Available online at www.sciencedirect.com
SCIENCE DIRECT®
DNA Repair 3 (2004) 1389–1407
DNA REPAIR
www.elsevier.com/locate/dnarepair

Invited mini review

Alkylation damage in DNA and RNA—repair mechanisms and medical significance

Finn Drabløs, Emadoldin Feyzi, Per Arne Aas, Cathrine B. Vaagbø, Bodil Kavli, Marit S. Bratlie, Javier Peña-Diaz, Marit Otterlei, Geir Slupphaug, Hans E. Krokan*

Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, N-7489 Trondheim, Norway

Received 20 February 2004

Available online 26 May 2004

Abstract

Alkylation lesions in DNA and RNA result from endogenous compounds, environmental agents and alkylating drugs. Simple methylating agents, e.g. methylnitrosourea, tobacco-specific nitrosamines and drugs like temozolomide or streptozotocin, form adducts at N- and O-atoms in DNA bases. These lesions are mainly repaired by direct base repair, base excision repair, and to some extent by nucleotide excision repair (NER). The identified carcinogenicity of O^6-methylguanine (O^6-meG) is largely caused by its miscoding properties. Mutations from this lesion are prevented by O^6-alkylG-DNA alkyltransferase (MGMT or AGT) that repairs the base in one step. However, the genotoxicity and cytotoxicity of O^6-meG is mainly due to recognition of O^6-meG/T (or C) mispairs by the mismatch repair system (MMR) and induction of futile repair cycles, eventually resulting in cytotoxic double-strand breaks. Therefore, inactivation of the MMR system in an AGT-defective background causes resistance to the killing effects of O^6-alkylating agents, but not to the mutagenic effect. Bifunctional alkylating agents, such as chlorambucil or carmustine (BCNU), are commonly used anti-cancer drugs. DNA lesions caused by these agents are complex and require complex repair mechanisms. Thus, primary chloroethyl adducts at O^6-G are repaired by AGT, while the secondary highly cytotoxic interstrand cross-links (ICLs) require nucleotide excision repair factors (e.g. XPF-ERCC1) for incision and homologous recombination to complete repair. Recently, *Escherichia coli* protein AlkB and human homologues were shown to be oxidative demethylases that repair cytotoxic 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) residues. Numerous AlkB homologues are found in viruses, bacteria and eukaryotes, including eight human homologues (hABH1-8). These have distinct locations in subcellular compartments and their functions are only starting to become understood. Surprisingly, AlkB and hABH3 also repair RNA. An evaluation of the biological effects of environmental mutagens, as well as understanding the mechanism of action and resistance to alkylating drugs require a detailed understanding of DNA repair processes.

Keywords: *N*-nitroso compounds; Alylating cytostatics; Alkyltransferase; Base excision repair; AlkB homologues; DNA repair modulation

1. Introduction

Humans are exposed to alkylating compounds produced endogenously [1,2] and in the environment [3–5]. In addition, alkylating agents are used as cytostatic drugs in cancer therapy in doses that by far exceed those humans are exposed to from other sources [6]. Alkylating agents are mutagenic and genotoxic. The focus of research has therefore been on their ability to damage DNA, and on the processing and consequences of such damage. However, alkylating agents also form adducts with RNA and protein and this is likely to contribute to the cytotoxicity. All organisms examined have repair systems for alkylation damage to DNA and the type of repair system involved depends on the type of lesion (reviewed in [7]).

In total, DNA damage from all endogenous sources give rise to some 20,000 lesions per day per cell. Of these, spontaneous depurination, damage from reactive oxygen species and deamination of bases represent the main sources [8], while the relative contribution of alkylation damage is not well established. However, the mere existence of several specific repair systems for alkylation damage indicates that endogenous and environmental exposures must be significant. In fact, all the main mechanisms of DNA repair (direct damage reversal, base excision repair, nucleotide excision repair (NER), mismatch repair and recombination repair [9]) are

* Corresponding author. Tel.: +47 73598695; fax: +47 7359880.
E-mail address: hans.krokan@ntnu.no (H.E. Krokan).

1568-7864/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.dnarep.2004.05.004

involved in the processing of either primary alkylation products (e.g. methylations and chloroethylations), or secondary damage from such products (e.g. abasic sites, strand breaks and interstrand cross-links (ICLs)).

It has only recently become evident that at least some types of RNA alkylations are also repaired. Whereas the significance of repair of alkylations in DNA is well established, the significance of RNA repair is less clear, but it does operate in vivo in *Escherichia coli* to increase the survival of alkylation-damaged RNA bacteriophages [10]. Considering the elaborate apparatus involved in quality control of RNA species at different levels during their synthesis [11–13], repair of RNA should possibly not come as a total surprise. Furthermore, proteins are repaired both at the amino acid level, and at the structural level, and failure to repair damaged amino acid residues in protein is fatal in mouse models [14–16] (reviewed in [17]). This underlines the significance of repair of macromolecules downstream of the DNA level.

Defective DNA repair generally causes cancer, developmental abnormalities and early aging (reviewed in [7]). It is now established from clinical syndromes and/or from mouse models that defects in each of the major DNA repair pathways are associated with increased risk of cancer (reviewed in [9,18]). Furthermore, modulation of DNA repair capacity may improve the usefulness of alkylating therapeutic drugs [19,20]. A detailed understanding of how cells handle macromolecular damage is therefore potentially very important for several aspects of clinical medicine.

The present mini review focuses on alkylation damage to macromolecules, their processing and their biological consequences in the cell. In this context, we also present an in depth analysis of properties, evolution and possible functions of the AlkB family of genes and their proteins. We also briefly discuss how modulation and chemical manipulation of such defense mechanisms can be exploited therapeutically.

2. Alkylating agents are formed endogenously and in the environment, and are used as anti-cancer drugs

Alkylating agents are ubiquitous. They are generated endogenously during the metabolism, and are found in the air, water and foods, although generally in low concentrations. Patients are exposed to alkylating agents in much higher doses from alkylating anti-cancer drugs. By selectively decreasing repair capacity in target tissues, or by enhancing repair in other tissues, improved therapy may be achieved. Some examples of such efforts will be discussed below. Alkylating agents may form adducts at all O- and N-atoms in nucleobases, as well as on O-atoms in phosphodiesters. The alkylation pattern depends on the agent, the position in DNA/RNA, and is different in single-stranded and double-stranded nucleic acids (Table 1). Base alkylations are both genotoxic and cytotoxic. In general, *O*-alkylations (O^6-alkylG and O^4-alkylT) are highly mutagenic and

Table 1
Methylation patterns of single-stranded vs. double-stranded nucleic acids upon reaction with MMS and MNU[a,b]

Site of methylation	MMS			MNU		
	ssDNA/RNA[c]	RNA[d]	dsDNA	ssDNA/RNA[c]	RNA[d]	dsDNA
Adenine						
N1-	18	11	3.8	2.8	2	1.3
N3-	1.4	1.5	10.4	2.6	nd[e]	9
N7-	3.8	nd	1.8	1.8	nd	1.7
Guanine						
N3-	∼1	nd	0.6	0.4	1	0.8
O^6-	–	0.2	0.3	3	4	6.3
N7-	68	72	83	69	80	67
Uracil/thymine						
O^2-	nd	nd	nd	nd	nd	0.11
N3-	nd	0.4	nd	nd	nd	0.3
O^4-	nd	nd	nd	nd	nd	0.4
Cytosine						
O^2-	nd	nd	–	nd	nd	0.1
N3-	10	7	<1	2.3	1.5	0.6
Diester	2	nd	0.8	∼10	nd	17

[a] Adopted from refs. [24,25].
[b] Expressed as percent of the total methylation.
[c] The values are from data compilation by [25], and originate from experiments with the ssDNA phage M13, and RNA from TMV, yeast, HeLa cells, animal ribosomes, and μ2 phage.
[d] These data are from the original work of [24] and indicate methylation pattern in RNA from bacteriophage R17. The values are percent of the total radioactivity analysed for methylation. 92% (MMS) and 89% (MNU) of the total radioactivity were recovered as identifiable and detectable residues (>0.5% of total radioactivity).
[e] nd: not determined.

genotoxic, whereas *N*-alkylations (e.g. 3-alkylA and 1-alkylA) are cytotoxic, but relatively less mutagenic [19,21–23]. The quantitatively dominating type of lesion in both DNA and RNA is 7-alkylG, which is considered as relatively harmless in itself. However, it is subject to rapid spontaneous depurination and enzymatic removal, thus creating cytotoxic abasic sites. Alkylation damage in RNA generally follow the same pattern as for single-stranded DNA [24]. Alkylating agents of the S_N1-type (e.g. *N*-methyl-*N*-nitrosourea, MNU) alkylate both oxygens and nitrogens in nucleic acids, whereas S_N2 reagents (e.g. methyl methanesulfonate, MMS) mainly alkylate nitrogens [25,26]. As discussed in more detail below, bifunctional alkylating agents, most commonly chloroethylating drugs, to a large extent generate more complex lesions, including interstrand cross-links. Monofunctional and bifunctional agents both generate toxic repair intermediates. Thus, even after induction of a simple lesion like 3-methyladenine (3-meA) in DNA using methyl-lexitropsin (Me-lex) that targets adenine in the minor groove of DNA, a wide range of effects is observed. These include sister chromatid exchange, DNA breaks, S-phase arrest, accumulation of p53, and apoptosis [27].

2.1. Endogenous alkylating agents

There are many possible sources of endogenous alkylating agents and several compounds have been identified, but their relative significance remains unknown. *S*-Adenosylmethionine, a methyl-donor in many biochemical reactions, is also a weak chemical methylating agent that has been shown to induce mutations in DNA. It mainly generates 7-meG and 3-meA, but little O^6-meG [1,28]. Nitrosated amines and related compounds in *E. coli* generate directly acting mutagenic alkylating agents, particularly in cells that are Ada- and Ogt-deficient and consequently deficient in repair of O^6-alkylG. This strongly indicates that the agents formed endogenously generate O-alkylating agents. The nitrosation reactions are predominantly enzymatic and occur to a higher degree in starving bacterial cells [2]. Probably such reactions will also take place in the large intestine. Several mutagenic *N*-alkyl-*N*-nitrosocompounds shown to cause alkylation of O^6-G in DNA are generated by nitrosation of glycine and glycine derivatives [29,30]. Furthermore, bile acids may be nitrosated to give O-alkylating mutagenic compounds [31]. Generation of endogenous alkylating agents by *E. coli* and from bile acids and amines is consistent with the significant content of O^6-meG and 7-meG found in DNA from colo-rectal tissues, although the source causing the adducts remains uncertain. Interestingly, the interindividual adduct levels vary at least 100-fold, but tend to be highest in the distal colon and rectum where cancers are most frequent (reviewed in [32]). Somewhat surprisingly, there does not seem to be a general correlation between adduct levels in individuals and frequency of cancer, whereas there is an inverse correlation between activity of the repair enzyme O^6-alkylguanine-DNA alkyltransferase (AGT or MGMT) and cancer risk [32].

2.2. Environmental alkylating agents

The most abundant atmospheric halocarbon is chloromethane (methyl chloride) gas which is generated in many terrestrial environments, e.g. by plants and fungi, but also industrially. Apparently, a large fraction is made abiotically in plant material [5]. Some 3.5–5 million tons of chloromethane is produced per year. In marine environments, bromomethane and other bromo-compounds are more abundant. Iodomethane is also formed in the environment [4]. Chloromethane and other halocarbons are present in the ambient air in small, but measurable concentrations [33]. Such compounds alkylate DNA, hence being mutagenic [34] and possibly carcinogenic [35]. However, presently there is inadequate evidence for the carcinogenicity of chloromethane to animals and humans [36].

Although not quantitatively dominating in the environment, *N*-nitroso compounds formed in tobacco smoke may represent the major environmental source of exposure to alkylating agents for humans. However, food compounds may also be nitrosated to agents that alkylate DNA, as discussed briefly below. Tobacco smoke contains both tobacco-specific *N*-nitroso compounds and several other nitroso compounds, e.g. *N*-nitrosodimethylamine [37]. Among the seven known tobacco-specific nitrosamines, the most carcinogenic in laboratory animals are 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), and *N*′-nitrosonornicotine (NNN). After metabolic activation they give rise to DNA alkylations (reviewed in [3]). NNK and its metabolite NNAL give rise to O^6-meG and other methylated bases. In addition, NNK and NNAL, as well as NNN, may give rise to the larger pyridyloxobutyl DNA adducts at O^6-G (Fig. 1) and other positions [3].

To monitor environmental exposure to tobacco-specific nitrosamines, several potential biomarkers have been evaluated. Among these NNAL currently seems best suited. NNAL and its glucuronides are metabolites of NNK and give specific and reasonably quantitative information on exposure to tobacco smoke. Whereas NNAL is also a carcinogen, the glucuronides are detoxification products. The ratio between NNAL and glucuronides varies over 10-fold between smokers and this ratio has been suggested as a useful indicator of cancer risk (reviewed in [38]). Tobacco-specific *N*-nitroso compounds are probably mainly formed from tobacco alkaloids during the curing and smoking of tobacco. However, at least some of the tobacco-specific nitrosamines may also be formed by endogenous nitrosation of tobacco alkaloids [39]. One of the tobacco alkaloids, myosmine, is also present in a variety of foods, including several fruits, vegetables and even milk, and it is easily nitrosated to DNA-alkylating compounds, including NNN [40]. It is also evident that different foods may contain nitrosating agents such as NO_x

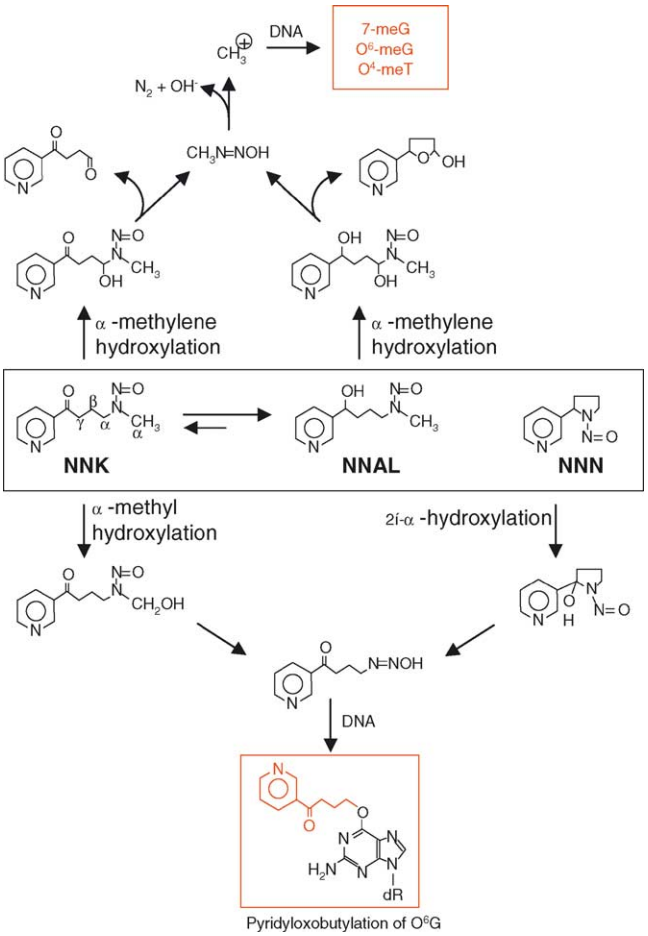

Fig. 1. Overview of metabolism and DNA adduct formation by the tobacco-specific nitrosamines NNK, NNAL and NNN. Results are conclusions from studies in laboratory animals and humans. Some of the metabolites and potential adducts are not shown.

and compounds, e.g. amines, that can be nitrosated to mutagenic compounds (reviewed in [41]). Furthermore, cured foods and fish may contain significant amounts of *N*-nitroso compounds and at least in some parts of the world, e.g. regions of China and Japan, this may be an important cause of gastric and esophageal cancers (reviewed in [41]).

2.3. Alkylating drugs

Alkylating drugs are the oldest anti-cancer drugs and remain important for treatment of several types of cancer (reviewed in [6,42]). The effects of these drugs are strongly modulated by DNA repair processes. Understanding and controlling repair processes may therefore allow development of new therapies by protecting normal tissues or by potentiating effects in target tissues [43]. Alkylating drugs are mostly methylating agents (e.g. temozolomide and streptozotocin, an antibiotic) or chloroethylating agents (e.g. carmustin, lomustine and fotemustine). In both cases O^6-G is an important target for alkylation, but several other sites are also alkylated. The chloroethylating agents have complex effects. First they generate O^6-chloroethylG and other chloroethyl-adducts. At the O^6-position this is followed by dechlorination, formation of an unstable 1-O^6-ethanoguanine adduct and subsequent cross-linking between the ring nitrogens in 1-position of guanine and 3-position of cytosine (1-(3-cytosinyl)-2-(1-guanosinyl)-ethane) (reviewed in [19,43]). Several other drugs also target the O^6-G position and generate O^6-meG or cross-links. However, some alkylating drugs exert their strongest effects through alkylation of other positions and other types of cross-links [43]. In contrast to many important environmental alkylating agents, the therapeutic drugs do not require metabolic activation, but are spontaneously decomposed in aqueous solutions.

3. Repair of alkylation lesions—more complex than the lesions

In *E. coli*, as well as in eukaryote cells, at least four different mechanisms are involved in repair of primary damage to DNA caused by simple monofunctional alkylating agents. These are direct base repair either by methyltransferases [44] (reviewed in [7]) or by oxidative demethylases [10,45–47], base excision repair initiated by DNA glycosylases (BER) [48] (reviewed in [22,23]) and nucleotide excision repair [49,50]. In addition, recombination repair may be significant, possibly mainly for repair of secondary lesions such as strand breaks (Fig. 2) [49]. Except after treatment with alkylating agents, the level of exposure is likely to be much smaller than the doses normally used in experimental settings. Possibly, some of these repair systems are mainly operative after such massive damage. Which mechanism(s) are most important after low level endogenous or environmental exposure remains undetermined, but would presumably be alkyltransferases repairing *O*-alkylations, and BER and oxidative demethylases repairing *N*-alkylations, with NER as a backup for at least some of these lesions.

Several types of primary alkylation products, e.g. O^6-chloroethylguanine, as well as secondary lesions, e.g. DNA–DNA cross-links, mismatches, strand breaks, and even protein–DNA cross-links, occur as main products after exposure to bifunctional agents [51]. AGT repairs O^6-chloroethylG, whereas other repair systems, e.g. mismatch repair and homologous recombination, are important for repair of secondary lesions (Figs. 2 and 3). The significance of homologous recombination and double-strand break repair may be limited to the types of damage inflicted by bifunctional alkylating agents, and after extensive experimental damage. Knowledge about repair of alkylation lesions in *E. coli* has been instrumental in the discovery of similar repair processes in eukaryotes. The inducible Ada regulon has been particularly important in this respect (reviewed in [52]), although the corresponding defense systems in mammalians are not regulated in the same way. The repair proteins encoded by genes in the Ada-regulon com-

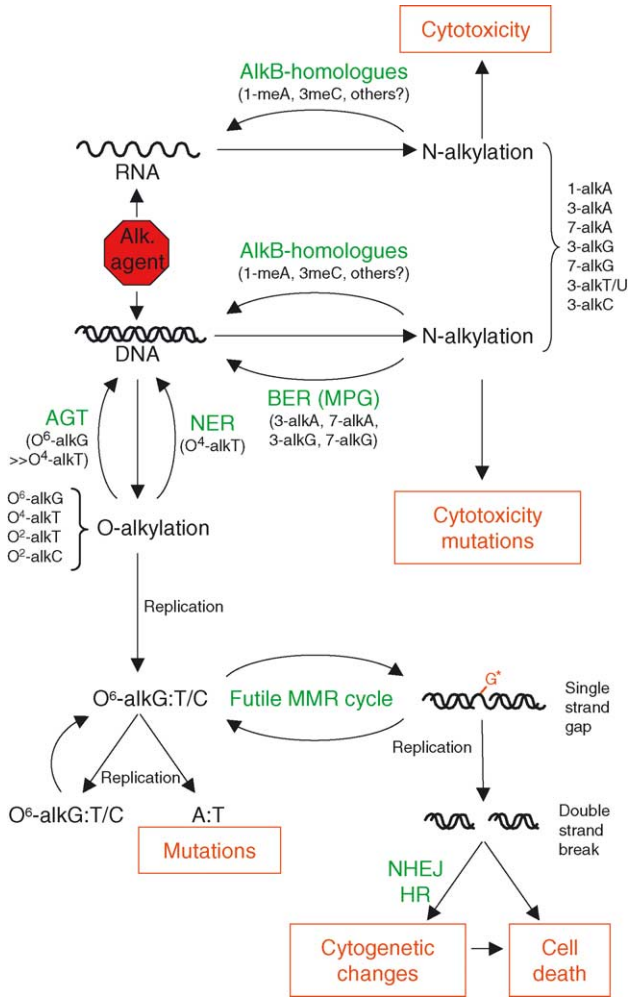

Fig. 2. Pathways for alkylation damage and repair of DNA and RNA bases. *N*-alkylations are repaired either directly by oxidative demethylases (DNA and RNA) or by BER (DNA). *O*-alkylations are repaired either directly by AGT or by NER. If not processed prior to replication, O^6-alkG:T mispairs may be recognized by the MMR system, thereby inducing futile repair cycles eventually leading to DSBs, cytogenetic changes and cell death.

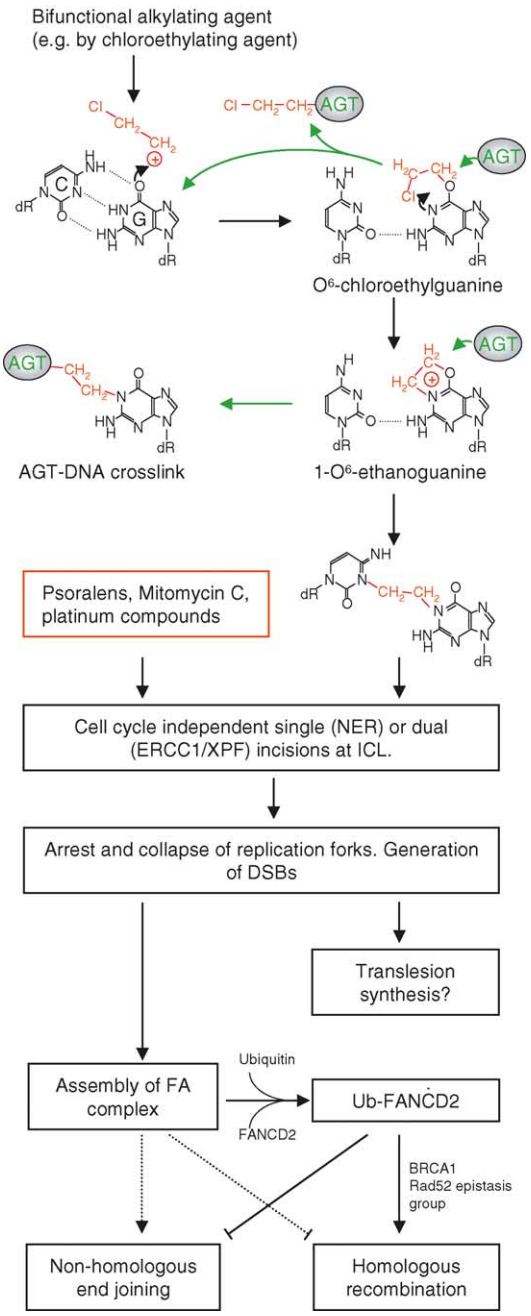

Fig. 3. Sources of DNA interstrand cross-links (ICLs) and potential routes for ICL processing in mammals. Formation of ICL from chloroethylating agents is shown as an example. Note that factors from several DNA metabolizing pathways may contribute, and the mechanism of choice apparently depends on the type of cross-linking agent, the cell type and the proliferative status of the cell.

prise the methyltransferase Ada, which after receiving the methyl-group is the principal transcription factor in the regulon, the DNA glycosylase AlkA, the oxidative demethylase AlkB and the less well understood AidB (reviewed in [52]). Human cells have homologues of Ada, AlkB and AidB, but apparently no AlkA homologue. Instead, mammalian cells have an unrelated broad specificity DNA glycosylase (MPG or ANPG) that repairs many of the same DNA lesions [22].

3.1. Repair of O^6-alkylG and O^4-alkylT in DNA by alkyltransferases

O^6-alkylG in DNA is probably the principal mutagenic damage formed by alkylating agents. In bacteria, yeast and eukaryotes this lesion is repaired in one step by transferases (Fig. 2). O^6-Alkyl transferases are highly conserved in bacteria [52], lower eukaryotes [53] to mammalian species [54]. *E. coli* has two genes for DNA repair methyltransferases, the inducible *ada* of the Ada regulon, and *ogt* which is constitutively expressed [7,52]. Ada and Ogt are *O*-alkylltransferase that transfer the methyl group to a cysteine residue in the repair protein itself, which is thereby permanently inactivated. Ada and Ogt repair O^6-meG and O^4-meT, and in addition Ada repairs alkylphosphotriesters.

The 39 kDa Ada protein contains a 19 kDa domain which repairs O^6-meG and O^4-meT and a 20 kDa N-terminal domain which repairs alkylphosphotriester (reviewed in [52,55]). The 19 kDa Ogt protein has properties similar to the 18 kDa domain of Ada, but is more efficient than Ada in removing O^4-alkylthymines [56,57] (reviewed in [55]). O^6-meG was shown to be repaired by a similar, irreversibly inactivated O^6-alkylguanine-DNA alkyltransferase (AGT or MGMT), in mice [58], rats and humans [59] (reviewed in [54]). Mammalian AGT repairs O^4-meT very inefficiently and possibly not phosphotriesters at all. Rather, it may bind to O^4-meT and shield it from repair by other repair systems, for instance NER, thereby actually enhancing mutations from this minor but strongly miscoding lesion [60,61]. AGT prefers lesions in double-stranded B-DNA and repair Z-DNA, DNA:RNA hybrids and single-stranded DNA more slowly (reviewed in [54]). Although AGT has preference for O^6-meG, it also removes longer alkyl groups such as ethyl-, *n*-propyl, *n*-butyl and 2-chloroethyl from the O^6-position. The latter is probably the most important primary adduct of chloroethylating anti-cancer drugs, such as chlorambucil, carmustine (BCNU), fotemustine and lomustine (CCNU) [51].

It follows from the irreversible inactivation of AGT after reaction with a substrate that any substrate for the protein will act as an irreversible inhibitor. This can in principle be exploited therapeutically to enhance sensitivity of a tumor to alkylating drugs. O^6-Benzylguanine is a candidate for this approach. It forms an *S*-benzylcysteine adduct at the active site of AGT and has been shown to sensitize several different cancer types to treatment with methylating and chloroethylating drugs [54,62,63].

The large gene (∼300 kb) for human AGT is located in chromosomal position 10q26 and contains five exons and a very large second intron (170 kb) [64]. The ∼24 kDa AGT protein contains the methyl-acceptor cysteine [59] in position 145 in a highly conserved PCRVV motif [65,66]. AGT-activity varies several-fold between different individuals [67], organs [65,67] and cell types within organs [68–70]. At the cellular level, the sensitivity to S_N1 alkylating agents inversely correlates with AGT content, but whether the large interindividual variation in AGT activity has a genetic or environmental basis remains poorly understood, in spite of considerable efforts [65].

Transgenic mice overexpressing human AGT were found to be protected from developing thymic lymphomas after exposure to MNU [71]. In contrast, knockout mice deficient in AGT activity are hypersensitive to the killing effect of alkylating agents and develop numerous lymphomas [72,73]. Loss of AGT activity also renders the mice sensitive to chemotherapeutic alkylating agents [74]. In cell lines, loss of AGT activity is relatively common due to promoter hypermethylation or mutations and these cells are sensitive to the killing effect of alkylating agents. However, loss of capacity for mismatch repair surprisingly abolished the alkylation sensitivity in such cell lines, suggesting that mismatch repair proteins are involved in the cytotoxicity from alkylating agents [75–77]. In the AGT-defective mice, inactivation of mismatch repair by mutation in a central mismatch repair gene (*MLH1*) restored alkylation resistance in terms of survival, but not in terms of tumorigenesis since the occurrence of lymphomas remained high [78]. These observations demonstrate the genotoxicity and tumorigenicity of O^6-meG, but also reveal the significance of the MMR system for genotoxicity. This is highly relevant for understanding the rather complex mechanism of cell killing and development of resistance after treatment with alkylating anti-cancer drugs.

As a consequence of studies on cell lines and mouse models, the current view regarding the mechanism of action of monofunctional alkylating cytostatic drugs, e.g. streptozotocin or temozolomide, is that replication over unrepaired O^6-meG:C results in an O^6-meG:T mismatch (or possibly an O^6-meG:C ambiguous pair). In the next round of replication this results in an A:T transition mutation, and again to an O^6-meG:C pair or an O^6-meG:T mismatch. The O^6-meG:T or C pair is recognized by the MutSα complex (hMSH2 and hMSH6) which initiates MMR that creates a gapped duplex after incision of the newly replicated strand. Since O^6-meG remains in the template, this process may be repeated in a "futile repair loop" that eventually results in highly toxic double-strand breaks that are intermediates both in apoptotic and recombination pathways [51] (Fig. 2).

3.2. Repair of O^6-chloroethylguanine, 1-O^6-ethanoguanine and interstrand cross-links

Bifunctional alkylating drugs have two reactive groups that react with DNA. Here the activated chloroethyl group attacks the O^6-G position forming O^6-chloroethylG that is substrate for direct repair by AGT (Fig. 3). If not repaired, spontaneous dechlorination is followed by formation of 1-O^6-ethanoguanine, which is a substrate for AGT, but the protein remains covalently bound to DNA through the linkage to the N1 atom (Fig. 3). Thus, AGT is inactivated and remains covalently bound as an AGT-DNA adduct [79,80] (reviewed in [19,51]).

If not acted upon by AGT, the 1-O^6-ethanoguanine adduct may react with cytosine in the complementary strand to yield a highly toxic DNA–DNA cross-link between position 1 in guanine and position 3 in cytosine (1-(3-cytosinyl)-2-(1-guanosinyl)-ethane) [81,82], reviewed in [51]. Fanconi anemia (FA) is a rare autosomal recessive genetic disorder characterized by congenital abnormalities, progressive bone marrow failure and increased risk of leukemia and other cancers. A marked chromosomal instability and hypersensitivity to interstrand cross-linking agents are cell biological hallmarks of the disease and are important for establishing the diagnosis [83]. FA cells have also been important for clarifying some aspects of the repair of interstrand cross-links, which have many different structures depending on the inducing agent (e.g. BCNU, *cis*-platin,

mitomycin C). The repair mechanism for this type of DNA interstrand cross-link seems to use a combination of FA proteins, nucleotide excision repair factors and factors used in homologous recombination (HR) [84]. Double-strand breaks occurring as a consequence of replication encountering single-strand regions may require non-homologous end-joining (NHEJ). Thus, NER, HR, NHEJ, and possibly even translesion synthesis, may all be involved in repair of secondary lesions, and may be operative in the vicinity of each other, in part on the same substrates [85]. Recent results indicate that different FA proteins are involved in positive and negative regulation of the choice of pathway (HR or NHEJ) [86]. However, the details in the repair process remain unclear and are not necessarily identical for all types of cross-links. Recent research indicates that initial steps in the repair pathway involve incisions at one or both sides of the lesion by the ERCC1/XPF complex and/or other NER factors. These steps are carried out independent of proteins involved in the FA/BRCA pathway and are cell cycle independent. However, the next steps are S-phase dependent and require factors in the FA/BRCA pathway and factors required for homologous recombination [87]. Based on data discussed above we present a model for repair of primary and secondary adducts resulting from chloroethylating drugs in Fig. 3.

3.3. Repair of alkylation lesions by base excision repair (BER)

After the original discovery of a glycosylase that removed 3-meA from DNA [48], such activity has been identified in most, if not all, organisms examined. While 7-meG is the quantitatively major adduct from S_N1 and S_N2 methylating agents in double-stranded and single-stranded DNA, 3-meA is thought to be a major cytotoxic adduct generated in double-stranded, and to a smaller extent in single-stranded DNA. Like DNA glycosylases that recognize oxidative damage, DNA glycosylases that recognize alkylated bases generally have a strong preference for double-stranded DNA substrate, although *E. coli* AlkA, but not Tag, works reasonably well on single-stranded DNA [88]. The broad specificity DNA glycosylase AlkA is inducible while the narrow specificity enzyme Tag is constitutively expressed. These enzymes are not related at the sequence level. While Tag recognizes only 3-meA and 3-meG, AlkA recognizes a wide variety of substrates, including alkylation products, the deamination product of adenine (hypoxanthine), ring opened purines, certain oxidation products and others, including some etheno-adducts and even O^2-alkylated pyrimidines [22,23,89]. Mammalian species apparently have only one glycosylase that recognizes alkylation lesions. The human methyl-purine DNA glycosylase (MPG, ANPG or AAG), is not related to AlkA or Tag at the sequence level, nor at the structural level, but has a broad specificity that resembles that of AlkA, except that it does not remove O^2-alkylated pyrimidines [22]. Apparently, MPG does not release products generated by chloroethylating agents [90]. Furthermore, targeted disruption of the MPG-gene in mouse, only resulted in a modestly increased sensitivity to MMS, but not to the chloroethylating agent BCNU [91]. Also expression of human MPG in CHO cells did not increase resistance to BCNU [92]. However, in one study targeted disruption of the mouse gene encoding MPG rendered the mouse cells sensitive to both BCNU and mitomycin C, so the possible function of MPG in repair of chloroethyl adducts is not excluded for all cell types [93], but must presumably have a minor role in repair of these lesions.

AlkA, MPG and the corresponding enzyme from *Saccharomyces cerevisiae* also promote release of intact guanines with very low, but significant frequencies. AlkA is in fact able to remove all the normal bases, although especially pyrimidines are removed at very low rates [94]. In agreement with these results, transformation of *E. coli* with plasmids expressing different levels of AlkA resulted in increased spontaneous mutation frequencies that correlated with the expression levels. It was therefore suggested that the release of these structurally highly diverse substrates depended on weakening of the glycosylic bond as a common property [94]. However, while there is a correlation between removal rates and weakening of the glycosylic bond, the mechanism is more complex [95]. Like several other enzymes acting on modified bases, base removal by MPG requires flipping of the damaged nucleotide from the interior of the helix into the catalytic cavity. Binding and release of the base then depend on the shape of the base, hydrogen bonding characteristics, aromaticity and salt bridging between the enzyme and flanking phosphates in DNA [96]. A water molecule is positioned and deprotonated by Glu125. The resulting nucleophile then attacks the glycosylic bond [96]. The k_{cat}/K_m is highest for 3-meA and 7-meG, but only two-fold higher than for hypoxanthine and some 10-fold higher than for ethenoadenine (εA). However, due to the positive charges on 3-meA and 7-meG these are lost spontaneously at very high rates, whereas the glycosylic bond of the uncharged hypoxanthine is three to four orders of magnitude more stable. The enzymatic rate enhancement is therefore by far highest for removal of hypoxanthine and it was suggested that MPG may have evolved in response to this lesion [95]. Etheno-adducts at guanine, adenine or cytosine result from lipid peroxidation and from exposure to vinyl chloride. A very recent study reported that εA, in contrast to εG, is not released, but that MPG binds strongly to this lesion when present in single-stranded or double-stranded DNA. This “hijacking” of MPG by εA inhibits BER of other substrates for MPG thus increasing cytotoxicity and possibly mutagenicity [97].

In contrast with effects of AGT, overexpression of MPG does not protect from the killing effect after exposure to alkylating agents, but rather enhances it. This effect was found to be more pronounced when MPG was engineered to be targeted to mitochondria. This may indicate that a balanced expression of MPG is important for its function in BER of alkylation damage [98]. This is consistent with the previ-

ously observed lack of correlation between MPG activity and alkylation resistance, whereas AGT activity is clearly important for resistance [99]. Furthermore, MPG-deficient myeloid progenitor cells in the bone marrow were more resistant to MMS and MeLex than wild-type cells [100]. However, embryonic stem cells from mice deficient in MPG have increased sensitivity to alkylating agents, in this case a chloroethylating agent [93]. These findings indicate that protection or sensitisation by MPG to the killing effects of alkylating agents may be tissue-specific and dependent upon the type of chemical. Furthermore, MPG only seems to offer protection within regulated levels of expression [101]. Similar findings were previously observed when overexpressing AlkA in *E. coli* [102]. This was also a conclusion after overexpression of AlkA, Tag or the yeast enzyme Mag1 in *E. coli*. Here the broad specificity enzymes Mag1 and AlkA were more toxic and mutagenic than *E. coli* Tag that only removes 3-meA and 3-meG [103]. High expression gives rise to a large number of abasic sites, from removal of both damaged and normal bases, and this repair intermediate may be the cause of toxicity and mutagenicity.

These observations indicate that a balanced expression of BER proteins is essential. Too rapid removal of damaged bases results in generation of too many cytotoxic and mutagenic abasic sites. It actually seems that MPG also has a built-in property that ensures that the rate is not too high. First, the single turnover rate of release of alkylated bases is generally relatively low, and some 10-fold lower for the most abundant lesion 7-meG, as compared with 3-meA (1.1 and 11 min^{-1}, respectively) [95]. Second, MPG has a strong affinity for the abasic site, thus shielding the fragile and toxic abasic site until BER can proceed [104]. Third, crude extracts from mammalian cells remove 7-meG more slowly than the less abundant 3-meA [91].

Deficiencies in the subsequent common steps in BER result in enhanced sensitivity to alkylating agents. Thus, fibroblasts from a mouse DNA polymerase β null mutant have strongly increased sensitivity to monofunctional alkylating agents, but not to BCNU [105]. Poly(ADP-ribose) polymerase-1 (PARP-1) is activated by strand breaks and participates in gap-sealing together with DNA ligase 3 and XRCC1. Inhibition of PARP-1 using a novel inhibitor, AG14361, restores sensitivity to the methylating drug temozolomide in mismatch-repair deficient cells, which have lost killing sensitivity to O^6-meG via the AGT/MMR pathway [106]. Like other monofunctional agents, temozolomide produces approximately 10-fold more 7-meG than O^6-meG, but under normal circumstances, the major cytotoxicity is presumably mediated via the AGT/MMR pathway and not the BER pathway. However, when the cytotoxicity from O^6-alkylation is abolished by inactivation of mismatch repair, the toxicity from substrates recognized by the BER pathway becomes important. Inhibition of DNA polymerase β or PARP-1 may be a means to overcome resistance that frequently develops due to selection of mismatch repair deficient cells during therapy.

4. Improved cancer therapy by modulation of alkylation sensitivity

The types of lesion, as well as the mechanisms of toxicity and resistance are clearly different for monofunctional and bifunctional agents. Identification of the most likely successful targets therefore depends on a thorough understanding of these factors. In principle, successful treatment can be achieved either by specifically sensitizing the cancer cells or enhancing resistance of non-target cells. Some of the strategies that are being explored are: (a) sensitizing target tissues by inhibiting the activity of AGT; (b) protection of non-target tissues by expressing Ada or related transferases; (c) direct application of alkylating agents to target tissues, with or without sensitization; (d) examining expression levels of AGT in tissues to guide dose of alkylating agent; (e) inactivating common steps in the BER pathway; (f) interfering with signaling pathways controlling interstrand cross-link repair; (g) development of novel cross-linkers that produce cross-links that are not efficiently repaired. Some examples of clinical results are described below. For more extensive and broader information, other reviews should be consulted [6,43,54,107].

AGT prevents killing by both methylating and chlorethylating drugs. The response to treatment has been shown to correlate inversely with the AGT activity in several studies [108–110]. Clinical phase I and II trials are underway, and some completed, with regard to modulation of AGT activity. Several studies have also been carried out with mouse models, either in animals with targeted gene disruptions, or by testing of inhibitors in animals carrying tumor xenografts. As mentioned, O^6-benzylguanine efficiently inactivates AGT by being a pesudosubstrate. Mice carrying xenografts have been shown to have improved therapeutic effects from alkylating agents when given O^6-benzylguanine, thus paving the way for clinical trials (reviewed in [111,112]). A phase II study on humans has now been completed for glioblastoma multiforme patients. This study comprised 58 patients at their first relapse after surgery and radiotherapy, but they had not previously received chemotherapy. The patients received O^6-benzylguanine, procarbazine, carmustine (BCNU) and vincristine (PCV). The progression-free survival at 6 months was better than historical data, but substantially more side effects were also observed than with the classical PCV regime [113].

Inactivation of BER as a target for enhanced therapy with alkylating drugs has been explored in cell lines and mice carrying xenografts, but to our knowledge not in humans. As discussed above, the initial step involving MPG does not seem promising as a target. However, inhibition of the common steps by methoxyamine that reacts with abasic sites sensitizes cells or xenografts to the methylating agent temozolomide [114]. Inhibition of PARP-1, which is involved in repair of single-strand breaks, with AG14361, potentiated the effect of temozolomide in cell lines. More importantly, the potentiating effect was relatively larger in

MMR-deficient cells that are resistant to alkylating agents. Thus, this could possibly be a means to overcome tolerance due to selection of MMR-deficient cells during chemotherapy [106].

Interstrand cross-links and double-strand breaks are highly cytotoxic and their mechanisms of repair are complex. Drugs that interfere with these processes, including signaling of such damage, could therefore be attractive drug candidates. The c-abl kinase inhibitor imatinab was recently shown to enhance the cytotoxic effect of chlorambucil (CLB) in chronic lymphatic leukemia cells in vitro, probably by decreasing Rad51-related DNA repair of CLB-induced DNA lesions [115].

5. Direct reversal of alkylation damage by AlkB homologues

The *E. coli* protein AlkB was recently shown to be an oxidative DNA demethylase that repairs the cytotoxic lesions 1-meA and 3-meC in DNA. This makes it an important component for maintenance of genetic information. A detailed mapping of homologues in the human genome has identified eight human AlkB homologues (hABH). In general, most multicellular eukaryotes seem to have several ABHs, whereas bacteria only have one. This is probably an important difference that is linked to the detailed function of each homologue. We know that the ABHs are directed to different cellular compartments, and the extra homologues may at least partly be a consequence of that. However, the actual reason for the difference is still unknown.

Although AlkB was first discovered more than 20 years ago, the most significant progress in the field has occurred during the last 3–4 years. The progress has generated novel knowledge in the field of DNA–RNA repair, which justifies a thorough overview and discussion at this stage. This section will give an introduction to AlkB-like proteins, and discuss the current status.

5.1. Twenty years of AlkB research—a brief overview

The *alkB* gene of *E. coli* was identified more than 20 years ago [116]. It was shown that *alkB* mutants had increased sensitivity towards the S_N2-type methylating agent MMS, and experiments based on the survival of alkylated λ phage in wild-type cells and *alkB* mutant cells indicated that it was involved in DNA repair. Subsequent studies showed that the *alkB* gene was adjacent to the *ada* gene on the *E. coli* chromosome [117], and that they constituted an operon [118]. It was later demonstrated that increased alkylation resistance was an intrinsic property of AlkB itself, as the expression of *E. coli* AlkB in human cell lines gave the same alkylation-resistant phenotype as in *E. coli* [119]. Shortly after this the first human AlkB homologue was identified [120]. A bioinformatics approach was used, searching with the *E. coli* AlkB sequence in a library of human EST sequences, and the prediction was confirmed by showing that the full-length hABH protein could partially rescue an *E. coli alkB* mutant from MMS-induced cell death. mRNA for hABH1was shown to be present in all human tissues tested, indicating that it could be a housekeeping gene playing a fundamental role in most tissues. Data from *Caulobacter crescentus* showed that the *alkB* gene there was cell cycle regulated, and the cycling pattern was similar to genes involved in DNA replication [121]. It was also shown that *E. coli alkB* mutants mainly were defect in processing of methylation damage in single-stranded DNA [122], and that the most likely substrates for AlkB were 1-meA and 3-meC. However, the actual mechanism of AlkB remained elusive until Aravind and Koonin [123] used a bioinformatics approach to show that the protein belongs to a large class of enzymes known as the 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily. They used a profile-based iterative search method to cluster together all sequences in the non-redundant NCBI protein sequence database that potentially shared the fold of this superfamily. The common feature in most of these proteins seems to be that they use Fe^{2+} and 2-oxogluarate (also known as α-ketoglutarate) to oxidize organic substrates with O_2 [124]. They showed that the essential residues for enzymatic activity were conserved, using a structural model of the catalytic domain. This was a breakthrough because a specific mechanism could then be tested. The analysis showed that AlkB homologues were found in most bacteria and eukaryotes as well as in some viruses, and that eukaryotes often had several homologues. They also suggested that AlkB homologues could be involved in RNA processing as it was found in RNA viruses. Shortly later two groups were able to confirm the predicted enzyme activity [45,46]. It was shown that AlkB could repair 1-meA and 3-meC in single- and double-stranded DNA using molecular oxygen, 2-oxoglutarate and Fe(II) to oxidize the methyl group resulting in direct reversal of the damage upon spontaneous release of formaldehyde (Fig. 4). Shortly after this two additional human AlkB homologues, hABH2 at chromosomal localization 12q24.1 and hABH3 at position 11q11, were reported. In addition, a corrected hABH sequence (hABH1) was reported [10,47]. hABH1 is similar to the previously published hABH, but with corrected N and C termini corresponding to a predicted conceptual splicing of genomic DNA from chromosome 14q24. However, previous suggestions that hABH1 is a functional AlkB homologue [120] could not be confirmed. It was shown that both hABH2 and hABH3 removed methyl groups from 1-meA and 3-meC in methylated polynucleotides in a 2-oxoglutarate-dependent manner, which was slightly stimulated by ascorbate and inhibited by EDTA. This is characteristic of 2OG-FeII oxygenase activity. hABH2 localized to cell nuclei, where it was present diffusely throughout nucleoplasm and accumulated in nucleoli, whereas hABH3 was present diffusely in nuclei and to a lesser extent in cytoplasm. It was also shown that hABH3, but not hABH2, repairs methylated RNA (Fig. 4)

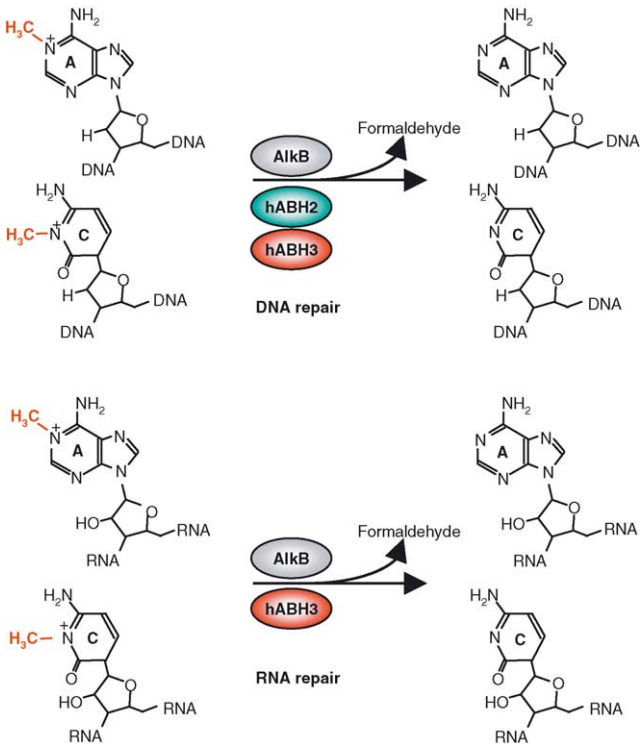

Fig. 4. *E. coli* AlkB and the human homologues hABH2 and hABH3 repair 1-meAand 3-meC in nucleic acids by oxidative demethylation. AlkB and hABH3 act upon both RNA and DNA, and have a preference for single-stranded substrate. hABH2 is specific for DNA, and prefers double-stranded substrate.

[10], and that hABH2 has a preference for double-stranded DNA, whereas AlkB and hABH3 acts more efficiently on single-stranded nucleic acids. It was shown that hABH2 relocates to replication foci in S-phase, which suggested that hABH2 repairs DNA close to replication forks, whereas hABH3 has a role in maintenance of nuclear single-stranded DNA and RNA, potentially with genes undergoing transcription as main targets. The preference of AlkB for ssDNA was confirmed using chemical cross-linking [125], only very small amounts of H131C-mutated AlkB cross-linked to dsDNA compared to ssDNA. An extensive bioinformatics mapping of bacterial AlkB proteins using iterative profile-based searches [126] showed that AlkB homologues can be classified in sub-families, based on phylogenetic properties, and that a small number of bacteria have an additional AlkB homologue. A similar approach was used to identify five additional human AlkB homologues [127], rising the total number of potential hABH sequences to eight.

5.2. Important properties of the 2OG-FeII oxygenase superfamily

The 2OG-FeII oxygenase superfamily is widespread in eukaryotes and bacteria [123], and is currently the largest known family of oxidising enzymes without a heme group [124]. These enzymes catalyze a variety of different reactions involving oxidation of an organic substrate using a dioxygen molecule. In almost all cases the oxidation of the substrate is coupled to the conversion of 2OG into succinate and CO_2, where one of the oxygen atoms from the dioxygen molecule is incorporated into the succinate. The 3D structure of several oxygenases is known, including isopenicillin N synthase [128], deacetoxycephalosporin C synthase [129], clavaminate synthase [130], proline 3-hydroxylase [131], anthocyanidin synthase [132] and taurine dioxygenase [133]. These structures share a common fold with a structurally conserved jelly roll β-sheet core with flanking helices. Very few residues are conserved across these structures, basically just two histidines and an aspartic acid involved in co-ordination of the Fe(II) ion and an arginine involved in co-ordination of the 2-oxoglutarate.

Although the requirement for 2OG seems to be a general feature, some enzymes are functionally active also in the absence of 2OG [124]. Two important examples are isopenicillin N synthase (IPNS) and 1-aminocyclopropane-1-carboxylic acid oxidase (ACCO). The ACCO activity is HCO_3^- dependent and strongly stimulated by ascorbate [134].

5.3. hABH1–8—human homologues of AlkB

The hABH sequences identified by Kurowski et al. are listed in Table 2 and an alignment of the conserved core of AlkB and hABH1–8 is shown in Fig. 5, with one important difference. As commented by Kurowski et al. the human transcript for ABH8 is split into two domains compared to orthologues from other genomes. However, this split goes right through the AlkB-domain, and seems to be a potential error. Osada et al. [135] used cDNAs from *Macaca fascicularis* to identify novel human genes, including the gene annotated as ABH8 here. They identified a potential full-length version of the human gene, although they also reported indications of alternative splicing. The protein sequence predicted by Osada et al. is identical to the sequence generated with the Genewise program [136] when the *Macaca* protein sequence (id BAB60797) is used to scan the human genome (F. Drabløs, unpublished). The predicted hABH8 sequence has 99% sequence identity to the *Macaca* sequence at the protein level. The difference between the sequence predicted from genomic data and the sequence found in the NCBI library is that exons 6, 9 and 10 are missing, and that there are frameshift errors in the (potentially alternative) splicing at these positions. The experimental human cDNA was identified from leiomyosarcoma tissue, which may indicate a possible defect in the mechanism for (alternative) mRNA splicing for this cell line. Therefore, the orthologous *M. fascicularis* sequence is used for this discussion.

Also the hABH6 transcript found in the NCBI nr database may be truncated. It is significantly shorter than the corresponding proteins from some other organisms in Table 2. Three different variants are available in the NCBI nr database; isoform 1 (NP_942567, 161 aa), isoform 2 (NP_116267, 140 aa) and slightly longer third form

Table 2
AlkB homologues in selected eukaryotes

Name	*Homo sapiens*[a,b]	*Mus musculus*[a]	*Rattus norvegicus*[a]	*Anopheles gambiae*[a]	*Drosophila melanogaster*[a]	*C. elegans*[a]
ABH1	NP_006011 gi5174385 299 aa 14q24	XP_127049 gi38050498 369 aa	XP_343093 gi34867478 301 aa	XP_310984 gi31204071 286 aa		NP_493970 gi17537375 169 aa
ABH2	XP_058581 gi17474993 261 aa 12q24.12	NP_778181 gi33239413 239 aa	XP_222273 gi27666578 239 aa			
ABH3	NP_631917 gi21040275 286 aa 11p11.2	XP_130317 gi25056889 286 aa	XP_342470 gi34856575 314 aa			
ABH4	NP_060091 gi8923019 302 aa 7q22.1	XP_132400 gi20846125 300 aa	XP_222071 gi27664044 301 aa	XP_318611 gi31231897 303 aa	NP_609314 gi24583140 304 aa	NP_741141 gi25148697 291 aa
ABH5	NP_060228 gi40353210 458 aa 17p11.2	NP_766531 gi31044423 395 aa	XP_220525 gi34871526 395 aa			
ABH6	*AAH07601* *gi14043223* 169 aa 19q13.13	NP_932144 gi38569508 238 aa	XP_214896 gi34855673 238 aa	XP_310425 gi31202953 217 aa	NP_609414 gi24583454 213 aa	NP_502522 gi17538328 160 aa
ABH7	NP_115682 gi14150066 221 aa 19p13.3	NP_079814 gi21313470 221 aa	XP_217314 gi27672438 221 aa	XP_311296 gi31204695 262 aa	NP_648511 gi24662930 206 aa	*CAE18012* *gi33300627* 227 aa
ABH8	(*BAB60797* *gi14388555* 664 aa)[c] NP_620130 gi20270315 238 aa 11q23.1	XP_134612 gi38089703 664 aa	XP_345767 gi34862443 688 aa	XP_320626 gi31240425 510 aa	NP_611690 gi24658267 615 aa	NP_497751 gi17552176 591 aa

[a] For each sequence the identifier, gi number and length (aa) is listed. Identifiers and gi numbers are taken from the NCBI nr database. In general, the entries are from the RefSeq reference sequence database subset [162], except for entries shown in *italics*.
[b] For *H. sapiens* chromosomal location is also listed.
[c] This is the *M. fascicularis* sequence, which is 98% identical to the corresponding full-length protein sequence encoded in the *H. sapiens* genome.

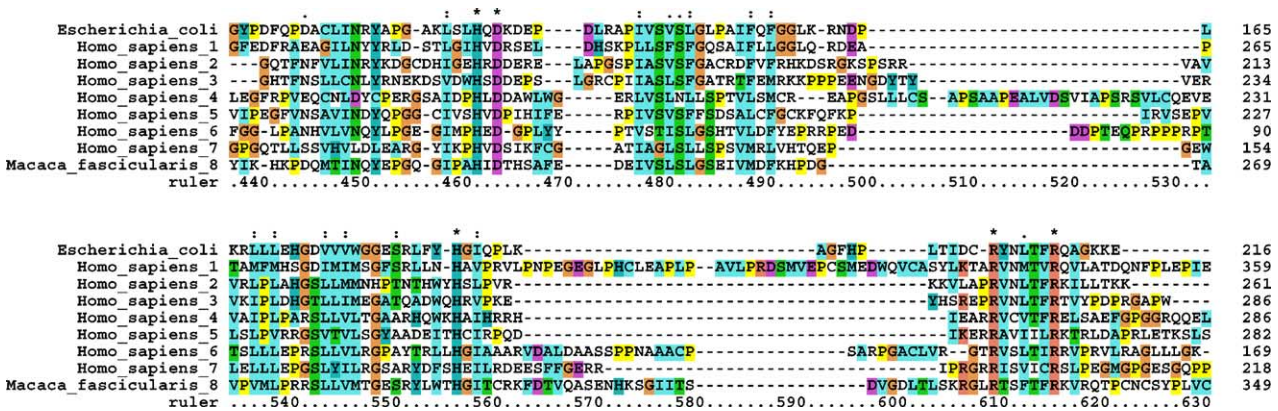

Fig. 5. Multiple alignment of AlkB homologues. The alignment is a subset of the alignment used for Fig. 6, and is numbered according to the full alignment. Sequences are numbered according to Kurowski et al. [127] (Table 2).

(AAH07601, 169 aa). Isoform 2 has a different (and shorter) C-terminal compared to isoform 1, possibly from using a different splice site. The third form shares 43 residues with isoform 1. It has a longer C-terminal region (including the AlkB-like domain), but is missing the C-terminal part of isoform 1 and 2. A combined sequence generated from these three forms is >90% identical to the full-length mouse and rat sequences.

Some of the human genes found by Kurowski et al. had been identified previously. Aravind and Koonin [123] identified hABH1, hABH4 and hABH5, although they focussed mainly on the larger 2OG-Fe(II) oxygenase superfamily. Duncan et al. [47] and Aas et al. [10] identified hABH2 and hABH3, and showed that they are enzymatically active. The hABH6, hABH7 and hABH8 seem to be novel AlkB homologues. Interestingly, whereas bacteria have just one or at most two *alkB*-like genes, eukaryotes may have as many as eight. This may reflect that eukaryotes are more complex, e.g. with an increased compartmentalization. We have studied some of these genes and the corresponding enzymes for some time, and at least hABH1, 2–4 and 6 show significant differences in subcellular localization. However, the increased number of homologues may also reflect that some alkB-genes have taken on new roles in eukaryotes.

An important property of enzymes that do demethylation by direct reversal, compared to glycosylases, is that they can work safely on single-stranded substrates without the risk of introducing strand breaks. This is consistent with the fact that AlkB targets 1-meA and 3-meC. These methylations interfere with A–T and C–G hydrogen bonding and occur more readily in single-stranded substrates. Repair capacity for single-stranded substrates (RNA and ssDNA) may therefore be important in several contexts, e.g. in transcription as well as in some replication processes.

However, in particular for the novel AlkB homologues (hABH4–8), the actual function is in general still unknown. Osada et al. [135] identified hABH8 as an RNA-binding methyl transferase, and this seems to be confirmed by Pfam [137] searches. Searching with the *M. fascicularis* sequence against the Pfam database identifies a 2OG-FeII oxygenase domain and an RNA recognition motif (RRM). However, the RRM motif has also been found in some single-strand DNA binding proteins [138]. The search also identifies a weak similarity to a DNA-binding helix-turn-helix motif and a putative S-AdoMet methyltransferase. It is possible that ABH8 is involved in control of normal methylation and demethylation of ssDNA or RNA. Experimental verification of activity is ongoing for all novel AlkB homologues.

5.4. Conserved properties of the AlkB homologues

The alignment in Fig. 5 shows 5 completely conserved positions; a HxD motif, a single H, and a RxxxxxR motif (given in single letter amino acid codes). The last motif may also be written as RhphphR, where h and p are hydrophobic and polar residues, respectively. This is consistent with a β-strand structure, where one side faces a polar substrate binding cleft whereas the other side faces the hydrophobic core of the protein. The positions that are conserved in this alignment are conserved also in complete alignments of all AlkB homologues identified by iterative searches. The H and D residues have previously been associated with co-ordination of the Fe(II) ion, whereas the first R most likely binds the oxoglutarate [47,123,129]. The last R may be involved in AlkB-specific substrate binding. The alignment made by Aravind and Koonin [123] shows that the R at this position is found almost exclusively in the AlkB sub-family, the other sub-families mainly have large aromatic residues like W, Y and F at this position. In the structure of clavaminic acid synthase in complex with proclavaminic acid [130] the Y found at this position points directly towards the substrate. Similarly, in the structure of anthocyanidin synthase in complex with trans-dihydroquercetin [132] there is a stacking interaction between the corresponding F and the substrate. One structure has an R at this position, corresponding to the AlkB sequence. This is the protein coded by the *gabT* gene in *E. coli* (entry 1jr7 [139] in PDB), which is part of the *gab* operon involved in the degradation of γ-aminobutyric acid (GABA). This sequence is not included in the alignment carried out by Aravind and Koonin. The structure does not include any substrate. However, it has been suggested that the GabT protein is involved in conversion of GABA (γ-aminobutyric acid) into succinic semialdehyde [140], and the R is therefore possibly important for co-ordination of the acidic group in the substrate. In the case of AlkB the conserved R may be involved in co-ordination of the substrate e.g. through hydrogen bonding or stacking interactions [141,142].

5.5. Phylogenetic distribution of AlkB homologues

A phylogenetic tree based on reliably aligned regions from an alignment of all non-viral ABH sequences identified through iterative searches (F. Drabløs, unpublished) is shown in Fig. 6. This tree consists of two sub-trees, one bacterial/eukaryote sub-tree that includes ABH1–3, and one eukaryote sub-tree with ABH4–8. The bacterial/eukaryote sub-tree can possibly be divided further into two additional sub-trees, where ABH1 is found in one, ABH2 and 3 is found in the other. The tree has to be interpreted with some care, as it is based on a relatively short alignment of just 60 positions. In general, a phylogenetic tree gets more robust when the length of the alignment is increased [143,144]. However, in this case, the degree of sequence conservation outside the active site region is low, therefore the full-length alignment will not be reliable. Focusing on the conserved regions of the alignment reduces the noise from mis-aligned residues, and the confidence in the major branching pattern of the tree seems to be good.

The split of the tree into eukaryote and mixed bacterial/eukaryote sub-trees is interesting, because it indicates that ABH1–3 are of relatively ancient origin, whereas the

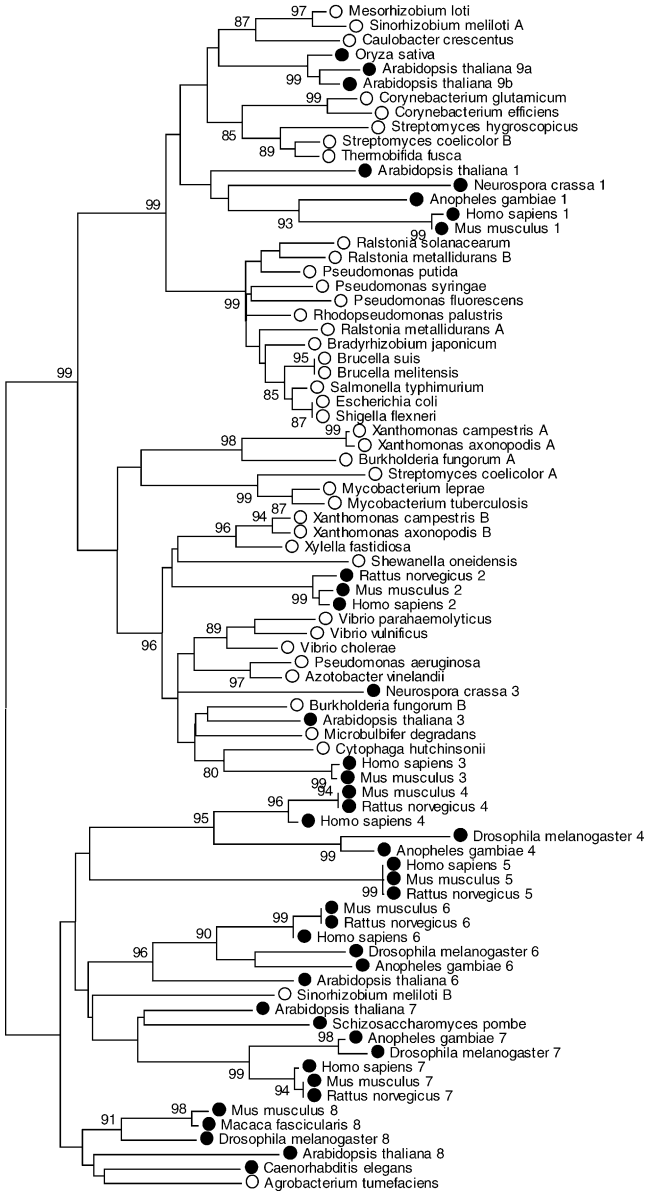

Fig. 6. Phylogenetic tree of AlkB homologues. Sequences with AlkB homologues were identified using PSI-Blast [163,164], and the tree is based on conserved regions in a ClustalX alignment [165] of these sequences (alignment positions 458–463, 476–490, 538–558 and 609–615, see Fig. 5). The phylogeny was computed in Mega2 [166] with the Minimum Evolution algorithm using the Poisson correction for amino acid distances, complete deletion of gap sites, and a Close-Neighbour-Interchange search starting from the Neighbour-Joining tree. Bacteria and eukaryotes are indicated with open and filled spheres, respectively. Numbers represent confidence probability (%) from a bootstrap interior branch test, showing values >80%. For species with more than one ABH-type gene this is indicated with letters (A, B) or numbers (1, 2, 3, etc.).

origin of ABH4–8 is more recent, possibly initiated by specific requirements in eukaryote organisms. However, there are two interesting exceptions to this pattern, as AlkB homologues from *Agrobacterium tumefaciens* and *Sinorhizobium meliloti* are found in the eukaryote sub-tree. These two bacteria are related, belonging to the Rhizobiaceae group of proteobacteria, and both live in a closely integrated relationship with plants. *S. meliloti* (formerly *Rhizobium meliloti)* is a soil bacterium that forms symbiotic nitrogen-fixing nodules on the roots of certain genera of leguminous plants. The genome consists of one chromosome (3.65 Mb) and two megaplasmids, pSymA (1.35 Mb) and pSymB (1.68 Mb) [145]. Previous investigation of the pSymB plasmid has shown that it most likely plays a role in adaption of the bacterium to an endosymbiotic lifestyle [146]. *A. tumefaciens* is the causative agent of crown gall disease in plants, which involves transfer of DNA from the bacterium to the plant (T-DNA), and integration of this DNA into the DNA of the plant. The genome consists of four replicons, a circular chromosome (2.84 Mb), a linear chromosome (2.08 Mb) and two plasmids, pAt (0.54 Mb) and pTi (0.21 Mb) [147,148]. The T-DNA is found on the Ti plasmid.

A. tumefaciens seems to have only one ABH gene, which is found on the circular chromosome. In *S. meliloti*, the "bacterium-like" ABH gene is found on the chromosome and the "eukaryote-like" gene on the pSymB plasmid. The fact that the copy on the plasmid is similar to eukaryote ABH sequences rather than bacterial may indicate that the plasmid ABH is important for host adaption of the bacterium. The adaption at a genetic level may have been achieved either by gene duplication and convergent evolution towards the function of the eukaryote analogues, or it may be a result of horizontal gene transfer.

5.6. Current status of research on AlkB homologues

It is a relevant question whether all active demethylases in the human genome now have been identified. From a sequence-similarity point of view it seems very likely that at least all AlkB homologues have been identified. The extensive searching by Kurowski et al. identified eight different homologues [127]. Searching the human proteome with Pfam-type profiles generated from all bacterial AlkB homologues and/or all eukaryote AlkB homologues do not identify any additional human homologues. Searching the human genome with each of the hABH sequences using all three reading frames in both directions also fails to identify any additional homologues (F. Drabløs, unpublished). However, from a functional point of view the answer is less clear. As shown in Table 2 the bacterial-like AlkB homologues, hABH1, 2 and 3 may be at least partly missing in non-vertebrates. It is a relevant question whether the activities of these proteins are maintained by other AlkB homologues or by completely different enzymes. In *E. coli*, the *alkB* gene is part of the ada regulon, as described above. This is part of the adaptive response to methylation. However, some of the genes involved in the adaptive response have functional counterparts that are constitutively expressed. For *ada*, the corresponding gene is *ogt* and for *alkA* it is *tag*. So far no constitutively expressed analog of AlkB has been identified, possibly indicating that AlkB type of activity is found only in the adaptive response in *E. coli*. However, the

situation is different in other bacteria, e.g. in *Caulobacter crecentus* transcription of the *alkB* homolog is cell cycle regulated [121], possibly linked to DNA replication, and not part of any adaptive response. This may indicate that AlkB is an "opportunistic" protein that is able to take on different roles. However, it also indicates that repair mechanisms are dynamic, where enzymes may be recruited into new roles as part of the adaption process to new challenges. The large number of AlkB homologues in eukaryotes may be consistent with this. Some bacteria have an extra copy of the *alkB* gene. These bacteria are in general either soil bacteria, where extra protection against harmful compounds may be useful, or plant-infecting bacteria that may need countermeasures against the post-transcriptional gene silencing (PTGS) mechanism of plants. The PTGS system includes methylation of foreign DNA and RNA as a signal for degradation, and the fact that a number of plant-infecting viruses also have *alkB* genes in their genomes is consistent with this [149]. However, only a small number of plant-infecting bacteria do have extra *alkB* genes. Some bacteria do not have any obvious AlkB homologues, and archaea seem to lack *alkB* genes completely although some of them live in very hostile environments. However, it is still relevant to look for alternative repair pathways for direct reversal of alkylation damage in ssDNA and RNA. It is also relevant in this context that the borderline between regulation and repair is not very clear in all cases. A relevant example is demethylation of the C5-cytosine in CpG. DNA methylation of CpG is important for long-term gene silencing [150]. It was recently shown that a silenced interleukin-2 gene may be activated by rapid demethylation of the promoter-enhancer region, and that this occurs by an active enzymatic mechanism [151]. However, the identity of this enzyme is unknown. It has been indicated that MbD2b may act as a demethylase for mCpG DNA [152], but the experimental support is not very strong. It has also been suggested that some of the AlkB homologues may have this type of activity [153–155], but this remains to be verified. However, this indicates that repair enzymes also may have regulatory functions (directly or via homologues), and vice versa.

To sum up, although we most likely have a good understanding of the distribution of AlkB homologues, we are far from understanding the function of all ABHs, and there may be unidentified demethylases that complement their function in various contexts.

6. Is there a threshold for carcinogenic effects of alkylating agents?

For risk assessment the question of threshold versus non-threshold for potential carcinogens is highly important, and still controversial. Traditionally, a non-threshold policy has been adopted in evaluation of risk. This is to a large degree a political decision, for safety reasons, due to the lack of adequate data for very low doses. The non-threshold approach implies linear extrapolation of data from animal experiments, using carcinogen doses that are mostly several orders of magnitude larger than those humans are exposed to. In general, a linear extrapolation would tend to overestimate risk if detoxification by metabolism and repair of lesions are reasonably efficient for a particular chemical. Linear extrapolation may be adequate for risk assessment of some chemicals, but almost certainly not for all. Thus, as an example there appears to be a threshold for a number of chemicals, including vinyl acetate (reviewed in [156]). For *N*-nitroso compounds, the situation has not been fully resolved, and may be dependent on the compound. In a study that involved 4080 rats, it was originally concluded that there was no evidence of a threshold for the carcinogenicity of *N*-nitrosodiethylamine [157]. However, a recent reanalysis of the same data concluded that there was a clear threshold at $10^{17.1}$ molecules/kg day) [158]. Also in other studies, a lack of dose–response relationship at low doses has been reported for *N*-nitrosamines and other chemicals, indicating a threshold [159,160] (reviewed in [161]). In conclusion, evidence is accumulating suggesting that for many chemicals there is a threshold below which many potential carcinogens are not carcinogenic. DNA repair processes are likely to contribute to this effect.

7. Concluding remarks

Alkylating agents are ubiquitously present in the environment and are produced endogenously, yet their significance for human disease remains obscure. However, it appears to be well established that *N*-nitroso compounds in some foods, as well as in main stream tobacco smoke, cause cancer. Available evidence strongly implicates mutagenic alkylations as the mechanism by which these agents cause cancer. The prevention of disease by DNA repair processes is also generally well documented. However, particularly with respect to repair of mutagenic O^6-alkylations there is a high, in some cases over 50-fold, variation in repair capacity in some organs. Since the repair of this lesion takes place by direct one-step repair, the variation is likely to be significant, but the cause of the variation remains unknown. This variation in repair capacity may also be significant for the response to treatment of cancer with alkylating drugs, and there is evidence for an inverse correlation between effect and repair capacity. The possibility to enhance effects of drugs by chemicals that inhibit repair is also presently being explored. The success of such approaches will presumably depend on the ability to improve the selectivity, since little is gained if the sensitivity of the tumor and non-target tissues, e.g. bone marrow, increases equally much.

The recently discovered oxidative demethylation of 1-meA and 3-meC by AlkB and human homologues expands the concept of alkylation repair, not only with respect to DNA, but also RNA as well. Although the biological significance of RNA repair remains to be proven, it seems

likely that it will be established on the growing list of mechanisms for macromolecular repair. The present review has explored in some detail the occurrence and evolution of AlkB homologues in viruses, prokaryotes and eukaryotes. It is also intriguing that while in humans there is apparently only one methyltransferase for repair of O^6-alkylations, and only one alkyl glycosylase, there are at least eight human AlkB homologues. These are located in different subcellular compartments. The functions of several of these are not yet well understood, but presently explored both in our laboratory and other laboratories. While some of the human AlkB homologues are already known to repair alkylation damage, others are likely to have different functions.

Acknowledgements

This work has been supported by the Research Council of Norway (project numbers 139617/140, 1470/310 and 149003/V40), The Norwegian Cancer Society, The Cancer Fund at the Regional Hospital in Trondheim and The Svanhild and Arne Must Fund for Medical Research.

References

[1] B. Rydberg, T. Lindahl, Nonenzymatic methylation of DNA by the intracellular methyl group donor *S*-adenosyl-L-methionine is a potentially mutagenic reaction, EMBO J. 1 (1982) 211–216.
[2] P. Taverna, B. Sedgwick, Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli*, J. Bacteriol. 178 (1996) 5105–5111.
[3] S.S. Hecht, DNA adduct formation from tobacco-specific *N*-nitrosamines, Mutat. Res. 424 (1999) 127–142.
[4] K. Ballschmiter, Pattern and sources of naturally produced organohalogens in the marine environment: biogenic formation of organohalogens, Chemosphere 52 (2003) 313–324.
[5] J.T. Hamilton, W.C. McRoberts, F. Keppler, R.M. Kalin, D.B. Harper, Chloride methylation by plant pectin: an efficient environmentally significant process, Science 301 (2003) 206–209.
[6] L.H. Hurley, DNA and its associated processes as targets for cancer therapy, Nat. Rev. Cancer 2 (2002) 188–200.
[7] E.C. Friedberg, G.C. Walker, W. Siede, DNA Repair and Mutagenesis, ASM Press, Washington, DC, 1995.
[8] T. Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709–715.
[9] J.H. Hoeijmakers, Genome maintenance mechanisms for preventing cancer, Nature 411 (2001) 366–374.
[10] P.A. Aas, M. Otterlei, P.O. Falnes, C.B. Vagbo, F. Skorpen, M. Akbari, O. Sundheim, M. Bjoras, G. Slupphaug, E. Seeberg, H.E. Krokan, Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA, Nature 421 (2003) 859–863.
[11] O. Gadal, U. Nehrbass, Nuclear structure and intranuclear retention of premature RNAs, J. Struct. Biol. 140 (2002) 140–146.
[12] T.H. Jensen, K. Dower, D. Libri, M. Rosbash, Early formation of mRNP: license for export or quality control? Mol. Cell 11 (2003) 1129–1138.
[13] M.F. Wilkinson, A.B. Shyu, RNA surveillance by nuclear scanning? Nat. Cell Biol. 4 (2002) E144–E147.
[14] E. Kim, J.D. Lowenson, D.C. MacLaren, S. Clarke, S.G. Young, Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 6132–6137.
[15] A. Yamamoto, H. Takagi, D. Kitamura, H. Tatsuoka, H. Nakano, H. Kawano, H. Kuroyanagi, Y. Yahagi, S. Kobayashi, K. Koizumi, T. Sakai, K. Saito, T. Chiba, K. Kawamura, K. Suzuki, T. Watanabe, H. Mori, T. Shirasawa, Deficiency in protein L-isoaspartyl methyltransferase results in a fatal progressive epilepsy, J. Neurosci. 18 (1998) 2063–2074.
[16] J. Moskovitz, S. Bar-Noy, W.M. Williams, J. Requena, B.S. Berlett, E.R. Stadtman, Methionine sulfoxide reductase (MsrA) is a regulator of antioxidant defense and lifespan in mammals, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 12920–12925.
[17] H.E. Krokan, B. Kavli, G. Slupphaug, Novel aspects of macromolecular repair and relationship to human disease, J. Mol. Med. 82 (2004) 280–297.
[18] G. Slupphaug, B. Kavli, H.E. Krokan, The interacting pathways for prevention and repair of oxidative DNA damage, Mutat. Res. 531 (2003) 231–251.
[19] G.P. Margison, M.F. Santibañez-Koref, A.C. Povey, Mechanisms of carcinogenicity/chemotherapy by O^6-methylguanine, Mutagenesis 17 (2002) 483–487.
[20] N.M. Martin, DNA repair inhibition and cancer therapy, J. Photochem. Photobiol. B 63 (2001) 162–170.
[21] M. Christmann, M.T. Tomicic, W.P. Roos, B. Kaina, Mechanisms of human DNA repair: an update, Toxicology 193 (2003) 3–34.
[22] M.D. Wyatt, J.M. Allan, A.Y. Lau, T.E. Ellenberger, L.D. Samson, 3-Methyladenine DNA glycosylases: structure, function, and biological importance, Bioessays 21 (1999) 668–676.
[23] H.E. Krokan, R. Standal, G. Slupphaug, DNA glycosylases in the base excision repair of DNA, Biochem. J. 325 (Pt 1) (1997) 1–16.
[24] K.V. Shooter, R. Howse, S.A. Shah, P.D. Lawley, The molecular basis for biological inactivation of nucleic acids. The action of methylating agents on the ribonucleic acid-containing bacteriophage R17, Biochem. J. 137 (1974) 303–312.
[25] B. Singer, D. Grünberger, Molecular Biology of Mutagens and Carcinogens, Plenum, New York, 1983.
[26] D.T. Beranek, Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents, Mutat. Res. 231 (1990) 11–30.
[27] B.P. Engelward, J.M. Allan, A.J. Dreslin, J.D. Kelly, M.M. Wu, B. Gold, L.D. Samson, A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome, J. Biol. Chem. 273 (1998) 5412–5418.
[28] L.R. Barrows, P.N. Magee, Nonenzymatic methylation of DNA by *S*-adenosylmethionine in vitro, Carcinogenesis 3 (1982) 349–351.
[29] P. Garcia-Santos Mdel, E. Calle, J. Casado, Amino acid nitrosation products as alkylating agents, J. Am. Chem. Soc. 123 (2001) 7506–7510.
[30] D.E. Shuker, G.P. Margison, Nitrosated glycine derivatives as a potential source of O^6-methylguanine in DNA, Cancer Res. 57 (1997) 366–369.
[31] W.F. Busby Jr., D.E. Shuker, G. Charnley, P.M. Newberne, S.R. Tannenbaum, G.N. Wogan, Carcinogenicity in rats of the nitrosated bile acid conjugates *N*-nitrosoglycocholic acid and *N*-nitrosotaurocholic acid, Cancer Res. 45 (1985) 1367–1371.
[32] A.C. Povey, A.F. Badawi, D.P. Cooper, C.N. Hall, K.L. Harrison, P.E. Jackson, N.P. Lees, P.J. O'Connor, G.P. Margison, DNA alkylation and repair in the large bowel: animal and human studies, J. Nutr. 132 (2002) 3518S–3521S.
[33] M.F. Mohamed, D. Kang, V.P. Aneja, Volatile organic compounds in some urban locations in United States, Chemosphere 47 (2002) 863–882.
[34] P. Vaughan, T. Lindahl, B. Sedgwick, Induction of the adaptive response of *Escherichia coli* to alkylation damage by the environmental mutagen, methyl chloride, Mutat. Res. 293 (1993) 249–257.
[35] H.M. Bolt, B. Gansewendt, Mechanisms of carcinogenicity of methyl halides, Crit. Rev. Toxicol. 23 (1993) 237–253.

[36] IARC MONOGRAPHS, Re-evaluation of Some Organic Chemicals, Hydrazine and Hydrogen Peroxide, Evaluation of carcinogenic risks to humans, vol. 71, World Health Organization, 1999.
[37] IARC MONOGRAPHS, On the Evaluation of Carcinogenic Risks to Humans Tobacco Smoking, vol. 38, World Health Organization, 1986.
[38] S.S. Hecht, Human urinary carcinogen metabolites: biomarkers for investigating tobacco and cancer, Carcinogenesis 23 (2002) 907–922.
[39] S.G. Carmella, A. Borukhova, D. Desai, S.S. Hecht, Evidence for endogenous formation of tobacco-specific nitrosamines in rats treated with tobacco alkaloids and sodium nitrite, Carcinogenesis 18 (1997) 587–592.
[40] S. Tyroller, W. Zwickenpflug, E. Richter, New sources of dietary myosmine uptake from cereals, fruits, vegetables, and milk, J. Agric. Food Chem. 50 (2002) 4909–4915.
[41] R. Goldman, P.G. Shields, Food mutagens, J. Nutr. 133 (Suppl. 3) (2003) 965S–973S.
[42] S.G. Chaney, A. Sancar, DNA repair: enzymatic mechanisms and relevance to drug response, J. Natl. Cancer Inst. 88 (1996) 1346–1360.
[43] M.R. Middleton, G.P. Margison, Improvement of chemotherapy efficacy by inactivation of a DNA-repair pathway, Lancet Oncol. 4 (2003) 37–44.
[44] M. Olsson, T. Lindahl, Repair of alkylated DNA in *Escherichia coli*. Methyl group transfer from O^6-methylguanine to a protein cysteine residue, J. Biol. Chem. 255 (1980) 10569–10571.
[45] P.O. Falnes, R.F. Johansen, E. Seeberg, AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*, Nature 419 (2002) 178–182.
[46] S.C. Trewick, T.F. Henshaw, R.P. Hausinger, T. Lindahl, B. Sedgwick, Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage, Nature 419 (2002) 174–178.
[47] T. Duncan, S.C. Trewick, P. Koivisto, P.A. Bates, T. Lindahl, B. Sedgwick, Reversal of DNA alkylation damage by two human dioxygenases, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 16660–16665.
[48] T. Lindahl, New class of enzymes acting on damaged DNA, Nature 259 (1976) 64–66.
[49] A. Memisoglu, L. Samson, Contribution of base excision repair, nucleotide excision repair, and DNA recombination to alkylation resistance of the fission yeast *Schizosaccharomyces pombe*, J. Bacteriol. 182 (2000) 2104–2112.
[50] B. Plosky, L. Samson, B.P. Engelward, B. Gold, B. Schlaen, T. Millas, M. Magnotti, J. Schor, D.A. Scicchitano, Base excision repair and nucleotide excision repair contribute to the removal of *N*-methylpurines from active genes, DNA Rep. (Amst.) 1 (2002) 683–696.
[51] G.P. Margison, M.F. Santibañez-Koref, O^6-Alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy, Bioessays 24 (2002) 255–266.
[52] B. Sedgwick, T. Lindahl, Recent progress on the Ada response for inducible repair of DNA alkylation damage, Oncogene 21 (2002) 8886–8894.
[53] W. Xiao, B. Derfler, J. Chen, L. Samson, Primary sequence and biological functions of a *Saccharomyces cerevisiae* O^6-methylguanine/O^4-methylthymine DNA repair methyltransferase gene, EMBO J. 10 (1991) 2179–2186.
[54] A.E. Pegg, Repair of O^6-alkylguanine by alkyltransferases, Mutat. Res. 462 (2000) 83–100.
[55] K. Kleibl, Molecular mechanisms of adaptive response to alkylating agents in *Escherichia coli* and some remarks on O^6-methylguanine DNA-methyltransferase in other organisms, Mutat. Res. 512 (2002) 67–84.
[56] M. Sassanfar, M.K. Dosanjh, J.M. Essigmann, L. Samson, Relative efficiencies of the bacterial, yeast, and human DNA methyltransferases for the repair of O^6-methylguanine and O^4-methylthymine. Suggestive evidence for O^4-methylthymine repair by eukaryotic methyltransferases, J. Biol. Chem. 266 (1991) 2767–2771.
[57] L.C. Harris, G.P. Margison, Expression in mammalian cells of the *Escherichia coli* O^6 alkylguanine-DNA-alkyltransferase gene ogt reduces the toxicity of alkylnitrosoureas, Br. J. Cancer 67 (1993) 1196–1202.
[58] J.M. Bogden, A. Eastman, E. Bresnick, A system in mouse liver for the repair of O^6-methylguanine lesions in methylated DNA, Nucleic Acids Res. 9 (1981) 3089–3103.
[59] B. Myrnes, K.E. Giercksky, H. Krokan, Repair of O^6-methylguanine residues in DNA takes place by a similar mechanism in extracts from HeLa cells, human liver, and rat liver, J. Cell. Biochem. 20 (1982) 381–392.
[60] L. Samson, S. Han, J.C. Marquis, L.J. Rasmussen, Mammalian DNA repair methyltransferases shield O^4MeT from nucleotide excision repair, Carcinogenesis 18 (1997) 919–924.
[61] S. Edara, S. Kanugula, A.E. Pegg, Expression of the inactive C145A mutant human O^6-alkylguanine-DNA alkyltransferase in *E. coli* increases cell killing and mutations by *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine, Carcinogenesis 20 (1999) 103–108.
[62] R. Stupp, C. Re, New drugs and combinations for malignant glioma, Forum (Genova) 13 (2003) 61–75.
[63] T. Kanzawa, J. Bedwell, Y. Kondo, S. Kondo, I.M. Germano, Inhibition of DNA repair for sensitizing resistant glioma cells to temozolomide, J. Neurosurg. 99 (2003) 1047–1052.
[64] Y. Nakatsu, K. Hattori, H. Hayakawa, K. Shimizu, M. Sekiguchi, Organization and expression of the human gene for O^6-methylguanine-DNA methyltransferase, Mutat Res. 293 (1993) 119–132.
[65] G.P. Margison, A.C. Povey, B. Kaina, M.F. Santibañez-Koref, Variability and regulation of O^6-alkylguanine-DNA alkyltransferase, Carcinogenesis 24 (2003) 625–635.
[66] D.S. Daniels, J.A. Tainer, Conserved structural motifs governing the stoichiometric repair of alkylated DNA by O^6-alkylguanine-DNA alkyltransferase, Mutat. Res. 460 (2000) 151–163.
[67] B. Myrnes, K.E. Giercksky, H. Krokan, Interindividual variation in the activity of O^6-methyl guanine-DNA methyltransferase and uracil-DNA glycosylase in human organs, Carcinogenesis 4 (1983) 1565–1568.
[68] S.A. Belinsky, M.A. Bedell, J.A. Swenberg, Effect of chronic ethanol diet on the replication, alkylation, and repair of DNA from hepatocytes and nonparenchymal cells following dimethylnitrosamine administration, Carcinogenesis 3 (1982) 1293–1297.
[69] T. Deilhaug, B. Myrnes, T. Aune, H. Krokan, A. Haugen, Differential capacities for DNA repair in Clara cells, alveolar type II cells and macrophages of rabbit lung, Carcinogenesis 6 (1985) 661–663.
[70] S.A. Belinsky, C.M. White, J.A. Boucheron, F.C. Richardson, J.A. Swenberg, M. Anderson, Accumulation and persistence of DNA adducts in respiratory tissue of rats following multiple administrations of the tobacco specific carcinogen 4-(*N*-methyl-*N*-nitrosamino)-1-(3-pyridyl)-1-butanone, Cancer Res. 46 (1986) 1280–1284.
[71] L.L. Dumenco, E. Allay, K. Norton, S.L. Gerson, The prevention of thymic lymphomas in transgenic mice by human O^6-alkylguanine-DNA alkyltransferase, Science 259 (1993) 219–222.
[72] K. Sakumi, A. Shiraishi, S. Shimizu, T. Tsuzuki, T. Ishikawa, M. Sekiguchi, Methylnitrosourea-induced tumorigenesis in MGMT gene knockout mice, Cancer Res. 57 (1997) 2415–2418.
[73] T. Iwakuma, K. Sakumi, Y. Nakatsuru, H. Kawate, H. Igarashi, A. Shiraishi, T. Tsuzuki, T. Ishikawa, M. Sekiguchi, High incidence of nitrosamine-induced tumorigenesis in mice lacking DNA repair methyltransferase, Carcinogenesis 18 (1997) 1631–1635.
[74] B.J. Glassner, G. Weeda, J.M. Allan, J.L. Broekhof, N.H. Carls, I. Donker, B.P. Engelward, R.J. Hampson, R. Hersmus, M.J. Hickman, R.B. Roth, H.B. Warren, M.M. Wu, J.H. Hoeijmakers, L.D. Samson, DNA repair methyltransferase (*Mgmt*) knockout mice are sensitive to the lethal effects of chemotherapeutic alkylating agents, Mutagenesis 14 (1999) 339–347.

[75] P. Branch, G. Aquilina, M. Bignami, P. Karran, Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage, Nature 362 (1993) 652–654.
[76] S. Ceccotti, G. Aquilina, P. Macpherson, M. Yamada, P. Karran, M. Bignami, Processing of O^6-methylguanine by mismatch correction in human cell extracts, Curr. Biol. 6 (1996) 1528–1531.
[77] O. Humbert, S. Fiumicino, G. Aquilina, P. Branch, S. Oda, A. Zijno, P. Karran, M. Bignami, Mismatch repair and differential sensitivity of mouse and human cells to methylating agents, Carcinogenesis 20 (1999) 205–214.
[78] H. Kawate, K. Sakumi, T. Tsuzuki, Y. Nakatsuru, T. Ishikawa, S. Takahashi, H. Takano, T. Noda, M. Sekiguchi, Separation of killing and tumorigenic effects of an alkylating agent in mice defective in two of the DNA repair genes, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5116–5120.
[79] T.P. Brent, J.S. Remack, D.G. Smith, Characterization of a novel reaction by human O^6-alkylguanine-DNA alkyltransferase with 1,3-bis(2-chloroethyl)-1-nitrosourea-treated DNA, Cancer Res. 47 (1987) 6185–6188.
[80] P.E. Gonzaga, P.M. Potter, T.Q. Niu, D. Yu, D.B. Ludlum, J.A. Rafferty, G.P. Margison, T.P. Brent, Identification of the cross-link between human O^6-methylguanine-DNA methyltransferase and chloroethylnitrosourea-treated DNA, Cancer Res. 52 (1992) 6052–6058.
[81] W.P. Tong, M.C. Kirk, D.B. Ludlum, Formation of the cross-link 1-[N3-deoxycytidyl),2-[N1-deoxyguanosinyl]ethane in DNA treated with *N,N′*-bis(2-chloroethyl)-*N*-nitrosourea, Cancer Res. 42 (1982) 3102–3105.
[82] C. Zlotogorski, L.C. Erickson, Pretreatment of normal human fibroblasts and human colon carcinoma cells with MNNG allows chloroethylnitrosourea to produce DNA interstrand cross-links not observed in cells treated with chloroethylnitrosourea alone, Carcinogenesis 4 (1983) 759–763.
[83] M. Grompe, A. D'Andrea, Fanconi anemia and DNA repair, Hum. Mol. Genet. 10 (2001) 2253–2259.
[84] P.J. McHugh, V.J. Spanswick, J.A. Hartley, Repair of DNA interstrand cross-links: molecular mechanisms and clinical relevance, Lancet Oncol. 2 (2001) 483–490.
[85] M.L. Dronkert, R. Kanaar, Repair of DNA interstrand cross-links, Mutat. Res. 486 (2001) 217–247.
[86] S.L. Donahue, R. Lundberg, R. Saplis, C. Campbell, Deficient regulation of DNA double-strand break repair in Fanconi anemia fibroblasts, J. Biol. Chem. 278 (2003) 29487–29495.
[87] A. Rothfuss, M. Grompe, Repair kinetics of genomic interstrand DNA cross-links: evidence for DNA double-strand break-dependent activation of the Fanconi anemia/BRCA pathway, Mol. Cell. Biol. 24 (2004) 123–134.
[88] S. Bjelland, E. Seeberg, Different efficiencies of the Tag and AlkA DNA glycosylases from *Escherichia coli* in the removal of 3-methyladenine from single-stranded DNA, FEBS Lett. 397 (1996) 127–129.
[89] E. Seeberg, L. Eide, M. Bjørås, The base excision repair pathway, Trends Biochem. Sci. 20 (1995) 391–397.
[90] Q. Li, S.E. Wright, Z. Matijasevic, W. Chong, D.B. Ludlum, M.R. Volkert, The role of human alkyladenine glycosylase in cellular resistance to the chloroethylnitrosoureas, Carcinogenesis 24 (2003) 589–593.
[91] R.H. Elder, J.G. Jansen, R.J. Weeks, M.A. Willington, B. Deans, A.J. Watson, K.J. Mynett, J.A. Bailey, D.P. Cooper, J.A. Rafferty, M.C. Heeran, S.W. Wijnhoven, A.A. van Zeeland, G.P. Margison, Alkylpurine-DNA-*N*-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate, Mol. Cell. Biol. 18 (1998) 5828–5837.
[92] J. Bramson, T. O'Connor, L. Panasci, Effect of alkyl-*N*-purine DNA glycosylase overexpression on cellular resistance to bifunctional alkylating agents, Biochem. Pharmacol. 50 (1995) 39–44.
[93] B.P. Engelward, A. Dreslin, J. Christensen, D. Huszar, C. Kurahara, L. Samson, Repair-deficient 3-methyladenine DNA glycosylase homozygous mutant mouse cells have increased sensitivity to alkylation-induced chromosome damage and cell killing, EMBO J. 15 (1996) 945–952.
[94] K.G. Berdal, R.F. Johansen, E. Seeberg, Release of normal bases from intact DNA by a native DNA repair enzyme, EMBO J. 17 (1998) 363–367.
[95] P.J. O'Brien, T. Ellenberger, Dissecting the broad substrate specificity of human 3-methyladenine DNA glycosylase, J. Biol. Chem. 279 (2004) 9750–9757.
[96] A.Y. Lau, M.D. Wyatt, B.J. Glassner, L.D. Samson, T. Ellenberger, Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 13573–13578.
[97] L. Gros, A.V. Maksimenko, C.V. Privezentzev, J. Laval, M.K. Saparbaev, Hijacking of the human alkyl-*N*-purine-DNA glycosylase by $3,N^4$-ethenocytosine, a lipid peroxidation induced DNA adduct, J. Biol. Chem. 279 (2004) 17723–17730.
[98] M.L. Fishel, Y.R. Seo, M.L. Smith, M.R. Kelley, Imbalancing the DNA base excision repair pathway in the mitochondria; targeting and overexpressing *N*-methylpurine DNA glycosylase in mitochondria leads to enhanced cell killing, Cancer Res. 63 (2003) 608–615.
[99] T. Grombacher, B. Kaina, Constitutive expression and inducibility of O^6-methylguanine-DNA methyltransferase and *N*-methylpurine-DNA glycosylase in rat liver cells exhibiting different status of differentiation, Biochim. Biophys. Acta 1270 (1995) 63–72.
[100] R.B. Roth, L.D. Samson, 3-Methyladenine DNA glycosylase-deficient *Aag* null mice display unexpected bone marrow alkylation resistance, Cancer Res. 62 (2002) 656–660.
[101] B. Kaina, K. Ochs, S. Grosch, G. Fritz, J. Lips, M. Tomicic, T. Dunkern, M. Christmann, BER, MGMT, and MMR in defense against alkylation-induced genotoxicity and apoptosis, Prog. Nucleic Acid Res. Mol. Biol. 68 (2001) 41–54.
[102] G. Evensen, E. Seeberg, Adaptation to alkylation resistance involves the induction of a DNA glycosylase, Nature 296 (1982) 773–775.
[103] L.M. Posnick, L.D. Samson, Imbalanced base excision repair increases spontaneous mutation and alkylation sensitivity in *Escherichia coli*, J. Bacteriol. 181 (1999) 6763–6771.
[104] F. Miao, M. Bouziane, T.R. O'Connor, Interaction of the recombinant human methylpurine-DNA glycosylase (MPG protein) with oligodeoxyribonucleotides containing either hypoxanthine or abasic sites, Nucleic Acids Res. 26 (1998) 4034–4041.
[105] J.K. Horton, D.F. Joyce-Gray, B.F. Pachkowski, J.A. Swenberg, S.H. Wilson, Hypersensitivity of DNA polymerase beta null mouse fibroblasts reflects accumulation of cytotoxic repair intermediates from site-specific alkyl DNA lesions, DNA Rep. (Amst.) 2 (2003) 27–48.
[106] N.J. Curtin, L.Z. Wang, A. Yiakouvaki, S. Kyle, C.A. Arris, S. Canan-Koch, S.E. Webber, B.W. Durkacz, H.A. Calvert, Z. Hostomsky, D.R. Newell, Novel poly(ADP-ribose) polymerase-1 inhibitor, AG14361, restores rensitivity to temozolomide in mismatch repair-deficient cells, Clin. Cancer. Res. 10 (2004) 881–889.
[107] S.L. Gerson, Clinical relevance of MGMT in the treatment of cancer, J. Clin. Oncol. 20 (2002) 2388–2399.
[108] K.A. Jaeckle, H.J. Eyre, J.J. Townsend, S. Schulman, H.M. Knudson, M. Belanich, D.B. Yarosh, S.I. Bearman, D.J. Giroux, S.C. Schold, Correlation of tumor O^6 methylguanine-DNA methyltransferase levels with survival of malignant astrocytoma patients treated with bis-chloroethylnitrosourea: a Southwest Oncology Group study, J. Clin. Oncol. 16 (1998) 3310–3315.
[109] Z.P. Chen, D. Yarosh, Y. Garcia, D. Tampieri, G. Mohr, A. Malapetsa, A. Langleben, L.C. Panasci, Relationship between O^6-methylguanine-DNA methyltransferase levels and clinical response induced by chloroethylnitrosourea therapy in glioma patients, Can. J. Neurol. Sci. 26 (1999) 104–109.

[110] M. Esteller, J. Garcia-Foncillas, E. Andion, S.N. Goodman, O.F. Hidalgo, V. Vanaclocha, S.B. Baylin, J.G. Herman, Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents, N. Engl. J. Med. 343 (2000) 1350–1354.
[111] M.E. Dolan, A.E. Pegg, O^6-Benzylguanine and its role in chemotherapy, Clin. Cancer Res. 3 (1997) 837–847.
[112] T.P. Spiro, S.L. Gerson, L. Liu, S. Majka, J. Haaga, C.L. Hoppel, S.T. Ingalls, J.M. Pluda, J.K. Willson, O^6-Benzylguanine: a clinical trial establishing the biochemical modulatory dose in tumor tissue for alkyltransferase-directed DNA repair, Cancer Res. 59 (1999) 2402–2410.
[113] A.A. Brandes, S. Turazzi, U. Basso, L.M. Pasetto, B. Guglielmi, L. Volpin, P. Iuzzolino, P. Amista, G. Pinna, R. Scienza, M. Ermani, A multidrug combination designed for reversing resistance to BCNU in glioblastoma multiforme, Neurology 58 (2002) 1759–1764.
[114] L. Liu, Y. Nakatsuru, S.L. Gerson, Base excision repair as a therapeutic target in colon cancer, Clin. Cancer Res. 8 (2002) 2985–2991.
[115] R. Aloyz, K. Grzywacz, Z.Y. Xu, M. Loignon, M.A. Alaoui-Jamali, L. Panasci, Imatinib sensitizes CLL lymphocytes to chlorambucil, Leukemia 18 (2004) 409–414.
[116] H. Kataoka, Y. Yamamoto, M. Sekiguchi, A new gene (*alkB*) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate, J. Bacteriol. 153 (1983) 1301–1307.
[117] H. Kataoka, M. Sekiguchi, Molecular cloning and characterization of the *alkB* gene of *Escherichia coli*, Mol. Gen. Genet. 198 (1985) 263–269.
[118] H. Kondo, Y. Nakabeppu, H. Kataoka, S. Kuhara, S. Kawabata, M. Sekiguchi, Structure and expression of the *alkB* gene of *Escherichia coli* related to the repair of alkylated DNA, J. Biol. Chem. 261 (1986) 15772–15777.
[119] B.J. Chen, P. Carroll, L. Samson, The *Escherichia coli* AlkB protein protects human cells against alkylation-induced toxicity, J. Bacteriol. 176 (1994) 6255–6261.
[120] Y.F. Wei, K.C. Carter, R.P. Wang, B.K. Shell, Molecular cloning and functional analysis of a human cDNA encoding an *Escherichia coli* AlkB homolog, a protein involved in DNA alkylation damage repair, Nucleic Acids Res. 24 (1996) 931–937.
[121] D. Colombi, S.L. Gomes, An *alkB* gene homolog is differentially transcribed during the *Caulobacter crescentus* cell cycle, J. Bacteriol. 179 (1997) 3139–3145.
[122] S. Dinglay, S.C. Trewick, T. Lindahl, B. Sedgwick, Defective processing of methylated single-stranded DNA by *E. coli* AlkB mutants, Genes Dev. 14 (2000) 2097–2105.
[123] L. Aravind, E.V. Koonin, The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases, Genome Biol. Res. 2 (2001) 0007.0001–0007.0008.
[124] C.J. Schofield, Z. Zhang, Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes, Curr. Opin. Struct. Biol. 9 (1999) 722–731.
[125] Y. Mishina, C. He, Probing the structure and function of the *Escherichia coli* DNA alkylation repair AlkB protein through chemical cross-linking, J. Am. Chem. Soc. 125 (2003) 8730–8731.
[126] P.O. Falnes, T. Rognes, DNA repair by bacterial AlkB proteins, Res. Microbiol. 154 (2003) 531–538.
[127] M.A. Kurowski, A.S. Bhagwat, G. Papaj, J.M. Bujnicki, Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB, BMC Genom. 4 (2003) 48.
[128] P.L. Roach, I.J. Clifton, C.M. Hensgens, N. Shibata, C.J. Schofield, J. Hajdu, J.E. Baldwin, Structure of isopenicillin *N* synthase complexed with substrate and the mechanism of penicillin formation, Nature 387 (1997) 827–830.
[129] K. Valegard, A.C. van Scheltinga, M.D. Lloyd, T. Hara, S. Ramaswamy, A. Perrakis, A. Thompson, H.J. Lee, J.E. Baldwin, C.J. Schofield, J. Hajdu, I. Andersson, Structure of a cephalosporin synthase, Nature 394 (1998) 805–809.
[130] Z. Zhang, J. Ren, D.K. Stammers, J.E. Baldwin, K. Harlos, C.J. Schofield, Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase, Nat. Struct. Biol. 7 (2000) 127–133.
[131] I.J. Clifton, L.C. Hsueh, J.E. Baldwin, K. Harlos, C.J. Schofield, Structure of proline 3-hydroxylase. Evolution of the family of 2-oxoglutarate dependent oxygenases, Eur. J. Biochem. 268 (2001) 6625–6636.
[132] R.C. Wilmouth, J.J. Turnbull, R.W. Welford, I.J. Clifton, A.G. Prescott, C.J. Schofield, Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*, Structure (Camb.) 10 (2002) 93–103.
[133] J.M. Elkins, M.J. Ryle, I.J. Clifton, J.C. Dunning Hotopp, J.S. Lloyd, N.I. Burzlaff, J.E. Baldwin, R.P. Hausinger, P.L. Roach, X-ray crystal structure of *Escherichia coli* taurine/α-ketoglutarate dioxygenase complexed to ferrous iron and substrates, Biochemistry 41 (2002) 5185–5192.
[134] A.M. Rocklin, K. Kato, H.W. Liu, L. Que Jr., J.D. Lipscomb, Mechanistic studies of 1-aminocyclopropane-1-carboxylic acid oxidase: single turnover reaction, J. Biol. Inorg. Chem. 9 (2004) 171–182.
[135] N. Osada, M. Hida, J. Kusuda, R. Tanuma, M. Hirata, M. Hirai, K. Terao, Y. Suzuki, S. Sugano, K. Hashimoto, Prediction of unidentified human genes on the basis of sequence similarity to novel cDNAs from cynomolgus monkey brain, Genome Biol. Res. 3 (2002) 0006.0001–0006.0005.
[136] E. Birney, R. Durbin, Using GeneWise in the *Drosophila* annotation experiment, Genome Res. 10 (2000) 547–548.
[137] A. Bateman, E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, E.L.L. Sonnhammer, The Pfam protein families database, Nucleic Acids Res. 30 (2002) 276–280.
[138] S.D. Johnston, J.E. Lew, J. Berman, Gbp1p, a protein with RNA recognition motifs, binds single-stranded telomeric DNA and changes its binding specificity upon dimerization, Mol. Cell Biol. 19 (1999) 923–933.
[139] M.R. Chance, A.R. Bresnick, S.K. Burley, J.S. Jiang, C.D. Lima, A. Sali, S.C. Almo, J.B. Bonanno, J.A. Buglino, S. Boulton, H. Chen, N. Eswar, G. He, R. Huang, V. Ilyin, L. McMahan, U. Pieper, S. Ray, M. Vidal, L.K. Wang, Structural genomics: a pipeline for providing structures for the biologist, Protein Sci. 11 (2002) 723–738.
[140] B.L. Schneider, S. Ruback, A.K. Kiupakis, H. Kasbarian, C. Pybus, L. Reitzer, The *Escherichia coli* gabDTPC operon: specific γ-aminobutyrate catabolism and nonspecific induction, J. Bacteriol. 184 (2002) 6976–6986.
[141] J. Allers, Y. Shamoo, Structure-based analysis of protein-RNA interactions using the program ENTANGLE, J. Mol. Biol. 311 (2001) 75–86.
[142] C. Biot, E. Buisine, J.M. Kwasigroch, R. Wintjens, M. Rooman, Probing the energetic and structural role of amino acid/nucleobase cation-π interactions in protein-ligand complexes, J. Biol. Chem. 277 (2002) 40816–40822.
[143] B. Bremer, R.K. Jansen, B. Oxelman, M. Backlund, H. Lantz, K.J. Kim, More characters or more taxa for a robust phylogeny: case study from the coffee family (*Rubiaceae*), Syst. Biol. 48 (1999) 413–435.
[144] G.V. Glazko, M. Nei, Estimation of divergence times for major lineages of primate species, Mol. Biol. Evol. 20 (2003) 424–434.
[145] F. Galibert, T.M. Finan, S.R. Long, A. Puhler, P. Abola, F. Ampe, F. Barloy-Hubler, M.J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R.W. Davis, S. Dreano, N.A. Federspiel, R.F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R.W. Hyman, T. Jones, D. Kahn, M.L. Kahn, S. Kalman, D.H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M.C. Peck, T.M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol, F.J. Vorholter, S. Weidner, D.H. Wells, K. Wong, K.C. Yeh, J. Batut, The composite genome of the

legume symbiont *Sinorhizobium meliloti*, Science 293 (2001) 668–672.
[146] T.M. Finan, S. Weidner, K. Wong, J. Buhrmester, P. Chain, F.J. Vorholter, I. Hernandez-Lucas, A. Becker, A. Cowie, J. Gouzy, B. Golding, A. Puhler, The complete sequence of the 1,683-kb pSymB megaplasmid from the N_2-fixing endosymbiont *Sinorhizobium meliloti*, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 9889–9894.
[147] B. Goodner, G. Hinkle, S. Gattung, N. Miller, M. Blanchard, B. Qurollo, B.S. Goldman, Y. Cao, M. Askenazi, C. Halling, L. Mullin, K. Houmiel, J. Gordon, M. Vaudin, O. Iartchouk, A. Epp, F. Liu, C. Wollam, M. Allinger, D. Doughty, C. Scott, C. Lappas, B. Markelz, C. Flanagan, C. Crowell, J. Gurson, C. Lomo, C. Sear, G. Strub, C. Cielo, S. Slater, Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58, Science 294 (2001) 2323–2328.
[148] D.W. Wood, J.C. Setubal, R. Kaul, D.E. Monks, J.P. Kitajima, V.K. Okura, Y. Zhou, L. Chen, G.E. Wood, N.F. Almeida Jr., L. Woo, Y. Chen, I.T. Paulsen, J.A. Eisen, P.D. Karp, D. Bovee Sr., P. Chapman, J. Clendenning, G. Deatherage, W. Gillet, C. Grant, T. Kutyavin, R. Levy, M.J. Li, E. McClelland, A. Palmieri, C. Raymond, G. Rouse, C. Saenphimmachak, Z. Wu, P. Romero, D. Gordon, S. Zhang, H. Yoo, Y. Tao, P. Biddle, M. Jung, W. Krespan, M. Perry, B. Gordon-Kamm, L. Liao, S. Kim, C. Hendrick, Z.Y. Zhao, M. Dolan, F. Chumley, S.V. Tingey, J.F. Tomb, M.P. Gordon, M.V. Olson, E.W. Nester, The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58, Science 294 (2001) 2317–2323.
[149] V. Anantharaman, E.V. Koonin, L. Aravind, Comparative genomics and evolution of proteins involved in RNA metabolism, Nucleic Acids Res. 30 (2002) 1427–1464.
[150] P.A. Jones, D. Takai, The role of DNA methylation in mammalian epigenetics, Science 293 (2001) 1068–1070.
[151] D. Bruniquel, R.H. Schwartz, Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process, Nat. Immunol. 4 (2003) 235–240.
[152] S.K. Bhattacharya, S. Ramchandani, N. Cervoni, M. Szyf, A mammalian protein with specific demethylase activity for mCpG DNA, Nature 397 (1999) 579–583.
[153] A. Bird, *Il2* transcription unleashed by active DNA demethylation, Nat. Immunol. 4 (2003) 208–209.
[154] J. Jiricny, DNA repair: bioinformatics helps reverse methylation damage, Curr. Biol. 12 (2002) R846–R848.
[155] G. Margison, A new damage limitation exercise: ironing (Fe(II)) out minor DNA methylation lesions, DNA Rep. (Amst.) 1 (2002) 1057–1061.
[156] J.G. Hengstler, M.S. Bogdanffy, H.M. Bolt, F. Oesch, Challenging dogma: thresholds for genotoxic carcinogens? The case of vinyl acetate, Annu. Rev. Pharmacol. Toxicol. 43 (2003) 485–520.
[157] R. Peto, R. Gray, P. Brantom, P. Grasso, Effects on 4080 rats of chronic ingestion of *N*-nitrosodiethylamine or *N*-nitrosodimethylamine: a detailed dose–response study, Cancer Res. 51 (1991) 6415–6451.
[158] W.J. Waddell, Threshold for carcinogenicity of *N*-nitrosodiethylamine for esophageal tumors in rats, Food Chem. Toxicol. 41 (2003) 739–741.
[159] S. Fukushima, H. Wanibuchi, K. Morimura, M. Wei, D. Nakae, Y. Konishi, H. Tsuda, N. Uehara, K. Imaida, T. Shirai, M. Tatematsu, T. Tsukamoto, M. Hirose, F. Furukawa, K. Wakabayashi, Y. Totsuka, Lack of a dose–response relationship for carcinogenicity in the rat liver with low doses of 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline or *N*-nitrosodiethylamine, Jpn. J. Cancer Res. 93 (2002) 1076–1082.
[160] M. Okamura, K. Yasuhara, T. Koujitani, M. Nasu, M. Hirose, K. Mitsumori, Dose-threshold for thyroid tumor-promoting effects of xylazine in rats, Toxicology 188 (2003) 275–283.
[161] H.M. Bolt, Genotoxicity-threshold or not? Introduction of cases of industrial chemicals, Toxicol. Lett. 140–141 (2003) 43–51.
[162] K.D. Pruitt, T. Tatusova, D.R. Maglott, NCBI Reference Sequence project: update and current status, Nucleic Acids Res. 31 (2003) 34–37.
[163] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[164] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[165] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.
[166] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, MEGA2: molecular evolutionary genetics analysis software, Bioinformatics 17 (2001) 1244–1245.