nature chemistry

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41557-018-0068-x

In the format provided by the authors and unedited.

A human MUTYH variant linking colonic polyposis to redox degradation of the $[4Fe4S]^{2+}$ cluster

Kevin J. McDonnell[1,9], Joseph A. Chemler[2,9], Phillip L. Bartels[3,9], Elizabeth O'Brien[3], Monica L. Marvin[4], Janice Ortega[5], Ralph H. Stern[6], Leon Raskin[7], Guo-Min Li[5], David H. Sherman[2,8*], Jacqueline K. Barton[3*] and Stephen B. Gruber[1*]

[1]University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA, USA. [2]Life Sciences Institute, University of Michigan, Ann Arbor, MI, USA. [3]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA. [4]Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA. [5]Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX, USA. [6]Division of Molecular Medicine and Genetics, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA. [7]Amgen Inc., Thousand Oaks, CA, USA. [8]Departments of Medicinal Chemistry, Chemistry and Microbiology & Immunology, University of Michigan, Ann Arbor, MI, USA. [9]These authors contributed equally to this work: Kevin J. McDonnell, Joseph A. Chemler, Phillip L. Bartels. *e-mail: davidhs@umich.edu; jkbarton@caltech.edu; sgruber@med.usc.edu

Supplementary Information

A Human MUTYH Variant Linking Colonic Polyposis to Redox Degradation of the $[4Fe4S]^{2+}$ Cluster

Kevin J. McDonnell[1,†], Joseph A. Chemler[2,†], Phillip L. Bartels[3,†], Elizabeth O'Brien[3], Monica L. Marvin[4,5], Janice Ortega[1], Ralph H. Stern[5], Leon Raskin[5], Guo-Min Li[1], David H. Sherman[2,6,*], Jacqueline K. Barton[3,*] and Stephen B. Gruber[1,*]

Table of Contents

I. Methods

A. Determination of Trans Chromosomal Configuration of MUTYH Gene Variants

Germline DNA was amplified using the polymerase chain reaction (PCR) to generate a 935 base pair amplicon that includes the open reading frame positions c.918C>G (p.C306W) and c.1187G>A (p.G396D). The PCR reaction used the forward primer 5′-CCA GGA GAT TTC AAC CAA GC-3′ and the reverse primer 5′ -AAG GGT CAA GGG GTT CAA AT-3′. The c.1187G>A mutation creates a unique BglII restriction endonuclease site which allowed generation of a 719 base pair fragment from the parent 935 base pair amplicon. The shorter 719 base pair fragment was resolved using agarose gel electrophoresis, purified and its DNA sequence determined (University of Michigan (U-M) Sequencing Core) to establish the identity of the c.918 position.

B. Identification of APC gene G:C → T:A Transversions

Tumor DNA was extracted from a formalin-fixed, paraffin-embedded colonic adenoma originating from the proband using the RecoverAll Total Nucleic Acid isolation kit (Ambion). A portion of the mutation cluster region of the *APC* gene[1] was amplified using PCR with the forward primer 5′-TGC CAC AGA TAT TCC TTC ATC A-3′ and the reverse primer 5′-CAT GGT TTG TCC AGG GCT AT-3′. The PCR product was subsequently sequenced (U-M Sequencing Core).

C. Cloning of wild type and Mutant MUTYH expression plasmids

A plasmid containing the open reading frame for the beta3 isoform of *MUTYH* (NM_001048174.1) was obtained from OriGene (Catalog #: RC201376-OR, Rockville, MD.) and used as a template for the cloning of derivative constructs. Mutant *MUTYH* construct synthesis was accomplished using PCR-based site-directed mutagenesis of wild type *MUTYH*. The wild type and mutants were then cloned as maltose binding protein (MBP) fusions into pMCSG19[2] between the KpnI and XbaI restriction sites to increase protein solubility[3]. The plasmids were further modified by removing the first fourteen codons encoding the mitochondrial recognition sequence, which alleviated heterologous protein toxicity. Furthermore, to decrease the capture of truncated heterologous protein, the N-terminus His_6 tag located between the MBP sequence and the TEV cleavage site was removed and a C-terminus His_{10} tag was attached with a flexible $(SG)_7$ linker to increase solvent exposure.

D. Preparation of Protein

Heterologous MUTYH proteins in *Escherichia coli* strain BL21(DE3) were initially purified in accordance with previously published protocols using nickel affinity chromatography with 1 mM DTT[4]. To improve yields and purity, the expression plasmids were transformed into the *E. coli* expression strain for toxic proteins, BL-AI (Invitrogen) also harboring the rare codon plasmid pRARE2-CDF[5]. One liter of fresh Terrific Broth modified with 4% glycerol and 50 µg/mL of antibiotics (ampicillin and streptomycin) in a three L baffled flask was inoculated with 25 mL of overnight cultures in the same medium. Cultures were grown at 37 ºC in a horizontal shaker at 175 rpm until the OD_{600} reached approximately 2.5. The temperature was adjusted to 15 ºC, and after 90 minutes, 0.25 mM IPTG and 0.2% arabinose were added. After 12-16 hours, cells were harvested by centrifugation, flash frozen in liquid nitrogen, and stored at -80 °C until processing. Cell pellets were thawed in an ice bath and re-suspended in 80 mL of ice cold 10% glycerol before the addition of 53 mg/mL of CelLytic Express (Sigma), one tablet of Protease Inhibitor Cocktail (Sigma) and 20 mM imidazole. The samples were clarified on a nutator for 30 minutes at 4ºC before the addition of 20 mM of β-mercaptoethanol. The crude cell lysate was passed through a 0.45 µm filter in preparation for nickel affinity chromatography. His-tagged proteins were loaded onto a 5 mL HisTrap HP column (GE Healthcare) at 2 mL/min using an AKTA Explorer FPLC instrument (GE Healthcare) at 4 ºC. The columns were first washed with 20 column volumes of 93% Buffer A (20 mM Tris-HCl, pH 7.4, 1 M NaCl, 20 mM β-mercaptoethanol, 10% glycerol) and 7% Buffer B (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 500 mM imidazole, 10% glycerol) and the proteins were eluted using a 7-100% Buffer B gradient over 10 column volumes. Fractions containing MBP-MUTYH protein (as determined by SDS-PAGE), were pooled. Typical yields of purified protein for MBP-MUTYH wild type, Y179C and G396D were between 10-20 mg from one liter cultures and 0.5-1.5 mg of soluble protein was obtained for MUTYH p.C306W. Monomeric protein was obtained by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) in Buffer C (20 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, 10% glycerol). Fractions eluting near the expected molecular weight (104 kD) were collected, partitioned into aliquots, flash frozen in liquid nitrogen and stored at -80 ºC until further use.

E. DNA Substrates for Glycosylase and Binding Assays

All oligonucleotides (Integrated DNA Technologies, Coralville, Iowa) were purchased PAGE purified. Duplexes were obtained by heating 50 µl 50 µM complementary strands at 85 ºC then decreasing the temperature by 0.5 ºC every 30 seconds until attaining room temperature. The FAM labelled 8-oxoG:A duplex used in the DNA glycosylase assay consisted of 5′-ACA AAG AAC TTA TAG CTC CTC CTT GAG CAC ACA GAG GTG TTC GAT GTA GTT G/A/C GCA GGA CGG GTT CAG T/6-FAM/-3′ and 3′-TGT TTC TTG AAT ATC GAG GAG GAA CTC GTG TGT CTC CAC AAG CAT GAT CAA C/8oxoG/G CGT CCT GCC CAA GTC A-5′. The biotin labeled 8-oxoG:A duplex used in the binding assay consisted of 5′-/BiotinTEG/AC AAA GAA CTT ATA GCT CCT CCT TGA GCA CAC AGA GGT GTT CAT GTA GTT G/A/C GCA GGA CGG GTT CAG T-3′ and the 8-oxoG oligomer.

F. DNA Glycosylase Assay

The DNA glycosylase assay was adapted as previously reported[4,6,7]. The activity was evaluated by providing 10 nM of DNA substrate containing a single 8-oxoG:A mismatch to wild type or mutant MBP-MUTYH proteins (0-1000 nM) at 37ºC in a buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1 mg/mL BSA). Reactions were quenched after 1 hour with 80 mM NaOH followed by heating the samples to 90 ºC for 4 minutes, cooled and then diluted with three volumes of formamide spiked with GeneScan™ 500 LIZ™ Size Standard (Invitrogen). DNA fragmentation was determined employing capillary electrophoresis (U-M DNA Sequencing Core, ABI 3730 DNA Analyzer). Traces were analyzed using Peak Scanner™ Software (version 1.0, Applied Biosystems). The percent of excised DNA was calculated as the ratio of the 6-FAM peak area migrating at 44 oligonucleotides to the total peak area (at 44 and 66 oligonucleotides).

G. Multiple Turnover assay: Active Site Titration

The multiple turnover assay was adapted as previously reported[4,6,7]. Reactions were analyzed for scission of 10 nM FAM-labeled 8-oxoG duplex DNA after the addition of MUTYH protein. The total protein concentrations, selected to give a burst amplitude in a detectible range, were 25, 2670, 500 and 25 nM of MUTYH wild type, Y179C, C306W and G396D, respectively. Samples were drawn over a 20 minute time course and processed as described above. The cleaved product concentration, [P], was fitted with Equation 1 to determine the amplitude of the burst (A_0), k_B (rate constant during the burst phase) and k_L (rate constant for the linear phase).

Equation 1: $[P] = A_0[1-\exp(-k_Bt)]+k_Lt$

The percent of active protein was calculated as a ratio of the A_0 to total protein concentration.

H. Binding Kinetics (Biolayer Interferometry)

All biolayer interferometry measurements were made on an Octet RED instrument (Pall ForteBio, Menlo Park, CA) using streptavidin (SA) biosensors[8]. Assays were performed in 96-well black microplates at 25 ºC and 1000 rpm. All volumes were 200 µL. All proteins were buffer exchanged using PD-10 columns (GE Healthcare) pre-equilibrated with PBS then serial diluted (25, 12.5, 6.25, 3.125, 0.78125 nM) into working volumes with 1X Kinetics Buffer (Pall FortBio;10 mM Phosphate, pH 7.4, 150 mM NaCl, 0.02% Tween-20, 1 mg/mL BSA). The biotinylated duplex DNA was first immobilized onto the SA biosensors for 300 seconds and then equilibrated in 1X Kinetics Buffer for 300 seconds. Protein association was performed for 150-300 seconds followed by dissociation into 1X Kinetics Buffer for 900 seconds. A reference sensor with immobilized dsDNA was subtracted from each data set. Shift data was analyzed with ForteBio's Analysis software (version 7.1). Kinetic parameters k_{on} and k_{off} and affinity (K_D) were determined from a global non-linear regression of association and dissociation binding kinetics using a 1:1 Langmuir binding model.

I. Fe Elemental Analysis

The presence of elemental Fe within MUTYH protein samples was determined using a Thermo Scientific Element2 ICP-HRMS[9]. Purified proteins were dialyzed overnight to remove glycerol and to allow equilibration with blank buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl and 1 mM DTT) using a 10K MWCO Slide-A-Lyzer dialysis cassettes (Thermo Scientific).

J. DNA Synthesis and Purification for Electrochemistry

Thiol, FA, and OG modified DNA strands were prepared on an automated DNA synthesizer (Applied Biosystems) and purified by HPLC on a PLRPS column (Agilent) as described in previously published protocols[10]; unmodified strands were ordered from IDT and purified by HPLC. For electrochemistry, 50 μL 50 μM complementary DNA strands were degassed and annealed in storage buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl) in equimolar amounts by a 5-minute incubation at 95^{o}C followed by slow cooling (1.5 hours) to RT on a thermocycler. Well-matched (WM) duplex DNA, DNA with an FA:OG lesion, and substrates containing an abasic site were all prepared in this way. Duplex sequences were as follows:

WM DNA

5′ – ACT GAA CCC GTC CTG GCT CAA CTA CAT GAA CAC CTC – 3′
3′ – TGA CTT GGG CAG GAC GCA GTT GAT GTA CTT GTG GAG – 5′ – C6 Thiol

FA:OG DNA

5′ – ACT GAA CCC GTC CTG GC**OG** CAA CTA CAT GAA CAC CTC – 3′
3′ – TGA CTT GGG CAG GAC GC**FA** GTT GAT GTA CTT GTG GAG – 5′ – C6 Thiol

Abasic DNA

5′ – ACT GAA CCC GTC CTG GCT CAA CTA CAT GAA C**Ab**C CTC – 3′
3′ – TGA CTT GGG CAG GAC GCA GTT GAT GTA CTT GTG GAG – 5′ – C6 Thiol

OG = 8-oxoguanine, FA = 2′-fluoroadenine, Ab = abasic site

K. Electrochemistry on DNA self-assembled monolayers (SAMs)

Electrochemical characterization of MUTYH was carried out on a multiplexed chip platform consisting of 16 individually-addressable gold electrodes separable into four quadrants[10]. Self-assembled DNA monolayers were formed by adding 25 μL 25 μM duplexed DNA in phosphate buffer (5 mM sodium phosphate, pH 7.0, 50 mM NaCl) to each quadrant of the chip and incubating overnight. After monolayer formation, gaps in the film were eliminated by backfilling for 45 minutes at RT with 1 mM 6-mercapto-1-hexanol in phosphate buffer with 5% glycerol. The surface was then extensively rinsed in phosphate buffer, followed by protein storage buffer (described below). To compare different monolayer morphologies, DNA was incubated in phosphate buffer as described to generate low-density DNA monolayers (surface coverage of ~15 pmol/cm^2) or in the presence of 100 mM $MgCl_2$ to form high-density monolayers (surface coverage of ~40 pmol/cm^2)[11,12]. Bulk electrolysis experiments were performed using gold rod electrodes in a custom-made electrochemical cell. Experiments were carried out in air unless otherwise noted; anaerobic experiments were performed in a glove bag (Coy) under a 95% N_2/5% H_2 atmosphere.

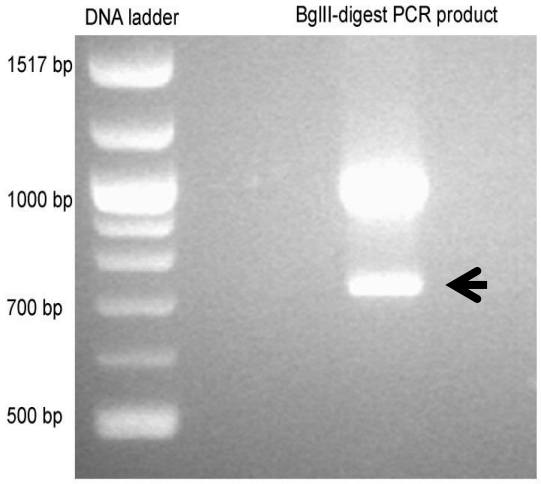
MUTYH concentration was determined by UV-visible spectroscopy, using an extinction coefficient of 17000 $M^{-1}cm^{-1}$ at 410 nm to determine [4Fe4S] cluster concentration and 102330 $M^{-1}cm^{-1}$ at 280 nm to determine total protein concentration. Cluster loading was determined by dividing [4Fe4S] cluster concentration by total protein concentration, and was typically around 15%. Initial characterization was carried out in Tris storage buffer (20 mM Tris, pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol v/v), while later electrochemical and spectroscopic experiments used a HEPES buffer (20 mM HEPES, pH 7.4, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v). MUTYH was transferred into HEPES using Amicon 10 kDa MW cutoff spin tubes (Millipore Biomedicals) at 4°C.

Once in an appropriate buffer, MUTYH was added to a multiplexed chip and incubated for several hours with cyclic and square wave voltammetry (CV and SQWV, respectively) scans taken once per hour. In typical experiments, CV scans were taken in a potential window of -0.188 to 0.412 V vs NHE at a scan rate of 100 mV/s, while SQWV scans were taken at a frequency of 15 Hz with 0.025 V amplitude. To plot the scan rate dependence of CV current, scans in the same window were carried out at 20, 50, 80, 100, 200, 500, 750, and 1000 mV/s. All experiments were performed on a CH Instruments potentiostat with an Ag/AgCl reference in 3 M NaCl and Pt wire counter electrode. Potentials were converted to NHE by adding 212 mV to the measured potentials, accounting for both the salt concentration (209 mV according to BASi®) and ambient temperature[13]. Bulk electrolysis was carried out at 0.412 mV versus NHE, and yields were estimated by subtracting the total charge passed with only buffer present from that passed when MUTYH was included. All buffer components were purchased from Sigma-Aldrich, the Ag/AgCl reference electrode was purchased from BASi®, and the Pt wire counter electrode was purchased from the Kurt J. Lesker Company.

L. EPR spectroscopy

Continuous wave X-band EPR was carried out at 10 K on a Bruker EMX instrument. Samples were prepared aerobically, using 150 μL 5 – 15 μM MUTYH in parallel with a storage buffer blank. Spectra were taken from the summation of 9 sweeps at 12.88 mW microwave power, 2 G modulation amplitude, and a receiver gain of 5.02 $x10^3$.

II. Establishment of Trans Chromosomal Configuration (Supplementary Figure 1)

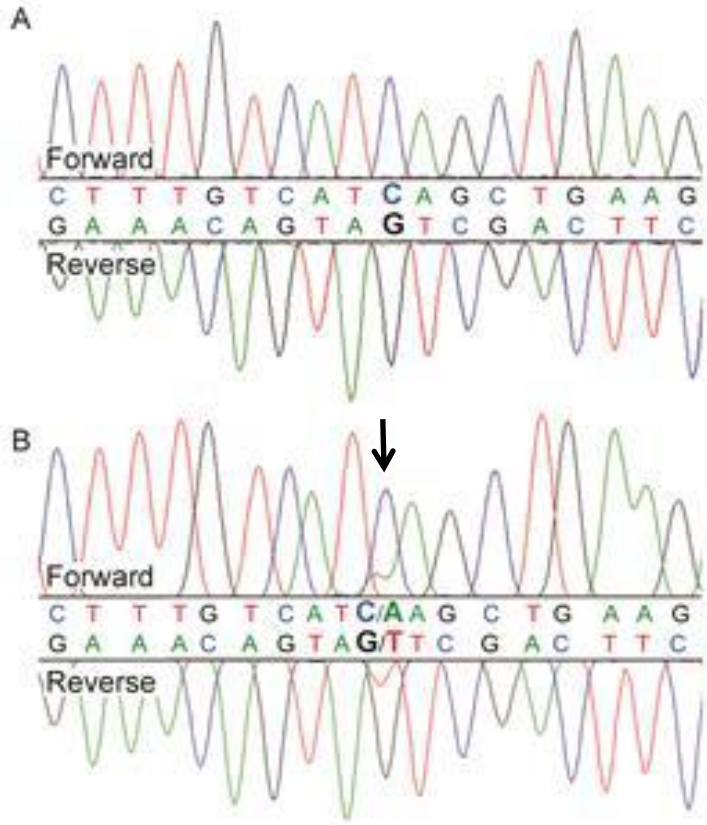


Supplementary Figure 1. ***The c.918C>G MUTYH variant is situated trans relative to the c.1187G>A MUTYH mutation.*** The c.1187G>A (p.G396D) *MUTYH* mutation creates a unique BglII restriction enzyme site which we employed to isolate the chromosomal DNA strand containing the c.1187G>A *MUTYH* mutation. A 957 bp fragment encompassing the *MUTYH* open reading frame nucleotide positions 918 and 1187 was generated with PCR and then restriction enzyme digested with BglII. The digestion products were resolved on 1.5% agarose gel and the lower molecular weight DNA band isolated (arrow). Sequencing of the lower molecular weight band demonstrated the wildtype cytosine nucleotide at position 918 consistent with a trans configuration of the c.1187G>A and c.918C>G *MUTYH* alterations.

III. Demonstration G:C to T:A Transversion

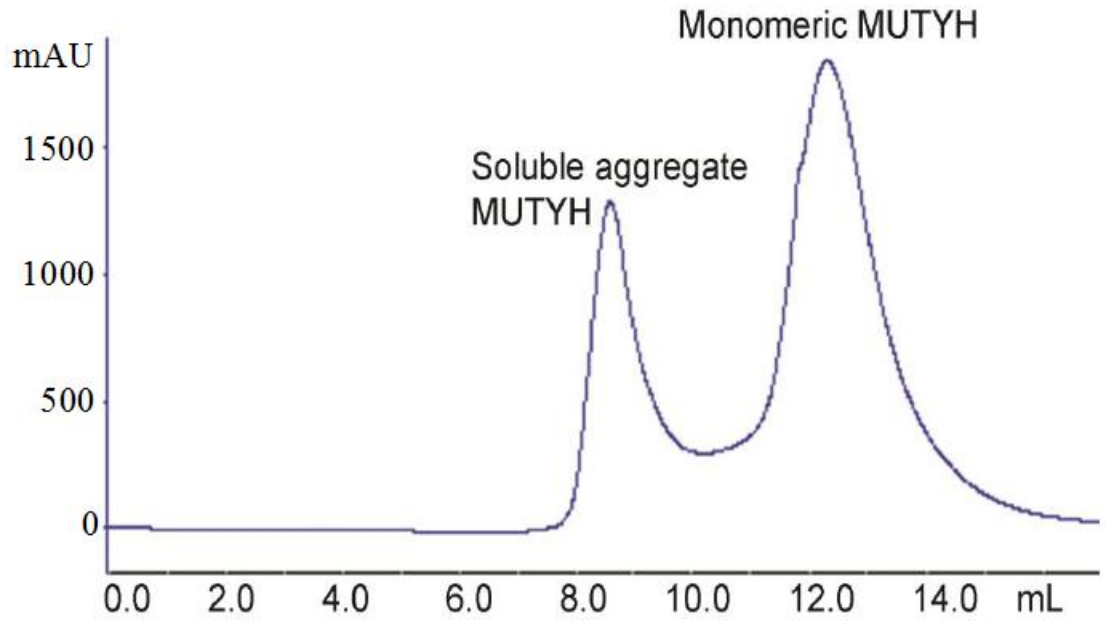
Normal Colonic Tissue

Colonic Adenoma



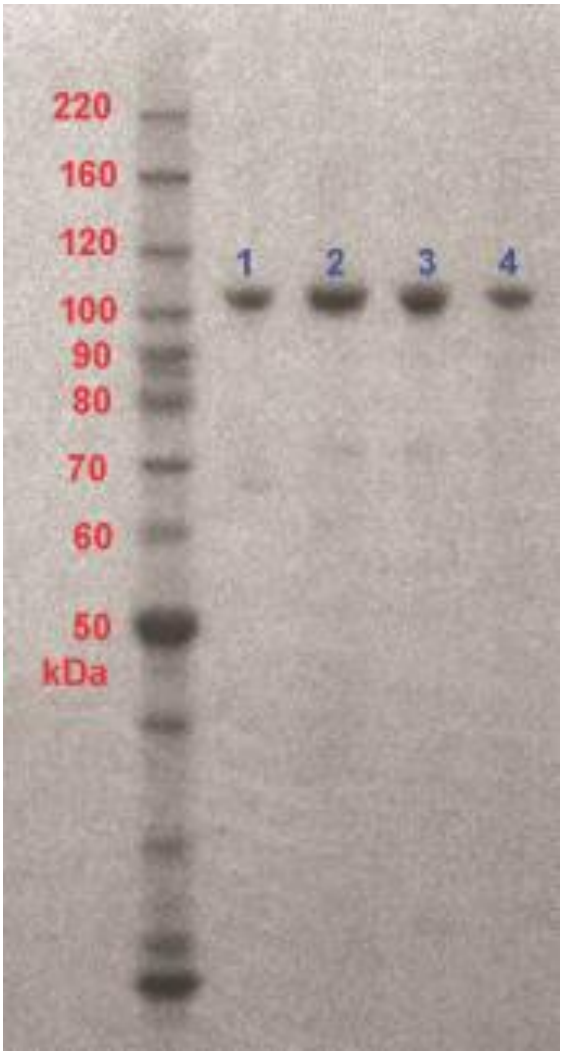
Supplementary Figure 2. ***Trans c.1187G>A (p.G396D) and c.918C>G (p.C306W) MUTYH Mutations are Associated with the Signature G:C to T:A Transversion Distinctive of Deficient MUTYH DNA Repair.*** From one of the patient's adenomatous polyps we extracted DNA to assess for the presence of the signature G:C to T:A nucleotide transversion which results from deficient MUTYH DNA repair. The figure depicts a portion of the germline nucleotide sequence from the mutation cluster region of the APC gene (upper tracing); this sequence reveals the normal wildtype *APC* sequence. In comparison, the lower tracing depicts the forward and reverse nucleotide sequence from DNA extracted from one of the patient's colonic polyps for this same mutation cluster region of the *APC* gene; there is identified a G:C to T:A transversion (arrow) resulting in a premature stop codon in the *APC* open reading frame and pathologic APC protein product.

IV. Size-Exclusion, Fast Protein Liquid Chromatography of MUTYH Protein



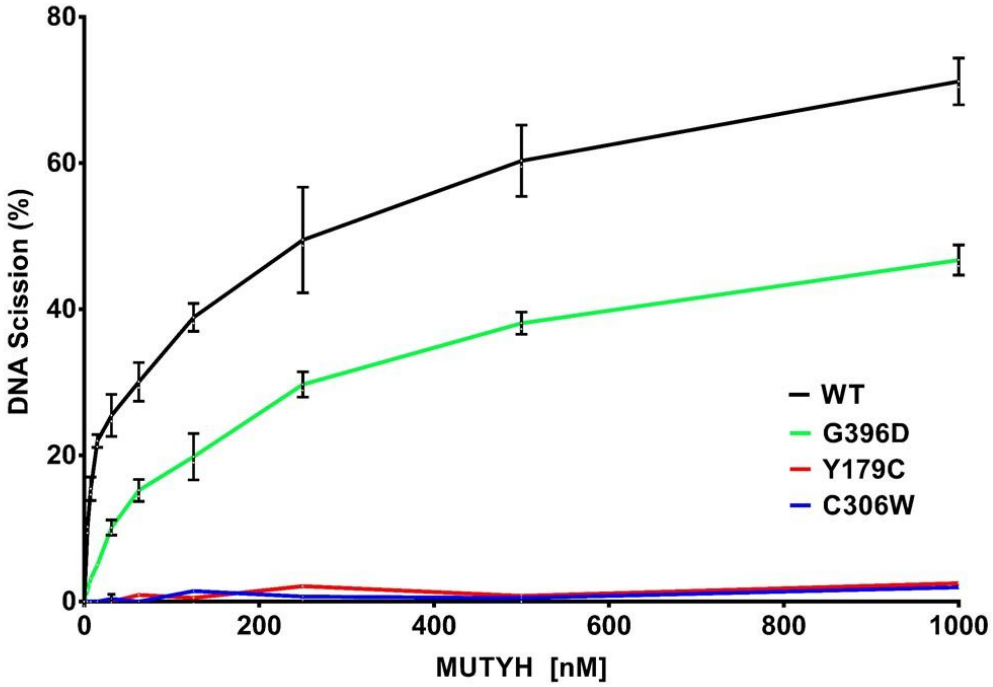
Supplementary Figure 3. SEC-FPLC UV_{260} trace of WT MUTYH following treatment with 20 mM β-mercaptoethanol. While some soluble aggregates still occur in the void volume, the majority of the purified protein is monomeric.

V. SDS-PAGE of Purified Monomeric MUTYH Proteins (Supplementary Figure 4).



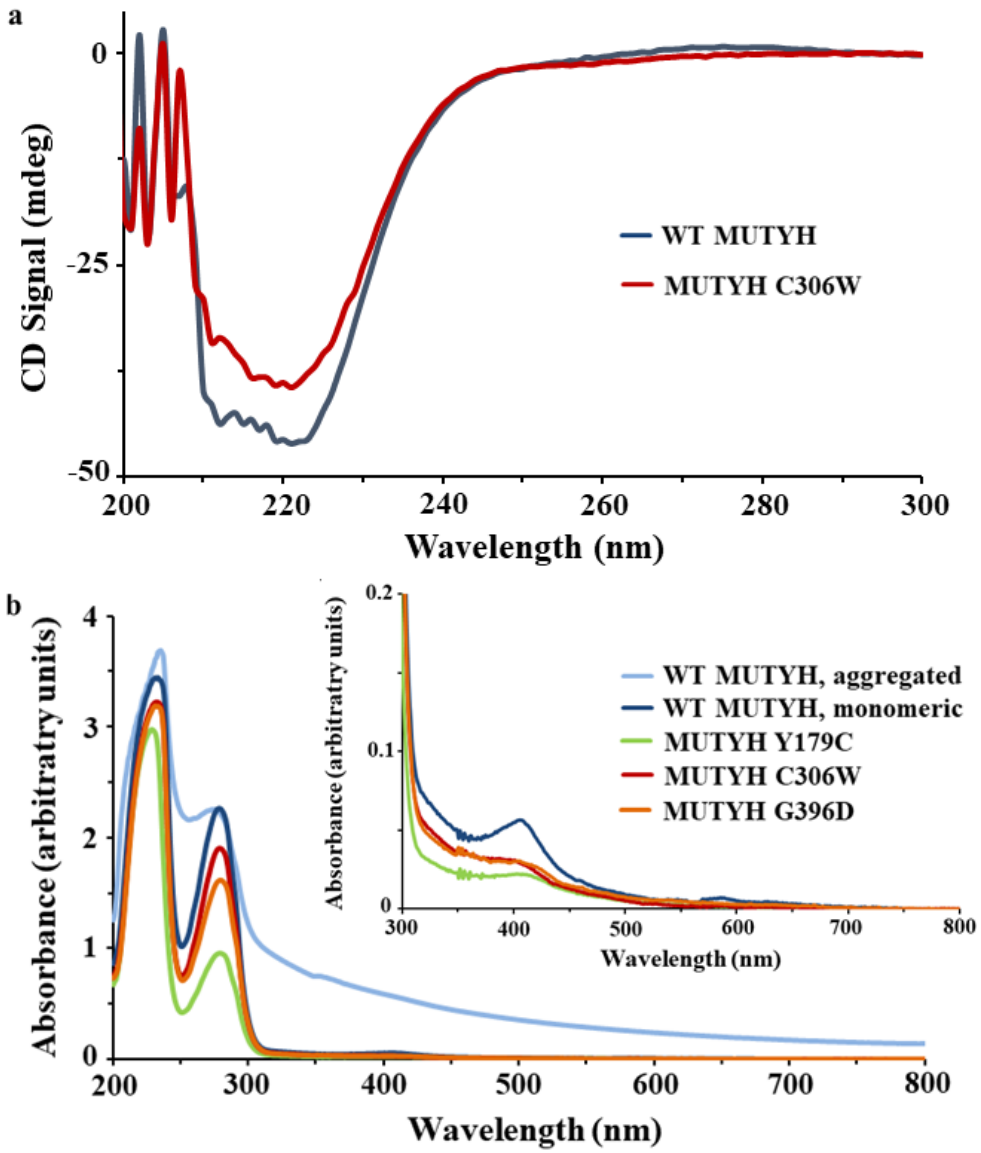
Supplementary Figure 4. SDS-PAGE gel of purified monomeric MUTYH proteins. The left lane contains the protein molecular weight ladder. The lanes 1-4 contain 1ug of MUTYH wild type, Y179C, C306W, and G396D, respectively.

VI. Glycosylase Assay with Aggregated Protein



Supplementary Figure 5. Glycosylase Assay Employing Aggregated MUTYH Proteins. The four MUTYH proteins, WT, G396D, Y179C, and C306W were overexpressed in *E. coli* and purified by nickel affinity chromatography followed by size exclusion FPLC. All of the MUTYH proteins eluted as soluble aggregates in the void fraction; evidence of aggregation in these samples was also clear from the UV-visible spectra. Aggregated WT, G396D and Y179C MUTYH proteins demonstrated glycosylase activity consistent with previous reports, with attenuated activity observed in the G396D and Y179C MUTYH proteins[4,6]; aggregated C306W protein lacked glycosylase activity. It was suspected that aggregated protein may not exhibit complete native protein activity, and thus, monomeric proteins were subsequently synthesized and employed in subsequent experiments.

VII. Circular Dichroism and UV-Visible Spectra of WT MUTYH and MUTYH C306W



Supplementary Figure 6. (a) Circular dichroism (CD) spectra of WT MUTYH and MUTYH C306W in HEPES (20 mM HEPES, pH 7.4, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v). The similarity in spectra suggest that the C306W mutation does not result in significant global structural perturbations. **(b)** UV-visible spectra of aggregated and monomeric WT MUTYH, and the three mutant variants, in Tris buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol v/v). Aggregated spectra are clearly distinct from those of monomeric protein, and the absorbance at 410 nm indicates that all monomeric MUTYH variants effectively bound [4Fe4S] cluster.

VIII. **Electrochemical Characterization of MUTYH in HEPES buffer**

To verify that the electrochemical properties of MUTYH were not altered in HEPES buffer, electrochemical analysis was performed with each variant as was done previously in Tris buffer (Supplementary Figure 7). In HEPES buffer, WT MUTYH underwent a slight decrease in midpoint potential from 105 ± 1 mV vs NHE in Tris to 93 ± 1 mV vs NHE in HEPES (Table 3), but the overall signal shape and properties remained unaltered. Unsurprisingly, the addition of more concentrated protein to the electrode (~5 μM for all electrochemical experiments in HEPES) gave substantially larger signals, with WT MUTYH yielding a reductive peak area of 7.6 ± 3 nC and an oxidative peak area of -5.6 ± 3 nC on low density monolayers. Remarkably, although the protein was only about twice as concentrated compared with previous experiments, the signal was over 10 times larger, indicating that some of this increase could be attributed to the distinct buffer conditions.

Like WT, all of the mutants showed ~10-fold larger signals when concentrated in HEPES, and their midpoint potentials decreased by a similar margin, placing all variants in the same potential window (Table 4). The signal from the MUTYH Y179C protein was still noticeably smaller than WT (reductive and oxidative peaks at 48 ± 18% and 33 ± 17% of WT), again most likely as a result of lower binding affinity of this mutant, while G396D was not significantly different from WT (reductive and oxidative peaks at 77 ± 32% and 76 ± 41% of WT). To ensure accurate comparison with EPR spectra, electrochemistry for the less stable C306W mutant was performed on the same day as EPR with the same sample stock. Consistent with the oxidation indicated in the UV-visible spectrum and the presence of a $[3Fe4S]^{+}$ cluster in the EPR spectra, the reversible signal in this sample was significantly smaller than any other mutant (reductive and oxidative peaks 35 ± 15% and 26 ± 14% of WT) and a very prominent secondary reductive peak was present centered at -88 ± 6 mV vs NHE (Figure S6). The larger signals made this peak much more readily quantifiable compared with Tris buffer, and the total charge of this irreversible peak was found to represent ~20% of either peak of the reversible $[4Fe4S]^{3+/2+}$ couple. The area in this case was lower than estimated for the C306W sample aerobically oxidized in Tris buffer, where the secondary peak was roughly equal in magnitude to the main peak, which suggested that the concentrated sample in HEPES was not 100% oxidized. Overall, UV-visible and EPR spectra obtained from the same MUYH C306W protein strongly supported the identification of the irreversible peak as a $[3Fe4S]^{+/0}$ redox couple. When compared to WT MUTYH, or even the Y179C and G396D mutants, the susceptibility of the C306W [4Fe4S] cluster to degradation upon oxidation during redox signaling was apparent, and the irreversibility of the degradation product signal indicates that the resultant $[3Fe4S]^{+/0}$ species does not bind DNA effectively.

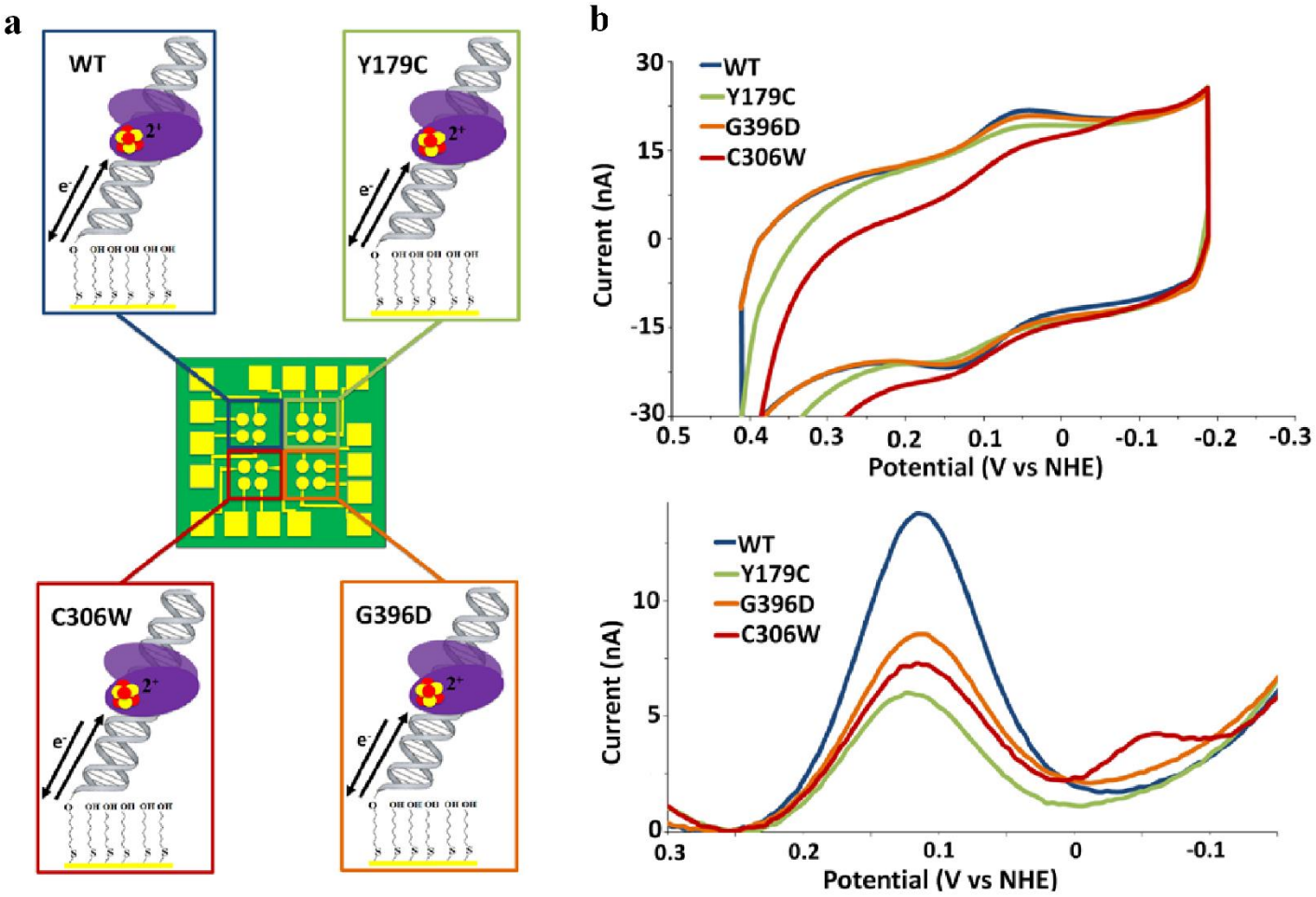
Based on the measured midpoint potentials in Tris buffer (Table 4), electrochemical signals from dilute MUTYH samples were most likely DNA-bound[10]; however, the small signal sizes made it impossible to determine if the signal was DNA-mediated. The substantially larger signals recorded in HEPES buffer with concentrated samples allowed this issue to be addressed, and chips containing half well-matched DNA and half DNA containing an abasic site were prepared for this purpose. Earlier work with DNA-mediated EndoIII signals showed a dependence on monolayer morphology[10]; thus, to see if this was also the case for MUTYH, the abasic site discrimination studies were carried out on both high and low density monolayers.

On abasic DNA, a maximum CV charge attenuation of 38% for the reductive peak and 46% for the oxidative peak was obtained on low density monolayers (Figure S7); in contrast, no appreciable discrimination was observed on high density monolayers. The observed abasic site

discrimination confirmed that MUTYH can take part in DNA-mediated signaling, and the differences in high and low density DNA monolayers emphasize the importance of surface accessibility to large proteins. The sterically hindered high density films clearly do not provide sufficient access to DNA in an appropriate conformation to observe such a signal, and the peaks seen in this case are likely attributable to DNA-bound proteins signaling directly through the monolayer surface[10].

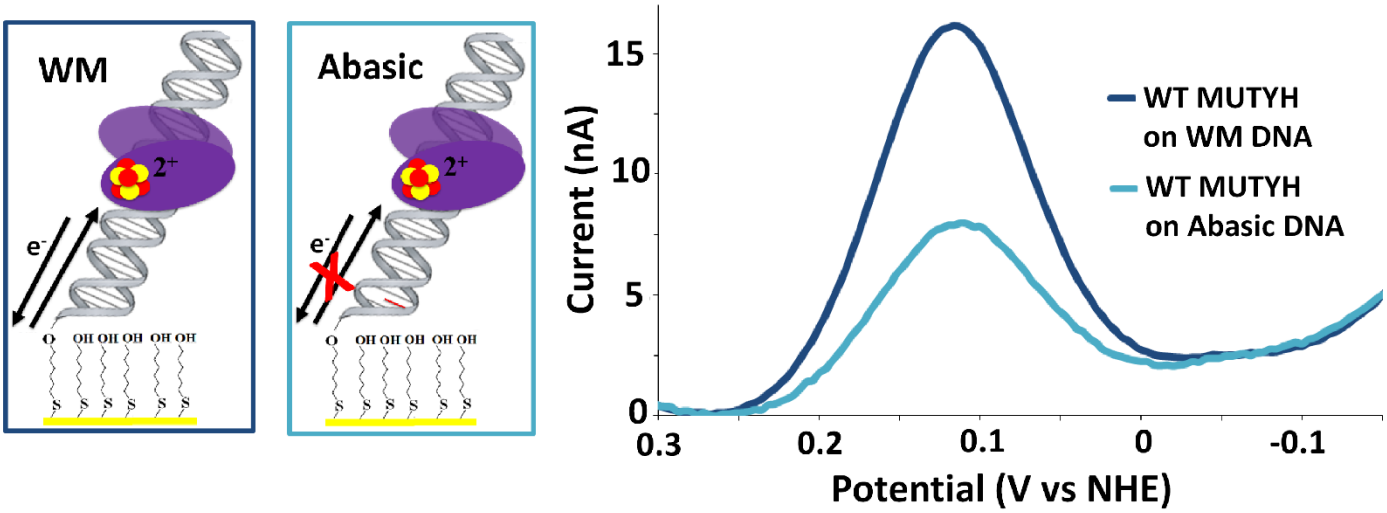
Interestingly, the pattern of abasic site discrimination observed here is opposite to previously published data for *E. coli* EndoIII[10]; this difference can be rationalized by considering two important factors. First, the DNA used in the EndoIII experiments was only 15 bp, while EndoIII has a binding footprint of 12 bp. Thus, the protein would have taken up much of the available space and the cluster could readily bypass the mismatch on fully accessible low density DNA[14]. In contrast, the DNA substrate chosen for MUTYH was 36 bp while the binding footprint is likely close to the 16 bp reported for mouse MUTYH[15], making bypass of the abasic site on low density monolayers less likely. Second, a considerable size discrepancy exists between the two proteins: unmodified MUTYH is 61 kDa, while EndoIII is only 24 kDa. Thus, steric hindrance in high density films would be expected to have a greater impact on MUTYH, a notion supported by the observation that the largest signals obtained on high density monolayers containing WM DNA were ~30% smaller than the equivalent signals on low density films.

IX. HEPES Buffer Electrochemical Characterization



Supplementary Figure 7. Electrochemical characterization of MUTYH variants in HEPES buffer. (**a**) Arrangement of MUTYH variants on a multiplexed chip. (**b**) CV (top) and SQWV (bottom) scans of 5 μM WT, Y179C, G396D, and C306W variants in HEPES storage buffer (20 mM HEPES, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4). Under these conditions, substantial signals occur even by CV. The MUTYH C306W secondary peak is readily apparent in both CV and SQWV. CV scans were taken at a scan rate of 100 mV/s, while SQWV scans were taken at a frequency of 15 Hz with 0.025 V amplitude. All scans shown are an average obtained from at least seven separate electrodes.

X. Abasic Site Discrimination by WT MUTYH



Supplementary Figure 8. Abasic site discrimination by WT MUTYH. 5.0 μM WT MUTYH was incubated in HEPES buffer (20 mM HEPES, 100 mM KCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol v/v, pH 7.4) on a single multiplexed chip with half of the available quadrants containing DNA with an abasic site (light blue) near the electrode and the other half containing well matched (WM) DNA (dark blue). As is apparent from the SQWV voltammetry, the signal on WM DNA was comparable in size to previously observed signals (Figure 2), but the total peak area on abasic DNA was decreased by ~40%, indicating that the signal was DNA-mediated. The SQWV voltammograms were obtained at 15 Hz, and signal shown is the average of at least seven separate electrodes from two quadrants.

XI. Coordinating Cysteine and Neighboring Arginine Residues Predicted/Reported to Occur with MUTYH Associated Polyposis

Supplementary Table 1: The four coordinating cysteines and surrounding arginines predicted or reported [16] to be associated with MAP.

Residue	Reported Variant
C290	C290W[17]
C297	not reported
C300	not reported
C306	C306W (this study)
R241	R241W[17,18] R241G[16]
R245	R245L[19] R245C[16] R245H[16,20]
R247	R247G[16]
R309	R309C[21,22]

XII. Somatic Alterations Associated with Coordinating Cysteine and Neighboring Arginine Residues Predicted/Reported to Occur with Colorectal and Other Cancers

Supplementary Table 1: The four coordinating cysteines and surrounding arginines reported with colorectal and other cancers from the cBioPortal database[23,24].

Residue	Reported Somatic Variant Colorectal Cancer (instances)	Reported Somatic Variant (other cancer, instances)
C290	not reported	not reported
C297	not reported	not reported
C300	not reported	not reported
C306	not reported	not reported
R241	not reported	R241L (liver, 1) R241Q (unknown primary, 1)
R245	R245C (2)	R245C (lung,2)
R247	not reported	R247Q (melanoma, 1)
R309	R309H (2)	R309H (glioblastoma, 1)

XIII. Supplementary References

1 Miyoshi, Y. *et al.* Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum Mol Genet* **1**, 229-233 (1992).

2 Donnelly, M. I. *et al.* An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr Purif* **47**, 446-454, (2006).

3 Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668-1674, (1999).

4 D'Agostino, V. G. *et al.* Functional analysis of MUTYH mutated proteins associated with familial adenomatous polyposis. *DNA Repair (Amst)* **9**, 700-707, (2010).

5 Whicher, J. R. *et al.* Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chem Biol* **20**, 1340-1351, (2013).

6 Kundu, S., Brinkmeyer, M. K., Livingston, A. L. & David, S. S. Adenine removal activity and bacterial complementation with the human MutY homologue (MUTYH) and Y165C, G382D, P391L and Q324R variants associated with colorectal cancer. *DNA Repair (Amst)* **8**, 1400-14, (2009).

7 Porello, S. L., Leyes, A. E. & David, S. S. Single-turnover and pre-steady-state kinetics of the reaction of the adenine glycosylase MutY with mismatch-containing DNA substrates. *Biochemistry* **37**, 14756-14764, (1998).

8 Rich, R. L. & Myszka, D. G. Higher-throughput, label-free, real-time molecular interaction analysis. *Anal Biochem* **361**, 1-6, (2007).

9 Profrock, D. & Prange, A. Inductively coupled plasma-mass spectrometry (ICP-MS) for quantitative analysis in environmental and life sciences: a review of challenges, solutions, and trends. *Appl Spectrosc* **66**, 843-868, (2012).

10 Pheeney, C. G., Arnold, A. R., Grodick, M. A. & Barton, J. K. Multiplexed electrochemistry of DNA-bound metalloproteins. *J Am Chem Soc* **135**, 11869-11878, (2013).

11 Boon, E. M., Salas, J. E. & Barton, J. K. An electrical probe of protein-DNA interactions on DNA-modified surfaces. *Nat Biotechnol* **20**, 282-286, (2002).

12 Kelley, S. O., Barton, J. K., Jackson, N. M. & Hill, M. G. Electrochemistry of methylene blue bound to a DNA-modified electrode. *Bioconjug Chem* **8**, 31-37, (1997).

13 Greeley, R. S., Smith, W. T., Stoughton, R. W. & Lietzke, M. H. Electromotive Force Studies in Aqueous Solutions at Elevated Temperatures .1. The Standard Potential of the Silver-Silver Chloride Electrode. *J Phys Chem-Us* **64**, 652-657, (1960).

14 O'Handley, S., Scholes, C. P. & Cunningham, R. P. Endonuclease III interactions with DNA substrates. 1. Binding and footprinting studies with oligonucleotides containing a reduced apyrimidinic site. *Biochemistry* **34**, 2528-2536 (1995).

15 Pope, M. A. & David, S. S. DNA damage recognition and repair by the murine MutY homologue. *DNA Repair (Amst)* **4**, 91-102, (2005).

16 Out, A. A. *et al.* Leiden Open Variation Database of the MUTYH gene. *Human Mutation* **31**, 1205-1215, (2010).

17 Fleischmann, C. *et al.* Comprehensive analysis of the contribution of germline MYH variation to early-onset colorectal cancer. *Int J Cancer* **109**, 554-558, (2004).

18 Bai, H. *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* **33**, 597-604, (2005).

19 Bai, H. *et al.* Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis. *Cancer Lett* **250**, 74-81, (2007).
20 Ali, M. *et al.* Characterization of mutant MUTYH proteins associated with familial colorectal cancer. *Gastroenterology* **135**, 499-507, (2008).
21 Sieber, O. M. *et al.* Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *N Engl J Med* **348**, 791-799, (2003).
22 Goto, M. *et al.* Adenine DNA glycosylase activity of 14 human MutY homolog (MUTYH) variant proteins found in patients with colorectal polyposis and cancer. *Human Mutation* **31**, E1861-1874, (2010).
23 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, (2012).
24 Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1, (2013).