three acidic residues to the solvent. The two final α-helices (I and J) of domain 2 are connected by a type II β-turn forming a conserved motif termed helix–hairpin–helix (HhH), which was first identified as the binding site for thymine glycol in crystals of EndoIII [160]. Residues at the C-terminal end of AlkA form three α-helices (K, L and M) that make up domain 3 together with α-helix C in the loop connecting domains 1 and 2. Interestingly, the structures of domains 2 and 3 can be superimposed on that of EndoIII, whereas domain 1 has no counterpart in EndoIII. Domain 1 is also absent from the *S. pombe* AlkA homologue Mag1 [125]. AlkA and Mag1 display 27 % identity and 63.5 % similarity, and are likely to be structurally related [91,129]. Although AlkA and EndoIII of *E. coli* are structurally related, sequence similarities are limited to the HhH region, and there is no strong reason to believe that they have a common ancestral origin. However, the amino acid sequences of EndoIII and MutY are 31 % and 22 % identical respectively to that of *Micrococcus luteus* pyrimidine-dimer glycosylase, and these are likely to have a common ancestral origin. In conclusion, structural data demonstrate substantial structural similarities between AlkA and EndoIII. Sequence information, as well as other data, indicate that MutY and the *M. luteus* pyrimidine-dimer glycosylase probably belong to the same structural family.

The conserved HhH motif and a unified catalytic mechanism for DNA glycosylases

Recent studies have provided compelling evidence that the HhH motif, together with a Pro/Gly-rich stretch and a conserved Asp residue C-terminal to it (GPD motif) [161], comprise the active site both in AlkA [158] and the bifunctional EndoIII [162]. The HhH motif has also been identified in several other DNA glycosylases [91,161,162] (Figure 8) and other DNA-binding proteins [163]. In both AlkA and EndoIII this HhH/GPD motif is located in the interdomain cleft, which is lined in AlkA with hydrophobic residues and in EndoIII with polar residues. In both enzymes a catalytically essential Asp residue protrudes into the cleft, and this Asp is invariant among the HhH/GPD-containing glycosylases (Figure 8). These observations, as well as other results, suggest a related mechanism for substrate recognition for monofunctional glycosylases and glycosylases/β-lyases, although the catalytic mechanisms are somewhat different [158,161] (Figure 9). For both glycosylases the target base is apparently flipped out of the dsDNA helix and accommodated in a substrate-binding pocket, which in AlkA is rich in hydrophobic residues and thus ideally suited to interact with a wide range of electron-deficient bases via Π-donor–acceptor interactions. In EndoIII this pocket is rich in hydrophilic residues that interact with flipped-out bases, such as thymine glycol, either directly or via water-mediated hydrogen bonds. The next step comprises a

HhH motif

No.	Organism	Position	N-helix C-helix GPD region	Fe–S	β-Lyase	Name/function
1	*Eco*	108	RAALEALPGVGRKTANVVLNTAFG--WPTIAVDTHI	+	+	EndoIII Nth
2	*Hin*	108	REALEALAGVGRKTANVVLNTAFG--HPTIAVDTHI	+	+	EndoIII homologue
3	*Bsu*	109	RDELVKLPGVGRKTANVVVSVAFG--VPAIAVDTHV	+	+	EndoIII homologue
4	*Sce*	231	INELLGLPGVGPKMAYLTLQKAWG-KIEGICVDVHV	–	+	EndoIII homologue NTG1
5	*Sce*	107	IEGILSLPGVGPKMGYLTLQKGWG-LIAGICVDVHV	+	+	EndoIII homologue NTG2
6	Hum	200	VAELVALPGVGPKMAHLAMAVAWG-TVSGIAVDTHV	+	+	EndoIII homologue
7	*Spo*	130	VEDLMTLPGVGPKMGYLCMSIAWN-KTVGIGVDVHV	+	+	EndoIII homologue Nth
8	*Cel*	107	LDGLCALPGVGPKMANLVMQIAWG-ECVGIAVDTHV	+	?	EndoIII-like
9	*Sce*	229	REHLMSYNGVGPKVADCVCLMGLH-MDGIVPVDVHV	–	+	8-OxoG-DNA glycosylase OGG1
10	*Mlu*	223	LEDLVALPGVGRKTAFVVLGNAFG--QPGITVDTHF	+	+	UV *N*-glycosylase Pdg
11	*Eco*	108	FEEVAALPGVGRSTAGAILSLSLG--KHFPILDGNV	+	+/–	A-mismatch-DNA glycosylase MutY
12	*Hin*	113	FEQVWALSGVGRSTAGAILSSVLN--QPYPILDGNV	+	?	MutY homologue
13	*Sty*	108	FAEIAALPGVGRSTAGAILSLALG--KHYPILDGNV	+	–	MutY homologue MutB
14	*Spo*	142	DEWAKGIPGVGPYTAGAVLSIAWK--QPTGIVDGNV	+	?	MutY homologue
15	*Mth*	114	RKAILDLPGVGKYTCAAVMCLAFG--KKAAMVDANF	+	–	MutY homologue Mig
16	Hum	192	ETLQQLLPGVGRYTAGAIASIAFG--QATGVVDGNV	+	?	MutY homologue hMYH
17	*Eco*	206	MKTLQTFPGIGRWTANYFALRGWQAKDVFLPDDYLI	–	–	3-MeA-DNA glycosylase AlkA
18	*Bsu*	226	EKNLIKIRGIGPWTANYVLMRCLRFPTAFPIDDVGL	–	?	AlkA homologue AlkA
19	*Sce*	177	ESLVTNVKGIGPWSAKMFLISGLKRMDVFAPEDLGI	–	–	AlkA homologue MAG
20	*Spo*	138	IERLTQIKGIGRWTVEMLLIFSLNRDDVMPADDLSI	–	?	AlkA homologue Mag1
21	*Eco*	158	LLDQAFLAGLGNYLRVEILWQVGLITGNHKAKDLNA	Zn	+	EndoVIII Nei
22	*Eco*	158	LMDNKLVVGVGNIYASESLFAAG------IHPDRLA	Zn	+	Fpg

Figure 8 Conserved HhH motif and catalytic residues of DNA glycosylases

Putative HhH motifs from various DNA glycosylases are shown. Conserved, small and/or hydrophobic residues apparently important for HhH structure are shaded grey. Conserved and charged residues apparently involved in catalysis are shaded pink. The Fpg protein is more distantly related than the others, but is included due to its overall identity with EndoVIII and the exact same positioning of the putative HhH in these two enzymes. Accession numbers/codes: 1, end3-ecoli; 2, end3-haein; 3, end3-bacsu; 4, ncbi 1510694; 5, yab5-yeast; 6, scyol043c; 7, U81285; 8, yaj7-schpo; 9, cer10e4; 10, U44855; 11, U22181; 12, muty-ecoli; 13, muty-haein; 14, muty-salty; 15, spac26a3-2; 16, gtmr-mettf; 17, U63329; 18, 3mg2-ecoli; 19, 3mga-bacsu; 20, mag-yeast; 21, U76637; 22, end8-ecoli, 23, fpg-ecoli. Abbreviations: *Bsu*, *Bacillus subtilis*; *Cel*, *Caenorhabditis elegans*; *Eco*, *Escherichia coli*; *Hin*, *Haemophilus influenzae*; Hum, human; *Mja*, *Methanococcus jannischii*; *Mlu*, *Micrococcus luteus*; *Mth*, *Methanobacterium thermoformicum*; *Sce*, *Saccharomyces cerevisiae*; *Spo*, *Schizosaccharomyces pombe*; *Sty*, *Salmonella typhimurium*.

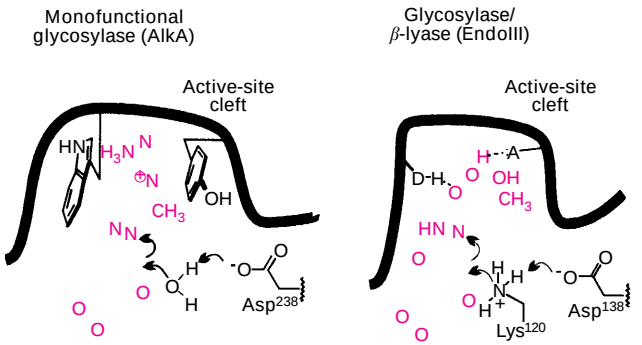


Figure 9 Extrahelical recognition of base lesions by mono- and bi-functional DNA glycosylases

The common location of a catalytic Asp residue in the active-site cleft of AlkA and EndoIII suggests that the two enzymes share a similar mode of nucleophilic activation. In AlkA, Asp^{238} may deprotonate water and activate it for nucleophilic attack at C-1 of the flipped-out nucleotide. In EndoIII, Asp^{138} may deprotonate Lys^{120}, which then attacks at C-1 to form a covalent enzyme–substrate intermediate. Adapted from [158] (copyright Cell Press) with permission.

nucleophilic attack involving a conserved Asp in the GPD region. In the bifunctional glycosylases this attack may take place by deprotonation of the ϵ-NH_3^+ group of a conserved Lys residue (Lys^{120} in EndoIII [162]), which then attacks the substrate anomeric carbon and causes release of the base. The covalent enzyme–substrate (Schiff-base) intermediate formed undergoes several transformations resulting in strand cleavage, degradation of deoxyribose and regeneration of free enzyme [164]. A Schiff-base intermediate has been demonstrated for a number of structurally different DNA glycosylases/β-lyases by borohydride-dependent enzyme–DNA cross-linking [161,164–169]. Covalent intermediates are not observed with the monofunctional DNA glycosylases such as AlkA, for which the attacking nucleophile may be a water molecule which is activated by Asp^{238} located at the same position as Asp^{138} in EndoIII (Figures 7 and 8). Interestingly, nucleophilic attack by water activated by an invariant Asp residue is also likely to occur in the proposed catalytic mechanism of UDG [83,84], although this enzyme does not contain a HhH motif and is structurally unrelated to AlkA and EndoIII.

5-MeC-DNA glycosylase

5-MeC in CpG contexts is important in embryogenesis and in the regulation of tissue-specific gene expression [170]. This implies that methylation and demethylation of cytosine must be regulated by complex mechanisms. 5-MeC is apparently removed and replaced by cytosine by an enzymic mechanism resembling the BER process [171]. While one report suggested that 5-meC is removed by an endonucleolytic process [172], others have demonstrated release of the free base (5-meC, which is subsequently

deaminated to thymine) by HeLa nuclear extracts [173]. Recently, release of 5-meC by a partially purified enzyme from HeLa cells [174] and by a highly purified enzyme from chick embryos [175] was reported, indicating that a DNA glycosylase activity is responsible for 5-meC removal. The 5-meC-DNA glycosylase purified from chick embryos had a molecular mass of 52.5 kDa and also displayed mismatch-specific TDG activity. It is not clear whether these enzymic activities are the properties of one protein or result from the co-purification of two enzymes of very similar sizes.

DNA GLYCOSYLASES RECOGNIZING OXIDIZED BASES

Normal aerobic metabolism including mitochondrial respiration, as well as ionizing radiation and certain drugs, generate oxygen-derived free radicals such as superoxide ($O_2^{-\bullet}$) and hydroxyl ($^{\bullet}OH$) radicals, as well as hydrogen peroxide (H_2O_2). Oxidative stress causes a large number of different lesions in DNA, as well as the formation of DNA–protein cross-links [176]. The steady-state level of oxidative damage in cellular DNA has been estimated to be as high as 0.5 fmol–1 pmol/μg of DNA [177], and is considered to be one of the most important causes of spontaneous mutations in humans. The primary cellular defence system against oxidatively damaged bases appears to be BER, with NER serving a back-up function. Several DNA glycosylases acting upon oxidized bases have been identified in both prokaryotes and eukaryotes. Although the substrate specificities of these glycosylases are generally relaxed and often overlapping, they may be classified into two subgroups. These are represented by *E. coli* EndoIII (Nth) and related enzymes that remove oxidized pyrimidines, and Fpg and related enzymes that remove oxidized purines (Figure 2).

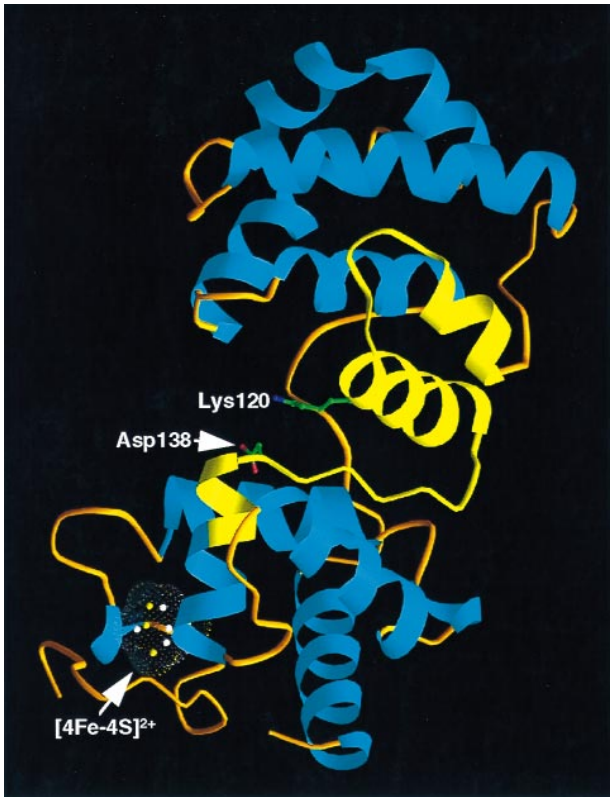


Figure 10 Three-dimensional structure and DNA-binding motifs of EndoIII

Ribbon diagram of EndoIII drawn with the program RIBBONS [232], and highlighting the conserved HhH/GPD motif (yellow), the catalytic residues Lys^{120} and Asp^{138} (both green) and the $[4Fe–4S]^{2+}$ cluster. Co-ordinates are from The Brookhaven Protein Data Bank (accession no. 2ABK).

EndoIII family

The EndoIII family of repair glycosylases constitutes a conserved class of enzymes that is apparently present throughout phylogeny. *E. coli* EndoIII (thymine glycol-DNA glycosylase; urea-DNA glycosylase) was originally identified as an endonucleolytic activity degrading heavily UV-irradiated DNA [178], but was subsequently shown to be a DNA glycosylase/β-lyase [179]. The enzyme displays a broad substrate specificity and excises ring-saturated, ring-opened and ring-fragmented pyrimidines (Figure 2). Overproduction of EndoIII protects against lethal effects of ionizing radiation and chemical oxidants [180], and the protein is considered to be a prime defence mechanism against oxidized pyrimidines in *E. coli*. EndoIII has been purified to physical homogeneity [181] and its crystal structure has been solved [160,162]. The 23.4 kDa protein (211 amino acids) is elongated and bilobal, with a central cleft separating a continuous six-α-helix barrel domain and a four-α-helix domain formed by the N-terminal helix and three C-terminal helices. The C-terminal loop contains a $[4Fe–4S]^{2+}$ cluster held in place by four conserved Cys residues. This motif is also found in several other DNA repair proteins (Figures 8 and 10), and is referred to as a $[4Fe–4S]^{2+}$-cluster loop. This cluster does not participate in redox chemistry [160], but is instead believed to have a structural function in positioning basic residues for DNA binding, and thus fulfils a novel role for metal ions in DNA repair [162]. The residues Lys^{120} and Asp^{138} are positioned at the mouth of the positively charged groove separating the two domains and directly above a solvent-filled pocket [160]. These residues reside in the HhH and GPD regions respectively (Figures 8 and 10), and are involved in the proposed catalytic mechanism described above [162]. Modelling of a 25-mer B-DNA against the crystallographic structure demonstrated that the DNA fits in the interdomain groove near Lys^{120} and that the solvent-filled pocket could accommodate a flipped-out base [162], similar to what was observed in UDG [83] and DNA cytosine-5-methyltransferases [85,86].

Early biochemical studies demonstrated enzyme activities similar to that of EndoIII in mammalian cells [182,183]. Subsequently, genes or cDNAs encoding homologues of EndoIII have been identified in *Bacillus subtilis* [184], *Haemophilus influenzae* [185], *Methanococcus jannaschii* [186], *S. cerevisiae* (NTG1) [187], *S. pombe* (Nth-Spo) [188], *Candidas elegans* [189], rat (EMBL/GenBank accession no. H33255) and human [190,191]. The gene for the human homologue of EndoIII is located in chromosome 16p13.2-3, close to the MPG gene [190,191]. Most notable is the strongly conserved nature of the HhH motif and the invariant Lys and Asp residues at the positions corresponding to the catalytic Lys^{120} and Asp^{138} in the *E. coli* enzyme. Interestingly, the $[4Fe–4S]^{2+}$-cluster motif is found in all of these proteins except for *S. cerevisiae* NTG1. In addition to excising thymine glycol and having β-lyase activity, the eukaryotic EndoIII homologues also remove formamidopyrimidines (L. Luna, M. Bjørås and E. Seeberg, personal communication), which in *E. coli* are recognized by the Fpg protein. The eukaryotic glycosylases do not, however, recognize 8-oxoG, which is a substrate for Fpg, and they thus share substrate specificities with both Nth and Fpg. Another homologue, NTG2, was recently identified in the yeast genome database and the purified enzyme has properties similar to those of NTG1 (I. Alseth, M. Bjørås, M. Pirovano, T. Rognes and E. Seeberg,

personal communication), and the corresponding protein expressed (I. Alseth, M. Bjørås and E. Seeberg, personal communication). Interestingly, NTG2 contains the conserved HhH motif and both the catalytic Asp and Lys residues of the bifunctional DNA glycosylases (Figure 8). In contrast with NTG1, however, NTG2 also contains the $[4Fe–4S]^{2+}$ binding motif of these enzymes. A 31 kDa EndoIII homologue from calf thymus displays DNA glycosylase activity against pyrimidine hydrates and thymine glycol, as well as β-lyase activity [169], and is closely related to the recently identified human homologue. The human homologue was shown to remove thymine glycol [190,191] and urea residues, and has an associated lyase activity [191].

EndoVIII and EndoIX

Two other *E. coli* DNA glycosylases, EndoVIII and EndoIX, share some functional properties with EndoIII [192]. EndoVIII has been purified and characterized [193] and its gene cloned [194]. The purified enzyme has a molecular mass of 28–30 kDa, and displays β-lyase activity as well as DNA glycosylase activity against thymine glycol and dihydrothymine. Interestingly, there is no significant sequence similarity between EndoVIII and the functionally related EndoIII, whereas significant similarity is observed in several regions between EndoVIII and Fpg. EndoIX is less well characterized. It recognizes both β-ureidoisobutyric acid and urea in DNA, but not thymine glycol or dihydrothymine. EndoVIII and EndoIX may thus serve a back-up function for EndoIII in *E. coli* [192].

Hydroxymethyluracil-DNA glycosylase

hmUra may be formed from thymine and 5-meC in DNA under conditions of oxidative stress, and has been assumed to be a cytotoxic lesion [195,196]. In *E. coli* hmUra is one of the many substrates removed by AlkA [143]. A DNA glycosylase activity removing hmUra has also been identified in and partially purified from mammalian cells [197,198]. Subsequent work with a mammalian cell line lacking hmUra repair activity (but proficient in uracil repair) has indicated that hmUra is not in fact highly toxic when incorporated opposite adenine, but that the toxicity is caused by extensive repair processes resulting from its incorporation. It was therefore proposed that the main function of this enzyme under normal conditions is to remove hmUra resulting from oxidation and subsequent deamination of 5-meC, a process that would result in a mutagenic mismatch of hmUra and G [199]. As mammalian DNA may comprise 2–5 % 5-meC, this would also justify the presence of a specific enzyme handling this lesion, although the oxidation of 5-meC to hmUra appears to be a minor lesion after oxidative stress in comparison with thymine glycol formation [200]. Still, failure to repair 5-meC oxidized either to thymine glycol or hmUra may contribute to the observed hypermutability of 5-meC sites [201].

5-Formyluracil-DNA glycosylase

5-Formyluracil in DNA is formed in yields comparable with those of 8-oxoG and 5-hmUra after γ-irradiation [202,203]. Computer modelling indicates that substitution of the thymine methyl group by formyl might interfere with base pairing to adenine, thus suggesting a mutagenic property of 5-formyluracil [203]. This is also supported by previous findings that the incorporation of 5-formyluracil during replication of *Salmonella typhimurium* resulted in A:T → G:C transitions [202]. Possibly as a consequence of labilization of the glycosylic bond, 5-formyluracil is excised by the *E. coli* AlkA enzyme, but not by EndoIII or Fpg, which remove other oxidation products [143]. In human cells a DNA glycosylase activity different from MPG and UDG removes 5-formyluracil, indicating that this activity may represent a previously uncharacterized glycosylase [203].

Fpg family and the oxoG system

The Fpg protein (also named MutM and Fapy-DNA glycosylase) in *E. coli* catalyses the excision of damaged purine bases such as the oxidation products 8-oxoG and 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine (fapy) from dsDNA. 8-OxoG is both miscoding and mutagenic, and is believed to be the major physiological substrate for Fpg [204]. 8-OxoG is also removed by the functional co-operation of two additional *E. coli* enzymes. In this 'GO (8-oxoG) system' [205], MutT (8-oxo-dGTPase) provides the first line of defence by eliminating 8-oxo-dGTP from the dNTP pool [206]. If 8-oxo-dGTP escapes MutT, 8-oxoG may be incorporated opposite either A or C, giving rise to G → T transversions unless it is removed. Here Fpg (MutM) constitutes a second line of defence, by removing 8-oxoG incorporated opposite C or formed by the oxidation of DNA guanine. 8-OxoG incorporated opposite A is poorly repaired by Fpg [207]. If both of these defence levels are bypassed, MutY provides a third level of defence by removing A incorporated opposite 8-oxoG formed by oxidation [208,209]. It should be noted, however, that if 8-oxoG is incorporated opposite A in the template, removal of A by MutY may lead to T → G transversion mutations due to subsequent incorporation of C opposite 8-oxoG.

The *fpg* gene encodes a protein of 30.2 kDa [210,211]. In addition to its N-glycosylase activity, this protein also has a nicking activity that cleaves both the 5′- and 3′-phosphodiester bonds at an AP site by successive β- and δ-elimination reactions, leaving both the 3′ and 5′ DNA ends phosphorylated [212]. In DNA incised by an AP-endonuclease, Fpg displays dRpase-like activity [10]. Recently, two distinct 8-oxoG-specific glycosylases/β-lyases were identified in *S. cerevisiae*. They differ in their preference for the base opposing 8-oxoG, and one of them is identical to the *OGG1* gene product [213]. *OGG1* encodes a 43 kDa protein that lacks a zinc-finger domain and the N-terminal PELPEVE sequence found in bacterial enzymes [213,214], but contains a HhH motif similar to those in EndoIII, MutY and AlkA [161]. Lys and Asp residues in OGG1 are found in positions corresponding to the putative catalytic residues Lys^{120} and Asp^{138} in EndoIII. Mutation of this Lys residue (Lys^{241} → Gln) in OGG1 abolishes all detectable borohydride trapping activity, as expected if it is involved in catalysis [161]. Despite being apparently structurally different, the bacterial and yeast enzymes are functionally remarkably similar [161,214]. They prefer 8-oxoG opposite pyrimidines, and repair 8-oxoG mispaired with A poorly. However, the rate of fapy excision by OGG1 is less than 10 % of the 8-oxoG excision rate, whereas the activities against these substrates are equal for Fpg [213]. This may, however, be compensated for by the yeast EndoIII homologue NTG1 and possibly by the recently discovered NTG2 (see above).

A second yeast 8-oxoG-DNA glycosylase/β-lyase, OGG2, has been identified by two groups [161,215]. The substrate specificity of the ~ 37 kDa OGG2 protein is different from that of OGG1, as it prefers 8-oxoG opposite purines. Perhaps a main function of OGG2 is to remove incorporated 8-oxoG from an 8-oxoG/A mispair. This would prevent mutations. In contrast, removal of 8-oxoG from an 8-oxoG/A mispair resulting from incorporation of A opposite a template 8-oxoG formed by oxidation would be mutagenic. This would be avoided if OGG2 were specific for the

repair of 8-oxoG in the nascent DNA strand. Recently, 8-oxoG glycosylase activity was demonstrated to be associated with a chromatin-bound form of the *Drosophila* ribosomal protein S3, and the protein also contained an associated β-lyase activity. This protein, which is unrelated to previously known non-ribosomal proteins, also complemented an *E. coli mutM* strain, thus demonstrating activity *in vivo* [216].

A mammalian counterpart of the *fpg* gene has not yet been cloned, although several reports support the presence of DNA glycosylases that remove 8-oxoG in mammalian cells [217–219]. It is, however, less clear what fraction of the repair is caused by MPG, which removes 8-oxoG in addition to a wide range of alkylated purines [130,145,146]. Support for MPG as a major repair activity for 8-oxoG is demonstrated by the ability of MPG to rescue *E. coli* lacking the Fpg protein [145]. Jaruga and Dizdaroglu [220] analysed base damage in human lymphoblasts after exposure to H_2O_2 and found that both fapy and 8-oxoG were excised as free bases by a process following first-order kinetics. Fapy is not a known substrate for MPG, indicating that mammalian cells contain glycosylases other than MPG that may remove oxidatively damaged purines.

OTHER DNA GLYCOSYLASES

Pyrimidine-dimer-DNA glycosylases

DNA glycosylases acting upon pyrimidine dimers in DNA are encoded by bacteriophage T4endoV and the *denV* gene of *M. luteus*. Pyrimidine-dimer-DNA glycosylase-like activities have also been partially purified from *Neisseria mucosa* extracts [221] and extracts of *S. cerevisiae* after extensive fractionation [222].

T4endoV

T4endoV is a small protein of 137 amino acids with N-glycosylase/β-lyase activity [164,223,224], and its three-dimensional structure has been determined by X-ray crystallography [225,226]. The protein consists of a single, compact domain containing three α-helices. The basic, concave surface of T4endoV interacts with dsDNA that is sharply kinked ($\sim 60\,^\circ$) at the central pyrimidine dimer, forming several interactions with phosphates on both strands [227] (Figure 11). The enzyme interacts with the pyrimidine dimer in the minor groove, as is also observed with UDG [83]. T4endoV does not, however, flip the target base(s) out of the dsDNA helix, but instead flips an adenine complementary to one of these thymines into a cavity on the protein surface. The flipping may allow the enzyme to discriminate between damaged and normal DNA, and to generate an empty space within the DNA helix, which can be occupied by catalytically important residues in the enzyme.

In the catalytic reaction, the amino group of the N-terminal Thr residue probably acts as the attacking nucleophile and forms a Schiff-base intermediate with C-1′ of the 5′ sugar in the thymine dimer. Two arginines (Arg^{22} and Arg^{26}) located within the helix probably secure the correct positioning of this thymine during the initial steps of the reaction, while Glu^{23} either stabilizes the Schiff base or protonates the pyrimidine ring and thus weakens the N-glycosylic bond. Glu^{23} may also participate in the subsequent β-elimination by abstracting a C-2′ hydrogen from the open deoxyribose [227].

Micrococcus luteus UV endonuclease

The *M. luteus* UV endonuclease functionally resembles T4endoV in that it exclusively cleaves the 5′-bond in dimerized pyrimidines, and contains β-lyase activity [228,229]. The recent purification of

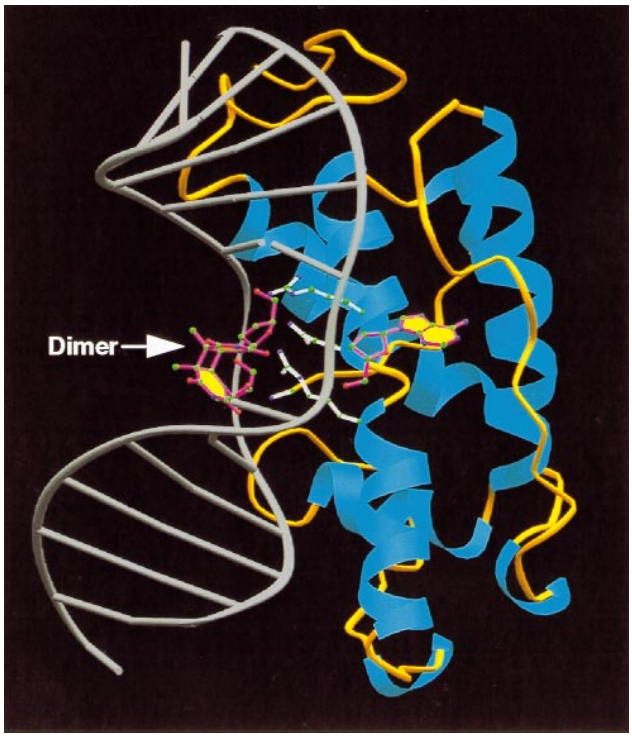


Figure 11 Three-dimensional structure of a T4endoV–(pyrimidine dimer)–DNA complex

Ribbon diagram of T4endoV ($Glu^{23} \rightarrow$ Gln mutant) bound to a pyrimidine-containing DNA substrate represented by grey tubes, drawn using the program RIBBONS [232]. The DNA is viewed perpendicular to the DNA helical axis, showing the sharp kink at the position of the pyrimidine dimer (pink). The extrahelical adenine (pink) is facing the DNA-binding groove of the enzyme. The proposed catalytically important residues Arg^{22}, Gln^{23} (wild-type Glu^{23}), Arg^{26} and Thr^2 (from top to bottom) are shown in white. Co-ordinates are from The Brookhaven Protein Data Bank (accession no. 1VAS).

M. luteus UV endonuclease and cloning of its corresponding gene (*pdg*) has revealed a protein of 31–32 kDa, with significant identity to the emerging family of DNA glycosylases containing a $[4Fe–4S]^{2+}$ cluster and a HhH motif [100]. A re-examination of the substrate specificity of the enzyme using various thymidylyl-(3′-5′)-thymidine photoproducts indicates that it exclusively cleaves the *cis–syn* thymine dimer and not *trans–syn*, (6–4) or Dewar dimers [100]. This substrate specificity is similar to that of T4endoV, except that the latter is also able to cleave *trans–syn* thymine dimers at a rate of $\sim 1\,\%$ that of cleavage of the *cis–syn* dimer [230].

CONCLUSIONS AND PERSPECTIVES

DNA glycosylases remove a large number of cytotoxic or mutagenic bases from DNA. Some of these enzymes have very narrow substrate specificities, whereas most remove a number of structurally different bases. The BER pathway apparently involves more protein factors (XRCC1 and possibly RPA) than those strictly required to produce the proposed DNA intermediates in the repair process. Many of the DNA glycosylases have overlapping substrate specificities and may serve back-up functions for each other. DNA glycosylases often contain a conserved HhH motif in the active site. Recent structural, biochemical and mutational data have established different catalytic mechanisms for simple DNA glycosylases and those

with an associated β-lyase activity. For both types, however, the catalytic residues may gain access to the glycosylic bond by flipping either the target (for UDG) or a complementary (for T4endoV) nucleotide out of the DNA helix. In UDGs a Leu side chain penetrates the helix and displaces uracil towards a narrow pocket that has strong affinity for the base. Similar mechanisms probably operate for other DNA glycosylases as well. In monofunctional DNA glycosylases, an Asp residue deprotonates water to produce an OH^- nucleophile that attacks the N-glycosylic bond. This bond may already be weakened, as is the case for most substrates of AlkA. Catalysis may take place at a low rate even though the pocket that binds the extrahelical base is relaxed to accommodate structurally different bases. For the highly efficient UDGs, very specific interactions between the tight-fitting active site and the target lesion weaken the glycosylic bond, allowing rapid hydrolysis by the nucleophile. In bifunctional DNA glycosylases, an Asp residue deprotonates a Lys residue, producing an amine nucleophile that forms a Schiff-base enzyme–sugar intermediate at C-1, thus expelling the base. This is followed by chain cleavage by β-elimination.

Eukaryotic DNA glycosylases contain N-terminal sequences not found in the bacterial homologues. These may be involved in subcellular targeting. Thus nuclear and mitochondrial forms of UDG are generated from the same gene (*UNG*) by transcription from two different promoters and alternative splicing. This results in enzyme species with different N-terminal sequences that target nuclear or mitochondrial import. This has established a new mechanism for subcellular targeting. A similar mechanism may target MPG to nuclei or mitochondria.

It has been argued that the spontaneous mutation rate in somatic cells is not sufficient to account for the multiple mutations observed in human cancer, and that cancer cells are genetically unstable, i.e. they express a 'mutator phenotype' [231]. Inaccurate DNA polymerases, dNTP pool imbalances and deficient DNA repair mechanisms are clearly among the factors that could cause such a mutator phenotype. In agreement with this, deficiencies in mismatch repair, a NER mechanism, have recently been shown to cause a form of hereditary colorectal cancer, as well as some other cancers [88]. Other hereditary deficiencies in NER cause different rare cancer-prone syndromes [1]. These observations provide evidence for the presence of a mutator phenotype in human cancers. However, a deficiency in BER as a cause of cancer has not yet been established, even though DNA damage of the type repaired by the BER pathway may be quantitatively dominant. This may be because deficiencies in BER have not really been thoroughly investigated yet. Alternatively, BER deficiencies may not be compatible with the development of complex organisms. The latter possibility is very likely to be investigated in the near future using knock-out mice. In the case of hereditary non-polyposis colorectal cancer, a mismatch-repair deficiency was suspected because of the microsatellite instability observed in this cancer form. BER deficiencies are not likely to leave specific 'fingerprints' indicating a particular deficiency in the DNA, making a search for a specific defect more difficult. Screening for low activities of particular DNA glycosylases or other enzymes in the BER pathway may not be fruitful. This is because the defect may be more subtle, affecting the fidelity of DNA polymerases, substrate specificities of DNA repair enzymes, intracellular transport of DNA repair proteins or the co-ordination of different repair steps, to mention just a few possibilities. The bright side is that the rapid progress in recent years has given clues as to what to look for and how to do it.

The rapid progress in our understanding of BER is likely to lead to elucidation of the functional significance of this process in the future. General progress in molecular biology and the large-scale sequencing of several genomes will lead rapidly to the identification of a number of new genes for BER enzymes. Knock-out techniques and biochemical characterization will be helpful in establishing their functional significance. Two-hybrid and structure analysis will identify how they interact with other proteins and target DNA. Interestingly, the progress in DNA repair, and the development of useful tools in general, have depended heavily upon simultaneous work on many different organisms, often 'simple' models such as *E. coli*. It is clearly fair to state that our understanding of human biochemistry and human disease at a molecular level would be in a very sad state if it were not for these small friends.

This work was supported by the Research Council of Norway and the Norwegian Cancer Society. We thank Dr. J.A. Tainer, Dr. C.D. Mol, Dr. Y. Guyan, Dr. T. Lindahl, Dr. E. Seeberg and Dr. G. Teebor for communicating unpublished results, Dr. C.D. Mol, Dr. J.A. Tainer, Dr. I. Sandlie and Dr. E. Seeberg for critical review of the manuscript, and Dr. T. Ellenberger for supplying Figure 9 as well as co-ordinates for AlkA. We regret that many relevant high-quality papers could not be cited due to space limitations.

REFERENCES

1 Friedberg, E. C., Walker, G. C. and Siede, W. (1995) DNA Repair and Mutagenesis, ASM Press, Washington, DC
2 Barnes, D. E., Lindahl, T. and Sedgwick, B. (1993) Curr. Opin. Cell Biol. **5**, 424–433
3 Gentil, A., Cabral Neto, J. B., Mariage Samson, R., Margot, A., Imbach, J. L., Rayner, B. and Sarasin, A. (1992) J. Mol. Biol. **227**, 981–984
4 Singhal, R. K., Prasad, R. and Wilson, S. H. (1995) J. Biol. Chem. **270**, 949–957
5 Dianov, G. and Lindahl, T. (1994) Curr. Biol. **4**, 1069–1076
6 Dianov, G., Price, A. and Lindahl, T. (1992) Mol. Cell Biol. **12**, 1605–1612
7 Wiebauer, K. and Jiricny, J. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 5842–5845
8 Barzilay, G. and Hickson, I. D. (1995) Bioessays **17**, 713–719
9 Kubota, Y., Nash, R. A., Klungland, A., Schär, P., Barnes, D. E. and Lindahl, T. (1996) EMBO J. **15**, 6662–6670
10 Graves, R. J., Felzenszwalb, I., Laval, J. and O'Connor, T. R. (1992) J. Biol. Chem. **267**, 14429–14435
11 Dianov, G., Sedgwick, B., Daly, G., Olsson, M., Lovett, S. and Lindahl, T. (1994) Nucleic Acids Res. **22**, 993–998
12 Caldecott, K. W., Aoufouchi, S., Johnson, P. and Shall, S. (1996) Nucleic Acids Res. **24**, 4387–4394
13 Nagelhus, T., Haug, T., Singh, K. K., Keshav, K. F., Skorpen, F., Otterlei, M., Bharati, S., Lindmo, T., Benichou, S., Benarous, R. and Krokan, H. E. (1997) J. Biol. Chem. **272**, 6561–6566
14 Matsumoto, Y., Kim, K. and Bogenhagen, D. F. (1994) Mol. Cell. Biol. **14**, 6187–6197
15 Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L. S., Lane, D. P., Abbondandolo, A. and Dogliotti, E. (1996) J. Biol. Chem. **271**, 9573–9578
16 Wu, X., Li, J., Li, X., Hsieh, C. L., Burgers, P. M. and Lieber, M. R. (1996) Nucleic Acids Res. **24**, 2036–2043
17 Mol, C. D., Kuo, C. F., Thayer, M. M., Cunningham, R. P. and Tainer, J. A. (1995) Nature (London) **374**, 381–386
18 Izumi, T., Ishizaki, K., Ikenaga, M. and Yonei, S. (1992) J. Bacteriol. **174**, 7711–7716
19 Agnez, L. F., Costa de Oliveira, R. L., Di Mascio, P. and Menck, C. F. (1996) Carcinogenesis **17**, 1183–1185
20 Sandigursky, M. and Franklin, W. A. (1992) Nucleic Acids Res. **20**, 4699–4703
21 Lindahl, T. (1974) Proc. Natl. Acad. Sci. U.S.A. **71**, 3649–3653
22 Lindahl, T. (1993) Nature (London) **362**, 709–715
23 Shen, J. C., Rideout, W. M. and Jones, P. A. (1992) Cell **71**, 1073–1080
24 Duncan, B. K. and Weiss, B. (1982) J. Bacteriol. **151**, 750–755
25 Impellizzeri, K. J., Anderson, B. and Burgers, P. M. (1991) J. Bacteriol. **173**, 6807–6810
26 Ingraham, H. A., Tseng, B. Y. and Goulian, M. (1980) Cancer Res. **40**, 998–1001
27 Zastawny, T. H., Doetsch, P. W. and Dizdaroglu, M. (1995) FEBS Lett. **364**, 255–258
28 Hatahet, Z., Kow, Y. W., Purmal, A. A., Cunningham, R. P. and Wallace, S. S. (1994) J. Biol. Chem. **269**, 18814–18820
29 Dizdaroglu, M., Karakaya, A., Jaruga, P., Slupphaug, G. and Krokan, H. E. (1996) Nucleic Acids Res. **24**, 418–422
30 Krokan, H. and Wittwer, C. U. (1981) Nucleic Acids Res. **9**, 2599–2613
31 Varshney, U. and van de Sande, J. H. (1991) Biochemistry **30**, 4055–4061
32 Bennett, S. E., Sanderson, R. J. and Mosbaugh, D. W. (1995) Biochemistry **34**, 6109–6119

33 Higley, M. and Lloyd, R. S. (1993) Mutat. Res. **294**, 109–116
34 Eftedal, I., Guddal, P. H., Slupphaug, G., Volden, G. and Krokan, H. E. (1993) Nucleic Acids Res. **21**, 2095–2101
35 Verri, A., Mazzarello, P., Spadari, S. and Focher, F. (1992) Biochem. J. **287**, 1007–1010
36 Nilsen, H., Yazdankhah, S. P., Eftedal, I. and Krokan, H. E. (1995) FEBS Lett. **362**, 205–209
37 Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N. M., Haug, T., Levine, D. W. and Krokan, H. E. (1995) Biochemistry **34**, 128–138
38 Gao, S., Drouin, R. and Holmquist, G. P. (1994) Science **263**, 1438–1440
39 Tornaletti, S. and Pfeifer, G. P. (1994) Science **263**, 1436–1438
40 Varshney, U., Hutcheon, T. and van de Sande, J. H. (1988) J. Biol. Chem. **263**, 7776–7784
41 Percival, K. J., Klein, M. B. and Burgers, P. M. (1989) J. Biol. Chem. **264**, 2593–2598
42 Walter, R. B. and Morizot, D. C. (1996) Adv. Struct. Biol. **4**, 1–24
43 Nilsen, H., Otterlei, M., Haug, T., Solum, K., Nagelhus, T., Skorpen, F. and Krokan, H. E. (1997) Nucleic Acids Res. **25**, 750–755
44 Olsen, L. C., Aasland, R., Wittwer, C. U., Krokan, H. E. and Helland, D. E. (1989) EMBO J. **8**, 3121–3125
45 Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C. et al. (1984) Nature (London) **310**, 207–211
46 Davison, A. J. and Scott, J. E. (1986) J. Gen. Virol. **67**, 1759–1816
47 McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E. and Taylor, P. (1988) J. Gen. Virol. **69**, 1531–1574
48 Worrad, D. M. and Caradonna, S. (1988) J. Virol. **62**, 4774–4777
49 Khattar, S. K., van Drunen Littel van den Hurk, S., Babiuk, L. A. and Tikoo, S. K. (1995) Virology **213**, 28–37
50 Dean, H. J. and Cheung, A. K. (1993) J. Virol. **67**, 5955–5961
51 Yoshida, S., Lee, L. F., Yanagida, N. and Nazerian, K. (1994) Virology **204**, 414–419
52 Sato, S., Yamamoto, T., Isegawa, Y. and Yamanishi, K. (1994) J. Gen. Virol. **75**, 2349–2354
53 McGeoch, D. J., Cook, S., Dolan, A., Jamieson, F. E. and Telford, E. A. R. (1995) J. Mol. Biol **247**, 443–458
54 Upton, C., Stuart, D. T. and McFadden, G. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 4518–4522
55 Goebel, S. J., Johnson, G. P., Perkus, M. E., Davis, S. W., Winslow, J. P. and Paoletti, E. (1990) Virology **179**, 247–266
56 Niles, E. G., Condit, R. C., Caro, P., Davidson, K., Matusick, L. and Seto, J. (1986) Virology **153**, 96–112
57 Shchelkunov, S. N., Blinov, V. M., Totmenin, A. V., Marennikova, S. S., Kolykhalov, A. A., Frolov, I. V., Chizhikov, V. E., Gytorov, V. V., Gashikov, P. V., Belanov, E. F., et al. (1993) Virus Res. **27**, 25–35
58 Tartaglia, J., Winslow, J., Goebel, S., Johnson, G. P., Taylor, J. and Paoletti, E. (1990) J. Gen. Virol. **71**, 1517–1524
59 Haug, T., Skorpen, F., Kvaloy, K., Eftedal, I., Lund, H. and Krokan, H. E. (1996) Genomics **36**, 408–416
60 Anderson, C. T. and Friedberg, E. C. (1980) Nucleic Acids Res. **8**, 875–888
61 Domena, J. D. and Mosbaugh, D. W. (1985) Biochemistry **24**, 7320–7328
62 Domena, J. D., Timmer, R. T., Dicharry, S. A. and Mosbaugh, D. W. (1988) Biochemistry **27**, 6742–6751
63 Wittwer, C. U. and Krokan, H. (1985) Biochim. Biophys. Acta **832**, 308–318
64 Slupphaug, G., Markussen, F. H., Olsen, L. C., Aasland, R., Aarsaether, N., Bakke, O., Krokan, H. E. and Helland, D. E. (1993) Nucleic Acids Res. **21**, 2579–2584
65 Tomkinson, A. E., Bonk, R. T. and Linn, S. (1988) J. Biol. Chem. **263**, 12532–12537
66 Burgers, P. M. and Klein, M. B. (1986) J. Bacteriol. **166**, 905–913
67 Karran, P., Cone, R. and Friedberg, E. C. (1981) Biochemistry **20**, 6092–6096
68 Bennett, S. E. and Mosbaugh, D. W. (1992) J. Biol. Chem. **267**, 22512–22521
69 Bennett, S. E., Schimerlik, M. I. and Mosbaugh, D. W. (1993) J. Biol. Chem. **268**, 26879–26885
70 Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W. and Tainer, J. A. (1995) Cell **82**, 701–708
71 Savva, R. and Pearl, L. H. (1995) Nature Struct. Biol. **2**, 752–757
72 Nagelhus, T. A., Slupphaug, G., Lindmo, T. and Krokan, H. E. (1995) Exp. Cell Res. **220**, 292–297
73 Koistinen, P. and Vilpo, J. A. (1986) Mutat. Res. **159**, 99–102
74 Koistinen, P. and Vilpo, J. A. (1986) Mutat. Res. **175**, 115–120
75 Vilpo, J. A. (1988) Mutat. Res. **193**, 207–217
76 Sirover, M. A. (1979) Cancer Res. **39**, 2090–2095
77 Gombar, C. T., Katz, E. J., Magee, P. N. and Sirover, M. A. (1981) Carcinogenesis **2**, 595–599
78 Gupta, P. K. and Sirover, M. A. (1981) Cancer Res. **41**, 3133–3136
79 Slupphaug, G., Olsen, L. C., Helland, D., Aasland, R. and Krokan, H. E. (1991) Nucleic Acids Res. **19**, 5131–5137
80 Haug, T., Skorpen, F., Lund, H. and Krokan, H. E. (1994) FEBS Lett. **353**, 180–184
81 Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E. and Tainer, J. A. (1995) Cell **80**, 869–878
82 Kavli, B., Slupphaug, G., Mol, C. D., Arvai, A. S., Petersen, S. B., Tainer, J. A. and Krokan, H. E. (1996) EMBO J. **15**, 3442–3447
83 Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E. and Tainer, J. A. (1996) Nature (London) **384**, 87–92
84 Savva, R., McAuley Hecht, K., Brown, T. and Pearl, L. (1995) Nature (London) **373**, 487–493
85 Klimasauskas, S., Kumar, S., Roberts, R. J. and Cheng, X. (1994) Cell **76**, 357–369
86 Reinisch, K. M., Chen, L., Verdine, G. L. and Lipscomb, W. N. (1995) Cell **82**, 143–153
87 Roberts, R. J. (1995) Cell **82**, 9–12
88 Kolodner, R. D. (1995) Trends Biochem. Sci. **20**, 397–401
89 Brown, T. C. and Jiricny, J. (1987) Cell **50**, 945–950
90 Neddermann, P., Gallinari, P., Lettieri, T., Schmid, D., Truong, O., Hsuan, J. J., Wiebauer, K. and Jiricny, J. (1996) J. Biol. Chem. **271**, 12767–12774
91 Seeberg, E., Eide, L. and Bjoras, M. (1995) Trends Biochem. Sci. **20**, 391–397
92 Horst, J. P. and Fritz, H. J. (1996) EMBO J. **15**, 5459–5469
93 Nghiem, Y., Cabrera, M., Cupples, C. G. and Miller, J. H. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 2709–2713
94 Radicella, J. P., Clark, E. A. and Fox, M. S. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 9674–9678
95 Lu, A. L. and Chang, D. Y. (1988) Genetics **118**, 593–600
96 Tsai-Wu, J. J., Liu, H. F. and Lu, A. L. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 8779–8783
97 Lu, A. L., Tsai-Wu, J. J. and Cillo, J. (1995) J. Biol. Chem. **270**, 23582–23588
98 Bulychev, N. V., Varaprasad, C. V., Dorman, G., Miller, J. H., Eisenberg, M., Grollman, A. P. and Johnson, F. (1996) Biochemistry **35**, 13147–13156
99 Michaels, M. L., Pham, L., Nghiem, Y., Cruz, C. and Miller, J. H. (1990) Nucleic Acids Res. **18**, 3841–3845
100 Piersen, C. E., Prince, M. A., Augustine, M. L., Dodson, M. L. and Lloyd, R. S. (1995) J. Biol. Chem. **270**, 23475–23484
101 Gogos, A., Cillo, J., Clarke, N. D. and Lu, A. L. (1996) Biochemistry **35**, 16665–16671
102 Manuel, R. C., Czerwinski, E. W. and Lloyd, R. S. (1996) J. Biol. Chem. **271**, 16218–16226
103 McGoldrick, J. P., Yeh, Y. C., Solomon, M., Essigmann, J. M. and Lu, A. L. (1995) Mol. Cell. Biol. **15**, 989–996
104 Yeh, Y. C., Chang, D. Y., Masin, J. and Lu, A. L. (1991) J. Biol. Chem. **266**, 6480–6484
105 Slupska, M. M., Baikalov, C., Luther, W. M., Chiang, J. H., Wei, Y. F. and Miller, J. H. (1996) J. Bacteriol. **178**, 3885–3892
106 Lindahl, T. (1976) Nature (London) **259**, 64–66
107 Klungland, A., Fairbairn, L., Watson, A. J., Margison, G. P. and Seeberg, E. (1992) EMBO J. **11**, 4439–4444
108 Klungland, A., Bjoras, M., Hoff, E. and Seeberg, E. (1994) Nucleic Acids Res. **22**, 1670–1674
109 Ristau, C., Bolt, H. M. and Vangala, R. R. (1990) Arch. Toxicol. **64**, 254–256
110 Working, P. K., Doolittle, D. J., Smith Oliver, T., White, R. D. and Butterworth, B. E. (1986) Mutat. Res. **162**, 219–224
111 Vaughan, P., Lindahl, T. and Sedgwick, B. (1993) Mutat. Res. **293**, 249–257
112 Mirvish, S. S. (1995) Cancer Lett. **93**, 17–48
113 IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans (1985) vol. 38, Tobacco Smoking, IARC, Lyon
114 Amin, S., Desai, D., Hecht, S. S. and Hoffmann, D. (1996) Crit. Rev. Toxicol. **26**, 139–147
115 Rydberg, B. and Lindahl, T. (1982) EMBO J. **1**, 211–216
116 Xiao, W. and Samson, L. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 2117–2121
117 Clarke, N. D., Kvaal, M. and Seeberg, E. (1984) Mol. Gen. Genet. **197**, 368–372
118 Nakabeppu, Y., Kondo, H. and Sekiguchi, M. (1984) J. Biol. Chem. **259**, 13723–13729
119 Steinum, A. L. and Seeberg, E. (1986) Nucleic Acids Res. **14**, 3763–3772
120 Pierre, J. and Laval, J. (1986) Gene **43**, 139–146
121 Morhoshi, F. and Munakata, N. (1995) Mutat. Res. **337**, 97–110
122 Berdal, K. G., Bjoras, M., Bjelland, S. and Seeberg, E. (1990) EMBO J. **9**, 4563–4568
123 Bjoras, M., Klungland, A., Johansen, R. F. and Seeberg, E. (1995) Biochemistry **34**, 4577–4582
124 Chen, J., Derfler, B. and Samson, L. (1990) EMBO J. **9**, 4569–4575
125 Memisoglu, A. and Samson, L. (1996) Gene **177**, 229–235
126 Santerre, A. and Britt, A. B. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 2240–2244

127 O'Connor, T. R. and Laval, F. (1990) EMBO J. **9**, 3337–3342
128 O'Connor, T. R. and Laval, J. (1991) Biochem. Biophys. Res. Commun. **176**, 1170–1177
129 Samson, L., Derfler, B., Boosalis, M. and Call, K. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 9127–9131
130 Chakravarti, D., Ibeanu, G. C., Tano, K. and Mitra, S. (1991) J. Biol. Chem. **266**, 15710–15715
131 Kielman, M. F., Smits, R. and Bernini, L. F. (1995) Mamm. Genome **6**, 499–504
132 Tatsuka, M., Ibeanu, G. C., Izumi, T., Narayan, S., Ramana, C. V., Kim, N. K., Kang, W., Roy, G. and Mitra, S. (1995) DNA Cell Biol. **14**, 37–45
133 Vickers, M. A., Vyas, P., Harris, P. C., Simmons, D. L. and Higgs, D. R. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 3437–3441
134 Bjelland, S., Bjoras, M. and Seeberg, E. (1993) Nucleic Acids Res. **21**, 2045–2049
135 Bjelland, S. and Seeberg, E. (1987) Nucleic Acids Res. **15**, 2787–2801
136 Bjelland, S. and Seeberg, E. (1996) FEBS Lett. **397**, 127–129
137 Habraken, Y., Carter, C. A., Kirk, M. C. and Ludlum, D. B. (1991) Cancer Res. **51**, 499–503
138 Matijasevic, Z., Sekiguchi, M. and Ludlum, D. B. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 9331–9334
139 Habraken, Y., Carter, C. A., Sekiguchi, M. and Ludlum, D. B. (1991) Carcinogenesis **12**, 1971–1973
140 Matijasevic, Z., Stering, A., Niu, T. Q., Austin-Ritchie, P. and Ludlum, D. B. (1996) Carcinogenesis **17**, 2249–2252
141 Mattes, W. B., Lee, C. S., Laval, J. and O'Connor, T. R. (1996) Carcinogenesis **17**, 643–648
142 Saparbaev, M. and Laval, J. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 5873–5877
143 Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G. and Seeberg, E. (1994) J. Biol. Chem. **269**, 30489–30495
144 Hang, B., Chenna, A., Rao, S. and Singer, B. (1996) Carcinogenesis **17**, 155–157
145 Bessho, T., Roy, R., Yamamoto, K., Kasai, H., Nishimura, S., Tano, K. and Mitra, S. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 8901–8904
146 Bessho, T., Tano, K., Kasai, H., Ohtsuka, E. and Nishimura, S. (1993) J. Biol. Chem. **268**, 19416–19421
147 Roy, R., Kumar, A., Lee, J. C. and Mitra, S. (1996) J. Biol. Chem. **271**, 23690–23697
148 Grombacher, T. and Kaina, B. (1996) DNA Cell Biol. **15**, 581–588
149 Male, R., Nes, I. F. and Kleppe, K. (1981) Eur. J. Biochem. **121**, 243–248
150 Pendlebury, A., Frayling, I. M., Koref, M. F. S., Margison, G. P. and Rafferty, J. A. (1994) Carcinogenesis **15**, 2957–2960
151 Chen, J. and Samson, L. (1991) Nucleic Acids Res. **19**, 6427–6432
152 Mitra, S. and Kaina, B. (1993) Prog. Nucleic Acid Res. Mol. Biol. **44**, 109–142
153 Grombacher, T. and Kaina, B. (1995) Biochim. Biophys. Acta **1270**, 63–72
154 Laval, F. (1990) Mutat. Res. **233**, 211–218
155 Laval, F. (1991) Biochem. Biophys. Res. Commun. **176**, 1086–1092
156 Gupta, P. K. and Sirover, M. A. (1980) Mutat. Res. **72**, 273–284
157 Krokan, H., Haugen, A., Myrnes, B. and Guddal, P. H. (1983) Carcinogenesis **4**, 1559–1564
158 Labahn, J., Scharer, O. D., Long, A., EzazNikpay, K., Verdine, G. L. and Ellenberger, T. E. (1996) Cell **86**, 321–329
159 Yamagata, Y., Kato, M., Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., Yasumura, K., Tomita, K., Ihara, K., Fujii, Y., et al. (1996) Cell **86**, 311–319
160 Kuo, C. F., McRee, D. E., Fisher, C. L., O'Handley, S. F., Cunningham, R. P. and Tainer, J. A. (1992) Science **258**, 434–440
161 Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Sponner, E., Lane, W. S. and Verdine, G. L. (1996) Curr. Biol. **6**, 968–980
162 Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P. and Tainer, J. A. (1995) EMBO J. **14**, 4108–4120
163 Doherty, A. J., Serpell, L. C. and Ponting, C. P. (1996) Nucleic Acids Res. **24**, 2488–2497
164 Dodson, M. L., Michaels, M. L. and Lloyd, R. S. (1994) J. Biol. Chem. **269**, 32709–32712
165 Kow, Y. W. and Wallace, S. S. (1987) Biochemistry **26**, 8200–8206
166 Dodson, M. L., Schrock, III, R. D. and Lloyd, R. S. (1993) Biochemistry **32**, 8284–8290
167 Sun, B., Latham, K. A., Dodson, M. L. and Lloyd, R. S. (1995) J. Biol. Chem. **270**, 19501–19508
168 Tchou, J. and Grollman, A. P. (1995) J. Biol. Chem. **270**, 11671–11677
169 Hilbert, T. P., Boorstein, R. J., Kung, H. C., Bolton, P. H., Xing, D., Cunningham, R. P. and Teebor, G. W. (1996) Biochemistry **35**, 2505–2511
170 Razin, A. and Kafri, T. (1994) Prog. Nucleic Acid Res. Mol. Biol. **48**, 53–81
171 Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. and Cantoni, G. L. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2827–2831
172 Steinberg, R. A. (1995) Nucleic Acids Res. **23**, 1621–1624
173 Vairapandi, M. and Duker, N. J. (1993) Nucleic Acids Res. **21**, 5323–5327
174 Vairapandi, M. and Duker, N. J. (1996) Oncogene **13**, 933–938
175 Jost, J. P., Siegmann, M., Sun, L. and Leung, R. (1995) J. Biol. Chem. **270**, 9734–9739
176 Halliwell, B. and Gutteridge, J. M. C. (1989) Free Radicals in Biology and Medicine, Clarendon Press, Oxford
177 Shigenaga, M. K., Aboujaoude, E. N., Chen, Q. and Ames, B. N. (1994) Methods Enzymol. **234**, 16–33
178 Radman, M. (1976) J. Biol. Chem. **251**, 1438–1445
179 Demple, B. and Linn, S. (1980) Nature (London) **287**, 203–208
180 Harrison, L., Skorvaga, M., Cunningham, R. P., Hendry, J. H. and Margison, G. P. (1992) Radiat. Res. **132**, 30–39
181 Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H. and Cunningham, R. P. (1989) Biochemistry **28**, 4444–4449
182 Doetsch, P. W., Henner, W. D., Cunningham, R. P., Toney, J. H. and Helland, D. E. (1987) Mol. Cell Biol. **7**, 26–32
183 Kim, J. and Linn, S. (1989) J. Biol. Chem. **264**, 2739–2745
184 Bruand, C. and Ehrlich, S. D. (1995) Microbiology **141**, 1199–1200
185 Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A. and Merrick, J. M. (1995) Science **269**, 496–512
186 Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., et al. (1996) Science **273**, 1058–1073
187 Eide, L., Bjoras, M., Pirovano, M., Alseth, I., Berdal, K. G. and Seeberg, E. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 10735–10740
188 Roldan-Arjona, T., Anselmino, C. and Lindahl, T. (1996) Nucleic Acids Res. **24**, 3307–3312
189 Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T. and Cooper, J. (1994) Nature (London) **368**, 32–38
190 Hilbert, T., Chaung, W., Boorstein, R., Cunningham, R. and Teebor, G. (1997) J. Biol. Chem. **272**, 6733–6740
191 Aspinwall, R., Rothwell, D. G., Roldan-Arjona, T., Anselmino, C., Ward, C. J., Cheadle, J. P., Sampson, J. R., Lindahl, T., Harris, P. C. and Hickson, I. D. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 109–114
192 Wallace, S. S., Ide, H., Kow, Y. W., Laspia, M. F., Melamede, R. J., Petrullo, L. A. and LeClerc, E. (1988) in Mechanisms and Consequences of DNA Damage Processing (Friedberg, E. C. and Hanawalt, P. C., eds.), pp. 151–157, Alan R. Liss, New York
193 Melamede, R. J., Hatahet, Z., Kow, Y. W., Ide, H. and Wallace, S. S. (1994) Biochemistry **33**, 1255–1264
194 Jiang, D. Y., Hatahet, Z., Melamede, R. J. and Wallace, S. (1995) J. Cell. Biochem. **274** (abstract)
195 Frenkel, K., Cummings, A., Solomon, J., Cadet, J., Steinberg, J. J. and Teebor, G. W. (1985) Biochemistry **24**, 4527–4533
196 Frenkel, K., Chrzan, K., Troll, W., Teebor, G. W. and Steinberg, J. J. (1986) Cancer Res. **46**, 5533–5540
197 Hollstein, M. C., Brooks, P., Linn, S. and Ames, B. N. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 4003–4007
198 Cannon Carlson, S. V., Gokhale, H. and Teebor, G. W. (1989) J. Biol. Chem. **264**, 13306–13312
199 Boorstein, R. J., Chiu, L. N. and Teebor, G. W. (1992) Mol. Cell. Biol. **12**, 5536–5540
200 Zuo, S. J., Boorstein, R. J. and Teebor, G. W. (1995) Nucleic Acids Res. **23**, 3239–3243
201 Rideout, III, W. M., Coetzee, G. A., Olumi, A. F. and Jones, P. A. (1990) Science **249**, 1288–1290
202 Kasai, H., Iida, A., Yamaizumi, Z., Nishimura, S. and Tanooka, H. (1990) Mutat. Res. **243**, 249–253
203 Bjelland, S., Eide, L., Time, R. W., Stote, R., Eftedal, I., Volden, G. and Seeberg, E. (1995) Biochemistry **34**, 14758–14764
204 Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P. and Nishimura, S. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 4690–4694
205 Michaels, M. L. and Miller, J. H. (1992) J. Bacteriol. **174**, 6321–6325
206 Tajiri, T., Maki, H. and Sekiguchi, M. (1995) Mutat. Res. **336**, 257–267
207 Tchou, J. and Grollman, A. P. (1993) Mutat. Res. **299**, 277–287
208 Michaels, M. L., Tchou, J., Grollman, A. P. and Miller, J. H. (1992) Biochemistry **31**, 10964–10968
209 Michaels, M. L., Cruz, C., Grollman, A. P. and Miller, J. H. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 7022–7025
210 Boiteux, S., O'Connor, T. R., Lederer, F., Gouyette, A. and Laval, J. (1990) J. Biol. Chem. **265**, 3916–3922
211 Boiteux, S., O'Connor, T. R. and Laval, J. (1987) EMBO J. **6**, 3177–3183
212 Bailly, V., Verly, W. G., O'Connor, T. and Laval, J. (1989) Biochem. J. **262**, 581–589
213 van der Kemp, P. A., Thomas, D., Barbey, R., de Oliveira, R. and Boiteux, S. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 5197–5202

214 Duwat, P., de Oliveira, R., Ehrlich, S. D. and Boiteux, S. (1995) Microbiology. **141**, 411–417
215 de Oliveira, R., van der Kemp, P. A., Thomas, D., Geiger, A., Nehls, P. and Boiteux, S. (1994) Nucleic Acids Res. **22**, 3760–3764
216 Yacoub, A., Augeri, L., Kelley, M. R., Doetsch, P. W. and Deutsch, W. A. (1996) EMBO J. **15**, 2306–2312
217 Chung, M. H., Kim, H. S., Ohtsuka, E., Kasai, H., Yamamoto, F. and Nishimura, S. (1991) Biochem. Biophys. Res. Commun. **178**, 1472–1478
218 Lee, Y. S., Lee, H. S., Park, M. K., Hwang, E. S., Park, E. M., Kasai, H. and Chung, M. H. (1993) Biochem. Biophys. Res. Commun. **196**, 1545–1551
219 Yamamoto, F., Kasai, H., Bessho, T., Chung, M. H., Inoue, H., Ohtsuka, E., Hori, T. and Nishimura, S. (1992) Jpn. J. Cancer Res. **83**, 351–357
220 Jaruga, P. and Dizdaroglu, M. (1996) Nucleic Acids Res. **24**, 1389–1394
221 Nyaga, S. G. and Lloyd, R. S. (1996) FASEB J. **10**, A966
222 Hamilton, K. K., Kim, P. M. and Doetsch, P. W. (1992) Nature (London) **356**, 725–728
223 Nakabeppu, Y., Yamashita, K. and Sekiguchi, M. (1982) J. Biol. Chem. **257**, 2556–2562
224 Nakabeppu, Y. and Sekiguchi, M. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 2742–2746
225 Morikawa, K., Matsumoto, O., Tsujimoto, M., Katayanagi, K., Ariyoshi, M., Doi, T., Ikehara, M., Inaoka, T. and Ohtsuka, E. (1992) Science **256**, 523–526
226 Morikawa, K., Ariyoshi, M., Vassylyev, D. G., Matsumoto, O., Katayanagi, K. and Ohtsuka, E. (1995) J. Mol. Biol. **249**, 360–375
227 Vassylyev, D. G., Kashiwagi, T., Mikami, Y., Ariyoshi, M., Iwai, S., Ohtsuka, E. and Morikawa, K. (1995) Cell **83**, 773–782
228 Gordon, L. K. and Haseltine, W. A. (1981) J. Biol. Chem. **256**, 8608–8616
229 Grafstrom, R. H., Park, L. and Grossman, L. (1982) J. Biol. Chem. **257**, 13465–13474
230 Smith, C. A. and Taylor, J. S. (1993) J. Biol. Chem. **268**, 11143–11151
231 Loeb, L. A. and Christians, F. C. (1996) Mutat. Res. **350**, 279–286
232 Carson, M. (1991) J. Appl. Crystallogr. **24**, 958–961