Published in final edited form as:

Gastroenterology. 2008 August ; 135(2): 499–507. doi:10.1053/j.gastro.2008.04.035.

Characterization of mutant MUTYH proteins associated with familial colorectal cancer

Mohsin Ali^{a,c}, Hyeja Kim^a, Sean Cleary^a, Claire Cupples^b, Steven Gallinger^{a,*}, and Robert Bristow^{c,*}

^a Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, ON, Canada

^b Dept. of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada

^c Applied Molecular Oncology, Ontario Cancer Institute/Princess Margaret Hospital (University Health Network) and Departments of Radiation Oncology and Medical Biophysics, University of Toronto, Toronto, ON, Canada

Abstract

Background & Aims—The human *mutyh* gene encodes a base excision repair (BER) protein that prevents G:C to T:A transversions in DNA. Biallelic mutations in this gene are associated with recessively inherited familial colorectal cancer. The aim of this study was to characterize the functional activity of mutant-MUTYH and SNP-MUTYH proteins involving familial colorectal cancer.

Methods—MUTYH variants were cloned and assayed for their glycosylase and DNA binding activities using synthetic double-stranded oligonucleotide substrates by analyzing cleavage products by polyacrylamide gel electrophoresis.

Results—In this study, we have characterized nine missense/frameshift mutants, and two SNPs, for their DNA binding and repair activity *in-vitro*. Two missense mutants (R260Q and G382D) were found to be partially active in both glycosylase and DNA binding, while three other missense mutants (Y165C, R231H and P281L) were severely defective in both activities. All of the frameshift mutants (Y90X, Q377X, E466X and 1103delC) were completely devoid of both glycosylase and DNA binding activities. One SNP (V22M) showed the same activity as wild type MUTYH protein, but the other SNP (Q324H) was partially impaired in adenine removal.

Conclusions—This study of MUTYH mutants suggests that certain SNPs may be as partially dysfunctional in BER as missense-MUTYH mutants and lead to colorectal carcinogenesis.

Keywords

MUTYH; base excision repair; glycosylase; familial colorectal cancer

Endogenous reactive oxygen species (ROS) are continually produced during normal cellular physiology due to the production of metabolite by-products. ROS are also produced

*Corresponding author: Robert Bristow MD PhD FRCPC, Radiation Medicine Program, Princess Margaret Hospital (UHN), 610 University Avenue, Toronto, Ontario, CANADA M5G2M9, Tel: 416-946-2223; Fax: 416-946-4586; rob.bristow@rmp.uhn.on.ca.

No conflicts of interest exist for all authors

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

exogenously by either ionizing radiation or chemical carcinogens. ROS can cause a variety of DNA damage including double strand breaks (DSB), single strand breaks (SSB) and DNA base lesions^{1–3} such that the repair of ROS-induced DNA lesions is important for preventing mutations and maintaining the stability of the genome.

One such ROS-induced base lesion is 8-hydroxy guanine (GO) which is generated from guanine (G) or dGTP and leads to G:C to T:A and T:A to G:C transversions, respectively.^{4,5} In human cells, two base excision repair (BER) proteins, OGG1 and MUTYH, initiate the repair of GO lesions through their DNA glycosylase activity, while another protein, MTH1, hydrolyses oxidized dGTP. These glycosylase activities remove the damaged base and leaves behind an apurinic/apyrimidinic (AP) site in the DNA. In subsequent processes, the long patch BER pathway completes the repair process and this requires the APE1, PCNA, pol $\delta(\epsilon)$, FEN1, and DNA ligase I proteins.⁶ The OGG1, MTH1 and MUTYH enzymes work in concert: OGG1 removes GO from the DNA when paired with cytosine (C), whereas MTH1 prevents incorporation of GO in the DNA strand from the nucleotide pool by hydrolyzing 8-oxo-dGTP to 8-oxo-dGMP. If GO escapes from initial OGG1 action, DNA polymerase incorporates adenine (A) opposite to GO allowing MUTYH to instead excise A from the A:GO pair, this allows for a secondary attempt by OGG1 to repair the GO lesion.^{7–9} MUTYH can also excise other DNA lesions including 8-hydroxadenine (AO; from AO:G) 2-OH-A (from 2-OH-A:G) and A from A:G mispairs to a lesser extent.^{10–14}

Recently, it has been reported that germ-line mutations of the MUTYH gene are associated with familial colorectal cancer (CRC). While studying a British family affected with multiple colorectal polyposis, Al-Tassan *et al.* found two compound heterozygotes in the MUTYH gene in affected patients which substituted tyrosine at residue 165 to cysteine, and glycine at residue 382 to aspartic acid.¹⁵ These authors concluded that these somatic G:C to T:A transversions constitute a genetic signature of defective MUTYH protein activity in these CRC patients. Later, other groups confirmed the finding and concluded that in MUTYH-associated polyposis (*i.e.* MAP), biallelic mutations of the MUTYH leading to G:C to T:A transversions in the APC gene could drive colorectal epithelial cell genomic instability and increased cell proliferation in the epithelium of the colon.^{16–24} To date, more than eighty mutations have been described in the MUTYH gene within MAP patients with the Y165C and G382D variants as the most common documented mutations in Caucasian populations.²⁵

One method to study the function of colorectal cancer-associated MUTYH variant proteins mutations is to express them as recombinant proteins and assay for BER activity *in vitro* using synthetic DNA substrates containing A:GO and AO:G base mismatches. This has led to biochemical characterization of a number of CRC-related MUTYH variant proteins including: Y165C, G382D, R227W, V232F and R231L. For example, Al-Tassan *et al.* showed that the Y82C and G253D MutY proteins of *E. coli* (analogous to Y165C and G382D variants of human MUTYH, respectively) were partially defective in removing adenine from A:GO pairs.¹⁵ Wooden *et al.* reported that bacterially-expressed recombinant Y165C and G382D MUTYH proteins were completely devoid of glycosylase activity.²⁶ In another study, the R227W and V232F MUTYH mutants were also found to be partially or severely defective in both DNA substrate binding and glycosylase activity. Yet, the latter mutants still bound to hMutS α (a heterodimer of the human mismatch repair proteins hMSH2 and hMSH6 and a complex that can enhance MUTYH activity). In addition, both of these mutants failed to complement bacterial MutY deficiency when expressed in *E. coli* cells *in vivo*.²⁷

Very recently, the R231L MUTYH mutant was shown to be severely defective in DNA substrate binding and in adenine removal activity. While this variant showed intact binding activity with hMutS α , it did not complement MutY-deficiency in *E. coli*.²⁸ Parker *et al.* reported that lysates derived from lymphoblastoid cell lines from MAP-patients that expressed the

mutants, Y165C, G382D and 1103delC had lowered DNA binding and adenine cleavage.¹⁰ However, frameshift mutants of MUTYH have not previously been studied *in-vitro*.

We have recently completed a large multi-site population-based colorectal cancer case-control study (Cleary et al.; *Germline MYH mutations in Colorectal Cancer: A Multi-site, Population-based Case-control study*; Submitted; 2008) on a total of 3835 CRC cases and 2889 controls in which all subjects were screened for 9 known germline MUTYH mutations using mass spectrometry. DNA from subjects with at least one mutation was screened further by dHPLC/WAVE and sequencing analysis of WAVE variance. We found that twenty-seven cases and one control subject carried either homozygous or compound heterozygous MUTYH mutations; carriers were at increased CRC risk (adjusted OR (95% CI) = 18.3 (2.5–133.6)). Heterozygous MUTYH mutations were identified in 88 CRC cases and 44 controls; carriers were at increased risk of CRC (adjusted OR (95% CI) = 1.47 (1.02–2.13)).

Herein, we have conducted an *in vitro* study of these clinically-relevant MUTYH mutations as they directly relate to human cancer risk. We have selected five missense, four nonsense or frameshift mutations and two SNPs of MUTYH protein for *in vitro* characterization of their glycosylase and DNA binding activities. The activities of six of these mutants (*e.g.* R231H, P281L, Y90X, Q377X, E466X and R260Q) have not been previously studied. We show that two frameshift mutants are partially active in DNA glycosylase and binding activities while other seven variants are totally devoid of both activities.

Materials and Methods

Construction of vectors of N-terminal GST-tagged MUTYH proteins

The MUTYH gene was amplified using PCR from a Hela cell cDNA library (Stratagene, La Jolla, CA) using the primers:

5'CATATTGAATTCATGACACCGCTCGTCTCC3' and 5'
CATACGTCGACTCACTGGGCTGCACTGTTGA3'

with Phusion high-fidelity DNA polymerase (Invitrogen, Carlsbad, CA). Polyacrylamide gel electrophoresis (PAGE)-purified oligonucleotides were purchased from Operon (Huntsville, AL). Gel-purified products were digested with *EcoRI* and *Sall* restriction enzymes (Invitrogen, Carlsbad, CA) and purified using nucleotide purification spin columns (Qiagen, USA). The doubly-digested products were ligated into the pGEX-4T-1 vector (GE Health Sciences) between *EcoRI* and *Sall* sites using T4 DNA ligase (Invitrogen). A final pGEX-4T-1-MUTYH (WT) construct was transformed and grown in DH5 α cells. Both strands of the extracted plasmid were sequenced for the entire open reading frame (ORF) of the cloned MUTYH gene and the DNA sequences were confirmed to be the same as previously published sequences.⁷

Mutations were generated in the cloned WT MUTYH gene in pGEX-4T-1 using a site-directed mutagenesis kit as per the manufacturer's instructions (Promega). The base sequences on the both strands of each mutant were confirmed by DNA sequence analysis.

Protein Induction and Purification

Expression vectors were transformed into the BL21 CodonPlus (DE3) RIL *E. coli* strain. Cells were grown in LB medium with 100 μ g/ml of chloramphenicol and ampicillin at 37°C to 0.6 OD at 600 nm and cooled on ice. Protein production was initiated by adding 0.4 mM IPTG to the cells and continuing incubation at room temperature for 2h. Cells were harvested, washed with ice cold PBS and stored at –80 °C until protein purification.

Cell pellets were re-suspended into ice cold PBS (1/40th of culture volume). Lysozyme (1 mg/ml), DTT (5 mM), and protease inhibitor, were added to the cell suspension and incubated on

ice for 30 min followed by three freeze/thaw cycles and ultrasonic disruption. Cell lysates were then centrifuged at 15K rpm for 15 min and the clear supernatant was saved. A 50% slurry of glutathione sepharose beads (GE Health Sciences) was added to the clear lysate and rocked overnight at 4°C. Beads were pelleted by centrifugation at 2K rpm for 5 min and washed 4 times with ice cold PBS. Finally protein-bound beads were diluted with ice-cold PBS to make a 75% slurry. The concentrations of the partially-purified proteins were estimated by bicinchoninic acid (BCA) method and proteins stored at -80 °C until use. The purity of induced proteins was estimated by quantifying the band at the relevant MW compared to all other bands in the lane in the scanned image by image quant software (see lane 3 in figure 1a, and supplementary Table 1).

MUTYH Glycosylase Assay

A 5'-cy5-labeled 39-mer oligonucleotide containing A at the 21-mer position from the 5'-end was hybridized with its complementary strand in a buffer containing 20 mM Tris-HCl (pH8), 10 mM EDTA (pH8) and 150 mM NaCl to make the following duplex DNA substrate containing a A:GO mismatch:

5'cyTGAGACTGGCCAGCTAACTGA~~A~~ACTGATCATGCCTAGCGT
ACTCTGACCGGTCGATTGACX~~T~~GACTAGTACGGATCGCA

Where, X = 8-hydroxyguanine (GO).

The glycosylase assay was performed according to the procedure used by Bai *et al.*²⁷ with slight modification. Briefly, 100 fmol 5'-cy5 labeled duplex was incubated with 2 µg wild type or mutant MUTYH at 37 °C for 30 min in 10 µl buffer containing 50 mM EDTA (pH8), 500 µM ZnCl₂, 250 mM HEPES (pH7) and 1.5% glycerol. The reaction was stopped by adding 10 µl denaturing PAGE gel loading buffer containing 10 mM EDTA (pH8), 98% formamide, 10 mg/ml blue dextran and 200 mM NaOH followed by heating at 90 °C for 30 min. Cleavage products were separated using a 14% denaturing PAGE gel running at 2000V and 100W for 1h and fluorescent bands on the gel were visualized using a Typhoon Variable Imager (Amersham Biosciences). A schematic of this assay is shown in Supplementary Figure 1a.

DNA Binding Assay

For the DNA binding assay, GST-tag MUTYH proteins were eluted from the glutathione sepharose beads by an elution buffer containing 10 mM reduced glutathione and 50 mM Tris-HCl, pH 8 and concentrated using a centrifugal filter (Millipore). 100 fmol of Cy5 labeled 39-mer duplex DNA substrate containing A:GO base pair was incubated with 2 µg partially purified, beads-free proteins at 37°C for 30 min in 18 µl buffer containing 10 mM Tris-HCl (pH 7.6), 0.5 mM DTT, 0.5 mM EDTA and 1.5% glycerol. The reaction mixture was supplemented with 2 µl of loading buffer (50% glycerol and 10 µg/µl blue dextran) and analyzed by 6% non-denaturing PAGE gel in TBE buffer running at 100 V at 4°C. The fluorescent bands on the gel were visualized using a Typhoon Variable Imager (Amersham Biosciences).

Results

Expression and Purification

MUTYH is a human DNA glycosylase that removes A preferentially from A:GO pairs in DNA to prevent G:C to T:A transversions. In this study we characterized the adenine removal and DNA substrate binding activities of a series of MUTYH variants (*i.e.* Y90X, Y165C, R231H, R260Q, P281L, Q377X, G382D, E466X and 1103delC) which are derived from MAP-phenotype patients.

Initially, we expressed these mutant proteins *in vitro* from a pTNT expression vector (Promega) using TNT SP6 Quick Coupled Transcription/Translation system (Promega) and significant amount of proteins were found to be induced. However, these proteins were found to be inactive. However, in subsequent studies using an *in vivo* approach, we were able to successfully induce N-termini GST-tag proteins from the pGEX-4T-1 expression vectors in BL21 CodonPlus RIL *E. coli* host cells and partially purify them with glutathione sepharose beads. We therefore expressed mutants and WT proteins, two SNPs (V22M and Q324H) and enzyme-active center mutant (D222N) as a positive control, a pseudo-positive control, and a negative control, respectively.

The SDS-PAGE analysis of the GST-tag recombinant MUTYH wild type protein and its variants showed that the full-length (535aa) proteins WT, V22M, D222N, Q324H, Y165C, R231H, R260Q, P281L and G382D all had the expected and similar molecular weights based on the number of expressed codons. The C-termini-truncated Y90X, Q377X, E466X and 1103delC MUTYH mutant proteins were expressed as proteins with expected lower molecular weights of 37, 68, 78 and 70 kDa, respectively. The specificity of protein induction was confirmed by Western blot analyses using polyclonal antibodies against MUTYH or GST (Figure 1a, b and c).

In order to ensure that the exogenous partially-purified proteins were not contaminated with endogenous bacterial homolog MutY, we purified the WT, V22M and D222N MUTYH proteins from un-induced and IPTG-induced cells and assayed their activities on duplex DNA substrates containing a A:GO mismatch. We observed that the proteins from un-induced cells harboring the plasmids pGEX-4T-1-MUTYH (WT) and pGEX-4T-1-MUTYH (V22M) possessed slight adenine removal activity (*c.f.* lane 4 and 6 in Supplementary Figure 1b). This activity could be from the co-eluted bacterial MutY protein from the *E. coli* host cells (*i.e.* the BL21CodonPlusRIL strain is not MutY-deficient) or from leaky MUTYH protein. However, the D222N variant is an inactive protein because of a mutated enzyme-active site²⁶ and it should not exhibit any glycosylase activity. Indeed, the D222N variant partially purified from un-induced or IPTG-induced cells was found to be completely inactive (*c.f.* lane 2 and 3 in Supplementary Figure 1b). This suggests that the glycosylase activity detected in the un-induced cells was from the leaky WT or V22M protein. Moreover, SDS-PAGE analysis of the partially purified proteins did not show any band at 39 kDa corresponded to mol. wt. of MutY. Therefore, we conclude that the N-termini GST-tag recombinant WT and mutant MUTYH proteins partially purified from the BL21 CodonPlus (DE3) RIL host cells are free from MutY contamination. This GST-tagged recombinant WT type MUTYH protein was almost as active as standard bacterial homolog MutY protein.

Adenine Removal Activity (glycosylase activity) Assay

Figure 2 shows a typical *in vitro* glycosylase assay profile of wild type or mutant MUTYH proteins on the synthetic duplex DNA substrates containing a A:GO pair. Similar to the bacterial homolog MutY control, the WT, V22M and Q324H MUTYH proteins were all able to cleave substrates containing a A:GO mismatch (*c.f.* lanes C4, C5 and C6, respectively with lane C3 in Figure 2). The major slower migrating band is a 20-mer α , β -unsaturated aldehyde generated by β -elimination of the AP site by NaOH treatment. The minor product migrating faster than the major product is a 20-mer product with 5'-phosphate and produced from the β -elimination product by δ -elimination. Ohtsubo *et al.* has reported similar products from the substrates incubated with wild type MUTYH.¹²

The intensity of the bands from the cleavage products generated by two missense MUTYH mutants (*i.e.* R260Q and G382D) was much less compared to the bands from wild type MUTYH protein (*c.f.* lanes C4 with 4 and 7 in Figure 2). The other missense mutants (Y165C, R231H and P281L) and the frameshift mutants (Y90X, Q377X, E466X and 1103delC)

generated no products (see lanes 1, 2, 3, 4, 5, 8 and 9 in Figure 2) even after two hours incubation.

The time course assay of A:GO repair activity of WT and SNP-V22M MUTYH protein indicated that the generated products reached a maximum within 16 min and plateaued at periods of up to 180 min (see Figure 3a). The rate constants k_2 for SNP-Q324H MUTYH and WT were found to be $4.288 \pm 0.7831 \text{ min}^{-1}$ and $4.507 \pm 0.5812 \text{ min}^{-1}$, respectively (Figure 3b). We reproducibly detected only 64 % activity in the SNP-Q324H MUTYH compare to the WT at 30 min. The glycosylase activities of both missense mutants (R260Q and G382D) were similar - only 21% of the WT at saturation - but their rate constants at the linear range of the activity curve were different: $1.28 \pm 0.122 \text{ min}^{-1}$ and $1.033 \pm 0.0979 \text{ min}^{-1}$, respectively (Figure 3a and b). In contrast, the missense mutants (Y165C, R231H and P281L) and frameshift mutants (Y90X, Q377X, E466X and 1103delC) produced no products, even after 10 hours of incubation.

DNA Binding Activity Assay

The binding of repair proteins to DNA substrates is a crucial step in the repair process. Human cells recruit MUTYH protein to the A:GO site where it binds tightly to DNA and catalyzes the removal of A from A:GO mismatch. MUTYH and its bacterial homolog MutY, remain bound to the substrate even after removal of A, until displaced by other proteins that are subsequently recruited to complete the repair process. In this work, we also examined the effect of amino acid substitution in mutant MUTYH on DNA binding activity. Figure 4a shows a typical gel shift assay for wild type and mutants of MUTYH.

The WT and SNP MUTYH proteins formed two complexes with the duplex DNA substrates containing A:GO mismatch, whereas the MutY bacterial homolog formed only one complex (see lanes 2, 3 and 4, in Figure 4a). The substrates bound to WT MUTYH protein as complexes I and II were found to be 52% (26 fmol/ μg protein) and 22% (11 fmol/ μg protein), respectively. The missense mutants (R260Q and G382D) also formed complexes I and II with the DNA substrates. The DNA substrates bound to R260Q protein were estimated to be 18% (9 fmol/ μg protein) and 8% (4 fmol/ μg protein), respectively. On the other hand, 22% (11 fmol/ μg protein) and 14% (7 fmol/ μg protein) substrates were bound with G382D as complexes I and II, respectively (lane 9 and 12 in Figure 4a and b). No such complexes were formed with the missense mutants (Y165C, R231H and P281L) or with the frameshift mutants (Y90X, Q377X, E466X and 1103delC) (Figure 4a, lane 6–8, 10, 11, 13 and 14).

Discussions

Compound heterozygotes in *mutyh* gene have been shown to be associated with familial colorectal carcinoma in human. In this work, we studied nine bacterially expressed mutant MUTYH proteins for their DNA glycosylase and binding activities. In vitro assay using synthetic DNA substrates revealed that missense mutants (R260Q and G382D) are partially active in glycosylase activity (rate constants k_2 , $1.28 \pm 0.122 \text{ min}^{-1}$ and $1.033 \pm 0.0979 \text{ min}^{-1}$, respectively, compare to $4.507 \pm 0.5812 \text{ min}^{-1}$ of WT) and DNA binding activity (26% and 36% of the substrates as complex I and II with R260Q and G382D, respectively.) whereas mutants (Y90X, Y165C, R231H, P281L, Q377X, E466X and 1103delC) are unable to generate any detectable cleavage products from the substrates containing A:GO mismatch or to bind to the substrates (See Figure 2, 3 and 4).

Previously, Al-Tassan *et al.* showed that the *E. coli* mutant Y82C (analog to human Y165C) exhibits barely detectable glycosylase activity, whereas the mutant G253D (analogous to the human G382D variant) cleaves adenine from an A:GO mismatch almost as efficiently as WT protein.¹⁵ And yet, in another study, the murine mutant G365D protein (also corresponding to

the human G382D MUTYH variant) was found to be fully active in removing A from a A:GO pair.¹³ Wooden *et al.* characterized bacterially expressed GST-tag mutants (Y165C and G382D) and found completely inactive in glycosylase activity.²⁶ When taken together, these previous data and our current data suggests strongly that frameshift mutants are completely defective in enzymatic activities due to loss of the C-terminal domain. Our study also shows that a mutation anywhere in the catalytic domain can impair enzyme activity (even at a distance from residue 222). This allows for the prediction of activity of clinical mutations in MYH based on sequence data.

Importantly, our DNA binding results are consistent with the relative glycosylase activity of mutant MUTYH proteins. Frameshift mutants (Y90X, Q377X, E466X and 1103delC) were unable to remove A from the substrates (although they possess intact catalytic domains), probably because of defective DNA-binding activity. Why the DNA-binding activity of the full-length (535aa) missense MUTYH variants (Y165C, R231H and P281L) was severely defective is unclear at this point and this deserves further study.

We observed that, in spite of similar substrates binding activity (74% and 68%, WT and Q324H, respectively), the SNP Q324H is only 64% active at saturation in excising A ($k_2 = 2.971 \pm 0.2172 \text{ min}^{-1}$) from the substrates compare to the WT ($k_2 = 4.507 \pm 0.5812 \text{ min}^{-1}$). In contrast, Shinmura *et al.* found Q324H to be fully active as WT.²⁹ Interestingly, Yuan *et al.* has just reported that SNP Q324H is strongly associated with familial colorectal cancer among African-Americans.³⁰ Therefore, on the basis of our results, we concluded in addition to the germ-line mutations, SNPs should also be studied for their possible involvement in MAP.

The MUTYH crystal structure has not yet been solved. However, the observed biochemical activity of the MUTYH variants in this study may be explained by the study of the amino acid sequence of the *E. coli* MutY protein, which is 41% identical with MUTYH.^{7,26–28} MutY has a catalytic domain consisting of helix-hairpin-helix (HhH), pseudo HhH and iron-sulfur cluster [4Fe-4S], and a characteristic C-terminal domain, that is not found in other HhH-superfamily BER proteins.^{31–33}

Positions of the amino acid substitution of the mutants studied in this work are shown in Figure 5. All the missense mutants, except G382D, would lie within this catalytic domain. These amino acid residues are highly conserved among the human, *E. coli*, murine and *S. pombe* MutY homologs. While the C-terminal domain of *E. coli* MutY protein is not needed for its adenine removal activity, it strongly interacts with the GO-containing strand to flip the target adenine out of the helix. In this way, A is buried into a pocket formed between 6-helix barrel module and [4Fe-4S] module and subsequently β -glycosidic bond between A and pentose sugar moiety is hydrolyzed by the protein's glycosylase activity.^{31,32} The mutation in the G382D variant lies within the C-terminal domain which is required for GO recognition. In support of this theory, it has been shown previously that the removal of this domain from MutY protein drastically reduces its adenine incision activity from A:GO pair, but not from A:G mismatch.³³ Therefore, the inactivity of our frameshift variants can be explained by the fact that the Y90X mutant loses both catalytic domain and the C-terminal domain whereas the other mutants (Q377X, E466X and 1103delC) lost only the catalytic domain (Figure 5). Indeed, it was previously shown that the C-terminal domain truncated MutY is not only defective in removing A but also in binding substrate containing A:GO.^{33,34}

A mutation in the catalytic domain might render the protein partially or fully inactive where as the lost of C-terminal domain could impaired binding to the substrate containing A:GO. Protein products of SNPs are generally active as WT, but we observed that the SNP Q324H is 36 % less active than WT. This finding is consistent with the recent observation by Yuan *et al.* who have reported that the Q342H variant SNP is strongly associated with familial

colorectal cancer among African-American population.³⁰ The second SNP V22M was found to be as active as the WT ($k_2 = 4.288 \pm 0.7831$ and $4.507 \pm 0.5812 \text{ min}^{-1}$, respectively). In summary, we have characterized *in vitro* a large series of frameshift, missense mutants and SNP of MUTYH in one laboratory setting that are clinically associated with increased CRC risk. The results are summarized in Table 2. The two missense variants (R260Q and G382D) were partially active in DNA binding and BER activities, while three missense variants (Y165C, R231H and P281L) and all four frameshift variants (Y90X Q377X, E466X and 1103delC) were dysfunctional in both activities.

Adding further complexity is that there may be cross-talk between repair pathways in preventing colorectal carcinogenesis. In human cells, three BER proteins (OGG1, MTH1 and MUTYH) and three MMR proteins (MSH2, MSH6 and MLH1) guard genome from mutagenic DNA base lesion 8-oxoG.³⁵ Their functions have been well studied both *in vitro* and *in vivo* and indeed MMR and MYH proteins interact biochemically.^{12,13,29,36–40} As such, mutations in the MUTYH catalytic domain that renders the protein fully or partially inactive could abrogate normal interactions between MYH and MMR proteins. In this study, we further confirmed that even mutations outside the catalytic domain can deactivate protein function in a partial or complete manner. This finding could explain the observed accumulation of G:C to T:A transversions in APC gene which are associated with CRC.

MUTYH protein is a second defence against deleterious 8-oxoG lesions that escape initial OGG1-based surveillance and repair. Our data suggests that mutations and SNPs in MUTYH could give rise to defective proteins that might affect genomic stability. At present, there are no reports of dysfunctional OGG1 or MTH1 as determinants of colorectal carcinogenesis. Given mutations within MMR genes are also linked to hereditary non-polyposis colorectal cancer (HNPCC),^{41,42} it will be of interest to closely study the interaction of wild type and mutant MUTYH proteins and MMR proteins in biochemical and cellular system to further delineate interactions between BER and MMR in colorectal carcinogenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Cancer Institute, National Institutes of Health under RFA #CA-95-011 grant number. U01-CA74783 to SG and grant number 1203-1 to RGB. The efforts of all the Ontario Familial Colon Cancer Registry (OFCCR) study staff are gratefully acknowledged. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the Consortium of Familial Registries. Work in CGC's lab was supported by a grant from the Canadian Institutes for Health Research. RGB is a Canadian Cancer Society Research Scientist.

Abbreviations

ROS	reactive oxygen species
DSB	double strand break
SSB	single strand break
GO	8-hydroxyguanine

BER	base excision repair
AP	apurinic/aprimidinic
AO	8-hydroxyadenine
CRC	colorectal cancer
APC	adenomatous polyposis coli
MAP	MUTYH-associated polyposis
PAGE	polyacrylamide gel electrophoresis
WT	wild type
SNP	single nucleotide polymorphism
aa	amino acid
ORF	open reading frame

References

1. Olinski R, Gackowski D, Foksinski M, et al. Oxidative DNA damage: assessment of the role in carcinogenesis, atherosclerosis, and acquired immunodeficiency syndrome. *Free Radic Biol Med* 2002;33:192–200. [PubMed: 12106815]
2. Frelon S, Douki T, Ravanat JL, et al. High-performance liquid chromatography--tandem mass spectrometry measurement of radiation-induced base damage to isolated and cellular DNA. *Chem Res Toxicol* 2000;13:1002–10. [PubMed: 11080049]
3. Dreher D, Junod AF. Role of oxygen free radicals in cancer development. *Eur J Cancer* 1996;32A:30–8. [PubMed: 8695238]
4. Cheng KC, Cahill DS, Kasai H, et al. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *J Biol Chem* 1992;267:166–72. [PubMed: 1730583]
5. Shibutani S, Takeshita M, Grollman AP. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* 1991;349:431–4. [PubMed: 1992344]
6. Parker A, Gu Y, Mahoney W, et al. Human homolog of the MutY repair protein (hMYH) physically interacts with proteins involved in long patch DNA base excision repair. *J Biol Chem* 2001;276:5547–55. [PubMed: 11092888]
7. Slupska MM, Baikalov C, Luther WM, et al. Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutY gene whose function is required for the repair of oxidative DNA damage. *J Bacteriol* 1996;178:3885–92. [PubMed: 8682794]
8. Ide H, Kotera M. Human DNA glycosylases involved in the repair of oxidatively damaged DNA. *Biol Pharm Bull* 2004;27:480–5. [PubMed: 15056851]

9. Michaels ML, Miller JH. The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J Bacteriol* 1992;174:6321–5. [PubMed: 1328155]
10. Parker AR, Sieber OM, Shi C, et al. Cells with pathogenic biallelic mutations in the human MUTYH gene are defective in DNA damage binding and repair. *Carcinogenesis* 2005;26:2010–8. [PubMed: 15987719]
11. Gu Y, Lu AL. Differential DNA recognition and glycosylase activity of the native human MutY homolog (hMYH) and recombinant hMYH expressed in bacteria. *Nucleic Acids Res* 2001;29:2666–74. [PubMed: 11410677]
12. Ohtsubo T, Nishioka K, Imaiso Y, et al. Identification of human MutY homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. *Nucleic Acids Res* 2000;28:1355–64. [PubMed: 10684930]
13. Ushijima Y, Tominaga Y, Miura T, et al. A functional analysis of the DNA glycosylase activity of mouse MUTYH protein excising 2-hydroxyadenine opposite guanine in DNA. *Nucleic Acids Res* 2005;33:672–82. [PubMed: 15681617]
14. Slupska MM, Luther WM, Chiang JH, et al. Functional expression of hMYH, a human homolog of the *Escherichia coli* MutY protein. *J Bacteriol* 1999;181:6210–3. [PubMed: 10498741]
15. Al-Tassan N, Chmiel NH, Maynard J, et al. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nat Genet* 2002;30:227–32. [PubMed: 11818965]
16. Sieber OM, Lipton L, Crabtree M, et al. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *N Engl J Med* 2003;348:791–9. [PubMed: 12606733]
17. Sampson JR, Dolwani S, Jones S, et al. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet* 2003;362:39–41. [PubMed: 12853198]
18. Jones S, Emmerson P, Maynard J, et al. Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C-->T:A mutations. *Hum Mol Genet* 2002;11:2961–7. [PubMed: 12393807]
19. Wang L, Baudhuin LM, Boardman LA, et al. MYH mutations in patients with attenuated and classic polyposis and with young-onset colorectal cancer without polyps. *Gastroenterology* 2004;127:9–16. [PubMed: 15236166]
20. Croitoru ME, Cleary SP, Di Nicola N, et al. Association between biallelic and monoallelic germline MYH gene mutations and colorectal cancer risk. *J Natl Cancer Inst* 2004;96:1631–4. [PubMed: 15523092]
21. Eliason K, Hendrickson BC, Judkins T, et al. The potential for increased clinical sensitivity in genetic testing for polyposis colorectal cancer through the analysis of MYH mutations in North American patients. *J Med Genet* 2005;42:95–6. [PubMed: 15635083]
22. Isidro G, Laranjeira F, Pires A, et al. Germline MUTYH (MYH) mutations in Portuguese individuals with multiple colorectal adenomas. *Hum Mutat* 2004;24:353–4. [PubMed: 15366000]
23. Venesio T, Molatore S, Cattaneo F, et al. High frequency of MYH gene mutations in a subset of patients with familial adenomatous polyposis. *Gastroenterology* 2004;126:1681–5. [PubMed: 15188161]
24. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell* 1996;87:159–70. [PubMed: 8861899]
25. Cheadle JP, Sampson JR. MUTYH-associated polyposis--from defect in base excision repair to clinical genetic testing. *DNA Repair (Amst)* 2007;6:274–9. [PubMed: 17161978]
26. Wooden SH, Bassett HM, Wood TG, et al. Identification of critical residues required for the mutation avoidance function of human MutY (hMYH) and implications in colorectal cancer. *Cancer Lett* 2004;205:89–95. [PubMed: 15036665]
27. Bai H, Jones S, Guan X, et al. Functional characterization of two human MutY homolog (hMYH) missense mutations (R227W and V232F) that lie within the putative hMSH6 binding domain and are associated with hMYH polyposis. *Nucleic Acids Res* 2005;33:597–604. [PubMed: 15673720]
28. Bai H, Grist S, Gardner J, et al. Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis. *Cancer Lett* 2007;250:74–81. [PubMed: 17081686]

29. Shinmura K, Yamaguchi S, Saitoh T, et al. Adenine excisional repair function of MYH protein on the adenine:8-hydroxyguanine base pair in double-stranded DNA. *Nucleic Acids Res* 2000;28:4912–8. [PubMed: 11121482]
30. Yuan, Z.; Shin, J.; Fordyce, K., et al. Q324H, A germline variant of the MUTY homolog (MYH) gene is strongly associated with early onset adenomatous colorectal polyp formation among African Americans. The 2nd biennial scientific meeting of international society for gastrointestinal hereditary tumours; Pacifico-Yokohama, Japan. 2007.
31. Guan Y, Manuel RC, Arvai AS, et al. MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat Struct Biol* 1998;5:1058–64. [PubMed: 9846876]
32. Fromme JC, Banerjee A, Huang SJ, et al. Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature* 2004;427:652–6. [PubMed: 14961129]
33. Gogos A, Cillo J, Clarke ND, et al. Specific recognition of A/G and A/7,8-dihydro-8-oxoguanine (8-oxoG) mismatches by *Escherichia coli* MutY: removal of the C-terminal domain preferentially affects A/8-oxoG recognition. *Biochemistry* 1996;35:16665–71. [PubMed: 8988002]
34. Noll DM, Gogos A, Granek JA, et al. The C-terminal domain of the adenine-DNA glycosylase MutY confers specificity for 8-oxoguanine. adenine mispairs and may have evolved from MutT, an 8-oxo-dGTPase. *Biochemistry* 1999;38:6374–9. [PubMed: 10350454]
35. Colussi C, Parlanti E, Degan P, et al. The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. *Curr Biol* 2002;12:912–8. [PubMed: 12062055]
36. Yeh YC, Chang DY, Masin J, et al. Two nicking enzyme systems specific for mismatch-containing DNA in nuclear extracts from human cells. *J Biol Chem* 1991;266:6480–4. [PubMed: 1848856]
37. Radicella JP, Dherin C, Desmaze C, et al. Cloning and characterization of hOGG1, a human homolog of the OGG1 gene of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 1997;94:8010–5. [PubMed: 9223305]
38. Rosenquist TA, Zharkov DO, Grollman AP. Cloning and characterization of a mammalian 8-oxoguanine DNA glycosylase. *Proc Natl Acad Sci U S A* 1997;94:7429–34. [PubMed: 9207108]
39. Sakumi K, Furuichi M, Tsuzuki T, et al. Cloning and expression of cDNA for a human enzyme that hydrolyzes 8-oxo-dGTP, a mutagenic substrate for DNA synthesis. *J Biol Chem* 1993;268:23524–30. [PubMed: 8226881]
40. McGoldrick JP, Yeh YC, Solomon M, et al. Characterization of a mammalian homolog of the *Escherichia coli* MutY mismatch repair protein. *Mol Cell Biol* 1995;15:989–96. [PubMed: 7823963]
41. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature* 2001;411:366–74. [PubMed: 11357144]
42. Kolodner RD. Mismatch repair: mechanisms and relationship to cancer susceptibility. *Trends Biochem Sci* 1995;20:397–401. [PubMed: 8533151]

a

Fig. 1. Analysis of induction and purification of recombinant N-termini tagged GST-MUTYH mutant

(a) Cell lysates were resolved by a 10% SDS-PAGE followed by coomassie-blue staining to visualize expressed proteins. For each MUTYH protein: lane 1, lysate from un-induced cells; lane 2, lysate from IPTG-induced cells; lane 3, partially-purified GST-tagged protein. (b) Western blot analysis of purified recombinant GST-tagged MUTYH and its mutants using rabbit polyclonal antibody against MUTYH (C-termini-truncated variants Y90X, Q377X, E466X and 1103delC were not detected as the antibody is corresponded to the amino acids 513–546 at the C-terminal of the MUTYH protein) or (c) mouse monoclonal antibody against GST. Fluor-conjugated secondary antibody bound to the antibody-antigen complex was

detected by infrared image analyzer (LICOR). In **(b)** and **(c)**, lane assignment is as follows: lane 1, WT; lane 2, V22M; lane 3, D222N; lane 4, Q324H; lane 5, Y90X; lane 6, Q377X; lane 7, 1103delC; lane 8, Y165C; lane 9, R231H; lane 10, P281L; lane 11, R260Q; lane 12, G382D; lane 13, E466X; arrows indicate positions of the induced protein bands.

Fig. 2.

Enzymatic activity of MUTYH mutants on 39 mer duplex DNA substrate containing A:GO mismatch. 100 fmol 5'-cy5 labeled duplex substrate was incubated with 2 μ g WT or mutants MUTYH at 37 °C for 30 min and analyzed in a denaturing PAGE to detect cleavage products. Lane assignment is as follows: lane S, 20p-mer standard; lane C1, substrate only; lane C2, GST beads; lane C3, MutY(positive control); lane C4, WT (positive control); lane C5, V22M (pseudo positive control); C6, Q324H (pseudo positive control); C7, D222N (negative control); lane 1, Y90X; lane 2, Y165C; lane 3, R231H; lane 4, R260Q; lane 5, P281L; lane 6, Q377X; lane 7, G382D; lane 8, E466X and lane 9, 1103delC.

Fig. 3.

Time course activity of the mutant MUTYH proteins was assayed to compare their glycosylase activity with WT. 1000 fmol 5'-cy5 labeled duplex was incubated with 10 μ g WT or mutant MUTYH in 100 μ l buffer at 37 $^{\circ}$ C, removed 10 μ l at 0, 1, 2, 4, 8, 16, 32, 60, 120 and 180 min, and immediately stopped the reaction by adding 10 μ l gel loading buffer with 200 mM NaOH followed by heating at 90 $^{\circ}$ C for 30 min. The reaction mix was fractionated and visualized by Typhoon variable imager. **(a)** Major product I was quantified by ImageQuant software. Product (mean $\pm$ standard deviation) of WT and mutants MUTYH was determined from three independent experiments. **(b)** Rate constants (mean $\pm$ standard deviation) determined by linear regression.

Fig. 4. Binding of WT and mutants MUTYH proteins with DNA substrates containing A:GO base pair. 100 fmol of 5'-Cy5 labeled 39-mer substrates was incubated with 2 μg partially purified beads-free WT or mutant MUTYH proteins at 37°C for 30 min in binding buffer. **(a)** DNA-protein complexes were analyzed by 6% non-denaturing PAGE. Lane 1, substrate only; lane 2, MutY; lane 3, WT; lane 4, Q324H; lane 5, D222N; lane 6, Y90X; lane 7, Y165C; lane 8, R231H; lane 9, R260Q; lane 10, P281L; lane 11, Q377X; lane 12, G382D; lane 13, E466X; lane 14, 1103delC. **(b)** DNA-protein Complex I and II were quantified by ImageQuant software. Product (mean ± standard deviation) of WT and mutants MUTYH was determined from three independent experiments.

MUTH	1	MTPLVSRSLRLMAIDRKFPRAAVSGHRKQAASQEGRRQKHAKNNISQAKPESACDGLARQPEEVVLQASVSSYHLFR-DVAEV	79
MutY	1	-----MQA	3
		Y90X	
MUTH	80	TAFRGSLLSWYDQE-KRDLPWPPRAEDMDLD-----RRAYAVWSEVNLQQTQVATVINYYTGWMQKMPILQQL	148
MutY	4	SQFSAQVLDWYCKYGRKTLNQ-----ID-----KTPYKWLSEVNLQQTQVATVIPYFEAFYARFPTVTUL	65
		DNA minor-groove reading motif	
		Y165C	
MUTH	149	ASASLE-EVWQLWAGLGYYSRGRRLQEGARKVVEELGGHMPRTASTLQQLLPQVGRYTAGAIIASIAFGQATGWDDQNVAR	227
MutY	66	ANAPLD-EVLHLNTGLGYVARARNLHQAQQVATLGGCKFPETFEVA-ALPGVGRSTAGAILSLSLGKHFPILQNVKR	143
		pseudo-HhH motif	
		HhH motif	
		active center	
		R231H	
		R260Q	
		P281L	
MUTH	228	VLCEVRAIGADPSSTLVSQQLWGLAQLVDPARPGDFNQAMELGATVCTPQRFPLCSQCPVESLCRARQVCEQILLASG	307
MutY	144	VLARCYAVSGWPGKCEVENKLSLSECTPANGVERFNQAMDLGAMICTRSKPKCSLCP LQNGQIA-----	210
		adenine recognition motif (4Fe-4S) binding motif	
		Q377X G382D	
MUTH	308	SLSGSPDVEECAPNTGQCHLCPLPSEPMQDTLGVVNFPRKSRKQPEESSATCVLEQPGALG-AQILIVQ-APNGLLN	385
MutY	211	-----AAND-----SMALYCKK--PKOT--LEETGY-FLLLQHD--EV-----LLAQ-RPPGGLNG	256
		1103delC	
MUTH	386	GLWEFPVTVWEPSEQLQKALLQELQR---WAGP---LPATHLRHLGEVVTFSHIKLYQVYGLALEGQTFVTIVPPG	458
MutY	257	GLYCFQCFADP--ES--LRQWLACRQ-----IAADNLQELAFRHTFSHFHLDIVPMNLVYS--SFTGCMDEG	256
		E466X	
MUTH	459	ARMLTQEEFHTAANSTAMSKYFVYQGQQPGTCMGSKASQVSPCSRNKPPMGQQVLNIFRSHISTDAHSLSNSAQ	535
MutY	319	NALWYNLAQ--PPSYGLAAP-VERLLQQLRTGAPV-----	350

Fig. 5. Alignment of MUTH and MutY amino acids sequences. Un-shaded, catalytic domain; shaded, C-terminal domain; red letter, conserved amino acid residue.

Table 1

Summary of activities of WT and variant MUTYH proteins

Mutant	Type of mutation	Glycosylase activity	DNA binding activity
Y90X	Frameshift	Severely defective	Severely defective
Q377X	Frameshift	Severely defective	Severely defective
E466X	Frameshift	Severely defective	Severely defective
I103 C	Frameshift	Severely defective	Severely defective
R260Q	Missense	Partial ($k_2 = 1.28 \pm 0.12 \text{ min}^{-1}$)	Partial 26%
G382D	Missense	Partial ($k_2 = 1.033 \pm 0.098 \text{ min}^{-1}$)	Partial 36%
Y165C	Missense	Severely defective	Severely defective
R231H	Missense	Severely defective	Severely defective
P281L	Missense	Severely defective	Severely defective
Q324H	SNP	Partial ($k_2 = 2.97 \pm 0.22 \text{ min}^{-1}$)	Same as WT
V22M	SNP	Full ($k_2 = 4.29 \pm 0.78 \text{ min}^{-1}$)	Not tested
WT		Full ($k_2 = 4.51 \pm 0.58 \text{ min}^{-1}$)	Full