We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,100
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
1. Introduction

Direct repair is defined as the elimination of DNA and RNA damage using chemical reversion that does not require a nucleotide template, breakage of the phosphodiester backbone or DNA synthesis. As such, the process of direct repair is completely error-free, granting a major advantage in preservation of genetic information. In mammalian cells, direct repair is utilized to repair specific types of DNA and RNA damage caused by ubiquitous alkylating agents. Only two major types of proteins conduct direct repair in mammalian cells, O6-methylguanine-DNA methyltransferase (MGMT or AGT) and ALKBH family Fe(II)/α-ketoglutarate dioxygenases (FeKGDs). In humans and mice, a single direct repair methyltransferase protein exists, MGMT. In contrast, ALKBH FeKGDs represent a family of nine homologs with conserved active site domains. Although the biochemical function of a number of ALKBH proteins and their biological roles require further investigation, several directly repair alkylation damage in DNA and RNA at base-pairing sites.

2. Direct repair substrates—DNA and RNA alkylation damage

Exposure to alkylating agents is major cause of DNA and RNA damage, generating adducts that can compromise genomic integrity. As a result, repair of alkylation adducts is mediated by a variety of DNA repair pathways, some with overlapping substrate specificity. However, direct DNA repair proteins utilize unique mechanisms to specifically eliminate damage at base-pairing sites. The frequency and site of DNA and RNA damage occurrence is dependent on the source and type of alkylating agent exposure, as discussed in this section.
3. Sources of alkylation damage

Alkylating agents are present environmentally and also generated within the cell via oxidative metabolism. They modify DNA and RNA, forming adducts that disrupt replication and transcription, trigger cell cycle checkpoints, and/or initiate apoptosis. If left unrepaired, some adducts formed by alkylation damage can be cytotoxic and/or mutagenic [1-3].

Environmental alkylating agents fall into two primary groups, nitrosoureas that generate primarily O-alkylations and methanesulfonates that cause mostly N-alkylations [1, 3] (Figure 1). These exogenous alkylating agents are present in air, water, plants and food, in the form of nitrosamines, chloro- and bromomethane gases, myosamines and halocarbons [4]. There are also industrially produced alkylating agents, including various chemotherapeutic agents [5, 6].

Enzymes involved in cellular metabolism are responsible for the majority of endogenous alkylation damage. Nitrosating agents are generated, resulting in amine nitrosation, and reactive oxygen species (ROS), which cause lipoperoxidation [7]. Additionally, a family of S-adenosyl methionine (SAM) methyltransferase enzymes is involved in more than 40 metabolic reactions using SAM as a methyl donor to modify nucleic acids, proteins and lipids [8, 9]. Four of those SAM methyltransferase enzymes participate in DNA and RNA modification in mammalian cells. DNMT1, DNMT3A, and DNMT3B catalyze methyl group transfer at the C5 position of cytosine in DNA CpG sequences [10], whereas TRDMT1 (DNMT2) methylates the C5 position of cytosine 38 in aspartic acid tRNA [11].

3.1. Types of alkylating agents

Alkylating agents can be categorized by their method of activation. Some alkylating agents react directly with DNA and do not require any activation, whereas many alkylating agents, in-
cluding many carcinogens, must undergo metabolic activation by the cytochrome P450 system to generate reactive species capable of modifying DNA [3, 12, 13]. In addition, alkylating agents are electrophilic compounds that possess either one or two reactive groups that can interact with the nucleophilic centers of DNA and RNA bases. Alkylating agents that can only react with one nucleophilic center are mono-functional, whereas bi-functional agents can react with two sites in DNA or RNA [1, 13]. Alkylating agents that are mono-functional primarily transfer alkyl groups to ring nitrogens, while agents that react in a bi-functional manner not only react with ring nitrogens, but can form cyclized DNA bases, by reacting with exocyclic nitrogen and oxygen groups [13] (Figure 2). In addition to methylating agents, larger alkylating agents also modify nucleic acids — bi-functional ethylating agents can form exocyclic ethano and etheno adducts at nitrogen and oxygen molecules in all DNA and RNA bases. Additionally, bi-functional alkylating agents can produce DNA inter- and/or intrastrand cross-links [13]. Some alkylating agents also react at phosphate residues to generate phosphotriesters, leading to potential single-strand breaks [13] (Figure 2). Two main pathways, characterized as $S_{N1}$ or $S_{N2}$, are defined based on the kinetics of the alkylation reaction, leading to the above mentioned modifications of DNA and RNA bases [2].

![Figure 2](http://dx.doi.org/10.5772/54449)

**Figure 2.** (A) Purple arrows indicate sites in DNA most often methylated by $S_{N1}$ alkylating agents. Green arrows indicate sites commonly modified by $S_{N2}$ alkylating agents, orange arrows indicate sites in single-stranded DNA. Blue arrows indicate exocyclic amino groups important in formation of cyclized DNA adducts. The location of the major and minor grooves in DNA are indicated. “R” is the attachment of the base to the deoxyribose and phosphodiester backbone. (B) Modified phosphodiester isoforms in the DNA backbone. $S_{N1}$ alkylating agents generally form more phosphotriester products than $S_{N2}$ agents. [2,14]
$S_{N1}$ agents act via a two step reaction involving a unimolecular nucleophilic substitution with a rate-limiting step that generates an intermediate carbonium ion electrophile that reacts with nucleophilic DNA sites. Thus, the reaction kinetics depend only on the formation of the carbonium ion intermediate (first-order). The trigonal planar conformation of the $sp^2$ hybridized carbon generated in the carbocation intermediate permits nucleophilic attack from either side, yielding a racemic mixture of reaction products at chiral centers [13] (Figure 3). Though agents that react via an $S_{N1}$ mechanism produce both N- and O-alkylations, increased amounts of modified oxygens are generated, compared to agents that react via an $S_{N2}$ mechanism.

![Figure 3. $S_{N1}$ and $S_{N2}$ nucleophilic substitution reactions. (A) Example of an $S_{N1}$ reaction. $S_{N1}$ reactions are dependent on formation of a carbonium ion intermediate that rate-limiting. Product chiral centres are a racemic mixture because the intermediate can be attacked by either side. (B) Example of an $S_{N2}$ reaction. Both reactants are required and there is direct attack by the nucleophile in $S_{N2}$ reactions. Chirality is maintained since a transition state is formed with the chiral center. [2,14]](image)

In contrast, $S_{N2}$ reaction mechanisms depend on both the alkylating agent and its target to define the kinetics (second-order). Using a one step reaction where both the electrophile and nucleophile are involved in the transition state, $S_{N2}$ alkylating agents proceed with direct attack by the nucleophile on an electron deficient center. The nucleophile attacks from the back of the electrophile, forming the carbon-nucleophile bond and breaking the carbon-leaving group bond. Simultaneous backside, nucleophilic attack and leaving group departure cause the incoming group to replace the leaving group. Because a transition state is formed with the chiral center, chirality is maintained, leading to a stereocenter (inversion) configuration [13] (Figure 3). Alkylating agents that react via an $S_{N2}$ mechanism cause primarily N-alkylations.

### 3.2. DNA and RNA alkylation damage

Modification sites of DNA bases are the same for all alkylating agents and include all the exocyclic nitrogens and oxygens, as well as ring nitrogens without hydrogen. Though all DNA nucleobase oxygen or nitrogen atoms can be alkylated, the type and frequency of specific damage varies depending on the type of alkylating agent, the structure of the substrate, and the position of the damage site [13] (Table 1). Generally, alkylation damage at nitrogen
molecules is less mutagenic than oxygen, though both types of alkylation damage are cytotoxic and genotoxic [14].

Common alkylations generated by exogenous alkylating agents include O⁶-alkylguanine and O⁴-alkylthymine adducts, as well as N⁷-alkylguanine, N³-alkyladenine, N¹-alkyladenine, and N³-alkylcytosine [13] (Figure 1). Moreover, the frequency of each adduct type depends on whether the DNA and RNA substrates are single- or double-stranded [13] (Table 1). For instance, nitrogen molecules involved in DNA base-pairing are less vulnerable to alkylation damage than the same base nitrogens in a single-stranded region arising during replication and transcription.
4. Direct repair proteins

Numerous cellular mechanisms have evolved to deal with various types of DNA damage and each DNA repair pathway is important to maintain genomic integrity. However, most repair mechanisms require DNA synthesis and therefore an intrinsic risk of causing mutation in executing the repair. In contrast, direct repair proteins, MGMT and ALKBH family proteins employ direct reversal mechanisms that result in complete restoration of DNA bases and are thus error-free mechanisms. Moreover, MGMT, ALKBH2, and ALKBH3 repair endogenous and exogenous DNA and RNA alkylation damage at critical base-pairing sites, facilitating proper replication of genetic information or transcription. This section will discuss each of these direct DNA repair enzymes in detail.

Figure 4. Major mechanisms of alkylation adduct repair. Direct repair pathways are indicated in green. Base and nucleotide excision repair pathways are indicated in blue [2,14].

4.1. Mechanisms of alkylation repair

Multiple mechanisms are employed to rid the genome of alkyl adducts, thereby preventing detrimental effects within the cell (Figure 4). Mismatch repair (MMR), base excision repair (BER) and nucleotide excision repair (NER) and direct repair (DR) pathways all participate in alkylation damage repair [15-24]. Specifically, BER and NER repair small alkylated base damage including 7-methylguanine (7-meG) and 3-methyladenine (3-meA) DNA adducts [25]. Although BER repairs the majority of small alkylated base damage (methyl and ethyl adducts) the NER system can also remove small, as well as bulky adducts larger than ethylated bases [24, 26]. As an alternative to NER, incomplete BER repair intermediates can be processed by homologous recombination (HR) [27]. However, BER, NER and HR repair pathways generate strand breaks during repair of alkyl adducts and could introduce muta-
tions or rearrangements [28]. On the contrary, DR mechanisms, provided by methyltransferase MGMT and ALKBH homologs, eliminate alkylation damage at DNA base-pairing sites, including O\textsuperscript{6}-methylguanine (O\textsuperscript{6}-meG), 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) and do not require a nucleotide template, result in phosphodiester backbone breakage, nor do they require DNA synthesis.

4.2. Methyl Guanine Methyl Transferase (MGMT) proteins

In mammals, methylguanine DNA methyltransferase (MGMT or AGT), can repair two types of DNA adducts: O\textsuperscript{6}-methylguanine (O\textsuperscript{6}-meG) and O\textsuperscript{4}-methylthymine (O\textsuperscript{4}-meT). O\textsuperscript{6}-meG adducts in DNA are extremely mutagenic [29, 30] and also block DNA polymerase extension, which is generally associated with cytotoxicity [31, 32]. The primary mutations observed when there is a failure to repair O\textsuperscript{6}-meG adducts prior to replication are G:C \rightarrow A:T transitions, whereas a failure to repair O\textsuperscript{4}-meT results primarily in T:A \rightarrow C:G transition mutations [29]. In mammals, elimination of O\textsuperscript{6}-meG by MGMT is preferred over O\textsuperscript{4}-meT, but the respective efficiency of each type of reversion is species dependent [29, 33-37].

Removal of O\textsuperscript{6}-meG and O\textsuperscript{4}-meT modifications are achieved via a one-step methyltransferase reaction, wherein MGMT accepts the alkyl adduct from the modified oxygen molecule, onto an internal residue, directly restoring the DNA base and inactivating the protein [38] (Figure 5). In addition to methyl groups, several other alkyl-adducts can also be transferred from guanine to MGMT, including ethyl-, propyl- butyl-, benzyl- and 2-chloroethyl-. However, the efficiency of the reaction is decreased for alkyl adducts greater than methylated bases [39]. Once modified, the protein is targeted for elimination via the proteasome [40].

4.2.1. Protein structure/active site organization

Alkyltransferase proteins are found in eukaryotic and prokaryotic organisms and have been identified in as many as 100 organisms [41]. Though sequences are not highly conserved between human MGMT and Eubacterial, Archea, and Eukaryotic DNA methyltransferase enzymes, structural domains and active site residues are almost identical [42-46].
In human MGMT, a conserved α/β roll structure, containing a three-stranded, anti-parallel β-sheet, followed by two helices, make up the N-terminus (residues 1-85). The MGMT C-terminus (residues 86-207) contains a short, two-stranded, parallel β-sheet, four α-helices and a $3_{10}$ helix [42, 47]. Found only in humans, a zinc ion stabilizes the interface between the N- and C-termini, binding Cys5, Cys24, His29 and His85 in a tetrahedral conformation to bridge three strands of the N-terminal β-sheet with the coil preceding the $3_{10}$ helix in the C-terminus [47].

The conserved active site cysteine motif (-PCHR-) is located in the C-terminus contained within the DNA binding channel, and the helix-turn-helix (HTH) DNA binding motif. Residues Try114-Ala121 form the first helix of the HTH motif and residues Ala127-Gly136 form the second, “recognition” helix, which interacts with DNA. Linked by an Asn-hinge (Asn137) that stabilizes the over-lapping turns by binding Val139, Ile143 and the Cys145 thiol, the -PCHR- active site is located near the “recognition” helix [42, 47, 48].

The active site of human MGMT is composed of at least ten residues that participate in substrate binding, enzyme structure and alkyl transfer. Residues Val155-Gly160 and Met134 generate a hydrophobic cleft in the active site loop, while residues Tyr114, His146, Val148, Ser159, and Glu172 participate in active site coordination and alkyl group transfer to residue Cys145. Not unexpectedly, mutation of residue Cys145 results in elimination of alkyl group transfer, however substrate binding is unaffected [49] (Figure 6).
4.2.2. Substrate recognition/repair mechanism

In repair, MGMT is unique in that one molecule is responsible for the removal of one O\textsuperscript{6}\textsubscript{meG} or O\textsuperscript{4}\textsubscript{meT} adduct. Unlike most enzymes with the capacity to catalyze multiple reactions, MGMT catalyzed reactions are stoichiometric and capable of only a single repair reaction [50]. As a result, removal of O\textsuperscript{6}\textsubscript{meG} and O\textsuperscript{4}\textsubscript{meT} alkyl adducts is dependent on both MGMT and the substrate concentrations (second-order reaction).

The recognition of guanine and thymine base methylation is accomplished by a highly conserved amino acid structure. The hydrophobic cleft of the active site loop and -PCHR- motif within the binding channel allow MGMT to bind to the minor-groove of DNA using residues Ala126, Ala127, Ala129, Gly131, and Gly132, of the HTH “recognition” helix [51, 52], which is followed by necessary conformational changes to orient the damaged base within the active site.

Identified based on bacterial Ada homology and human MGMT structures, following substrate recognition, the target base is repaired using a base flipping mechanism [53-58]. In the MGMT repair reaction, the damaged base undergoes a residue Tyr114-mediated, sterically enforced 3’ phosphate rotation into the active-site pocket. The hydrophobic cleft formed by the active site loop easily accepts the extra-helical base, causing the DNA minor groove to widen [51]. The arginine finger residue, Arg128, intercalates between the DNA bases and interacts with the unpaired cytosine, via a charged hydrogen bond [55], maintaining an appropriate DNA duplex conformation (Figure 6).

Once bound within the MGMT active site, numerous residues participate in the methyltransferase reaction. A hydrogen bond network, conserved in AGTs, is formed between Glu172, His146, water and Cys145. His146 acts as a water-mediated base that deprotonates Cys145, converting Cys145 to a cystine thiolate anion and generating an imidazolium ion.
that is stabilized by Glu172 [35, 59]. Residues, Val148 and Cys145 carbonyls accept guanine exocyclic amine hydrogen bonds and nitrogen atoms of residues Tyr114 and Ser159 donate protons to N3 and O6 of O6-meG, respectively. The deprotonated Tyr114 residue abstracts a proton from Lys165, simultaneously transferring the alkyl group from the O6 position of guanine to the thiolate anion of the Cys145 residue [35]. Transfer of the alkyl group generates a thioether, S-alkylcysteine, and results in complete restoration of the guanine base, as well as irreversible inactivation of the methyltransferase enzyme (Figure 5). While many DNA repair proteins have a specific requirement for double-stranded DNA, MGMT can also bind to single-stranded DNA [60].

4.2.3. Gene expression/protein regulation

Removal of O6-meG modifications by MGMT has a major role in cell cycle checkpoint control, proliferation, and differentiation [61]. As a result, MGMT is a house-keeping gene that is expressed in all tissues; though expression varies depending on cell type [62]. MGMT expression in an individual cell or tissue type is dependent on a variety of factors, including numerous types of stimuli and promoter regulator elements. However, the relationship between factors that mediate MGMT expression and the regulation of its function is not well-understood. The lack of understanding regarding the consequences of MGMT regulation is illustrated by the fact that MGMT expression is silenced in some cancers, but expression is up-regulated in others [62, 63].

MGMT is a single gene on chromosome 10q26, spanning approximately 300kb [64]. The gene has five exons, but the first is non-coding [65, 66]. The promoter of MGMT is a non-TATA-box promoter that contains a GC-rich CpG island of 780 bp that includes 97 CpG dinucleotides [67]. CpG islands are commonly associated with promoter regions of constitutively expressed genes, from which transcription is initiated from a single promoter site [68-70]. Additionally, the promoter contains six transcription consensus binding sites (SP1, AP1, and AP2), three upstream and three downstream of the transcription start site, a glucocorticoid-responsive element, and a 3’ enhancer element [62, 67, 69, 71]. Though unmethylated in normal cells, promoter CpG island methylation-induced silencing of MGMT is found in various cancer types and MGMT-deficient cell lines and is one mechanism that regulates MGMT expression [72-76]. However, whether MGMT promoter methylation disables transcription factor binding or contributes to chromatin reorganization remains uncertain [71, 75].

In addition to numerous transcription factor binding sites that surround the MGMT promoter transcription start site, the MGMT promoter CpG islands exhibit a chromatin structure that mediates interaction with transcription factors. The MGMT gene is organized around five or more nucleosomes in a manner that positions 300 bp region of the promoter sequence, which contains known MGMT transcription factor binding sites, so that it does not lie within the nucleosomes, and therefore does not maintain a higher-order chromatin structure [62, 72, 77]. Such nucleosomal positioning facilitates an “open” stretch of DNA that enables constitutive interaction of transcription factors with the promoter.
Methylation of the CpG island surrounding the transcription factor binding sites contributes to lack of transcription factor binding, but could also effect nucleosomal positioning of the MGMT promoter [62, 71], suggested by histone H3 Lys9 (H3K9) di-methylation, exhibited in relationship to MGMT silencing [78, 79]. Further, deacetylation of histones H3 and H4 could also be associated nucleosome organization that is more condensed, resulting in transcription inactivation. Therefore, the chromatin structure of the MGMT promoter, as well as CpG island methylation, mediate transcription factor access to the promoter and are important for MGMT expression.

4.2.4. Protein localization and cell type dependence

Immunofluorescence studies indicate MGMT nuclear localization at discrete nuclear regions [80]. Although a nuclear localization signal (NLS) for MGMT has not been identified, the small size of MGMT, 23 kDa, may not require an active translocation signal to traverse nuclear pores [53]. However, a –PKAAR- sequence within the DNA binding domain of MGMT is necessary for DNA interactions to facilitate nuclear retention [81]. The highest MGMT expression levels are found in the liver, where high levels of endogenous nitrosating agents are present, but MGMT is also expressed at high levels in the lung, kidney and colon. MGMT expression is heterogeneous in the brain and the lowest levels are observed in the pancreas, hematopoietic cells, lymphoid tissues [62, 67, 82-86].

4.2.5. Post-translational modification

Once MGMT has transferred a methyl group to its Cys145 residue, no further reactions are catalyzed, so the protein must be eliminated. The degradation of MGMT is an ubiquitination-dependent process that has been evaluated using inactivation of the protein by O-6-BzG, BCNU, or NO-generating agents at position Cys145 [40, 87, 88]. Conformational changes in the protein structure after alkyl group transfer target MGMT for ubiquitination and proteasomal degradation [40, 89]. Two sites within MGMT, Lys125 and Lys178, have been identified as ubiquitination targets in B lymphocyte (NCI-H929) or 293T, and myeloid (MV4-11) cells, respectively. Additionally, examination of potential MGMT modification sites using predictive software also identifies Lys104 as an ubiquitination target. Furthermore, predictions also indicate post-translational modification sites for methylation (Arg128, Arg135), acetylation (Lys8, 125, 178, 193), and sumoylation (Lys75, 205, 18, 107), as well as numerous phosphorylation sites (Ser36, 56, 130, 182, 202, 206, 208; Thr37; Tyr91, 115) [90-93], which all merit further consideration. Notably, phosphorylation of residues Thr10 and Thr11 was also noted in HeLa cells [92], and phosphorylation of Ser201 is observed in B lymphocyte cells (DG75 and GM00130), KGI myeloid cells, and HeLa cervical cancer cells. Importantly, crystallographic data suggests that modification of Ser201 could disrupt interaction with DNA [48, 51, 55].

4.3. AlkBh Fe(II)/α-ketoglutarate-dependent dioxygenases

In mammals, repair of cytosine and adenine base methylation at base-pairing positions is specifically associated with the AlkB family dioxygenase proteins [92, 94-96]. Discovered
first in *Escherichia coli* (*E. coli*) in 1983 [96] alkylation protein B (AlkB) belongs to a superfamily of Fe(II)/α-ketoglutarate-dependent dioxygenases (FeKGDs), with roles in histone de"methylation [97-99], proline hydroxylation [95] and in the case of AlkB, the ability to directly remove alkyl adducts generated in DNA residues as a result of exposure to S\textsubscript{N}2 alkylating agents [94, 100]. Originally predicted to act on 1-methyladenine (1-meA) and 3-methylcytosine (3-meC), bacterial AlkB has been shown to repair a variety of DNA and RNA adducts, including 1-meA, 3-meC, 1-meG, 3-meT, 1-etA, as well as aromatic ethyl, 3-etC, and etheno adducts, 1,N\textsuperscript{6}-ethenoadenine (εA) and 3,N\textsuperscript{4}-ethenocytosine (εC) [94, 100-108] (Figure 7). Using bioinformatics, nine human ALKBH family enzymes, ALKBH1-8 and FTO, were identified, of which only four have been reported to have DNA repair activity, ALKBH1 – ALKBH3 and FTO [109, 110]. Though all of the ALKBH homologs contain conserved catalytic domain residues, none entirely encompass the enzymatic activity of AlkB [15, 103, 104, 111-114]. Removal of alkyl adducts from DNA is only accomplished by three ALKBH proteins, ALKBH1-3, known to remove 1-meA and 3-meC adducts. However, ALKBH1 is reportedly a mitochondrial protein [115], therefore in the nucleus ALKBH2 and ALKBH3 proteins are employed to remove specific adducts in single- or double-stranded DNA or in RNA [104]. Lesions that are repaired by ALKBH proteins generally interfere with base-pairing and block replication and transcription, triggering cell cycle checkpoints and apoptosis [92, 95, 96, 110, 115]. In *E. coli* AlkB mutants, as well as in Alkbh2- or Alkbh3-deficient mouse embryonic fibroblasts, cells exhibit increased sensitivity to alkylating agents, particularly the S\textsubscript{N}2 type, and increased mutant frequency [101, 116-119].

![Figure 7](image)

**Figure 7.** ALKBH protein substrates. (A) DNA methyl adducts repaired by ALKBH proteins. (B) DNA etheno adducts repaired by ALKBH proteins.
4.3.1. Protein structure/active site organization

Similar to MGMT, the sequences of human ALKBH proteins do not contain a high percentage of sequence homology in regions other than active sites and conserved domains, but do have conserved secondary structures [109, 110, 114, 120-122]. In AlkB family proteins, the catalytic core is composed of three major components, the double-stranded β-helix (DSBH), the nucleotide recognition lid (NRL) and the N-terminal extension (NTE) (Figure 8). The DSBH is comprised of eight β-strands in the C-terminal portion that form two β-sheets to create a central core jelly-roll fold. Within the major and minor β-sheets of the DSBH lie conserved catalytic residues RxxxxxR and HxDxₙH, respectively [120, 121, 123]. The HxD dyad is near the amino terminal end and is located in a flexible loop that follows the first strand, stacking with the minor β-sheet. The carboxy-terminal histidine of the conserved HxDxₙH residues is associated with the beginning of the sixth strand and together these residues coordinate iron (His171, Asp173 and His236—Alkbh2; His191, Asp193 and His258—Alkbh3) [114, 120, 121, 123, 124]. The histidine and aspartic acid residues (Asp248 and Asp254—ALKBH2; Asp269 and Asp275—ALKBH3), conserved in the DSBH minor β-sheet, coordinate Fe(II), α-ketoglutarate and the DNA or RNA repair substrate within the catalytic core. A conserved Arg residue in the C-terminal β-strand (Arg254—ALKBH2 and Arg275—ALKBH3) sets AlkB family proteins apart from other α-ketoglutarate-dependent dioxygenases within the Fe(II)/α-ketoglutarate dioxygenase superfamily, forming the base of the substrate binding pocket [110, 120, 121, 123].

Figure 8. Structure of human AlkB homolog DNA repair proteins. Two looped structures (flip1 and flip2) generated by anti-parallel β-sheets create the nucleotide recognition lid (NRL) and are involved in DNA base flipping. (A) Structure of ALKBH2 (PDBid3BTX). ALKBH2 double-strand DNA substrate specificity is facilitated by residues in loops L1 and L2. (B) Structure of ALKBH3 (PDBidZIUV). β-sheets 4 and 5 form the β-hairpin motif in ALKBH3. Part of loop 1, involved in ALKBH substrate specificity, was omitted due to electron density problems. [121]

The N-terminal extension (NTE) and Nucleotide Recognition Lid (NRL) are formed by the β-hairpin motifs that extend from the DSBH jelly-roll, forming a substrate binding groove.
that covers the active site until bound. Ninety residues are contained within two looped structures, forming “flips” that lie between a single β-sheet and two α-helices in the N-terminal portion of the catalytic core [120, 121]. Secondary structures are of similar size, but possess different characteristics important for substrate specificity and DNA activity. In ALKBH2, the first flip is 20 residues that make up a β-hairpin and short α-helix, creating a hydrophobic binding groove. In contrast, the first flip in ALKBH3 is a β-hairpin made up of 17 residues that form a hydrophilic, positively charged binding groove, more suitable for single-stranded DNA or RNA substrates [15, 120]. The characteristics of the second flip are also unique. Flip two of ALKBH2 spans 24 residues that is made up of three β-sheets, with numerous sites for DNA substrate interaction. The orientation of the three β-sheets, which fold back towards the C-terminal end of the first α-helix, is also unique only to ALKBH2 [114, 121]. However, flip 2 of ALKBH3 is only 12 residues and contains a single β-sheet [114]. The N-terminal regions of each ALKBH homolog are more variable and hypothesized to play roles in sub-cellular sorting and protein-protein interactions [114, 115] (Figure 8).

In addition to the conserved catalytic dioxygenase residues, some human ALKBH proteins also contain additional catalytic residues and domains [104, 109, 110, 113, 125] (Figure 9). Structural analysis of bacterial AlkB and human ALKBH homologs provides insight into substrate preferences and repair capabilities. For instance, ALKBH2 contains three unique motifs that facilitate enhanced activity on double-stranded DNA [121]. A long, flexible β-sheet hairpin loop that contains DNA binding residues Arg198, Gly204 and Lys205, a short loop that contains the RKK motif (Arg241-Lys243) and an aromatic finger residue (Phe102) are used to make contacts with both DNA strands, rotate and take the place of the damaged base in duplex DNA molecules. On the other hand, the number and organization of the catalytic domains in ALKBH3 result in differential manipulation of the DNA backbone, explaining the preference for single-strand substrates. Lack of an aromatic finger residue and RKK motif in ALKBH3, the damaged base is squeezed on either side, forcing it to rotate, and the immediate 5’ and 3’ bases to stack against one another. However, structural analysis of ALKBH3 has identified residue Arg122, specifically the arginine side chain length, as important for double-stranded DNA substrate activity, possibly mimicking the base-flipping and stacking activities of ALKBH2 residue Phe102 [114, 121].

Unfortunately, extensive biochemical analysis or structural studies have not been conducted on ALKBH homologs 4-8. However, it is apparent that differences in the number and organization of catalytic residues, as well as secondary structures play a large role in the diversity of ALKBH family protein substrate specificities and enzymatic activities [113]. For instance, although single- or double-strand DNA repair activity has not been established for ALKBH8, the presence of RNA binding and methyltransferase domains in ALKBH8 (Figure 9) suggested that this homolog plays a role in maintenance of methylation patterns. Investigation of such activities led to the identification of ALKBH8 tRNA methyltransferase activity, necessary in the biogenesis of wobble uridine modifications utilized in translational decoding [126, 127].
4.3.2. Substrate recognition/repair mechanism

Initially, it was predicted that AlkB family proteins directly repaired alkylation adducts by hydroxylating methyl groups and removing the resultant hydroxymethyl groups via an oxidative reaction that directly restores the undamaged base [94, 109, 112, 124, 128, 129]. However, specific investigation of the AlkB family dealkylation mechanism [130] determined that the direct repair reaction mediated by AlkB family proteins involves several intermediate steps that had not yet been identified. Regardless, dealkylation catalyzed by AlkB and its human homologs occurs via transformation of α-ketoglutarate into succinate, formaldehyde release, and restoration of the undamaged base [94, 100, 111, 130, 131] (Figure 10).
Figure 10. ALKBH protein repair reactions. (A) ALKBH methyl adduct repair reaction. (B) ALKBH ethyl adduct repair reaction. (C) ALKBH etheno adduct repair reaction. Repair of ethyl and etheno adducts requires the same co-factors, but displaces acetaldehyde or water and glyoxal as byproducts of the repair reaction, respectively, instead of formaldehyde [100,102,103].

First, Fe(II) and three water molecules must be coordinated within the conserved catalytic core, stimulating α-ketoglutarate (KG) binding in the catalytic pocket. Binding of α-KG into the catalytic pocket chelates Fe(II) by displacing two water molecules to create the Fe(II)/α-KG active-site complex. Ligation of dioxygen to the Fe(II) molecule displaces the remaining water molecule, generating a ferric-superoxido species that undergoes self-redox and nucleophilic attack on the α-keto group. This nucleophilic attack is necessary to decarboxylate α-KG, releasing succinate and generating a ferryl-oxo intermediate. Reorientation of this intermediate facilitates removal of a hydrogen atom from the methyl adduct. Finally, radical rebound hydroxylation of the methylene group results in decomposition of the hydroxymethyl nucleobase, yielding formaldehyde and the repaired nucleobase. Though two co-factors were noted initially, α-ketoglutarate and Fe(II), ascorbate also plays a role, helping to convert the Fe(III) to Fe(II), thereby regenerating the original oxidative state of iron in the Alkbh proteins that permits enzymatic cycling [94, 111, 112, 122, 124, 130].

The major methylated bases repaired by ALKBH proteins are 1-methyladenine (1-meA) and 3-methylcytosine (3-meC), however homologs have also been reported to repair ethylated, and some etheno and exocyclic bases [102-105, 107, 131, 132]. Similar mechanisms are proposed for repair of ethano and exocyclic etheno (ε) adducts, though the final steps of these
reactions result in release of acetaldehyde and glycol, respectively [130] (Figure 10). However, additional biochemical studies are needed to confirm these mechanisms in similar detail to removal of methyl adducts from DNA.

4.3.3. Gene expression/protein regulation

Human AlkB DNA repair homologs, ALKBH2 and ALKBH3 are single genes on chromosomes 12q24 and 11p11, respectively. Expression of human AlkB homologs has been reported in a variety of normal tissue samples, including ALKBH homologs 4-8, despite the lack of DNA repair activity in the literature [133]. Expression of ALKBH family proteins varies depending on cell types. Protein expression levels in the various tissue types vary depending on the homolog evaluated. Little is known of ALKBH protein regulation mechanisms and is an area in need of further study.

4.3.4. Protein localization and cell type dependence

Differences amongst AlkB homolog proteins in their biological roles are partially ascribed to their sub-cellular localizations. ALKBH2 and ALKBH3 homolog proteins are expressed at the highest levels in the testis and ovary, however detectable expression of all AlkB homolog proteins is exhibited in the spleen, pancreas, lung, kidney, prostate and brain [133]. Although ALKBH1 activity is confined to mitochondria [115], immunofluorescence imaging indicates that the protein is cytoplasmic and nuclear [133]. Similarly, AlkB homolog proteins ALKBH3, 4, 6, and 7 are also present in the nucleus and cytoplasm [133], though ALKBH3 is the only homolog reported to possess repair activity [1, 104, 111]. Localization of ALKBH3 in both the nucleus and cytoplasm is consistent with identified interactions with helicase enzymes to facilitate DNA repair [134] and roles in mRNA repair [131]. ALKBH2 is present only in the nucleus and exhibits diffuse as well as localized, punctate staining, supporting pre-established co-localization with PCNA at replication foci during S phase [111, 131, 133], suggesting a role in replication- and transcription-related repair, as well as genome maintenance. On the contrary, AlkB homolog proteins ALKBH5 and 8 are present only in the cytoplasm [133], which supports known ALKBH8 tRNA methyltransferase activity [126, 127].

4.3.5. Post-translational modification

Unlike MGMT, ALKBH proteins are not suicide enzymes and a single protein can catalyze multiple direct repair reactions, requiring only ascorbate to regenerate the Fe(II) active site center [135]. Therefore, immediate degradation of ALKBH proteins following repair is not required, as it is for MGMT. Other possible post-translational modifications in ALKBH2 and ALKBH3 include candidate sites for phosphorylation and acetylation. Mass-spectrometric analysis of a curated database of cell lines revealed that both ALKBH2 and ALKBH3 proteins undergo post-translational modification of specific residues present in various cancer types [92].

Post-translational modifications curated for ALKBH2 include acetylation of residue Lys34 and Lys104 in various colorectal cancer cell types (HCT116, HT29, XY3-92-T and XY3-68-T),
as well as phosphorylation of residue Thr252 in esophageal cancer cell line XY2-E111N [92].

Though the exact effects of these modifications are unknown, it is important to state that Lys34 is within the variable region of the N-terminus that is thought to provide protein specificity. Similarly, Lys104 is between two residues that make contact with the complimentary DNA strand during double-strand DNA repair and Thr230 is a residue in the most C-terminal α-helix of the active site [92]. Examination of potential ALKBH2 modification sites using predictive software shows possible post-translational modification sites for methylation (Arg128, 135), sumoylation (Lys75, 205), and ubiquitination (Lys104), along with other possible phosphorylation sites (Ser36, 56, 130, 182, 202, 206, 208; Thr37; Tyr91, 115) [90-93]. All of those possible post-translational modifications merit further consideration.

Post-translational modifications were also present in ALKBH3, corresponding to various disease states. Phosphorylation of Thr126 and Tyr127 residues in the β-hairpin of the NRL, as well as residue Try229 in the ALKBH3 active site, was present in acute myelogenous, chronic myelogenous and/or T-cell leukemia [92]. Additionally, phosphorylation of Tyr127 was exhibited in lung and non-small cell lung cancer cell lines. Phosphorylation of residue Tyr143, which precedes the first residue of the second β-hairpin in the NRL, was also noted in the gastric carcinoma cell line MKN-45, as well as phosphorylation of residues T212 and T214, within the ALKBH3 active site, was found in liver cancer tissue samples [92]. Examination of potential ALKBH3 modification sites using predictive software shows possible post-translational modification sites for acetylation (Lys43, 116, 219, 220), and sumoylation (Lys57, 236), along with other possible phosphorylation sites (Ser32, 50, 187, 192, 208, 265; Thr29, 41; Tyr78, 127, 229) [90-93]. All of those possible post-translational modifications merit further consideration.

5. Biological significance of direct repair in mammalian cells

Normal cells depend on direct repair to eliminate damage that is possibly cytotoxic or mutagenic. Our knowledge of the biological significance of direct repair proteins in mammalian cells is based on the evaluation of effects on cell cytotoxicity, replication, transcription and subsequent mutagenic consequences observed in the absence of each protein of interest. Recent investigations performed in model system organisms, most prominently in mice, to assess the impact of the absence of Mgmt or Alkbh family proteins will be highlighted in this section. These studies also provide insight into the function and importance of direct repair proteins in humans.

5.1. Knock-out animal models

It is important to remember that a number of DNA repair systems are implicated in the elimination of DNA lesions formed by exposure to alkylating agents. Therefore, dysfunction of repair systems can lead to pathologies that include cancer development. However, without use of a model organism to assay the effects, the consequences to the organism as a whole cannot be assessed. Knock-out animal models are a valuable tool for understanding
the overall physiological effects of genes on an organism, and provide insight into disease research and therapeutic development.

Murine Mgmt models have been studied by multiple groups to evaluate sensitivity to alkylating agents commonly used in chemotherapeutics [5, 6, 82, 86, 136-139]. Though Mgmt repairs DNA damage that is known to be mutagenic, Mgmt-deficient mice surprisingly lack any overt phenotype. However, these mice are significantly more sensitive to treatment with N-methyl-N-nitrosourea (MNU), 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU), 1-(4-amino-2-methyl-5-pyrimidinyl)methyl-3-(2-chloroethyl)3-nitrosourea (ACNU), streptozotocin, temozolomide, and dacarbazine alkylating agents [5, 136, 137, 139-142]. Mgmt knock-out mice treated with various chemotherapeutic agents also show ablation of hematopoietic tissues at the stem cell level [38, 141, 143] and are prone to development of thymic lymphomas [144] and lung adenomas [82, 138, 144, 145]. Similarly, mouse embryonic stem (ES), embryonic fibroblasts (MEFs) and bone marrow cells deficient in Mgmt also exhibit a significant increase in sensitivity (~10-fold) to MNU and BCNU [83, 141, 146]. However, mice heterozygous for Mgmt do not display a significant reduction in survival following treatment with nitrosoureas or increased tumorigenesis, compared to their wild-type counterparts.

Although in vitro DNA repair activity has been established for ALKBH1, studies conducted in murine models lacking Alkbh1 suggest roles involved in transcription. Mice deficient in Alkbh1 exhibit apoptosis in adult testis, sex-ratio distortion and unilateral eye defects, as well as impaired differentiation of specific trophoblast lineages in the developing placenta [147, 148]. Though the specific activity and function of ALKBH1 remains to be determined, ALKBH1 biological roles seem linked to spermatogenesis and embryonic development.

On the other hand, Alkbh2- and/or Alkbh3-deficient murine models do not manifest any obvious phenotype or histopathological changes [116, 119, 132]. However, over time mice lacking Alkbh2 accumulate significant levels of 1-meA, confirming a role in removing endogenous DNA alkyl adducts. In a recent study, Alkbh2,Alkbh3,Aag knock-out mice (Aag also known as Mpg, a DNA glycosylase in the BER pathway) were viable, but underwent rapid death when exposed to a chemically-induced colitis treatment [119]. Similarly, primary mouse embryonic fibroblasts (MEFs) derived from mice lacking functional Alkbh2 exhibited significantly increased cytotoxicity and mutagenesis following exposure to the S\textsubscript{2} alkylating agent methyl methanesulfonate (MMS) [116, 118, 119]. Survival of Alkbh3-deficient MEFs exposed to MMS was reduced by ~50% compared to wild type MEF sensitivity, though mutant frequency did not significantly increase [116].

5.2. Replication and transcription defects

Though not all lesions generated by exposure to alkylating agents cause defects in replication and transcription, DNA and RNA adducts that are specifically removed via a direct repair mechanism interfere with replication and transcription machinery. The presence of O\textsuperscript{6}-meG in DNA impedes polymerization by DNA and RNA polymerases [31, 32, 149, 150]. Polymerase beta (β), involved in base excision repair (BER) of alklation adducts, is completely blocked by O\textsuperscript{6}-meG adducts [150]. Polymerase delta (δ) is able to replicate past, but insertion of the correct base opposite O\textsuperscript{6}-methylguanine is very
inefficient. However, these adducts can be bypassed using polymerase eta (η) [149], a member of the Y-family DNA translesion synthesis (TLS) polymerases, but TLS polymerases are notorious for being error-prone. Interestingly, when replicating past O^6^-meG DNA adducts, TLS polymerase, Polη is twice as efficient at inserting cytosines opposite O^6^-meG as replicative polymerase, Pol δ [32].

1-meA and 3-meC lesions that are repaired by Alkbh2 and Alkbh3 are at DNA base-pairing positions and hinder proper base insertion [101]. During replication, this can lead to arrest of nucleotide synthesis, resulting in replication fork collapse [151]. Similarly, 1-meA and 3-meC adducts can also cause stalling of transcription. Correspondingly, Alkbh2 co-localizes with replication foci during S-phase [111, 131, 133] and Alkbh3 has a role in removal of alkyl adducts from mRNA [1, 15, 108, 115, 131, 152]. However, a TLS polymerase that is linked to 1-meA and/or 3-meC DNA adduct bypass has not been identified.

5.3. Cell cytotoxicity

Treatment with alkylating agents introduces a variety of adducts into DNA and RNA (Figure 2, Table 1). In the absence of direct repair proteins, those lesions can lead to cell death or damage tolerance, which allows for cell survival, but can introduce mutations into the genome that could have detrimental effects [101, 116, 142, 153]. As exhibited in Mgmt- and Alkbh-deficient murine models, lack of direct repair proteins correlates with a significant increase in cell death following treatment with S_{1} or S_{2} alkylating agents, respectively [116, 118, 140, 141].

5.4. Mutagenesis

When a modified nucleoside can form at least two hydrogen bonds, transcription and replication templates and translation of messengers are active [13]. O^6^-meG, 1-meA, and 3-meC are all involved in DNA base-pairing. Modification at O^6^-meG and 3-meC still allow for formation of two hydrogen bonds, while 1-meA results in only a single hydrogen bond between paired bases [13]. However, the exocyclic amino group of 1-meA can rotate so that both amino group hydrogens can generate the necessary base-pairing bonds, though a slight distortion of the double-strand DNA helix does occur [13]. The addition of a methyl group to O^6^-G, N1-A, or N3-C interferes with normal replication, and could recruit DNA translesion synthesis (TLS) polymerases to bypass the DNA adducts. The size and organization of the Y-family TLS polymerase active sites is variable and allows for accommodation of numerous adducts. However, not only are TLS polymerases inherently error-prone [154, 155], the number and type of hydrogen bonds that can be made with the modified bases has been altered. Those factors can produce insertion of an erroneous base during bypass that accompanies replication or transcription.

O^6^-meG mutagenicity has been established in bacterial and mammalian systems [29, 30]. O^6^-meG is mutagenic and primarily gives rise to G:C→A:T mutations. A mis-insertion of thymine is thought to occur due to mis-identification of O^6^-meG as adenine, as hydrogen bonding can occur with the N1 and exocyclic amino group of O^6^-meG [13].
Unfortunately, studies evaluating the mutagenicity of a site-specific 1-meA, 3-meC, 1-meG, or 3-meT adducts have not been conducted in mammalian systems, but studies in *E. coli*, show that 1-meA adducts are only slightly mutagenic, whereas 3-meC, 1-meG, and 3-meT adducts are much more mutagenic [101]. Work evaluating the anti-mutagenic role of Alkbh2 and Alkbh3 in a murine model showed increased mutant frequency, specifically for mouse embryonic fibroblast (MEF) cells deficient in either Alkbh2 or Alkbh3 [116]. Those Alkbh-deficient cells exhibited increased amounts of C:G→A:T C:G→T:A mutations, respectively. Additionally, when treated with MMS, Alkbh2-deficient MEFs displayed an increased frequency of C:G→T:A and T:A→A:T mutations. Similarly, Alkbh3-deficient MEFs also exhibited an increased frequency of T:A→A:T mutations, as well as an increased frequency of A:T→G:C mutations, in response to MMS treatment. Like O\(^6\)-meG, misidentification of the modified DNA bases due to the presence of two sites for hydrogen bond formation could arise if 1-meG or 3-meC is recognized as thymine and an adenine is paired with the two remaining hydrogen bond acceptors. Furthermore, T:A→A:T mutations could arise if 3-meT becomes recognized as adenine and a thymine is paired via hydrogen bonds between thymine O\(^4\) and O\(^2\) and adenine N-3 exocyclic amino group nitrogen. It is likely that 1-meA is rarely mutagenic in *E. coli*, deficient in AlkB, because 1-meA can utilize the C6 exocyclic amine and N7 as an alternative binding site providing two sites for hydrogen bond formation with thymine N-7 and O\(^4\) molecules, using Hoogsteen base-pairing [156].

6. Medical significance of direct repair proteins in humans

Genetic and epigenetic controls that regulate *MGMT*, *ALKBH2*, and *ALKBH3* gene expression and influence how these proteins directly repair DNA are critical factors that can lead to a better understanding of cancer development. In addition, comprehension of factors that cause variations in the direct DNA repair activities of cancer cells will provide important progress toward formulating cancer therapeutics that target MGMT or ALKBH proteins. Understanding the impact of direct DNA repair proteins will eventually result in treatments that can be tailored to achieve better therapeutic results or to predict treatment and/or disease outcomes.

6.1. Epigenetic and transcriptional regulation

Epigenetic modifications are stable alterations of DNA that are heritable in the short term, but do not involve mutations of the DNA itself, and are mediated by DNA methylation and histone modifications. The stable alterations that are involved in epigenetics have a major role in exerting control on gene expression. Endogenous cell signaling as well as external influences, including diet and other lifestyle choices, can alter gene expression mediated by changes in epigenetic modifications [157, 158]. Methylation of cytosines at transcription factor recognition sites can interfere with binding and/or function and repress transcription of that gene [159, 160]. Alternatively, protein recruitment that binds methyl CpG islands can block transcription machinery or alter chromatin structure [161, 162]. Transcriptional silencing also is connected to histone deacetylation [163, 164]. Methyl CpG binding domain
(MBD) family proteins direct histone deacetylases to remove acetyl groups from lysines in the amino terminal histone tails, stabilizing DNA-histone interactions, and condensing chromatin so that transcription factor binding sites are inaccessible.

Though unmethylated in normal cells, transcriptional silencing of MGMT, associated with promoter CpG island methylation has been reported in a variety of cancer cell types and MGMT-deficient cell lines [82, 138]. Additionally, in a glioma mouse model a subpopulation of glioma cells with stem cell properties were identified [165] that are capable of re-establishing tumor growth following temozolomide treatment. Although Mgmt promoter CpG methylation or protein levels were not determined in that study, when MGMT transcript levels were evaluated in glioma patients [166], those with MGMT CpG promoter methylation had increased response to temozolomide, but also maintained a subset of glioma cells with stem cell-like character and MGMT promoter methylation. Interestingly, mRNA levels of DNMT1 and DNMT3b methyltransferases are increased in a number of human glioma patients, but there does not appear to be a link to MGMT expression levels [167]. Moreover, MGMT promoter CpG methylation levels and DNA methyltransferase levels alone do not account for patient response to alkylating agent therapy. However, whether MGMT promoter methylation disables transcription factor binding or contributes to chromatin reorganization remains uncertain [71, 72, 74]. Therefore, regulation of MGMT expression is still unclear and merits intense scrutiny.

The inability to establish direct connections among MGMT expression, CpG methylation, and response to alkylating agent therapy indicates that other mechanisms contribute in regulating MGMT levels. Studies evaluating MGMT expression and microRNAs in patient samples have established a modest inverse correlation between the levels of MGMT transcript and miR-181d [168]. Moreover, expression of mi-181d in A1207 glioblastoma cells, results in abnormal sensitivity to temozolomide. However, expression of MGMT cDNA, restores the survival to levels close to that of the A1207 parental line. These results suggest that identification of other miRNAs involved in regulating MGMT expression will help elucidate the mechanisms that control the gene transcript levels.

In addition to control at the DNA and transcript levels, histone modifications can also control the epigenetic state and direct expression. Acetylated histone H3 and H4 levels also increase in cell lines expressing MGMT, compared to cell lines deficient in MGMT [169], which would facilitate nucleosomal positioning that enables transcription factor interactions. Further, binding of MBD proteins in the MGMT promoter of was greater in MGMT-silenced cells, implicating MBD proteins in recruitment of histone deacetylases that remove lysine acetylation from the amino-terminal tails of histones H3 and H4, resulting in more condensed chromatin and transcription inactivation [73, 79, 170]. Therefore, epigenetic and/or enzymatic CpG island methylation at the MGMT promoter influences transcription factor access, as well as chromatin structure that are important for MGMT expression.

ALKBH2 and ALKBH3 both have CpG islands in their promoters, but epigenetic regulation and/or gene silencing has not been reported for either homolog. However, mutations that alter protein expression have been observed [171], but it is likely that methylation of CpG islands near any of the seven transcription factor binding sites in the promoter of ALKBH2
or the single transcription factor binding sites within the promoter region of ALKBH3, would repress transcription factor binding and possibly gene expression. Because data on the function of ALKBH promoters are less abundant compared to those available for the MGMT promoter, examination of the promoter function for those genes is an area that would benefit from further investigation.

6.2. Links to cancer

Dysregulation of numerous DNA repair pathways are involved in tumor development, progression, diagnosis, treatment and prognosis, including direct DNA repair proteins [82, 159, 172-179]. Over-expression of direct repair proteins is generally associated with a protective effect against cell death that would otherwise be induced by alkylating agent treatment. However, down-regulation or silencing of direct repair protein expression is associated with increased mutagenesis that precedes tumorgenesis. Therefore expression profiles could be used to predict potential resistance or enhanced sensitivity to therapeutics.

MGMT has been implicated in many types of human tumors. Numerous MGMT polymorphisms have risk associations with breast, lung, colon, and head and neck cancers [63, 82, 180-186]. Decreased MGMT expression is also found in glioma, lymphoma, retinoblastoma, breast (including triple-negative breast cancer) and prostate cancer [82, 138, 187] [188]. Moreover, lack of MGMT is associated with enhanced outcomes using alkylating agent therapies [5, 62, 67, 82, 86, 139, 180, 181, 183, 189]. Though MGMT silencing occurs in a variety of tumor types, increased levels have also been observed in non-Hodgkin lymphoma, myeloma and glioma, as well as in some colon, pancreatic, breast, and lung cancers [63, 183, 184].

Mutations in ALKBH2 and 3 have been associated with an enhanced expression of these proteins in glioma cells and pediatric brain tumors [171, 190]. Similarly, over-expression of ALKBH3 has been associated with human rectal carcinoma [191] and prostate cancer, as well as, lung adenocarcinoma and non-small-cell lung cancer [134] [192]. On the contrary, down regulation of ALKBH2 has been observed in gastric cancer, promoting growth of gastric cancer cells [193]. Although down regulation of ALKBH2 in gastric cancer cells caused increased proliferation, ALKBH2 silencing in H1299 lung cancer cells had the opposite effect, increasing cisplatin sensitivity. Similarly, ALKBH3 silencing induced senescence and sensitivity to alkylating agents in human adenocarcinoma and prostate cancer cells [134, 193]. Therefore, further study of the role of ALKBH2 and 3 in both normal and tumor cells is necessary to elucidate their biological role(s).

6.3. Therapeutic targets

Understanding the mechanism of proteins involved in various DNA repair pathways is crucial for developing new chemotherapeutic targets and eventually new drugs. DNA alkylating agents and ionizing radiation (IR) are often used as chemotherapeutic treatments because of ability to control the dose administered and area of treatment, as well as the major cytotoxic effects of both agents at high doses. However, in addition to generation of cyto-
toxic adducts that cause apoptosis, alkylating agents and IR also form adducts that can be mutagenic and as a result can cause initiation of secondary cancers. Although DNA repair deficiencies are associated with increased cancer risk and formation, cancer cells proficient in DNA repair can reduce therapeutic efficacy. Currently, combination cancer treatment regimens are being explored that utilize chemotherapy or IR and target specific DNA repair proteins with pharmacological agents to enhance treatment efficacy and eliminate resistance to treatment regimens exhibited in some patients [189].

6.3.1. MGMT

Chemotherapeutic drugs such as temozolamide (TMZ) and bis-(2-chloroethyl)-nitrosourea (BCNU) generate some lesions repaired via the direct methyltransferase mechanism. Combination treatment with MGMT inhibitors prevents repair and resistance to methylating and chloroethylating agents [1, 38, 137] and has also been shown to reverse cisplatin drug resistance [194].

Understanding cellular regulation of MGMT expression will allow for selective down regulation and sensitization of tumors to alkylating agent chemotherapies. Studies have evaluated manipulation of MGMT expression and protein levels. Initial experiments evaluating MGMT inhibitors identified O⁶-benzyl guanine (BG) as an efficacious inhibitor of MGMT activity, a single, micromolar dose depleting greater than 99% of MGMT activity in human cells for 24-hours following drug removal [195]. Moreover, treatment with BG lacks any mutagenic or cytotoxic effects [195-197]. Clinical trials combining BG and BCNU treatment have been conducted in colon cancer, sarcoma, melanoma and myeloma, as well as studies evaluating combination of BG and TMZ [138]. Since synthesis of BG, additional BG-like inhibitors have been developed [196], including O⁶-(4-bromothenyl) guanine, which has been evaluated in patients with glioma [187]. Similarly, targeting of MGMT along with combination of platinum drugs, including cis- and carboplatinum [198], as well as topoisomerase I inhibitors has been investigated in various clinical trials [86].

Another approach to regulate MGMT that holds great, essentially untapped therapeutic potential is strategies utilizing RNA interference-mediated gene silencing to target MGMT [168, 199, 200]. For instance, if anti-sense molecules can specifically target MGMT mRNA translation, and degradation is also inhibited, depletion of MGMT is sustainable for long periods of time [62]. As seen in glioblastoma patients, expression levels of various miRNA markers correlate with prognosis [168, 199, 200]. Therefore, one potential new treatment could use miRNAs, such as miR-181d, to decrease MGMT levels, thus increasing sensitivity to alkylating agents [168]. Similarly, targeting regions of the MGMT promoter that is accessible to transcription factors could interfere with binding and down-regulate MGMT transcription. However, non-specific targeting of MGMT inhibitors in all cells increases chemotherapeutic toxicity. Therefore, mutant forms of MGMT that are resistant to BG-like inhibitors are also being evaluated to limit myelosuppression, affording hematopoietic progenitor cells protection from BG and BCNU or temozolomide treatment [201-204].
6.3.2. Alkbh homologs

Similar to MGMT, the role of ALKBH2 and ALKBH3 in repair of DNA alkylation damage at base-pairing sites is anti-carcinogenic. However, investigations indicate that over-expression of ALKBH proteins in various cancer cell lines shields those cells against methylating agent toxicity and would thereby protect against some chemotherapeutic treatments [134, 171, 192]. Additionally, because loss of ALKBH2 and/or ALKBH3 leads to disruption of replication, inhibition of ALKBH2 and/or ALKBH3 is a strong target for the development of novel chemotherapeutic agents. Some specific inhibitors of these proteins have already been identified [135, 205, 206], as well as generic α-KG/dioxygenase inhibitors including dimethyl oxalylglycine (DMOG) and α-ketoglutarate derivatives such as oxoglutarate. Studies have addressed the application of DNA aptamers as inhibitors of ALKBH proteins [207]. However, to date no studies have been conducted in mammalian models that evaluate the combination of ALKBH inhibitors with chemotherapeutic alkylating agents.

7. Summary

Direct repair proteins represent a unique class of enzymes that remove DNA damage without a dependence on DNA synthesis. In the future, better comprehension of how these proteins function and are produced in cells will lead to understanding their roles in formation of mutations that cause cancer. Eventually, that knowledge will foster the development of drugs to target these proteins and/or to regulate their expression to improve patient outcomes.

Author details

Stephanie L. Nay1,2 and Timothy R. O’Connor2

1 Irell and Manella Graduate School of Biological Sciences, USA
2 Department of Cancer Biology, Beckman Research Institute, Duarte, CA, USA

References


[127] Songe-Moller L, van den Born E, Leihne V, Vagbo CB, Kristoffersen T, Krokan HE, et al. Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the


Roos WP, Christmann M, Fraser ST, Kaina B. Mouse embryonic stem cells are hypersensitive to apoptosis triggered by the DNA damage O(6)-methylguanine due to high E2F1 regulated mismatch repair. Cell Death Differ. 2007;14(8):1422-32.


