Published online 3 February 2016

Nucleic Acids Research, 2016, Vol. 44, No. 8 ***3713–3727***
doi: 10.1093/nar/gkw059

An interplay of the base excision repair and mismatch repair pathways in active DNA demethylation

Inga Grin[1,2,3,4] and Alexander A. Ishchenko[1,2,*]

[1]Laboratoire ≪Stabilité Génétique et Oncogenèse≫ CNRS, UMR 8200, Univ. Paris-Sud, Université Paris-Saclay, Equipe Labellisée Ligue Contre le Cancer, F-94805 Villejuif, France, [2]Gustave Roussy Cancer Campus, F-94805 Villejuif, France, [3]SB RAS Institute of Chemical Biology and Fundamental Medicine, 8 Lavrentieva Avenue, Novosibirsk 630090, Russia and [4]Department of Natural Sciences, Novosibirsk State University, 2 Pirogova Street, Novosibirsk 630090, Russia

Received June 8, 2015; Revised January 20, 2016; Accepted January 22, 2016

ABSTRACT

Active DNA demethylation (ADDM) in mammals occurs via hydroxylation of 5-methylcytosine (5mC) by TET and/or deamination by AID/APOBEC family enzymes. The resulting 5mC derivatives are removed through the base excision repair (BER) pathway. At present, it is unclear how the cell manages to eliminate closely spaced 5mC residues whilst avoiding generation of toxic BER intermediates and whether alternative DNA repair pathways participate in ADDM. It has been shown that non-canonical DNA mismatch repair (ncMMR) can remove both alkylated and oxidized nucleotides from DNA. Here, a phagemid DNA containing oxidative base lesions and methylated sites are used to examine the involvement of various DNA repair pathways in ADDM in murine and human cell-free extracts. We demonstrate that, in addition to short-patch BER, 5-hydroxymethyluracil and uracil mispaired with guanine can be processed by ncMMR and long-patch BER with concomitant removal of distant 5mC residues. Furthermore, the presence of multiple mispairs in the same MMR nick/mismatch recognition region together with BER-mediated nick formation promotes proficient ncMMR resulting in the reactivation of an epigenetically silenced reporter gene in murine cells. These findings suggest cooperation between BER and ncMMR in the removal of multiple mismatches that might occur in mammalian cells during ADDM.

INTRODUCTION

Erasure of DNA methylation is essential for reprogramming and for maintenance of pluripotency during embryonic development (1). DNA demethylation (DDM) can occur either passively (via inhibition of methylation in newly synthesized DNA after DNA replication) or via an active process where 5mC residues are removed from non-replicating DNA by enzymatic action. The rapid loss of DNA methylation within a single cell cycle is suggestive of the existence of mechanisms that can actively modify and/or eliminate 5mC residues from DNA. Recent advances in the research on active DDM (ADDM) in mammals revealed several key players. The ten-eleven translocation (TET) family of proteins (TET1, TET2 and TET3) are Fe^{2+}-dependent and 2-oxoglutarate-dependent dioxygenases that catalyze conversion of 5mC in DNA to 5-hydroxymethylcytosine (5hmC) (2,3). Furthermore, the TET enzymes were shown to further oxidize 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (4,5). Subsequently, the 5fC and 5caC residues are removed from DNA by thymine-DNA glycosylase (TDG) in the base excision repair (BER) pathway, suggesting that BER is directly involved in the next step of ADDM (4,6). Accordingly, it was shown that the *Tdg* knockout is embryonically lethal in mice due to *de novo* DNA methylation that affects the expression of developmental genes (7).

In zebrafish, an alternative mechanism was proposed that is based on deamination of 5mC to thymine; this reaction is catalyzed by activation-induced deaminase (AID or AICDA). This enzymatic conversion results in the mismatched base pair G•T, which in turn is repaired by a *methyl-CpG binding domain protein 4* (MBD4) (8). It should be noted that deamination of 5mC may be carried out not only by proteins of the AID/APOBEC family but also by the Dnmt3a and Dnmt3b DNA methyltransferases in the absence of the methyl donor S-adenosylmethionine (9). There is a concern about reproducibility of these results in zebrafish (10), AID-deficient mice and humans, having no obvious developmental defects (11,12); nevertheless, in mice, changes in the litter size and abnormal DDM in germ cells are suggestive of a possible role of AID in the first stages of development (13). Moreover, AID-deficient cells show inefficient formation of inducible pluripotent cells (14)

*To whom correspondence should be addressed. Tel: +33 142 115 405; Fax: +33 142 115 008; Email: alexander.ishchenko@gustaveroussy.fr

and a lack of demethylation of the *OCT4* and *NANOG* promoters during reprogramming of heterokaryons (mouse embryonic stem cells fused with human fibroblasts) toward pluripotency (15). Furthermore, overexpression of AID/APOBEC deaminases in neural cells promotes the removal of 5hmC residues from DNA, and simultaneous overexpression of TET1 and AID leads to global accumulation of 5-hydroxymethyluracil (5hmU), an oxidation product of thymine or a deamination product of 5hmC (16). Conversion of 5hmC to 5hmU gives rise to 5hmU•G mismatches, which are in turn excised by the TDG, MBD4, single-strand selective monofunctional uracil-DNA glycosylase 1 (SMUG1) or much less efficiently by endonuclease VIII-like glycosylase 1 (NEIL1) (7,16,17). At present, the direct involvement of AID/APOBECs in the formation of 5hmU *in vivo* via coupled oxidation-deamination of 5mC residues in DNA (7,16) remains unclear because the purified AID/APOBECs proteins failed to deaminate 5hmC-containing DNA substrates *in vitro* (18). Human cell-free extracts drive very efficient repair of 5hmU•G mispairs, 60 times faster than that of 5hmU•A base pairs (19); this finding is suggestive of possible physiological relevance of the processes that generate 5hmU residues in methylated DNA.

Recently, it was proposed that cytidine deaminase-induced damage can lead to indirect DDM in the context of the long-patch DNA repair pathways, such as the long-patch BER and mismatch repair (MMR) (20,21). Classical MMR is a replication-linked 'long-patch' repair process implying mismatch recognition that is mediated primarily by the MSH2/MSH6 (MutSα) complex, recruitment of proliferating cell nuclear antigen (PCNA) and MLH1/PMS2 (MutLα) nicking endonuclease at a preexisting strand discontinuity in the nascent strand, excision of this strand by EXO1 exonuclease up to 150 nucleotides (nt) past the mispair, and resynthesis of the excised tract (up to 1 Kb) by a high-fidelity replicative DNA polymerase (22). In contrast, so-called non-canonical MMR (ncMMR) is an S phase-independent process where the MMR pathway is activated by various DNA lesions (uracil, O6-methylguanine, 5-hydroxyuracil and probably others) rather than by regular base mispairs (23,24). The ncMMR lacks strand directionality, triggers PCNA monoubiquitination and promotes recruitment of the error-prone polymerase η to chromatin (23,24). It was suggested that ncMMR is required for AID-dependent somatic hypermutation (SHM) of the immunoglobulin (Ig) genes during antibody maturation in B cells (23). In the long-patch BER and ncMMR repair pathways, 5mC does not have to be targeted for deamination but can be replaced by an unmethylated cytosine in the process of repair of the neighboring lesions. In order to maintain genome integrity, DDM via generation of the highly promutagenic dU•G, T•G or 5hmU•G intermediates would require high efficiency and accuracy to coordinate deamination and DNA repair processes, especially in densely methylated DNA regions. Molecular mechanisms of the coordination of AID-mediated DDM and the role of long-patch repair in this process remain unknown.

Here, using a specially constructed circular heteroduplex DNA substrate, we reconstituted the removal of 5mC and its various modifications in cell-free extracts of mouse embryonic fibroblasts (MEF) and human Burkitt's lymphoma (BL2) cells. We demonstrate that a significant fraction of deaminated cytosines is removed together with a distant methylation mark in the ncMMR-dependent pathway. Despite the competition between DNA glycosylases and MMR targeting the same base mispairs, efficient BER promotes MMR-dependent removal of multiple distant or adjacent mispairs via generation of a DNA strand nick as a strand discrimination signal for initiation of MMR. The possible roles of the MMR pathway and cytidine deaminases in ADDM are discussed toward the end of the 'Discussion' section.

MATERIALS AND METHODS

Reagents and proteins

Restriction enzymes, T4 DNA polymerase, T4 DNA ligase, hSMUG1 and NEBuffers were purchased from New England Biolabs (Evry, France). The purified DNA glycosylases were from the laboratory stock. All oligonucleotides containing modified bases and their complementary strands were purchased from Eurogentec (Seraing, Belgium). Sequences of the oligonucleotides that were used for construction of plasmids in the present work are shown in Supplementary Tables S1 and S2.

The cell line, culture conditions and preparation of cell extracts

MEF wild-type (WT) and *Msh2*$^{-/-}$ cell lines were from the laboratory stock; the *Tdg*$^{-/-}$ MEF cell line was kindly provided by P. Schar (University of Basel, Switzerland). Type I Burkitt's lymphoma BL2 cells were kindly provided by Saïd Aoufouchi's group (UMR 8200, Villejuif, France). The BL2 cell line carries the variant *cMYC-IgL* translocation of Burkitt's lymphomas and has a deficient p53 pathway due to overexpression of its inhibitor, Mdm2 (25). WT, *Tdg*$^{-/-}$ and *Msh2*$^{-/-}$ MEFs were grown in DMEM (Dulbecco's modified Eagle's medium) supplemented with sodium pyruvate, penicillin/streptomycin (Invitrogen, Carlsbad, CA) and 10% fetal calf serum (FCS) in a humidified atmosphere containing 5% of CO_2. Type I Burkitt's lymphoma BL2 cells were cultured in RPMI 1640 supplemented with GlutaMAX (Invitrogen, Carlsbad, CA), penicillin/streptomycin and 10% FCS.

Cell-free MEF and BL2 extracts were prepared as described previously, with minor modifications (24). All procedures were carried out at 4°C. A pellet of 3×10^9 cells was washed twice in cold phosphate buffered saline, followed by one wash in a hypotonic buffer (1 mM dithiothreitol [DTT], 1.5 mM $MgCl_2$, 20 mM HEPES-KOH pH 7.9, 5 mM KCl, cOmplete protease inhibitors [Roche]) containing 250 mM sucrose. The cells were resuspended in 0.54 volumes of the hypotonic buffer without sucrose, were allowed to swell for 10 min on ice, and then were lysed by 35 strokes of a tight-fitting Dounce homogenizer. After that, 0.039 volumes of 3 M KCl were added to the suspension of nuclei, and the resulting lysate was centrifuged at 15 000 × *g* for 20 min. Protein concentration in the supernatant was determined by the Bradford assay (was usually ∼15 mg/ml), and the aliquots were stored at −80°C.

Construction of circular heteroduplex DNA substrates

Circular DNA substrates were prepared as we described previously (26) with minor changes. Briefly, heteroduplex DNA substrates containing a 5mC•G, 5hmC•G, 5caC•G, 5hmU•G, dU•G or a G•T mismatch within different restriction sites of the pMM2 plasmid (a pBluescript II SK[+] derivative) were constructed by primer extension, using a lesion-containing oligonucleotide (Supplementary Table S1) as the primer and single-stranded phagemid DNA as a template. The strand discrimination signal (8-nt gap) 71 bp upstream of the methylated dC in the hemimethylated SmaI site was introduced by incubation with Nt.BbvCI and Nb.BsmI nicking endonucleases (Figure 1). Nicked forms of Mult-X-pMM2 plasmids (where X = 5hmU[1–6] or dU5[6]; Supplementary Table S1) were obtained by omission of DNA ligation during the primer extension reaction. The desired covalently closed circular (ccc), nicked or gapped circular heteroduplex substrates were purified on a 0.8% agarose gel and by means of the MinElute Gel Extraction Kit (Qiagen). Contamination with the nicked form in ccc DNA did not exceed 5% of the total DNA substrate purified (Supplementary Figure S1). After measurement of the concentration, the plasmids were stored in aliquots at −80°C.

Construction of luciferase reporter plasmids

We used plasmids containing either the firefly luciferase reporter gene (pCMVGL3 and pCMVIG) or the *Renilla* luciferase reporter gene (pRL-CMV), which were expressed under control of the cytomegalovirus (CMV) promoter. The pCMVGL3 and pRL-CMV plasmids are a gift from A. Sarasin (27). These plasmids can replicate in bacteria but not in mammalian cells. The CMV promoter of pCMVGL3 was modified by insertion of two BbvCI restriction sites and deletion of a 227-bp DNA fragment at the beginning of the CMV promoter using the standard site-directed mutagenesis approach with a pair of complementary primers: 5′-CGGTTTTGGCAGTACATCAATGGGCTGAGGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGT CTCCACCCCATTGACGTCAATGGGAGCTGAGG TTTGTTTTGGCACCAAAATCAACG-3′ and 5′-CGATGTACGGGCCAGATATACGCGTTGTGGCA GTACATCAAGTGTATCATATGC-3′ for two simultaneous insertions and a deletion, respectively. The resulting pCMVIG plasmid was used for preparation of single-stranded phagemid DNA, for annealing with the modified oligonucleotides and for subsequent primer extension and purification as described above for the pMM2 plasmid, with minor changes. In order to attain complete modification of the truncated CMV promoter region, two or more primers containing different modifications were used in the primer extension reaction (Supplementary Table S2). The list of the resulting pCMVIG-derived plasmids is presented in Supplementary Table S3. Levels of hemimethylation (≥95%) of the resulting ccc plasmids were controlled by digestion with the ZraI restriction enzyme.

The BER and MMR reconstitution assay based on plasmid DNA and cell extracts

The BER/MMR reconstitution assays were performed using the following reaction mixture (total volume 20 μl): 20 mM HEPES-KOH, pH 7.6; 105 mM KCl; 5 mM $MgCl_2$; 1 mM glutathione; 50 μg/ml BSA; dNTPs (0.1 mM each); 1.5 mM ATP; 24 ng of pMM2- or 48 ng of pCMVIG-derived substrates; and 3 mg/ml cell-free extracts from MEF WT, $Tdg^{-/-}$, $Msh2^{-/-}$ cells or 1 mg/ml cell-free extracts from BL2 cells. After 30 min of incubation at 37°C, the reactions were terminated by the addition 35 μl of 250 μg/ml RNase A in NEBuffer 3 and by incubation for 40 min at 37°C, followed by addition of 5 μl of the sodium dodecyl sulphate (SDS) stop solution (5% SDS, 125 mM ethylenediaminetetraacetic acid (EDTA), 0.56 μg/ml Proteinase K). In negative and positive control experiments (marked as STOP in the figures), a modified or unmodified plasmid DNA substrate, respectively, was added after the RNase A treatment and addition of the SDS stop solution. Next, the reaction mixture was incubated for 10 min at 37°C and 40 min at 50°C. After phenol extraction with 60 μl of phenol, the DNA was purified using the Qiagen MinElute Reaction Cleanup Kit.

In order to assess site-specific base lesion repair or 5mC erasure, the following modification-discriminating restriction enzymes were used: SmaI and ZraI for 5mC erasure; BbvCI, BsmI, XhoI and XbaI for 5hmU•G repair to C•G; EcoRI for dU•G, 5caC•G or 5hmU•G repair to C•G and HindIII for 5hmU•G repair to C•G or G•T repair to A•T, depending on the type and position of the modification on a plasmid substrate (see Figure 1 and Supplementary Table S1). Restriction digestion reactions were performed in NEBuffer 4 with 500 U/ml AlwNI for linearization of a pMM2-derived plasmid or 500 U/ml BamHI for linearization of a pCMVIG-derived plasmid in the presence of 50 μg/ml RNase A and one appropriate DNA modification-discriminating the restriction enzyme: BbvCI (50 U/ml), BsmI (75 U/ml), XhoI (75 U/ml), SmaI (75 U/ml), HindIII (120 U/ml), EcoRI (100 or 50 U/ml in the case of mC-14-dU-pMM2), XbaI (100 U/ml) or ZraI (250 U/ml) for 40 min at 37°C. Under the reaction conditions used, each restriction enzyme efficiently digested at the corresponding modification site only after the modification had been corrected (Supplementary Figure S1). Incision of the plasmid substrates by the restriction enzymes was monitored on a 0.8% UltraPure Agarose (Invitrogen) gel in Tris-acetate EDTA buffer (Millipore) stained with GelRed Nucleic Acid Gel Stain (Biotium, USA) and analyzed on a Syngene Imaging system (Syngene, UK). The efficiency of the incision (E_{exp}) was calculated assuming that the intensities of the upper (2078-bp for SmaI) and the lower (865-bp) product bands were respectively 1.42- and 3.4-fold lower than the intensity of the 2943-bp substrate band at the same molarity. In many cases the intensity of the 865-bp band was low against the background, thus we used only the more prominent upper band, multiplying its intensity by 1.42 (for SmaI) to calculate the percentage of substrate cleavage. To minimize the background effect, we used the 'rolling ball' background subtraction in ImageQuant TL (GE Healthcare) image analysis software, which gave us the most re-

liable and reproducible results in calibration experiments. Efficiency of DNA damage repair (E_r) and 5mC erasure at a particular restriction site was calculated by subtracting non-specific background activity of the corresponding restriction enzyme in negative control experiments (E_{neg}) and by normalization to the maximum of restriction enzyme activity on unmodified pMM2 plasmid in positive control (E_{pos}) experiments, according to Equation (1). E_{neg} and E_{pos} were measured in the control STOP experiments for each type of substrates and extracts used (Supplementary Figure S1).

$$E_r = (E_{exp} - E_{neg})/(E_{pos} - E_{neg}) \cdot 100\% \quad (1)$$

DNA repair assays on duplex oligonucleotides

The oligonucleotides 5′-TGACTGCATAXGCATGTA GACGATGTGCAT-3′ (where X is either 5caC, alpha-anomeric 2′-deoxyadenosine (αdA), tetrahydrofuranyl (THF) or 5hmU) were 5′-end-labeled by T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP (3000 Ci/mmol, Perkin Elmer Life Science Research) and then annealed with the corresponding complementary strands as described previously (17). The resulting duplex oligonucleotides were named as X/Y-RT, where X is a modified residue and Y is an opposite base in the complementary strand, respectively. The standard reaction mixture (20 μl) for DNA repair assays contained 20 nM ^{32}P-labeled duplex oligonucleotide; 20 mM Tris–HCl (pH 7.6); 50 mM KCl; 1 mM EDTA; 1 mM DTT; 100 μg/ml BSA; and either 0.3 mg/ml MEF cell-free extracts or 50 nM hTDG, or 25 U/ml hSMUG1, unless stated otherwise. For each experiment, we used two WT cell lines and a mix of three or more batches of the independently prepared extracts. After incubation at 37°C for 30 min, the samples were treated either with 0.1 M NaOH for 3 min at 99°C and then neutralized by 0.1 M HCl in order to cleave at AP sites left after excision of damaged bases. Further, 5 μl of a Proteinase K solution, containing 0.5% SDS, 10 mM EDTA and 0.7 mg/ml Proteinase K (Sigma-Aldrich), was added and the reaction mixtures were incubated for 10 min at 50°C. To analyze reaction products, the samples were desalted using hand-made spindown columns filled with Sephadex G25 (Amersham Biosciences) equilibrated in 7.5M urea and the cleavage fragments were separated by electrophoresis in denaturing 20% (w/v) polyacrylamide gels (7 M urea, 0.5× TBE, 42°C). The gels were exposed to a Fuji FLA-3000 Phosphor Screen, then scanned with Typhoon FLA-9500, and analyzed using Image Gauge V4.0 software.

DNA glycosylase activity assays on covalently closed dumbbell-shaped oligonucleotide (DSO) substrates

The 22-mer oligonucleotides 5′-CACTTCGGAXGGTG ACTGATCC-3′, where X is either 5caC or A, were 5′-[^{32}P]-labeled and annealed to the corresponding complementary 50-mer oligonucleotide 5′-GGTACGTTCGTACCGGAT CAGTCACAYTCCGAAGTGGCCGTCTTGACGGC-3′, where Y is an opposite base (G or T) to X in the 22-mer oligonucleotide as described above. Oligonucleotide samples were ligated in the presence of 7 units/μl of T4 DNA ligase (New England BioLabs) at 16°C for 4 h, under the standard manufacturer's conditions. The resulting dumbbell-shaped oligonucleotides (DSOs) were desalted using Sephadex G25 columns equilibrated in water and stored at −20°C. The standard reaction mixture for the DNA glycosylase activity assay on DSO substrates (10 μl) contained 10 nM of [^{32}P]-labeled DSO, 12 mM HEPES-KOH (pH 7.6), 80 mM KCl, 0.5 mM EDTA, 1 mM DTT, 43 μg/ml BSA, 3% glycerol, cOmplete protease inhibitors (Roche) and either 3 mg/ml MEF cell-free extracts or 400 nM hTDG. All assays were performed at 37°C for 30 min, unless specified otherwise. Reactions were stopped by adding 1 μl of a SDS stop solution (5% SDS, 125 mM EDTA, 0.56 μg/μl Proteinase K) and the mixtures were incubated for 10 min at 37°C and then for 40 min at 50°C. After phenol:chloroform (1:1) extraction of DNA products of the treatment they were precipitated with 100 μl of 2% $LiClO_4$ in acetone and washed twice in 70% EtOH. Restriction digestion reactions (20 μl) were performed in NEBuffer 2.1 with 500 U/ml HaeIII for 60 min at 37°C. The resulting [^{32}P]-labeled reaction products were analyzed as described above.

The reporter assay for demethylation of the CMV promoter

For an *ex vivo* transient transfection assay, WT and *Msh2*$^{-/-}$ MEF cells were seeded in 12-well plates (50 000/well) and one day later were cotransfected with 75 ng of pCMVIG-derived plasmids and 75 ng of the pRL-CMV plasmid, which was used as an internal control of transfection efficiency, using the ExGen 500 reagent (Euromedex, Souffelwryersheim, France). Dual-reporter assays were performed 18 h after the transfection using the Dual-Glo Luciferase Assay System (Promega, France). Firefly and *Renilla* luminescence were measured in a VICTOR3 V Multilabel Counter (Perkin Elmer) and the data were processed in the Wallac 1420 Workstation software. In all the experiments, the transfection and luciferase assays were performed in triplicate. The firefly luciferase activity was normalized to the *Renilla* luciferase activity; for each pCMVIG-based construct, this relative luciferase activity was normalized to the firefly/*Renilla* ratio obtained with the unmodified pCMVIG plasmid. The final ratio was expressed as a fold change.

RESULTS

NcMMR is directly involved in the repair of dU●G but not of 5caC●G or 5hmC●G pairs in DNA

We aimed to discriminate the DNA repair pathways involved in processing of the derivatives of C and 5mC residues in DNA and to measure the efficiency of concomitant removal of adjacent 5mC residues. For these purposes, we constructed circular plasmid DNA substrates containing various base lesions within the recognition sites of restriction endonucleases downstream and upstream of the 5mC residue that was placed in the SmaI sequence and followed their processing by cell-free extracts from DNA repair-deficient cell lines (see 'Materials and Methods' section and Figure 1). Repair of the modified bases in the

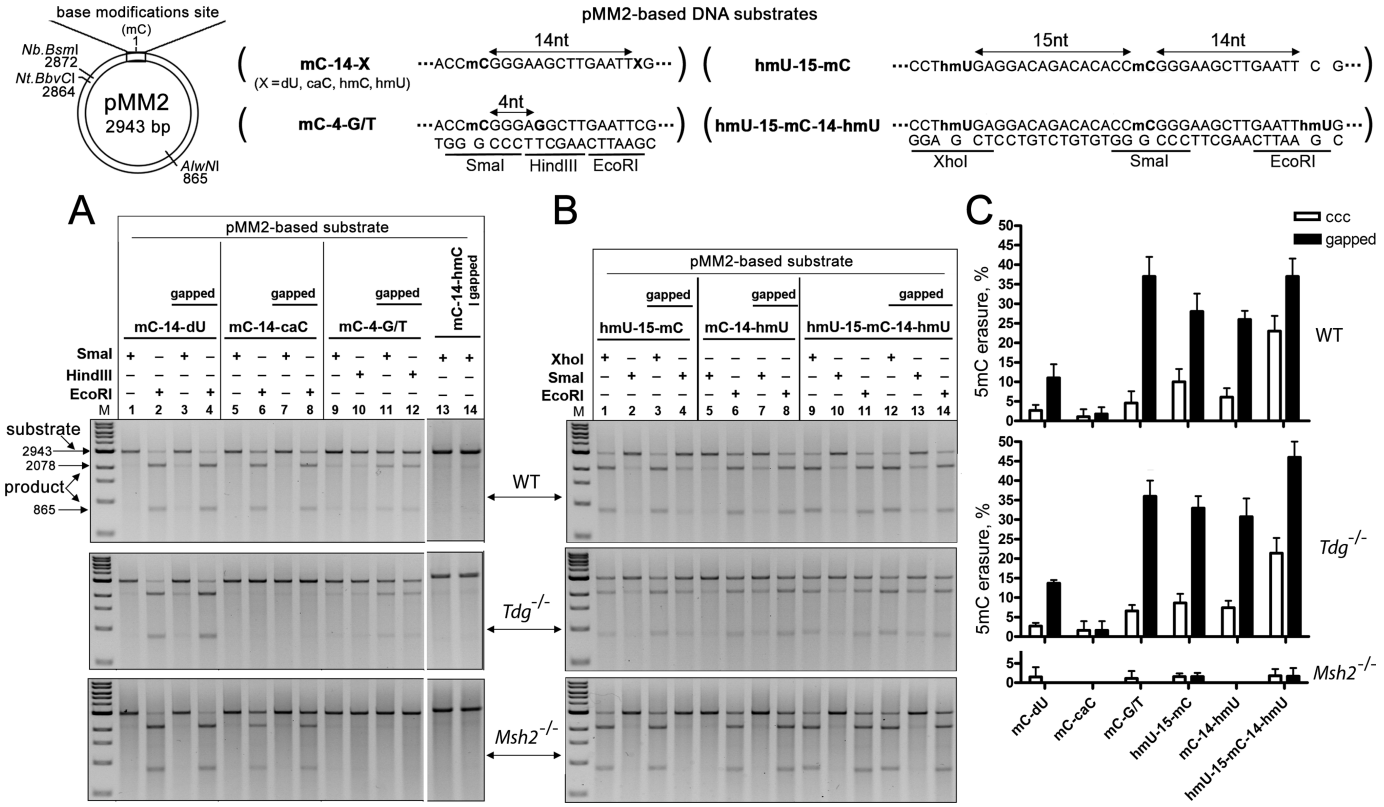

Figure 1. Processing of the oxidation/deamination products of deoxycytosine and 5mC by BER and ncMMR pathways in cell extracts from MEF cells. (**A**) *In vitro* reconstitution of repair of plasmids containing stand-alone dU•G, 5caC•G, G•T or 5hmC•G sites. (**B**) *In vitro* reconstitution of repair of plasmids containing 5hmU•G mispairs. (**C**) Comparison of demethylation efficiency at the hemimethylated SmaI reporter site in the experiments presented in panels A and B. The error bars represent the standard deviation (SD) ($n \geq 3$).

in vitro reconstitution assay was assessed by means of restriction endonucleases. Each restriction enzyme efficiently digests DNA at the corresponding modification site only after the modification has been corrected (Supplementary Figure S1). Hence, the cleavage level of the plasmids by SmaI or other modification-discriminating restriction enzymes reflects the level of 5mC erasure or corresponding DNA base repair, respectively. This approach allowed us to examine the length of repair patches and the roles of the BER and MMR pathways in both the removal of DNA base damage and in ADDM. It should be noted that a single-stranded gap (5′) upstream of the 5mC and damaged sites was introduced by specific restriction enzymes Nb.BsmI and Nt.BbvCI in order to ensure functionality of the MMR machinery when it was acting upon a mismatch.

As shown in Figure 1A, digestion of the ccc and gapped dU•G-containing plasmid DNA after incubation with cell-free extracts by EcoRI revealed (lanes 2 and 4), as expected, that most of the dU residues were removed in all extracts tested (*WT*, *Tdg*$^{-/-}$ and *Msh2*$^{-/-}$ MEFs). Efficiency of the damage repair (92–98% for dU in mC-14-dU-pMM2, Table 1) was calculated as described in 'Materials and Methods' section, taking into account the efficiency of digestion of the control unmodified pMM2 plasmid by the corresponding restriction enzyme under the experimental conditions used (Supplementary Figure S1). In the absence of a gap, repair of the dU•G pair occurs only via DNA glycosylase-mediated BER (Figure 1A, lanes 1 and 2), whereas removal of 1.5–2.7% of 5mC in the ccc plasmids that was revealed by SmaI digestion in both the MMR-proficient and deficient extracts (Table 1) can be considered background level variations. The presence of a gap significantly stimulated erasure of neighboring 5mC, with ∼11 and 14% of 5mC residues removed after incubation of a gapped dU•G plasmid in WT and *Tdg*$^{-/-}$ extracts, respectively (Figure 1A and C). No stimulation of the removal of 5mC was observed in the *Msh2*$^{-/-}$ extract; this result was suggestive of direct involvement of the MMR pathway in the repair of the dU•G mismatch in a gap/nick-directed manner (Figure 1A and C).

The 5caC residues in both ccc and gapped plasmid DNA molecules were removed with good efficiency (75–81%) by the *WT* MEF extract (Figure 1A, lanes 6 and 8, and Table 1). On the contrary, the repair of 5caC in *Tdg*$^{-/-}$ MEF extracts was abrogated in both ccc and gapped substrates (lanes 6 and 8). These results are in agreement with other studies showing that TDG is the only known DNA glycosylase that acts on 5caC in mammalian cells (4). Recently, it was demonstrated that the MutSα complex of MMR can recognize 5caC•G pairs in DNA (25), but in the present study, in contrast to dU•G (Figure 1A, lane 3), the 5caC•G in gapped plasmid DNA did not lead to the removal of the neighboring 5mC even in the absence of TDG (Figure 1A, lanes 5 and 7, and Table 1), suggesting that MMR does not

Table 1. Restriction analysis of 5-methylcytosine (5mC) removal and lesion repair by cell-free extracts of mouse embryonic fibroblasts (MEFs)*

Modification site	5mC (removal,%)		Lesion 3′ to mC (repair,%)		Lesion 5′ to mC (repair,%)		
pMM2-based substrate	ccc	gapped/nicked	ccc	gapped	ccc	gapped	
mC-14-dU	**2.7**	**11**	**92**	**98**			*WT*
mC-14-caC	1.1	1.8	**81**	**75**			
mC-4-G/T	**4.6**	**37**	**4.5**	**37**			
hmU-15-mC	**10**	**28**			**88**	**88**	
mC-14-hmU	**6.1**	**26**	**83**	**89**			
hmU-15-mC-14-hmU	**23**	**37**	**89**	**92**	**84**	**93**	
Mult-hmU1	**43**	**46**					
Mult-hmU2	**30**	**37**					
Mult-hmU3	**25**	**35**					
Mult-hmU5	**30**	**46**					
Mult-hmU6	**33**	**39**					
Mult-dU5	**33**	**32**					
Mult-dU6	**28**	**29**					
mC-14-dU	**2.7**	**14**	**92**	**93**			$Tdg^{-/-}$
mC-14-caC	1.6	1.6	ND	ND			
mC-4-G/T	**6.6**	**36**	**4.2**	**38**			
hmU-15-mC	**9.6**	**33**			**49**	**67**	
mC-14-hmU	**7.4**	**31**	**52**	**65**			
hmU-15-mC-14-hmU	**21**	**46**	**51**	**71**	**53**	**72**	
Mult-hmU1	**38**	**48**					
Mult-hmU2	**24**	**35**					
Mult-hmU3	**22**	**36**					
Mult-hmU5	**26**	**38**					
Mult-hmU6	**26**	**36**					
mC-14-dU	1.5	ND	**96**	**95**			$Msh2^{-/-}$
mC-14-caC	ND	ND	**52**	**45**			
mC-4-G/T	1.1	ND	ND	ND			
hmU-15-mC	1.6	1.6			**77**	**75**	
mC-14-hmU	ND	ND	**74**	**73**			
hmU-15-mC-14-hmU	1.8	1.7	**74**	**68**	**82**	**79**	
Mult-hmU[1–6]	ND	ND					
Mult-dU5[6]	ND	ND					

*The data are presented as means. Standard deviation was ≤5% ($n \geq 3$). ND: not detectable (<1%).

act on 5caC•G. As expected, no removal of 5mC residues was detected in the plasmids containing 5hmC residues (Figure 1A, lanes 13 and 14), confirming that the 5hmC•G pair is not processed by MMR.

To detect and evaluate efficiency of MMR in different cell extracts, we used the plasmid DNA containing a G•T mismatch in a non-CpG context. As expected, the repair of the G•T mispair to the A•T pair (36–37% efficiency; Table 1) was accompanied with concomitant removal of the 5mC residues to the same extent and required the presence of both a gap and the MSH2 protein (Figure 1A, lanes 9 and 10 versus 11 and 12) but was not dependent on the presence of TDG (Figure 1A, lane 11 versus 12, and Table 1). This result suggests that in presence of a nick MMR is a major pathway responsible for the repair of the G•T mismatch in a non-CpG context with possible ADDM of the repair patch region.

The presence of 5hmU residues stimulates the erasure of 5mC via ncMMR

Deamination of dC and its methylated derivatives such as 5mC and 5hmC results in formation of dU•G, T•G and 5hmU•G mismatches in duplex DNA, respectively, which in turn are potential substrates for the MMR and BER pathways. Several DNA glycosylases including SMUG1, TDG and MBD4 recognize and efficiently excise 5hmU from the 5hmU•G mispair (17).

Here we studied the effect of the presence of 5hmU•G mismatches (that were placed in an XpG context within EcoRI and XhoI sites) on the 5mC erasure when a heteroduplex DNA substrate was incubated in cell-free extracts; subsequently, the involvement of MMR was also explored. As shown in Figure 1B, in extracts from WT MEFs, a single 5hmU located either 15 nt upstream (lanes 1–4) or 14 nt downstream (lanes 5–8) of 5mC in the ccc and gapped plasmid DNA was repaired with high efficiency (∼90%), which is similar to the rate of dU repair (Table 1).

Replacement of dU by 5hmU in the mC-14-dU plasmid triggered 2.5-fold more efficient erasure of 5mC, resulting in a 26% rate of SmaI cleavage in the gapped mC-14-hmU plasmid, in contrast to the 11% rate in the gapped mC-14-dU plasmid DNA (Figure 1C, Table 1). When the 5hmU residue was located 15 nt (5′) upstream of 5mC (the hmU-15-mC plasmid), we observed a rate of 5mC erasure (28%) similar to that in the plasmid with 5hmU located 3′ of 5mC (Figure 1C), thus confirming that excision and subsequent resynthesis go beyond the mismatch site.

It should be noted that under the experimental conditions involving silica membrane-based isolation of ccc plasmids from agarose gels, we observed formation of ∼5% of a nicked form; this form can arise spontaneously during multiple freeze-thaw cycles required for storage and use (Supplementary Figure S1D). Thus, removal of 5mC in reactions with the ccc plasmids containing a single mismatch was somewhat (∼5%) overestimated because of contamination with the nicked form of these plasmids, which is subject to MMR repair. Indeed, incubation of the same ccc plasmids in MMR-deficient *Msh2*$^{-/-}$ extracts did not lead to significant 5mC removal (Figure 1C).

Next, we examined the effect of the presence of two 5hmU residues—located on the 5′ and 3′ side of 5mC in the hmU-15-mC-14-hmU plasmid—on the 5mC erasure (Figure 1B). Two 5hmU residues, when present in the same plasmid, were repaired with ∼90% efficiency in WT MEF extracts, just as a single 5′ 5hmU and single 3′ 5hmU residue (Table 1). Remarkably, the presence of two 5hmU residues increased the rate of removal of 5mC up to 23 and 37% in the ccc and the gapped plasmid, respectively, in comparison with 6–10% and 26–28% of 5mC in the ccc and the gapped plasmid with a single 5hmU residue, respectively (Table 1). As expected, no significant removal of 5mC in the hmU-15-mC-14-hmU plasmids was observed in *Msh2*$^{-/-}$ extracts (Figure 1B, *Msh2*$^{-/-}$ panel). Overall repair of 5hmU•G in the *Msh2*$^{-/-}$ extracts was only slightly less efficient (∼1.2-fold) in comparison with the MMR-proficient WT extract (Table 1).

In contrast to dU-containing plasmids, 5hmU-plasmids showed a ∼1.5-fold decrease in 5hmU•G repair in *Tdg*$^{-/-}$ versus WT extracts (Figure 1B and Table 1), suggesting that TDG is required for fully functional 5hmU•G repair but not for dU•G repair in MEFs. Simultaneously, we observed a significant increase in the MMR-dependent repair of 5hmU•G in the gapped 5hmU-plasmids in the absence of TDG because the relative ratio of 5mC removal (via MMR) to the total 5hmU repair (via both MMR and BER) increased ∼1.6-fold (from 0.32 to 0.49 for hmU-15-mC, from 0.29 to 0.48 for mC-14-hmU and from 0.40 to 0.64 for the hmU-15-mC-14-hmU gapped plasmids; Table 1). Thus, in the presence of a gap, ∼30 and 49% of repaired single 5hmU•G mismatches (as opposed to 11 and 14% in the case of dU•G mismatches) were processed through the MMR pathway in WT and *Tdg*$^{-/-}$ extracts, respectively. Moreover, in the absence of TDG the difference between the rate of MMR-dependent 5mC removal in ccc versus gapped hmU-15-mC-14-hmU plasmids was decreased from 25% in WT to 14% in *Tdg*$^{-/-}$ extracts, which presents ∼1.8-fold decrease (Figure 1 C and Table 1). These results suggest that the presence of two 5hmU residues results in efficient removal of 5mC in a ccc plasmid by MMR in a BER-dependent manner, and imply that the BER-generated nick in ccc DNA could be utilized by MMR as a strand discrimination signal.

Taken together, these results suggest that the BER-dependent removal of 5hmU•G is less efficient than that of dU•G, and this drawback in turn may provide more space and time for the MMR machinery to recognize and repair the 5hmU•G mismatch, and consequently 5hmU•G stimulates more active removal of 5mC located either 5′ upstream or 3′ downstream of the lesion site. In agreement with this notion, a noticeable increase in MMR-dependent 5mC erasure was observed in the case of downregulated BER on 5hmU•G in the hmU-17–5mC-6-hmU and 5mC-6-hmU plasmids containing two and one 5hmU residue, respectively, in a non-CpG context (Table 1 and Supplementary Figure S2).

Cells from an *Msh2* knockout mouse show imbalanced BER

In addition to the MMR pathway, the MSH2 protein is involved in a variety of different pathways, which participate in the maintenance of genome integrity, e.g. nucleotide excision repair (28), non-homologous end joining (29), BER [through interaction with hMYH glycosylase (30)], cell cycle checkpoints and apoptosis (31). Previously, we demonstrated that cell-free extracts from *Msh2*$^{-/-}$ MEFs show a dramatic (3- to 5-fold) decrease in repair of 8oxoG and tandem 8oxoG-5ohU (24). Here we demonstrate that repair of 5caC•G is totally dependent on TDG and that Msh2-deficient MEF extracts exhibit a significant decrease in the repair of 5caC•G (rate of repair 45−52%) compared to WT extracts (rate of repair 75–81%; Table 1).

To examine the effects of the Msh2 deficiency in MEF cells only on DNA glycosylase and apurinic/apyrimidinic (AP) endonuclease activities of BER, we used ATP- and Mg-free conditions and employed as DNA substrates 5′-^{32}P-labeled short (30mer) X/Y-RT oligonucleotide duplexes containing 5caC•G, αA•T, THF•G or 5hmU•G base pairs as well as a single-stranded oligonucleotide containing 5hmU residues. After incubation with the extracts, the oligonucleotides were treated with a hot alkali to cleave DNA at the AP sites left over after the base excision.

As shown in Supplementary Figure S3A, incubation of αA/T-RT and THF/G-RT duplexes with extracts resulted in formation of 10mer fragments containing 3′ hydroxyl group (3′-OH) termini (lanes 2–3, 6–7 and 10–11) characteristic of the action of a major AP endonuclease, APE1, on these substrates. Cleavage of the 5′-labeled 5hmU/G-RT duplex in extracts produced two closely migrating 10mer fragments: a major one containing 3′ OH and a minor one, which migrated faster and had 3′-P termini (lanes 4, 8 and 12). This result suggested that the majority of AP sites that were generated after 5hmU excision were cleaved by AP endonucleases, and the remaining AP sites were cleaved by the hot alkaline treatment. As in the experiments with plasmid DNA substrates containing single 5hmU, we observed only a slight (∼1.3-fold) decrease in the cleavage of 5hmU/G-RT in *Tdg*$^{-/-}$ and *Msh2*$^{-/-}$ MEF as opposed to WT MEF extracts (Figure 2A). Moreover, in the WT extracts, we were unable to detect TDG activity toward an oligonucleotide duplex containing 5caC•G under the reaction conditions used; the increased protein concentrations led instead to non-specific oligonucleotide degradation (Supplementary Figure S3B). These results suggest that the TDG glycosylase activity plays only minor role in 5hmU•G repair in MEF cell-free extracts. In agreement with these observations, recently, it was shown that SMUG1 is the dominant 5hmU-DNA glycosylase in mouse cell extracts when the researchers used an oligonucleotide duplex containing single 5hmU•G in a non-CpG context and standard reaction con-

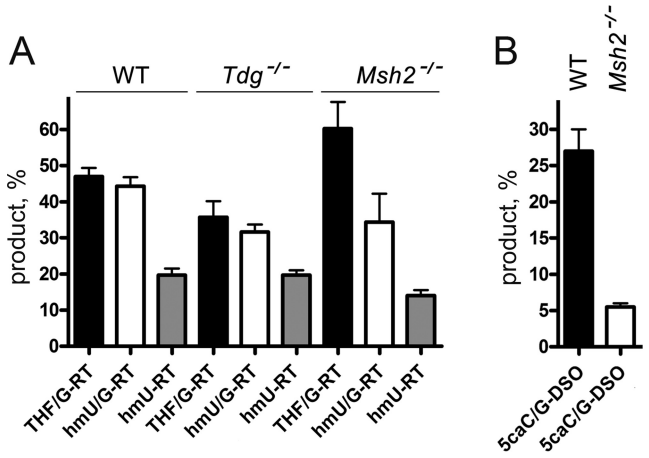

Figure 2. BER activities toward oligonucleotide substrates containing a single modified base in MEF extracts. (**A**) 5hmU DNA glycosylase and AP endonuclease activities on double- and single-stranded 30 mer RT oligonucleotide substrates containing 5hmU or tetrahydrofuranyl (THF) residues. (**B**) 5caC DNA glycosylase activity on covalently closed DSO substrates containing 5caC•G base pair. The error bars represent SD ($n = 3$). For details, see 'Materials and Methods' section and Supplementary Figure S3.

ditions in the DNA glycosylase assay (32). It should be emphasized that the 5hmU-DNA glycosylase activity of human TDG is not strongly dependent on the DNA sequence context, in contrast to the T•G mispair (17).

To specifically measure activity of SMUG1 in the extracts, we used ssDNA containing 5hmU, known to be the best substrate for this enzyme (33). Single-stranded 5hmU-RT oligonucleotide incubated with extracts followed by hot alkali treatment generated 10mer fragments with 3′-P termini (Supplementary Figure S3A; lanes 5, 9 and 13), suggesting that the AP site in ssDNA was cleaved by the hot alkali. As expected, the repair of 5hmU in ssDNA did not depend on TDG, with the same 20% rate of cleavage in both WT MEF and *Tdg*$^{-/-}$ MEF extracts (Figure 2A). Again, as in the case of 5hmU/G-RT, we observed a 1.4-fold decrease in the repair of 5hmU-ssDNA in *Msh2*$^{-/-}$ MEF (14% of cleavage) in comparison with WT and *Tdg*$^{-/-}$ MEF extracts, suggesting that either a decreased level of SMUG1 or an increased level of some unidentified inhibitor(s) was present in the *Msh2*$^{-/-}$ MEF extracts.

As shown in Figure 2A, AP endonuclease activity was ∼1.3-fold higher in *Msh2*$^{-/-}$ MEF extracts (60% of cleavage) than in WT and *Tdg*$^{-/-}$ extracts (47 and 36%, respectively), indicating, together with decreased 5hmU-DNA glycosylase activities, a significant deregulation of BER in MEF *Msh2*$^{-/-}$ cells. We believe that the observed increase in the AP endonuclease activity may have some 'compensatory' effect. Indeed, the 3′ exonuclease activity of APE1 can efficiently edit 3′ mispaired nucleotides generated during DNA synthesis (34), thereby providing a back-up function to MMR in terms of elimination of misincorporated bases.

To detect TDG action on 5caC-containing oligonucleotide duplexes and to prevent their non-specific degradation by exonucleases in cell-free extracts, we constructed a linear ^{32}P-labeled 72mer covalently closed DSO molecule containing a 5caC•G pair in the middle of the stem (Supplementary Figure S3C). Cleavage of the DSO substrates at the site of the modified base after incubation with the extracts was quantified using additional digestion of the substrate with the HaeIII restriction endonuclease. When the 5caC/G- and control A/T-DSO were incubated with the extracts, we observed a 5-fold higher 5caC-DNA glycosylase activity in the WT MEF extract (Figure 2B and Supplementary Figure S3C, lanes 1–2) than in *Msh2*$^{-/-}$ MEF extracts (lanes 5–6); this finding confirmed our data from DNA plasmid substrates. Immunoblotting revealed that the TDG level in *Msh2*$^{-/-}$ MEFs was slightly (∼1.6-fold) lower than in WT extracts (Supplementary Figure S3D), suggesting that, indeed, the decreased TDG activity in *Msh2*$^{-/-}$ MEF extracts is associated, with a lower level of the TDG protein in *Msh2*$^{-/-}$ than in WT MEF extracts and possibly with some other factors that affect the efficiency of the DNA glycosylase step in BER.

Taken together, these results show that BER is significantly imbalanced in *Msh2*$^{-/-}$ MEFs, showing considerable elevation of an AP endonuclease and a marked decrease in TDG and SMUG1 DNA glycosylase activities.

The presence of multiple 5hmU and dU residues promotes efficient MMR and subsequent 5mC removal in MEF cell-free extracts via interplay with the BER pathway

Here we determined whether the presence of multiple, up to four, 5hmU residues per plasmid molecule would stimulate MMR-dependent removal of 5mC marks in the plasmid incubated with cell-free extracts (Table 1 and Supplementary Table S1 and Figure S4). As shown in Figure 3, the increase in the percentage of 5mC removed by MMR correlated with the increase in the number of 5hmU residues per plasmid DNA, particularly in the ccc plasmid. Indeed, we observed an almost twofold increase (from 23% for hmU-15-mC-14-hmU to 43% for Mult-hmU1) of the MMR-dependent erasure of 5mC marks in WT extracts in the case of the ccc Mult-hmU1 plasmid, which has two additional 5hmU•G mispairs 69 and 78 nt upstream of the 5mC residue, in contrast to the hmU-15-mC-14-hmU plasmid. In the ccc Mult-hmU2 plasmid, the 5mC erasure was reduced to 30% but remained higher than that of the ccc hmU-15-mC-14-hmU plasmid (23%); this plasmid was derived from Mult-hmU1 but contains only three 5hmU•G mispairs because one of the 5hmU residues located 15-nt upstream of 5mC in Mult-hmU1 was replaced by matching the C base.

When two 5′ upstream 5hmU residues in Mult-hmU2 were replaced by one tandem 5hmU–5hmU site in the ccc Mult-hmU3 plasmid, we observed a similar level of 5mC erasure (25%) in comparison with the ccc hmU-15-mC-14-hmU plasmid (23%), suggesting that tandem 5hmU–5hmU is repaired as a single site and does not stimulate MMR to the same extent as do two stand-alone separate 5hmU•G sites. The Mult-hmU5 plasmid (derived from Mult-hmU2) contains all three separate 5hmU residues in the non-CpG context, whereas, Mult-hmU6 contains one 5hmU-G-5hmU-G-5hmU-G cluster, which replaced two 5hmU residues in Mult-hmU2 5′ upstream of 5mC mark (XbaI site) generating all four 5hmU residues in the CpG context. In the Mult-hmU5 and Mult-hmU6 plasmids, the level of the MMR-dependent 5mC removal in ccc plasmids

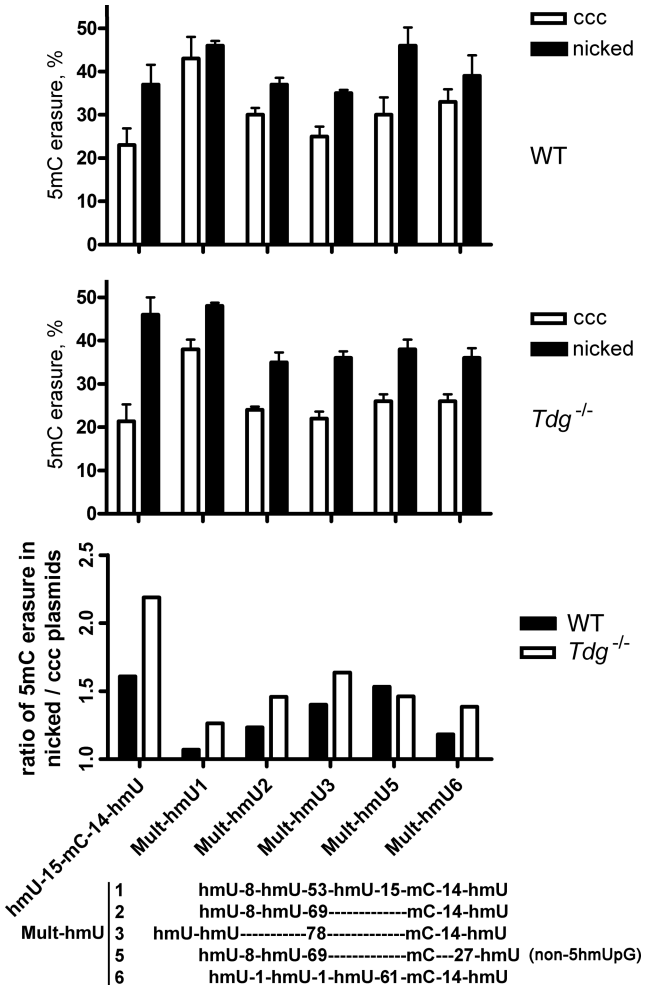

Figure 3. Demethylation of the pMM2 plasmid substrates containing multiple 5hmU•G mispairs in MEF extracts. The error bars represent SD (n = 3).

did not change significantly (30 and 33%, respectively), suggesting that placing 5hmU residues in the CpG context has no significant impact on the efficiency of MMR and competition between MMR and BER during removal of multiple 5hmU residues. It should be noted that the level of 5mC erasure by WT MEF extracts in the gapped version of Mult-hmU[1–6] plasmids, varies slightly from 35 to 46%, reflecting the maximal level of MMR activity under the reaction conditions used. A twofold increase in protein concentrations in the reaction mixture resulted in higher MMR activity, with a 1.15- to 1.5-fold increase in 5mC erasure (Supplementary Figure S5), suggesting that in live cells, MMR may be a major pathway for repair of multiple deaminated residues in DNA. Results obtained with the ccc and nicked Mult-hmU[1–6] plasmids incubated in *Tdg*$^{-/-}$ MEF extracts were quite similar to those obtained with WT MEFs (Figure 3). Importantly, the ratios of the MMR-dependent 5mC removal in nicked/gapped to ccc plasmids containing multiple 5hmU residues in CpG context notably increased in *Tdg*$^{-/-}$ compare to WT MEF extracts (Figure 3), suggesting that efficient removal of 5mC in a ccc plasmid containing multiple base lesions by MMR requires fully functional BER pathway. This observation supports the cooperative role of BER in the MMR-dependent removal of 5mC by generation of nicks essential for the MMR strand discrimination and excision.

We sought to increase the difference in the MMR activity on ccc versus nicked/gapped DNA substrates even further by inhibiting BER in MEF *Tdg*$^{-/-}$ extracts using CRT0044876, the low-molecular-weight inhibitor of APE1, or by 5hmU-containing oligonucleotide competitors of DNA glycosylases. However, these attempts were unsuccessful, apparently due to a quite high cell extract concentration required for MMR. The high extract concentration stipulated extremely high concentration of CRT0044876 and induced nuclease degradation of the 5hmU-oligonucleotides, which in turn led to non-specific inhibition of MMR in both cases (Supplementary Figure S6).

As expected, the incubation of ccc Mult-dU5 and Mult-dU6 plasmids (containing multiple dU instead of 5hmU residues) in WT MEF extracts has resulted in high 33 and 28% of the 5mC mark erasure, respectively, similar to the rates (32 and 29%) observed with nicked Mult-dU5[6] plasmids (Table 1). Taken together, these results suggest that the presence of multiple 5hmU•G or dU•G mispairs stimulates DNA damage repair and concomitant erasure of 5mC marks via the MMR pathway. Furthermore, multiple types of DNA damage stimulate the MMR pathway even in the absence of a preexisting discriminatory signal such as a single-stranded gap, suggesting that DNA glycosylase-catalyzed excision of 5hmU or dU residues generates DNA strand breaks or gaps, which in turn may trigger the removal of the damaged DNA strand by the MMR machinery.

MMR-dependent 5mC removal in human BL2 Cell-free extracts

The non-targeted action of cytosine deaminases on cellular DNA can be genotoxic to the cell; therefore, the expression of these enzymes is subject to tight regulation. One of the well-studied examples of a high level of endogenous AID expression is Burkitt's lymphoma (BL) 2 cells (35). Here we used BL2 cells to characterize the MMR-dependent 5mC removal in human cells and to compare our data with those from the MEF cell extracts.

As shown in Figure 4A, MMR was highly proficient in the extracts of BL2 cells due to the observed ∼60% of 5mC erasure in control plasmid substrates containing the G•T mismatch (lane 11) at a 3-fold lower protein concentration (1 mg/ml) in comparison with MEF extracts (3 mg/ml). In the gapped mC-14-dU plasmid, ∼20% of the 62% of repaired dU residues were processed by the MMR pathway (lanes 3 and 4), suggesting that in BL2 cells, MMR can efficiently compete with the BER pathway when acting upon deaminated cytosines. Indeed, in the absence of MMR on ccc mC-14-dU plasmid (lane 1) we observed only ∼37% dU•G repair in the BER pathway (lane 2). In contrast to dU, under the reaction conditions used, 5caC residues were removed with very low efficiency (only 5–7%) in both ccc and gapped plasmid DNA (lanes 6 and 8), pointing to a low level of the TDG protein in BL2 cells.

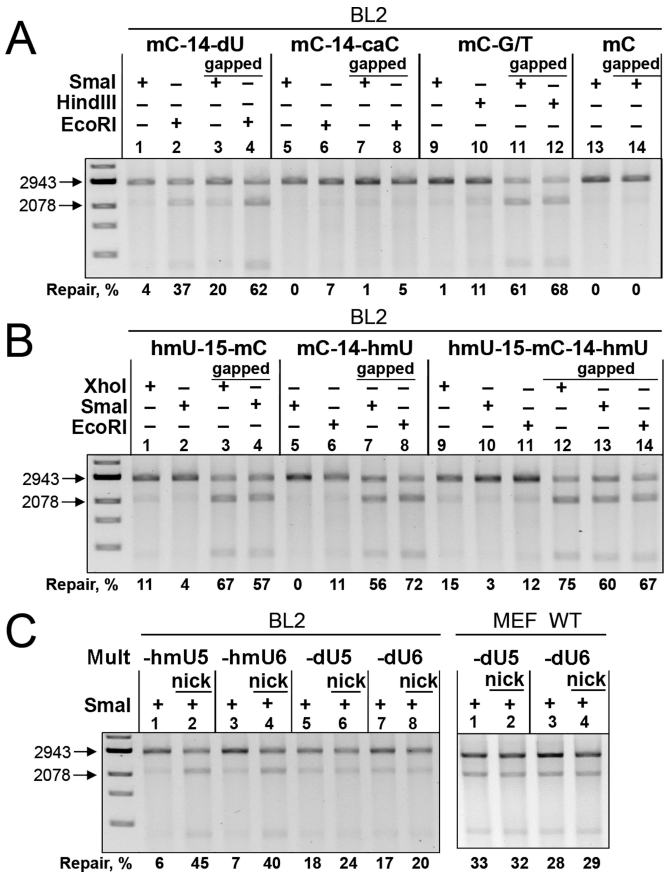

Figure 4. Reconstitution of the BER and MMR pathways on pMM2 plasmid substrates containing oxidation/deamination products of deoxycytosine and 5mC in cell extracts from BL2 and MEF cells. (**A**) Efficiency of repair of dU•G, 5caC•G, G•T and 5hmC•G plasmid substrates in BL2 extracts. (**B**) Efficiency of repair of plasmids containing 5hmU•G mispairs in BL2 extracts. (**C**) Efficiency of demethylation at the SmaI site of Mult-hmU5[6] and Mult-dU5[6] plasmids in BL2 and MEF cell extracts.

It was shown elsewhere that the UNG2 protein is the major and most abundant uracil DNA glycosylase (UDG) in human cells, whereas SMUG1 and TDG are the major and abundant UDGs in mouse cells (36). In agreement with these observations, only 11–12% of the 5hmU residues were repaired in ccc plasmids that contained a single 5hmU•G mispair (Figure 4B, lanes 1 and 6), whereas up to 66–73% of 5hmU and 56% of 5mC were removed from the gapped hmU-15-mC and mC-14-hmU plasmids, indicating that ∼80% of the repaired 5hmU residues in the gapped plasmids are removed in the MMR pathway. Similar results were obtained with the ccc and gapped hmU-15-mC-14-hmU plasmids containing two 5hmU residues per plasmid (Figure 4B, lanes 9–14). Contrary to MEF, BL2 extracts contained very little or no detectable MMR activity on the ccc hmU-15-mC-14-hmU plasmid (lane 10), suggesting that the BER-mediated nick formation is a rate-limiting step for MMR in human BL2 cells. Nevertheless, the presence of multiple 5hmU and dU residues in ccc Mult-hmU5[6] and Mult-dU5[6] plasmids (Figure 4C) stimulates MMR activity toward 5hmU residues in BL2 extracts; this phenomenon resulted in concomitant removal of 5mC marks: up to 7% in ccc Mult-hmU5[6] (lanes 1 and 3) and 18% in Mult-dU5[6] plasmids (lanes 5 and 7), respectively. These results are in agreement with the data from Jiricny's group, who used a similar plasmid-based assay in BL2 and 293T cell extracts and demonstrated efficient MMR in a gapped plasmid containing a single dU•G mispair and the interference between MMR and BER pathways during processing of two adjacent dU•G mispairs (23,37).

The presence of nicks in the Mult-hmU5[6] plasmid greatly stimulated (6- to 7-fold) MMR (up to 45% of 5mC removal; lanes 1 and 3 versus 2 and 4). This was not the case for the plasmids containing uracil residues: only 1.2- to 1.3-fold stimulation of MMR in the Mult-dU5[6] plasmid in the presence of a nick was observed, with only 20–24% of 5mC removal (lanes 5 and 7 versus 6 and 8). This finding suggests that dU residues more efficiently induce single-strand breaks via BER in comparison with 5hmU residues in BL2 extracts; both BER and MMR compete for dU removal. Importantly, a strong decrease (up to 20-fold) in MMR activity on ccc versus nicked/gapped hmU-15-mC-14-hmU and Mult-hmU5[6] plasmids in extracts with low 5hmU-DNA glycosylase activity fully support our observations in MEF extracts, namely, that BER generated the nicks that in turn act as the MMR strand discrimination signals. Taken together, these results suggest that the balance between BER and MMR pathways is strongly dependent on the cell type; because of low 5hmU-DNA glycosylase activity in BL2 cells, the MMR-dependent DDM is more efficient on multiple dU than on multiple 5hmU residues.

Long-patch BER is involved in 5mC removal from DNA substrates containing a 5′ adjacent dU residue

To examine the potential roles of different DNA repair pathways (which do not directly target 5mC marks) in ADDM, we characterized the efficiency of long-patch BER in the *in vitro* reconstitution assay. Long-patch BER inserts 2–13 nt via coordinated action of polymerase(s) β/δ/ϵ, PCNA, flap endonuclease 1 (FEN1) and DNA ligase I (38). The choice between a single-nucleotide (short-patch) BER pathway and the long-patch BER pathway depends on several factors including the type of DNA glycosylase that initiates BER, the cell type, the proliferation status of the cells and availability of downstream BER factors (38).

To determine the efficiency of long-patch BER, we incubated the ccc DNA plasmids containing a single dU residue one nt (dU-1-mC) and 5 nt (dU-5-mC) 5′ upstream of the 5mC residue in cell extracts and measured the loss of the methylation marks, that should be removed in at least a 3- and 7-nt repair patch, respectively. It should be stressed that the repair of a single dU•G pair in the ccc mC-14-dU-pMM2 plasmid in MEF and BL2 cell extracts occurs only via DNA glycosylase-mediated BER (Table 1 and Figure 4). As shown in Figure 5A, incubation of dU-1-mC and dU-5-mC ccc DNA plasmids in BL2 and MEF cell extracts resulted in efficient erasure of the 5mC mark through the long-patch BER pathway: 24 and 37% in BL2 extracts (lanes 2 and 4), 62 and 45% in WT MEF extracts (lanes 6 and 8), and 53 and 34% in MEF *Msh2*$^{-/-}$ extracts (lanes 10 and 12), respectively. Taking into account the reaction conditions used with the overall efficiency of dU•G BER

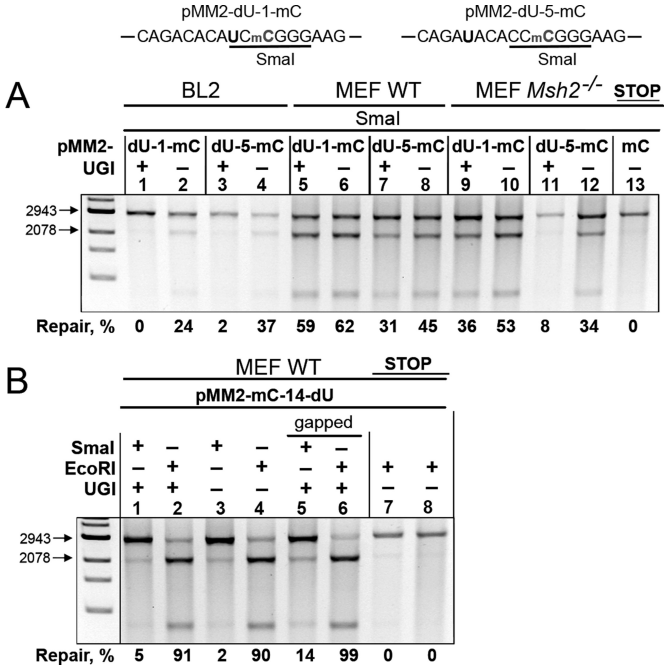

Figure 5. Long-patch BER at a single dU•G in cell-free extracts of MEF and BL2 cells. (**A**) Long-patch BER dependent on 5mC removal in covalently closed circular (CCC) plasmid substrates containing a 5′ adjacent dU residue. (**B**) Effects of a uracil-DNA glycosylase inhibitor (UGI) on total repair of a single dU•G in MEF extracts.

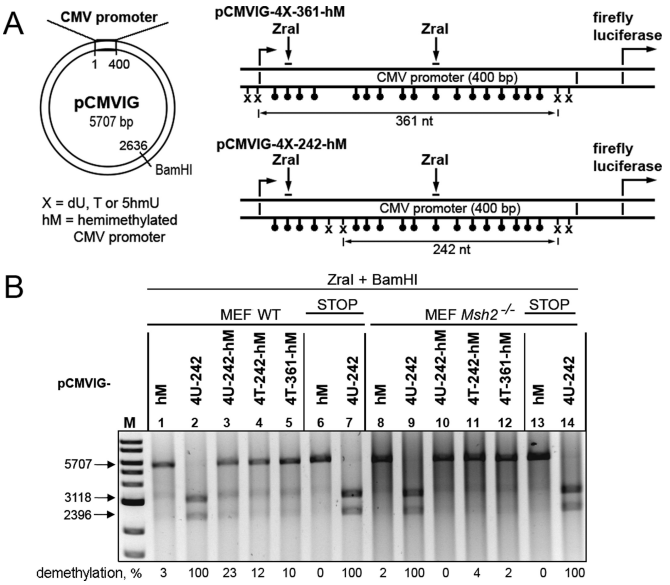

Figure 6. MMR-dependent demethylation of DNA plasmids containing multiple 5mC residues and mispairs in MEF extracts. (**A**) Schematic representation of the pCMVIG luciferase reporter plasmid containing custom-made lesions and/or methylated CpG in the promoter region. (**B**) MMR-dependent demethylation of the hemimethylated pCMVIG plasmid containing multiple dU•G and T•G mispairs. For details, see 'Materials and Methods' section and Supplementary Material.

pathway on ccc mC-14-dU constituted ∼37% for BL2 (Figure 4A, lane 2) and 90–96% for MEF extracts (Table 1), the observed efficiency of the long-patch BER pathway on the ccc dU-1(5)-mC plasmids (24–37% in BL2 extracts and 45–62% in WT MEF extracts) suggest that the majority (65–100%) and a half (47–67%) of the BER-dependent removal of dU•G in BL2 and in WT MEF extracts, respectively, occurred through the long-patch BER pathway, which implies a repair synthesis patch at least 3–7 nt long. Previously, it was shown that UNG2, the major nuclear UDG in mammalian cells, removes ∼80% of uracil from dU•T mispairs via the long-patch BER pathway (39). Indeed, addition of a uracil-DNA glycosylase inhibitor (UGI) that totally inhibits UNG2-dependent BER, resulted in a complete loss of long-patch BER on dU in BL2 extracts (Figure 5A, lanes 1 and 3), but only a partial decrease in WT and *Msh2*$^{-/-}$ MEF extracts (lanes 7, 9 and 11 versus 8, 10 and 12, respectively). At the same time, we did not observe significant changes in the efficiency of long-patch BER on dU-1-mC substrate in WT MEF extracts (lanes 5 and 6), possibly due to either a more efficient backup of UNG2 by SMUG1 enzyme for 3-nt than for 7-nt long-patch BER or because of a saturation effect since WT MEF extracts on dU-1-mC exhibit overall stronger long-patch BER activity. It should be noted that mouse SMUG1 displayed an almost six-fold higher dU•G excision activity as compared to human SMUG1 (36), providing a major uracil–DNA glycosylase activity on dU•G mispairs in MEF extracts (40). In agreement with these observations, addition of UGI to the WT MEF extracts did not influence the total repair of dU•G in ccc mC-14-dU (Figure 5B, lanes 2 and 4).

These results suggest that the UNG2 protein is essential for the long-patch BER pathway acting on dU•G in BL2 cells but not in MEFs, in which long-patch BER can be initiated by other uracil-DNA glycosylases, most likely by mouse SMUG1.

Multiple 5hmU and dU residues inside the hemimethylated promoter region increase reporter gene expression in an MMR-dependent manner

The *in vitro* data suggested that multiple 5hmU•G and dU•G mismatches promote efficient MMR, which in turn can induce ADDM in the regions surrounding the lesions. To determine whether the data obtained *in vitro* can be applied to living cells, we used the plasmid-based approach described above, to set up a reporter system that allowed us to measure gene expression in MEF cells. As shown in Figure 6A, the plasmid vectors pCMVIG-4U-242-hM, pCMVIG-4T-242-hM and pCMVIG-4T-361-hM contained a truncated (up to 400 nt) hemimethylated CMV promoter and four 5hmU•G, dU•G or T•G mismatches at different positions in the coding DNA strand, respectively. Before transfection, we demonstrated that all three plasmids underwent partial DDM when incubated in WT but not *Msh2*$^{-/-}$ MEF cell-free extracts (Figure 6B). The extent of DDM was measured by the cleavage of methylated sites by the ZraI restriction endonuclease, which is sensitive to cytosine methylation.

Overall efficiency of MMR-dependent demethylation of ccc pCMVIG-4U-242-hM constructs containing multiple dU•G mispairs in WT MEF extracts (23%, Figure 6B, lane 3) was similar to that of the ccc Mult-dU5[6] pMM2-based constructs (28–33%) (Table 1). Replacement of dU•G by

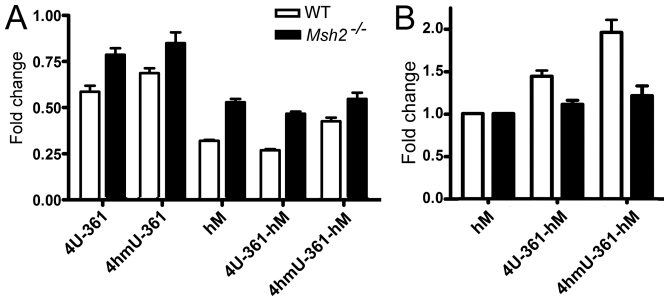

Figure 7. Expression of the reporter pCMVIG plasmids with a hemimethylated promoter containing multiple dU•G and 5hmU•G mispairs after transfection in MEFs. (**A**) Relative changes in the reporter expression of pCMVIG-derived plasmids ($L_{pCMVIG\text{-}xxx}$) that was normalized to the expression level of unmodified pCMVIG (bars represent $L_{pCMVIG\text{-}xxx}/L_{pCMVIG}$ ratio). (**B**) Relative reporter expression of hemimethylated plasmids containing dU•G and 5hmU•G mispairs ($L_{pCMVIG\text{-}4X\text{-}361\text{-}hM}$); the data are normalized to the level of the reporter expression in the presence of the corresponding mispairs in unmethylated plasmids and to the expression levels of hemimethylated pCMVIG-hM and unmodified pCMVIG plasmids (bars represent $L_{pCMVIG\text{-}4X\text{-}361\text{-}hM}/L_{pCMVIG\text{-}4X\text{-}361}/L_{pCMVIG\text{-}hM}/L_{pCMVIG}$ ratio). Original data of the pCMVIG-derived plasmids expressions shown in Supplementary Figure S7.

T•G mispairs in the pCMVIG-4T-242-hM substrate resulted in a 2-fold decrease in the efficiency of demethylation (Figure 6, lane 3 versus lane 4). This result is in agreement with above findings that efficiency of MMR on ccc pMM2-based constructs containing multiple mispairs depends on the efficiency of the BER pathway. It should be noted that dU•G is a more preferred DNA substrate than T•G mispair for the mismatch-specific thymine/UDGs (33), thus BER would be more efficient on the former DNA substrate. Next, pCMVIG-4X-361 constructs containing a hemimethylated promoter of the firefly luciferase reporter gene, which was flanked by dU or 5hmU residues, were transfected into WT and *Msh2*$^{-/-}$ MEF cells and changes in the promoter methylation status were monitored by means of the luciferase assay. In addition, the cells were cotransfected with the pRL-CMV vector expressing *Renilla* luciferase for control of the transfection efficiency. As shown in Figure 7A, hemimethylation of the pCMVIG promoter (pCMVIG-hM) results in a significant decrease in expression of the gene reporter in comparison with the control unmethylated vector, down to 32 and 53% in WT and *Msh2*$^{-/-}$ MEF cells, respectively.

Insertion of four dU•G or 5hmU•G mispairs in the unmethylated vector also resulted in a less efficient reporter expression (down to 59–69% efficiency in WT MEFs and 79–85% in *Msh2*$^{-/-}$ MEFs), indicating that DNA damage repair in the promoter region can interfere with transcription. When taking into account the negative effect of DNA repair on transcription, we normalized the data obtained for the plasmids containing both a hemimethylated promoter and mispairs (Figure 7B), to avoid summing up the effects of damage repair and promoter methylation. We found that the presence of dU•G and especially 5hmU•G mispairs in the hemimethylated promoter region resulted in a 1.4- and 1.9-fold higher expression level of the reporter, respectively, in WT but not in *Msh2*$^{-/-}$ MEF cells. Taken together, these results suggest that multiple base mispairs generated by deamination of cytosine or its analogs can induce ADDM of a hemimethylated reporter gene in an MMR-dependent manner, thus confirming our *in vitro* results. These findings support involvement of the ncMMR pathway in ADDM via a long-patch DNA excision-resynthesis mechanism with concomitant removal of epigenetic marks.

DISCUSSION

Molecular and genetic studies in mammals have shown the waves of genome-wide ADDM that occur in the germline and during early embryogenesis (41). In the last two decades, a growing amount of evidence of the locus-specific ADDM in somatic cells has been reported as well. In mature neurons, lymphocytes and in breast cancer cell lines, local DDM that leads to specific gene activation has been detected after stimulation by hormones (42).

Recently, Petersen-Mahrt's group proposed that DDM might proceed via long-patch DNA excision/resynthesis repair pathways (long-patch BER and MMR) (21). It was shown that targeting of AID to certain loci induces *in vitro* and *in vivo* (in a transgenic mouse model) local processive (at least 1 kb) DDM (21). In this scenario, mCpG sites do not have to be targeted for demethylation because a single DNA lesion can induce multiple DDM events. A major source of the lesions that may initiate ADDM is the products of deamination of C•G and 5mC•G in duplex DNA such as the dU•G, 5ohU•G, T•G and 5hmU•G mismatches. These mispairs are possible substrates for both the MMR and long-patch BER pathways, the former being a longer-range mechanism. Recently, it was demonstrated that certain phases of active demethylation in mouse zygotes possibly involve AID-mediated cytosine deamination coupled with UNG2-initiated long-patch BER (20). It is widely accepted that deamination of cytosines and the interplay between BER and MMR pathways during removal of uracil residues play essential roles in SHM in the Ig locus of activated B cells (43). Nevertheless, the role of such an interplay and the role of each DNA repair pathway in genome-wide or in locus-specific ADDM are still unknown.

In the present work, using a plasmid-based approach, we examined the roles of short-patch and long-patch BER and non-canonical MMR pathways in the removal of deaminated and oxidized derivatives of cytosine and in the non-targeted erasure of 5mC marks. The results showed that the 5hmU•G and dU•G mispairs are efficiently removed not only by the short-patch BER but also by the MMR and long-patch BER pathways with concomitant removal of adjacent and distant 5mC residues, respectively, in MEF and BL2 cell extracts. The presence of only two lesion-containing mispairs in ccc DNA was sufficient for the efficient MMR-dependent removal of lesions and local demethylation of plasmid DNA (~23% rate of 5mC erasure in hmU-15-mC-14-hmU in MEF extracts, Table 1). These findings support the role of nicking in plasmid DNA; this process is initiated during BER-mediated removal of one of the lesions during the MMR activation to remove remaining mismatch(es) (37). Therefore, two deamination events could lead to the removal of 5mC residues across a stretch of up to 1 kbp of DNA, characteristic of a repair

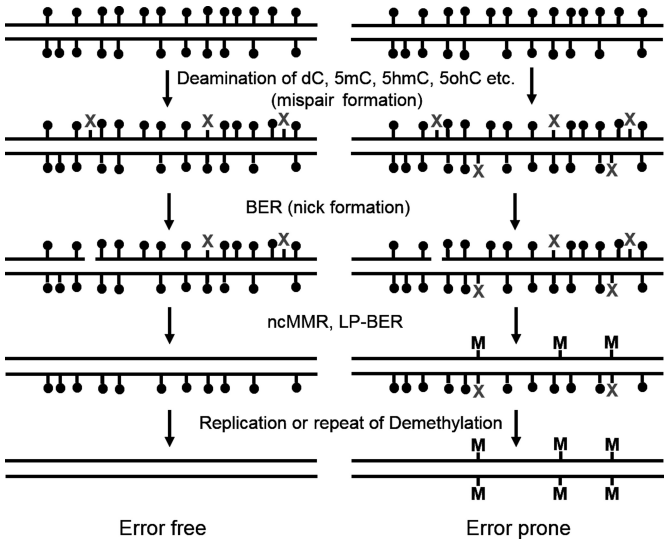

Figure 8. Schematic representation of the interplay between BER and MMR during long-patch DDM. X: deaminated cytosines, their derivatives and other mispairs; M: mutations.

patch in the MMR pathway. Furthermore, the presence of three to four 5hmU•G or dU•G mispairs in the same MMR nick/mismatch detection region enables MMR to become fully proficient and makes BER's action on the same lesions synergistic with that of MMR rather than making these two pathways compete. Such cooperation results in a long-patch repair synthesis, which in turn leads to demethylation of the entire DNA region. In agreement with this notion, in BL2 cell extracts that lack proficient BER toward 5hmU residues, MMR inefficiently removes 5mC in the ccc plasmids that contain multiple 5hmU•G mispairs (Figure 4). This cooperative action between BER and MMR strongly depends on the cell type; the balance between these repair pathways may be important for long-patch DDM.

According to our data, about a half of mispairs generated by cytosine deamination were processed through the long-patch BER mechanism, which resynthesized at least a 3–7-nt patch in all the extracts tested (Figure 5), suggesting that even in the absence of functional MMR, cytosine desamination may initiate non-targeted ADDM albeit much more local. On the basis of our results, we propose a model of ADDM in which the interplay between deamination, BER and MMR pathways leads to non-targeted 5mC erasure (Figure 8). In this model, the deamination of multiple C, 5mC and/or 5hmC residues in one or both DNA strands generates mismatches, which are recognized by BER and MMR. At the next step, introduction of a single-strand break into the DNA duplex by BER enables MMR to selectively remove one of the methylated DNA strands in the duplex, after which, the long-gap-filling DNA synthesis generates hemimethylated duplex DNA. If the methylation pattern of the demethylated strand after DNA repair of the deaminated cytosines is not maintained by DNMT1, then the complementary methylated strand can lose methylation marks either passively after DNA replication or via the active process mediated by a variety of established mechanisms [for review, see (42,44)]. It should be stressed that the deamination of cytosines in both strands of a methylated DNA duplex will not only cause erasure of 5mC in both strands but will also lead to multiple mutations during removal of the damaged cytosines in both the BER and MMR pathways (Figure 8, right panel).

Our model is supported by the reports that ncMMR can also be activated by alkylation DNA damage (O6-methyl-G) (23) and/or by oxidative DNA damage (5-hydroxy-dU) (24); this is another source for initiation of nontargeted DDM. It is reasonable to assume that to induce DDM at specific genomic loci, DNA damage should be targeted to the region near the methylated/hydroxymethylated CpG sites. The molecular mechanisms of targeted DNA damage (including the AID-initiated SHM and TET-mediated demethylation) remain largely unknown. A single round of long-range DDM can both generate the stretches of hemimethylated CpG sites and convert existing hemimethylated sites into unmethylated ones in DNA. Clusters of hemimethylated sites have been observed in the satellite 2 DNA repeat region in carcinomas; this phenomenon may represent the intermediates of ADDM that occur during carcinogenesis (45). Further studies are needed to determine which mechanisms of ADDM are at work *in vivo* at different genomic loci and regions, in different cell types, and at different stages of development.

An important question arises: 'How is genome integrity maintained during ADDM if cytosine deamination is rampant, especially when both C and 5mC residues are targeted?' Indeed, cytosine deamination can be highly mutagenic and lead to formation of DSBs, during SHM and class switch recombination in the Ig locus of activated B cells. To decrease genotoxicity of the deamination of cytosines, the process should be coordinated in the cell in such a way that both DNA strands are not deaminated simultaneously to avoid creation of mutations during repair or synthesis opposite to dU residues and formation of a DSB. Furthermore, largely replication-independent ncMMR can be error prone because its activation promotes PCNA ubiquitination, which appears to be directly responsible for recruitment of Pol-H (and possibly additional error-prone polymerases) to chromatin (23,24). Moreover, it was shown that the ssDNA regions generated by MMR become susceptible to AID/APOBEC-mediated mutagenesis (46). Modrich's group showed that during *in vitro* reconstitution of MMR in cell-free extracts, DNA synthesis is catalyzed by Pol-δ (47), but their *in vivo* reconstitution experiment could not rule out the situation where MMR-generated single-stranded gaps may be filled either by error-prone or high-fidelity polymerases, depending on the phase of the cell cycle, availability of dNTPs and some other unknown factors. Further studies are needed to determine the key factors that regulate error-free and error-prone MMR.

Spontaneous and enzymatically catalyzed hydrolytic deamination of 5mC yields a mutagenic C→T transition at the CpG methylation sites that is frequently observed in inherited human disorders and in the p53 gene in cancer cells (48). Gastrointestinal (esophageal, colorectal and gastric) tumors that are frequently associated with MMR deficiencies have a high frequency of transition mutations in CpG dinucleotides that may reflect increased DNA methylation in these cancers (49). Cervical, bladder, breast and

some head and neck cancers frequently have mutations of cytosines in the context of TpC dinucleotides; such mutations are often caused by the APOBEC family of cytidine deaminases (50). According to these observations and our data, we believe that an interplay between MMR and BER (when they are acting upon deaminated cytosines) involves long-patch ADDM (physiological or pathological), which may contribute to tumorigenesis.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We thank S. Aoufouchi, D. Zharkov and M. Saparbaev for critical review of the manuscript and helpful discussions.

FUNDING

Fondation de France [2012-00029161 to A.A.I.]; Russian Science Foundation [14-24-00093 to I.G.]; Fondation ARC Postdoctoral Fellowship [PDF20101202141 to I.G.]. Funding for open access charge: La Ligue Nationale Contre le Cancer [EL2016.LNCC/MuS].
Conflict of interest statement. None declared.

REFERENCES

1. Feng,S., Jacobsen,S.E. and Reik,W. (2010) Epigenetic reprogramming in plant and animal development. *Science*, **330**, 622–627.
2. Ko,M., Huang,Y., Jankowska,A.M., Pape,U.J., Tahiliani,M., Bandukwala,H.S., An,J., Lamperti,E.D., Koh,K.P., Ganetzky,R. *et al.* (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, **468**, 839–843.
3. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
4. He,Y.F., Li,B.Z., Li,Z., Liu,P., Wang,Y., Tang,Q., Ding,J., Jia,Y., Chen,Z., Li,L. *et al.* (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, **333**, 1303–1307.
5. Ito,S., Shen,L., Dai,Q., Wu,S.C., Collins,L.B., Swenberg,J.A., He,C. and Zhang,Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
6. Maiti,A. and Drohat,A.C. (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.*, **286**, 35334–35338.
7. Cortellino,S., Xu,J., Sannai,M., Moore,R., Caretti,E., Cigliano,A., Le Coz,M., Devarajan,K., Wessels,A., Soprano,D. *et al.* (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, **146**, 67–79.
8. Rai,K., Huggins,I.J., James,S.R., Karpf,A.R., Jones,D.A. and Cairns,B.R. (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell*, **135**, 1201–1212.
9. Chen,C.C., Wang,K.Y. and Shen,C.K. (2013) DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. *J. Biol. Chem.*, **288**, 9084–9091.
10. Shimoda,N., Hirose,K., Kaneto,R., Izawa,T., Yokoi,H., Hashimoto,N. and Kikuchi,Y. (2014) No evidence for AID/MBD4-coupled DNA demethylation in zebrafish embryos. *PLoS One*, **9**, e114816.
11. Muramatsu,M., Kinoshita,K., Fagarasan,S., Yamada,S., Shinkai,Y. and Honjo,T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553–563.
12. Revy,P., Muto,T., Levy,Y., Geissmann,F., Plebani,A., Sanal,O., Catalan,N., Forveille,M., Dufourcq-Labelouse,R., Gennery,A. *et al.* (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, **102**, 565–575.
13. Popp,C., Dean,W., Feng,S., Cokus,S.J., Andrews,S., Pellegrini,M., Jacobsen,S.E. and Reik,W. (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature*, **463**, 1101–1105.
14. Kumar,R., DiMenna,L., Schrode,N., Liu,T.C., Franck,P., Munoz-Descalzo,S., Hadjantonakis,A.K., Zarrin,A.A., Chaudhuri,J., Elemento,O. *et al.* (2013) AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature*, **500**, 89–92.
15. Bhutani,N., Brady,J.J., Damian,M., Sacco,A., Corbel,S.Y. and Blau,H.M. (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature*, **463**, 1042–1047.
16. Guo,J.U., Su,Y., Zhong,C., Ming,G.L. and Song,H. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, **145**, 423–434.
17. Morera,S., Grin,I., Vigouroux,A., Couve,S., Henriot,V., Saparbaev,M. and Ishchenko,A.A. (2012) Biochemical and structural characterization of the glycosylase domain of MBD4 bound to thymine and 5-hydroxymethyuracil-containing DNA. *Nucleic Acids Res.*, **40**, 9917–9926.
18. Nabel,C.S. and Kohli,R.M. (2011) Molecular biology. Demystifying DNA demethylation. *Science*, **333**, 1229–1230.
19. Rusmintratip,V. and Sowers,L.C. (2000) An unexpectedly high excision capacity for mispaired 5-hydroxymethyluracil in human cell extracts. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 14183–14187.
20. Santos,F., Peat,J., Burgess,H., Rada,C., Reik,W. and Dean,W. (2013) Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. *Epigenetics Chromatin*, **6**, 39.
21. Franchini,D.M., Chan,C.F., Morgan,H., Incorvaia,E., Rangam,G., Dean,W., Santos,F., Reik,W. and Petersen-Mahrt,S.K. (2014) Processive DNA demethylation via DNA deaminase-induced lesion resolution. *PLoS One*, **9**, e97754.
22. Jiricny,J. (2006) MutLalpha: at the cutting edge of mismatch repair. *Cell*, **126**, 239–241.
23. Pena-Diaz,J., Bregenhorn,S., Ghodgaonkar,M., Follonier,C., Artola-Boran,M., Castor,D., Lopes,M., Sartori,A.A. and Jiricny,J. (2012) Noncanonical mismatch repair as a source of genomic instability in human cells. *Mol. Cell*, **47**, 669–680.
24. Zlatanou,A., Despras,E., Braz-Petta,T., Boubakour-Azzouz,I., Pouvelle,C., Stewart,G.S., Nakajima,S., Yasui,A., Ishchenko,A.A. and Kannouche,P.L. (2011) The hMsh2-hMsh6 complex acts in concert with monoubiquitinated PCNA and Pol eta in response to oxidative DNA damage in human cells. *Mol. Cell*, **43**, 649–662.
25. Shibutani,T., Ito,S., Toda,M., Kanao,R., Collins,L.B., Shibata,M., Urabe,M., Koseki,H., Masuda,Y., Swenberg,J.A. *et al.* (2014) Guanine- 5-carboxylcytosine base pairs mimic mismatches during DNA replication. *Sci. Rep.*, **4**, 5220.
26. Fischer,F., Baerenfaller,K. and Jiricny,J. (2007) 5-Fluorouracil is efficiently removed from DNA by the base excision and mismatch repair systems. *Gastroenterology*, **133**, 1858–1868.
27. Pastoriza-Gallego,M., Armier,J. and Sarasin,A. (2007) Transcription through 8-oxoguanine in DNA repair-proficient and Csb(-)/Ogg1(-) DNA repair-deficient mouse embryonic fibroblasts is dependent upon promoter strength and sequence context. *Mutagenesis*, **22**, 343–351.
28. Bertrand,P., Tishkoff,D.X., Filosi,N., Dasgupta,R. and Kolodner,R.D. (1998) Physical interaction between components of DNA mismatch repair and nucleotide excision repair. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 14278–14283.
29. Smith,J.A., Waldman,B.C. and Waldman,A.S. (2005) A role for DNA mismatch repair protein Msh2 in error-prone double-strand-break repair in mammalian chromosomes. *Genetics*, **170**, 355–363.
30. Gu,Y., Parker,A., Wilson,T.M., Bai,H., Chang,D.Y. and Lu,A.L. (2002) Human MutY homolog, a DNA glycosylase involved in base excision repair, physically and functionally interacts with mismatch repair proteins human MutS homolog 2/human MutS homolog 6. *J. Biol. Chem.*, **277**, 11135–11142.
31. Pabla,N., Ma,Z., McIlhatton,M.A., Fishel,R. and Dong,Z. (2011) hMSH2 recruits ATR to DNA damage sites for activation during DNA damage-induced apoptosis. *J. Biol. Chem.*, **286**, 10411–10418.

32. Kemmerich,K., Dingler,F.A., Rada,C. and Neuberger,M.S. (2012) Germline ablation of SMUG1 DNA glycosylase causes loss of 5-hydroxymethyluracil- and UNG-backup uracil-excision activities and increases cancer predisposition of Ung-/-Msh2-/- mice. *Nucleic Acids Res.*, **40**, 6016–6025.
33. Krokan,H.E., Drablos,F. and Slupphaug,G. (2002) Uracil in DNA–occurrence, consequences and repair. *Oncogene*, **21**, 8935–8948.
34. Chou,K.M. and Cheng,Y.C. (2002) An exonucleolytic activity of human apurinic/apyrimidinic endonuclease on 3′ mispaired DNA. *Nature*, **415**, 655–659.
35. Aoufouchi,S., Faili,A., Zober,C., D'Orlando,O., Weller,S., Weill,J.C. and Reynaud,C.A. (2008) Proteasomal degradation restricts the nuclear lifespan of AID. *J. Exp. Med.*, **205**, 1357–1368.
36. Doseth,B., Visnes,T., Wallenius,A., Ericsson,I., Sarno,A., Pettersen,H.S., Flatberg,A., Catterall,T., Slupphaug,G., Krokan,H.E. *et al.* (2011) Uracil-DNA glycosylase in base excision repair and adaptive immunity: species differences between man and mouse. *J. Biol. Chem.*, **286**, 16669–16680.
37. Schanz,S., Castor,D., Fischer,F. and Jiricny,J. (2009) Interference of mismatch and base excision repair during the processing of adjacent U/G mispairs may play a key role in somatic hypermutation. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 5593–5598.
38. Krokan,H.E. and Bjoras,M. (2013) Base excision repair. *Cold Spring Harb. Perspect. Biol.*, **5**, a012583.
39. Bennett,S.E., Sung,J.S. and Mosbaugh,D.W. (2001) Fidelity of uracil-initiated base excision DNA repair in DNA polymerase beta-proficient and -deficient mouse embryonic fibroblast cell extracts. *J. Biol. Chem.*, **276**, 42588–42600.
40. Nilsen,H., Haushalter,K.A., Robins,P., Barnes,D.E., Verdine,G.L. and Lindahl,T. (2001) Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. *EMBO J.*, **20**, 4278–4286.
41. Wu,S.C. and Zhang,Y. (2010) Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.*, **11**, 607–620.
42. Franchini,D.M., Schmitz,K.M. and Petersen-Mahrt,S.K. (2012) 5-Methylcytosine DNA demethylation: more than losing a methyl group. *Annu. Rev. Genet.*, **46**, 419–441.
43. Di Noia,J.M. and Neuberger,M.S. (2007) Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.*, **76**, 1–22.
44. Wu,H. and Zhang,Y. (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell*, **156**, 45–68.
45. Shao,C., Xiong,S., Li,G.M., Gu,L., Mao,G., Markesbery,W.R. and Lovell,M.A. (2008) Altered 8-oxoguanine glycosylase in mild cognitive impairment and late-stage Alzheimer's disease brain. *Free Radic. Biol. Med.*, **45**, 813–819.
46. Chen,J., Miller,B.F. and Furano,A.V. (2014) Repair of naturally occurring mismatches can induce mutations in flanking DNA. *Elife*, **3**, e02001.
47. Constantin,N., Dzantiev,L., Kadyrov,F.A. and Modrich,P. (2005) Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J. Biol. Chem.*, **280**, 39752–39761.
48. Pfeifer,G.P. and Besaratinia,A. (2009) Mutational spectra of human cancer. *Hum. Genet.*, **125**, 493–506.
49. Cancer Genome Atlas Network. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, **487**, 330–337.
50. Burns,M.B., Temiz,N.A. and Harris,R.S. (2013) Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.*, **45**, 977–983.