UC Davis

UC Davis Previously Published Works

Title

The Zinc Linchpin Motif in the DNA Repair Glycosylase MUTYH: Identifying the Zn2+ Ligands and Roles in Damage Recognition and Repair

Permalink

https://escholarship.org/uc/item/8649h1kv

Journal

Journal of the American Chemical Society, 140(41)

ISSN

0002-7863

Authors

Nuñez, Nicole N
Khuu, Cindy
Babu, C Satheesan
et al.

Publication Date

2018-10-17

DOI

10.1021/jacs.8b06923

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2019 October 17.

Published in final edited form as:

J Am Chem Soc. 2018 October 17; 140(41): 13260–13271. doi:10.1021/jacs.8b06923.

The Zinc Linchpin Motif in the DNA Repair Glycosylase MUTYH: Identifying the Zn^{2+} Ligands and Roles in Damage Recognition and Repair

Nicole N. Nuñez[†], **Cindy Khuu**[†,††], **C. Satheesan Babu**[‡], **Steve J. Bertolani**[†,#], **Anisha N. Rajavel**[†], **Jensen E. Spear**[†], **Jeremy A. Armas**[†], **Jon D. Wright**[‡], **Justin B. Siegel**[†,#,##], **Carmay Lim**[‡,*], and **Sheila S. David**[†,*]

[†]Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, California, 95616, USA

[††]Biochemistry, Molecular, Cellular and Developmental Graduate Group, University of California, Davis, 95616, USA

[#]Genome Center, University of California, Davis, One Shields Avenue, Davis, California, 95616, USA

[##]Department of Biochemistry and Molecular Medicine, University of California, Davis, One Shields Avenue, Davis, California, 95616, USA

[‡]Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan R. O. C.

Abstract

The DNA base excision repair (BER) glycosylase MUTYH prevents DNA mutations by catalyzing adenine (A) excision from inappropriately formed 8-oxoguanine (8-oxoG):A mismatches. The importance of this mutation suppression activity in tumor suppressor genes is underscored by the association of inherited variants of MUTYH with colorectal polyposis in a hereditary colorectal cancer syndrome known as MUTYH-associated polyposis, or MAP. Many of the MAP variants encompass amino acid changes that occur at positions surrounding the two-metal cofactor-binding sites of MUTYH. One of these cofactors, found in nearly all MUTYH orthologs, is a $[4Fe\text{-}4S]^{2+}$ cluster coordinated by four Cys residues located in the N-terminal catalytic domain. We recently uncovered a second functionally relevant metal cofactor site present only in higher eukaryotic MUTYH orthologs: a Zn^{2+} ion coordinated by three Cys residues located within the extended interdomain connector (IDC) region of MUTYH that connects the N-terminal adenine excision and C-terminal 8-oxoG recognition domains. In this work, we identify a candidate for the fourth Zn^{2+} coordinating ligand using a combination of bioinformatics and computational modeling. In addition, using *in vitro* enzyme activity assays, fluorescence polarization DNA binding assays,

[*]Corresponding Authors ssdavid@ucdavis.edu, carmay@gate.sinica.edu.tw.

Author Contributions

The manuscript was written through contributions of all authors and all authors have given approval to the final version of the manuscript.

The authors declare no competing financial interests.

ASSOCIATED CONTENT

Supporting Information. Files regarding the bioinformatics and modeling studies, as well as further results from *in vitro* and cell based assays. This material is available free of charge via the Internet at http://pubs.acs.org."

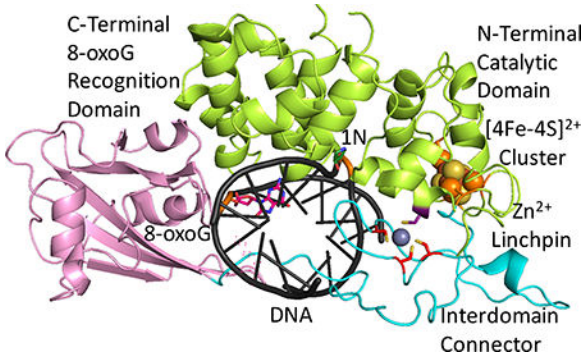
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

circular dichroism spectroscopy and cell-based rifampicin resistance assays, the functional impact of reduced Zn^{2+} chelation was evaluated. Taken together, these results illustrate the critical role that the "Zn^{2+} linchpin motif" plays in MUTYH repair activity by providing for proper engagement of the functional domains on the 8-oxoG:A mismatch required for base excision catalysis. The functional importance of the Zn linchpin also suggests that adjacent MAP variants or exposure to environmental chemicals, may result in compromised Zn^{2+} coordination and the ability of MUTYH to prevent disease.

SYNOPSIS TOC:

QM/MM energy-minimized structure of human MUTYH highlighting the four Cys residues interactions with Zn^{2+}. A representative structure of human MUTYH from an MD simulation in explicit solvent after QM/MM minimization of the Zn^{2+}-binding site. Color coding is as follows: C-terminal 8-oxoG recognition domain, pink; N-terminal adenine removal domain, green; interdomain connector, light blue; three previously identified Cys ligands (Cys318, Cys325 and Cys328 in human MUTYH), red; newly identified fourth Cys ligand (Cys230 in human MUTYH), purple; Cys residues coordinating the Fe-S cluster, orange, with the Fe-S cluster designated by orange (Fe) and yellow (S) spheres.

Graphical Abstract



INTRODUCTION

The human base excision repair (BER) glycosylase MUTYH and orthologs (such as MutY in *Escherichia coli*, and Mutyh in *Mus musculus*) are responsible for the removal of adenine (A) when base paired opposite a common product of oxidative DNA damage, 8-oxo-7,8-dihydroguanine (also known as 8-oxoG or OG) (Figure 1).[1–2] The absence of MUTYH-mediated DNA repair of 8-oxoG:A mismatches results in the accumulation of G:C to T:A transversion mutations.[1] Inheritance of biallelic mutations in the *MUTYH* gene has been correlated with the early onset of colorectal polyposis resulting from the accumulation of mutations in the tumor suppressor gene *adenomatous polyposis coli* (*APC*), leading to its inactivation; this hereditary colorectal cancer syndrome is referred to as MUTYH-associated polyposis, or MAP.[3–5] Clinically identified MAP variants include missense substitutions throughout the protein, many of which are located in the regions surrounding the two required metal cofactor-binding sites of MUTYH.[6–7]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The $[4Fe\text{-}4S]^{2+}$ (Fe-S) cluster cofactor is highly conserved among MUTYH orthologs from bacteria to humans, and is also found in many other structurally related BER glycosylases, such as Endonuclease III (EndoIII), and its human ortholog, NTH1.[1] The Fe-S cluster in MUTYH is coordinated by four Cys residues (Cys-X_6-Cys-X_2-Cys-X_5-Cys) located in the N-terminal catalytic domain and its presence is essential for 8-oxoG:A mismatch recognition and adenine excision.[8,9] Fe-S clusters are versatile cofactors, offering unique structural, catalytic and redox roles.[10–11] MUTYH does not require redox chemistry for catalysis of glycosidic bond hydrolysis (Figure 1)[12]; however, the Fe-S cluster is amenable to oxidation in the presence of DNA.[13] Oxidation of the Fe-S cluster in the bacterial glycosylase EndoIII, as well as DNA polymerase δ and primase, has been shown to result in increased DNA affinity.[14–16] The redox-dependent affinity and similarity of redox potentials of several BER glycosylases and other DNA binding enzymes has prompted the hypothesis that DNA-mediated charge transfer between Fe-S cluster containing enzymes may serve as a means to coordinate DNA damage search and repair processes in cells.[17–18]

Recently, our laboratory identified a second metal-binding site in mammalian MUTYH orthologs: a Zn^{2+} ion coordinated by three additional highly conserved Cys residues (Cys-X_6-Cys-X_2-Cys) within the interdomain connector (IDC) region (Figure 2).[19] In the *Geobacillus stearothermophilus* (*Gs*) MutY structure, the IDC wraps around the DNA major groove, linking the C-terminal 8-oxoG recognition and the N-terminal catalytic domains.[20–22] However, mammalian MUTYH enzymes possess an extended IDC region containing the three Cys residues in a sequence location not found amongst their prokaryotic orthologs.[19] Somewhat surprisingly, the IDC region of the X-ray crystal structure of the MUTYH N-terminal domain (PDB: 3N5N) is disordered and lacks the coordinated Zn^{2+} ion (Figure 2).[19–23] This may be due to the lack of the C-terminal domain or loss of Zn^{2+} during purification. Nevertheless, this structure illustrates that the IDC projects away from the DNA, poised to serve as the docking site for protein partners involved in DNA maintenance and cell cycle regulation.[7] Notable interactions within the MUTYH IDC include associations with proteins such as Hus1 of the Rad9-Rad1-Hus1 (911) DNA damage sensing complex, and AP endonuclease (APEI) that processes the abasic site product formed by MUTYH glycosylase activity.[23–25]

Extended X-ray absorption fine structure (EXAFS) spectroscopy studies of the Zn^{2+} K-edge of truncated mouse Mutyh (MutyhΔ333–515) fit best to a model of one Zn^{2+} coordinated to four sulfur ligands with an average bond distance of ~2.3 Å.[19] Replacement of two of the three originally proposed Zn^{2+} Cys ligands with Ser (Cys307 and Cys310 in mouse Mutyh or Cys325 and Cys328 in human MUTYH, respectively) completely ablate Zn^{2+} chelation, attenuate the adenine glycosylase activity and reduce the ability to suppress mutations in cells.[19] Replacement of the third proposed Zn^{2+} ligand (Cys300 in mouse Mutyh or Cys318 in human MUTYH) with Ser, maintained WT Zn^{2+} levels and only slightly reduced 8-oxoG:A repair activities; however, its conservation in MUTYH orthologs and the proximity of Cys318 to the other two Cys ligands strongly suggests its role as a coordinating Zn^{2+} ligand.[19] In addition, it was previously demonstrated that the Zn^{2+} ion in Mutyh may be removed via chelation with 1,10-phenanthroline resulting in almost complete loss of *in vitro* adenine glycosylase activity.[19] Taken together these results established the importance of this previously unrecognized Zn^{2+} ion coordinated within the IDC of MUTYH, and led us to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

propose that this “Zinc Linchpin” motif serves to structurally organize the IDC to coordinate the OG recognition and adenine excision functions of the C and N-terminal domains.

A candidate for the fourth Zn^{2+} ligand has been elusive: no other Cys residue was readily identified that would be located nearby to the other three Cys in the IDC.[19] The EXAFS data also fit well to a model with 3 S and 2 O/N (bond distances of 2.32 and 2.08Å); however, site-directed mutagenesis (SDM) of several adjacent Asp and Glu residues within the IDC did not reduce Zn^{2+} levels or enzymatic activity.[19] Structural and computational studies of proteins have demonstrated overwhelming statistical evidence that Zn^{2+} bound by two or more Cys is typically tetra-coordinated.[26–28] Indeed, examples of 5-coodinate Zn^{2+} sites with three Cys and 2 O/N ligands are comparatively more rare.[26] In addition, proteins with Zn^{2+} bound by multiple Cys residues are characteristic of structural rather than catalytic sites, with Cys residues being the most likely ligands participating in Zn^{2+} metal binding. These arguments, taken together with the EXAFS results, strongly suggested that the elusive 4th ligand would be a Cys residue.

To identify candidates for the fourth Zn^{2+} ligand, we performed a sequence profile search and alignment followed by computational modeling of the protein. This led to the identification of a Cys residue within the N-terminal domain and provided a testable hypothesis for the four Cys residues expected to be coordinated to Zn^{2+}. The participation of the proposed fourth Cys ligand (Cys230 in human MUTYH,) in Zn^{2+} binding along with the previously implicated third Zn^{2+} ligand (Cys318 in human MUTYH) was established via single and double replacements of these two Cys residues with Ser in human MUTYH (or corresponding residues in mouse Mutyh). Reduction in Zn^{2+} levels correlated with reduced glycosylase activity and ability to suppress DNA mutations. Furthermore, we establish that the absence of Zn^{2+} coordination within the IDC does not significantly alter overall folding of MUTYH; however, the presence of Zn^{2+} is absolutely essential for high affinity binding to the damaged 8-oxoG:A mismatch required for catalysis of adenine excision and DNA mutation suppression. Notably, participation of Cys230 in the structural models generated herein also aids in rationalizing the requirement of Zn^{2+} coordination for high affinity substrate binding by MUTYH. Indeed, Zn^{2+} coordination mediated by Cys230, along with the three Cys residues in the IDC, may help to fold the IDC in a manner that allows for stabilization of the MUTYH-DNA binding interface.

EXPERIMENTAL SECTION

Sequence Searching, Alignment & Filtering.

A search of the NCBI non-redundant sequence database (1/1/2015) using the *Homo sapiens* MUTYH (sp_Q9UIF7–3α) sequence on the HMMER3 server yielded ~9000 related protein sequences.[29] These sequences were aligned using PROMALS3D[30] and then filtered to remove any sequences that did not contain the two key MUTYH catalytic residues (Glu120, Asp222) and the two critical Zn^{2+} binding residues (Cys325, Cys328). Species annotations are associated for every sequence name and spans three different databases to obtain species information, with one of the following three annotations as examples:

1. >533170473|kingdom_Eukaryota|species_Chinchilla_la nigera

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2. >533170469|EntrezAnnot_[Chinchilla_lanigera]

3. >M4BEK6_HYAAE|UniProtAnnot_[Hyaloperonospora_ar abidopsidis]

where annot1 is the NCBI GI number, followed by the kingdom, followed by the species, annot2 is the NCBI GI number, followed by the NCBI Entrez species annotation, and annot3 is the UniProt accession code, followed by the UniProt Species annotation.

Computational Modeling of MUTYH.

Homology modeling of human MUTYH was performed using the program Modeller[31] based on the crystal structures of the human MUTYH N-terminal fragment (PDB ID: 3N5N)[23] and *Geobacillus stearothermophilus* full-length MutY (PDB ID: 5DPK)[12], as well as the human MUTYH IDC amino acid sequence. The resulting structure positions the C- and N-terminal domains in correct orientations to surround the DNA. However, the IDC, which wraps around the DNA, shows irregular and non-native structural features and appears as a long, disordered loop due to the difference in sequence length between human MUTYH and bacterial MutY in this region.

The ability of the implicated Cys to participate as a ligand to the Zn^{2+} ion was assessed using a type of steered molecular dynamics (MD), akin to umbrella sampling.[32] The modelled full structure of human MUTYH was first energy minimized and then subjected to MD simulations in FACTS implicit solvent model[33] using the CHARMM27 force field[34–35] for the protein and the Babu-Lim force field for interactions with Zn^{2+}.[36] For the Fe-S cluster, the topology of the $[4Fe\text{-}4S]^{2+}$ moiety was built using distances, angles and dihedrals from the crystal structure. Partial charges were distributed among the 4 Fe, 4 inorganic S and 4 Cys thiolates such that the total charge of the 4Fe-4S-4Cys cluster is −2e. All simulations were performed at pH 7, a mean temperature of 298 K and 1 atmosphere using the CHARMM39 program.[37]

As Cys230 is the only residue in the N-terminal domain that is similarly conserved to the other Cys ligands and spatially positioned to coordinate the Zn^{2+} ion, it was chosen as the target of the umbrella potentials. Initially, all the protein residues, except those belonging to IDC, were fixed, and umbrella potentials of strength 10 kcal/mol/Å were imposed between S atoms of Cys230 and other three S atoms of Cys318, Cys325 and Cys328 in human MUTYH. All three S atoms were brought to within 7Å of the Cys230 S atom.

The Zn^{2+} ion was placed at a coordination distance of 2.3Å from the Cys230 S atom. Constraint dynamics was repeated with umbrella potentials acting between Zn^{2+} and each of the three S atoms of Cys318, Cys325 and Cys328 while keeping Zn^{2+} and Cys230 S atoms frozen. After many trial steps, we were successful in bringing the three S atoms within the Zn^{2+} coordination sphere. In the next stage, all constraints were released, and the system was re-equilibrated followed by 5 ns of production dynamics. The final structure showed that Zn^{2+} can interact favorably with all four Cys residues; however, it also interacted with a carbonyl O of neighboring residues forming a square pyramidal structure. Such a 5-corodinate geometry, which is atypical of similar Zn^{2+}-sites in X-ray structures, likely stem from inadequacies of the implicit solvent model used and the classical force field for Zn^{2+} (see Results).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Consequently, the final structure was solvated in explicit TIP3P water molecules.[38] The entire protein was centered in a cubic box with an edge length 90Å, containing 27,000 TIP3P water molecules corresponding to an experimental water density of 1gm/cc. Water molecules whose oxygen atoms fell within 2.7 Å from protein heavy atoms were removed to avoid overlaps. The resulting final system contained 24,418 water molecules. Nonbonded interactions were truncated by a force switching function between 10 and 14Å. All bonds involving hydrogens were constrained by the SHAKE algorithm.[39] After energy minimization and 1-ns equilibration, a 2-ns production trajectory was generated. In the final structure, one of the Zn^{2+}-bound Cys (Cys328) drifted away from the first shell and was replaced by a water molecule from bulk solvent (see Results).

Since the classical force field used for Zn^{2+} does not include charge transfer and polarization effects,[40–41] we next employed electronic structure methods to account for such effects. Therefore, we subjected the final structure from explicit water simulations to a cycle of quantum mechanical/molecular mechanical (QM/MM) energy minimizations using a protocol similar to that used on metallothioneins.[28] The QM region consisting of Zn^{2+}, the four Cys ligands and a bound water molecule were treated using the S-VWN/SDD method. [42] It was connected to the MM system consisting of the rest of the protein and water molecules using hypothetical H link atoms between the C^{α} and C^{β} of each of the four Cys residues. The solvent and protein atoms were frozen, and the Zn^{2+} and its ligands were subjected to QM/MM energy minimizations in the frozen field of MM charges. This was followed by a short equilibration run and a 1 ns of production dynamics of the MM part with the QM part frozen. This process consisting of QM/MM minimization, followed by equilibration and a 1 ns of production dynamics was repeated five times. The QM/MM minimizations utilized the CHARMM39/GAUSSIAN 09 interface with the QM calculations performed with Gaussian 09 program.[42]

Mutyh Expression and Purification.

This construct was previously created by inserting the *Mus musculus* Mutyh gene lacking the first twenty eight N-terminal amino acids into the *pET28a* expression vector (EMD Biosciences) equipped with an N-terminal six His tag to aid in purification.19 The antisense primer 5’ – GGCCCGGACACGGCTCAGCACCCGTAAAAC - 3’ was designed to create the Cys215Ser single point mutation. Furthermore, the double variant (Cys215Ser/ Cys300Ser) was created using antisense primer for position Cys300 using the sequence 5’ – GGCACTGTCTAGTGTTGAGAGCAGACTCCTCTATGTCAGGACGGCC - 3’ in the Mouse Mutyh gene followed by Kunkel mutagenesis (Transcriptic).

The following protein purification is based on a previous protocol,[19, 43–44] with the following optimizations. BL21 DE3 (NEB) competent cells co-expressing a *pRKISC* plasmid were transformed with the Mutyh *pET28a* plasmid by a 45 second heat shock. Transformed cells were streaked onto LB agar plates with appropriate antibiotic selection (in this case, 15 μg/ml tetracycline and 34 μg/ml kanamycin) and incubated overnight at 37 °C. A single colony was used to inoculate 50 ml of LB starter media supplemented with the same antibiotics. Following overnight incubation at 37 °C and 220 rpm, the culture was expanded to 2 L and incubated at 37 °C and 150 rpm until OD_{600nm} reached 0.8.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Overexpression was induced with 1 m*M* IPTG and 0.1g/L $FeCl_2$ and the temperature reduced to 30 °C. Following overexpression for 6 h, the culture was centrifuged at 7,500 rpm for 15 mins at 4 °C, and the pellet was resuspended in resuspension buffer consisting of 20 m*M* sodium phosphate buffer pH 7.5, 10 % glycerol, and 1 protease inhibitor cocktail tablet (Roche) and stored at −80 °C.

Pelleted cells were thawed at 37 °C and 150 rpm for 30 mins and lysed via sonication on ice for 6 mins in 30 sec cycles using a Branson Sonifier 250, followed by centrifugation at 12,000 rpm for 15 mins at 4 °C. Imidazole and NaCl (20 m*M* and 1 *M* final concentrations respectively) were added to the resuspended pellet and batch bound to Ni^{2+}-NTA resin (Qiagen) for 1 h at 4 °C with rotation. The slurry was poured over a PD10 column and allowed to flow through via gravity. The protein-loaded resin was washed with 10 ml wash buffer (20 m*M* sodium phosphate buffer pH 7.5, 10 % glycerol, 1 *M* NaCl and 20 m*M* imidazole) followed by 3 ml elution buffer (20 m*M* sodium phosphate buffer pH 7.5, 10 % glycerol, 300 m*M* NaCl and 500 m*M* imidazole).

The elutant was then concentrated using an Amicon ultrafiltration cell with a 10,000 MWCO filter while stirring and diluted 10-fold with heparin buffer A (20 m*M* sodium phosphate buffer pH 7.5, 5 % glycerol, 1 m*M* EDTA). 2 U thrombin (Novagen) per mg Mutyh was added and incubated for 16–20 h to cleave the His tag. A final concentration of 1 m*M* PMSF was added and incubated for 2 h. Imidazole and NaCl (20 m*M* and 1 *M* final concentration respectively) were added after incubation, followed by batch binding to Ni^{2+}-NTA resin for 1 h, then poured over a PD10 column along with 2 ml wash buffer. The solution was diluted 10-fold in heparin buffer A, filtered with a 0.2 μm filter and added to a 5 ml Pharmacia Hi-Trap heparin column on an AKTApurifier FPLC system. Mutyh was eluted via linear gradient in heparin buffer A to 100 % heparin buffer B (heparin buffer A plus 1 *M* NaCl), over 20 column volumes with resulting associated fractions combined.

Fractions were concentrated (same as above) and buffer exchanged into Mutyh concentration buffer (20 m*M* sodium phosphate buffer pH 7.5, 25 % glycerol, 1 m*M* EDTA, and 200 m*M* NaCl) and stored as single use aliquots at −80 °C. Purity of Mutyh samples were confirmed via 12 % SDS PAGE stained with SYPRO Ruby stain as according to the manufacturer's protocol. The concentration of total protein was determined using ε_{280} = 86,120 $M^{-1}cm^{-1}$, and the concentration of the iron sulfur cluster was estimated using ε_{410} = 17,000 $M^{-1}cm^{-1}$.

Inductively Coupled Plasma-Mass Spectrometry.

Samples were prepared as previously described,[19] where samples and method blanks were prepared in a range of 500 – 250 μl Mutyh concentration buffer (same components as above). Samples and blanks were submitted in triplicate to the UC Davis Interdisciplinary Center for Inductively-Coupled Plasma Mass Spectrometry (ICP-MS). Data are represented as the mole ratio of metal ion to enzyme ([E] by A_{280nm}, where ε_{280} = 86,120 $M^{-1}cm^{-1}$).

Preparation of Oligonucleotide Substrates.

DNA substrates containing 8-oxoG (8-oxo-7,8-dihydroguanosine) or FA (2'-deoxy-2'-fluoroadenosine) were synthesized at the University of Utah DNA and Peptide Synthesis

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Core Facility and their complementary strands were purchased from Integrated DNA Technologies (IDT). A 30 bp oligonucleotide (5′ - CGATCATGGAGCCACXAGCTCCCGTTACAG - 3′) and its complement, (5′ - CTGTAACGGGAGCTYGTGGCTCCATGATCG - 3′) with a central X (8-oxoG) and Y (A or FA) were generated for *in vitro* studies as previously described.[43] DNA used in fluorescence polarization studies also incorporated a 5' 6-FAM (6-carboxyfluorescein) on the FA containing strand. Oligonucleotides containing the central 8-oxoG or FA were deprotected and cleaved from the solid support column by incubation in NH_4OH, with the addition of 2-mercaptoethanol to 8-oxoG samples. The cleaved DNA substrates were dissolved in H_2O, filtered with a 0.2 μm filter, and HPLC purified using a Beckman Gold Nouveau system with a Waters AP1DEAE 8-HR column with a 100–10% gradient of 90:10 H_2O/acetonitrile with 2 *M* NH_4Ac. Isolated fractions were lyophilized and desalted with a SEP-PAK C18 column, and DNA integrity confirmed via MALDI-MS. Concentration of purified DNA strands were quantified by ε_{260} values determined by IDT.

Glycosylase Assays.

The adenine glycosylase assay was performed as previously described.[43–45] Under multiple-turnover (MTO) conditions, the reaction is initiated by addition of 5 to 10 n*M* Mutyh, wild-type or variant, to 20 n*M* DNA in 20 m*M* Tris-HCl, pH 7.6, 1 m*M* EDTA, 0.1 mg/ml BSA, and 30 m*M* NaCl at 37 oC. At time points from 20 secs to 1 h, an 8 μl reaction aliquot was quenched with 2 μl of 1 *M* NaOH followed by heating at 90 oC for 3 mins. An equal volume of formamide loading dye was added and incubated up to 5 mins at 90 °C. Samples were loaded onto a 15% (19:1) denaturing polyacrylamide gel and electrophoresed at 1,200 to 1,500 V for 2 h in 1X TBE buffer. Following electrophoresis, the gel is wrapped and exposed on a phosphor screen overnight for image capture. The screen was scanned on a Typhoon imager and quantified using ImageQuaNT (v.5.2), and data is graphed employing GraFit (v.5). Under single-turnover (STO) conditions, the assay is carried out as just described except for the active enzyme is in excess [active E] > [DNA], where the [DNA] is 20 n*M* and [E] = 100 n*M*.

Circular Dichroism.

Circular Dichroism (CD) Spectroscopy was employed to determine the secondary structure of WT mouse Mutyh and corresponding Zn^{2+} linchpin variants (Cys215Ser, Cys300Ser, Cys307Ser, Cys310Ser and double variant Cys215Ser/Cys300Ser), alongside BSA as a standard control with known secondary structure.[46–47] Samples were buffer exchanged in a Amicon Ultra-4 centrifugal unit with a 10,000 MWCO filter using an optimized buffer containing 10 m*M* potassium phosphate and 50 m*M* sodium sulfate. Samples were concentrated to approximately 0.1 mg/ml at a final volume of 500 μl, filtered with a 0.2 μm filter, followed by exact concentration determination utilizing UV/vis. Samples were ran using a 1 mm CD quartz cell at room temperature, utilizing a CD spectrophotomer (Jasco J720) and scanned at a range of 190–240 nm. Acquired data was converted from millidegrees to delta epsilon ($\Delta\varepsilon$, $M^{-1}cm^{-1}$) by normalization of the sample concentration and mean residue weight.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Fluorescence Polarization Assay.

Fluorescence polarization assays were performed to determine the dissociation constant (K_d) using a 6-FAM (6-carboxyfluorescein) 5'- labeled 30 oligonucleotide duplex DNA with a central 8-oxoG:FA base pair (see *oligonucleotide substrates* above). A DNA master mix solution was made consisting of 20 m*M* Tris-HCl, pH 7.5, 1 m*M* EDTA, 100 m*M* NaCl, 0.1 mg/ml BSA, 1 m*M* DTT and 4 n*M* labeled duplex DNA. The DNA master mix was aliquoted out to be added in equal volumes to tubes containing WT or variant Mutyh with concentrations ranging from 10 μ*M* to 0 μ*M*. The DNA and enzyme were incubated for 30 mins at 25 oC and were loaded in triplicate into a 384-well plate alongside a no-enzyme control. Samples were ran on a BMG Labtech Clariostar multimode plate reader with an excitation of 482–16, Dichroic F:LP 504, emission set to F:530–40 with top optic and a gain adjustment set to 60. MARS data analysis software was used to determine extent of polarization at each enzyme concentration, followed by blank subtraction and normalization to the polarization observed at maximum enzyme concentration. Average binding curves include three to five separate trials each and were established by plotting percent bound versus enzyme concentration using a one site-binding model in the graphing software GraFit (v.5).

Rifampicin Resistance Assay.

For site directed mutagenesis experiments, the mutagenesis primer 5' - GGCAACGTAGCACGGGTGCTGAGCCGTGTCCGAGCCATTGGTGC - 3' and its complement were used to create a Cys to Ser single point mutation at position 230 in human MUTYH. Briefly, 50 ng of the WT and Cys318Ser human MUTYH plasmids[19] were used in the primer extension reaction to create the mutation, along with 1.1 μl of a 25 μ*M* stock of each primer, 1X final dNTP mixture, and 1 μl polymerase with its corresponding GC buffer. The primer extension reactions were performed in a thermocycler under the following conditions: 95 °C × 30 sec; 18 cycles at 95 °C × 30 sec, 55 °C × 1 min, and 72 °C × 4.5 min; with a final extension of 72 °C × 6 min, followed by a 4 °C hold. The primer extension products were subjected to a DpnI treatment for 1 h at 37 °C to degrade any methylated DNA. The resulting DNA was then transformed into the *E. coli* strain DH5α and plated on LB selection plates supplemented with 100 μg/ml ampicillin and grown overnight at 37 °C. DNA was isolated using the Promega PureYield™ Plasmid Prep System and sequenced to confirm the correct human MUTYH Cys230Ser mutation in each plasmid.

To perform the assay as previously described,[48] the human MUTYH WT, Cys230Ser, Cys318Ser and Cys230Ser/Cys318Ser double variant plasmids were individually transformed into the *E. coli* strain GT100 mutY:mini-Tn10 mutM and incubated at 37 °C overnight. A minimum of 15 single colonies were picked for each variant to start 1 ml overnight cultures in LB media containing 100 μg/ml ampicillin and 15 μg/ml tetracycline. Cultures were grown overnight for 16–18 h at 37 °C at 220 rpm.

To determine the number of viable cells, 100 μl of a 10^{-7} fold dilution of the 1 ml culture was plated on LB/Amp/Tet plates. To determine the number of rifampicin revertant cells, 100 μl of each overnight culture was plated on LB/Amp/Tet/Rif plates. LB/Amp/Tet/Rif selection plates were poured with the above ampicillin and tetracycline concentrations, and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

supplemented with 100 μg/ml rifampicin. Plates were incubated at 37 °C overnight, and the resulting colonies on each set of plates counted. The median number of resistant colonies from LB/Amp/Tet/Rif plates was divided by the average number of viable colonies from LB/Amp/Tet plates to calculate the mutation frequency (f) for each variant. Mutation frequencies for each clone were calculated and plotted on a boxplot. Outliers were removed if they were outside of the interquartile range.

RESULTS

Sequence Alignment.

To reveal a fourth ligand candidate, we performed a sequence search and alignment of approximately 9,000 MUTYH orthologs. These sequences were filtered for two strictly conserved MUTYH catalytic residues (Glu120 and Asp222 in human MUTYH) and the two vital Zn^{2+}-coordinating Cys ligands of the C-X_2-C motif (Cys325 and Cys328 in human MUTYH).[12, 19] This resulted in 149 sequences, all identified as eukaryotic orthologs. The resulting list of sequences includes the two genes used in this paper, *Homo sapiens* sp_Q9UIF7–3 and *Mus musculus* sp_Q99P21. Human MUTYH Cys325 and Cys328 were identified to be present in even primitive eukaryotes, such as green algae, amoebas, fungi, flatworms and seventeen protist species. Interestingly, the third Cys ligand of the Zn^{2+} linchpin, Cys318, was found to be only 54% conserved compared to Cys325 and Cys328; Cys318 was demonstrated to first originate from the phylum of Echinodermata. These species and higher eukaryotes were shown to possess the seven highly conserved Cys residues: four for the Fe-S cluster, and three for the Zn^{2+} linchpin (Figure S1). The sequence alignment revealed a 4th ligand candidate in human MUTYH, Cys230, whose conservation is similar to Cys318: Cys230 and Cys318 are 52% and 54% retained, respectively, in eukaryotes compared to the highly conserved Cys325 and Cys328 ligands in human MUTYH (Figure S1). However, unlike Cys318, Cys230 is not present in invertebrates such as the acorn worm or starlet sea anemone, nor ghost sharks that diverged from modern day sharks nearly 400 million years ago. Cys230 is located within the N-terminal domain close to the Fe-S cluster cofactor, 87 amino acids before the previously identified MUTYH Zn^{2+} coordinating ligands (Cys318, Cys325 and Cys328) that are located in the IDC. Other conserved potential metal coordinating residues were also identified but were ruled out based on their location in the homology model structure (vide infra) or previous mutagenesis studies.[19]

Modeling MUTYH and Metal Sites.

We first generated a homology model structure of human MUTYH (Uniprot sp_Q9UIF7–3α) and this structure shows the spatial possibility of the four proposed Cys residues participating in the coordination of Zn^{2+} (Figures S2 and S3). Notably, no other plausible ligand (O, N or S containing), with sequence conservation similar to the previously identified Zinc Linchpin residues, are within a reasonable distance to the Zn^{2+} site in the MUTYH modeled structure (Figure S4). In the N-terminal fragment crystal structure of human MUTYH (PDB ID: 3N5N)[23], the region containing the Zn^{2+} ligands, Cys318, 325 and 328 within the IDC was disordered, suggesting that this region of the IDC is unfolded in the absence of Zn^{2+}. Therefore, we constructed a Zn^{2+}-binding coordination model that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

involves the three Cys residues in the IDC and the proposed fourth ligand Cys230, located within the N-terminal domain (see Methods).

For efficient sampling, we first examined Zn^{2+} binding in an implicit solvent model that predicted coordination of all four Cys residues to Zn^{2+}; however, in this structure a square-pyramidal geometry was observed with axial coordination by oxygen of an amide backbone carbonyl (see Figure S5A). Subsequent solvation with explicit water molecules yielded a tetrahedral coordination sphere with Zn^{2+} bound directly to Cys230, Cys318, Cys325 and a water molecule that mediated the interaction of Cys328 which had drifted to the outer-sphere (see Figure S5B). Attempts to restrain Cys328 in the inner shell and subsequent reequilibration/dynamics of the system resulted in both Cys318 and Cys325 drifting from the inner-shell, while neighboring carbonyl O atoms and a water molecule became bound to Zn^{2+} forming an octahedral complex, as shown in Figure S5C. Therefore, the simulations in explicit water do not favor Cys328 in the inner-sphere likely due to residual strain present in the starting homology modelled structure and the classical force field used for Zn^{2+}, which neglects charge transfer and polarization effects that are important for Zn^{2+}-Cys^{-} interactions.[40–41] Nevertheless, Cys230 is seen coordinated to Zn^{2+} in both implicit and explicit solvent models, indicating that this residue is likely a key component of the Zn^{2+} coordination sphere.

To alleviate the inadequacies of the classical force field used, the final structure from explicit water simulations shown in Figure S5B was subjected to a cycle of quantum mechanical/molecular mechanical (QM/MM) energy minimization, as described in Methods. The final QM/MM energy-minimized structure is indicative of Zn^{2+}-mediated structural stabilization near the DNA-binding region of MUTYH (Figure 3, Figure S6). An overlay with the *Geobacillus stearothermophilus* MutY X-ray crystal structure (PDB ID: 5DPK)[12] shows similarity to the QM/MM structure of MUTYH in both the N and C-terminal domains, with the greatest difference between the orthologous structures noted just before and within the IDC region (Figure S6). The QM/MM structure is consistent with Cys230 serving as the fourth Zn^{2+} ligand in human MUTYH. Cys230 is bound to Zn^{2+} at a distance of 2.3 Å, typical of Cys-bound to tetra-coordinated Zn^{2+} in protein structures. However, Cys328 remained in the outer-sphere (as noted above) at a distance of 4.4 Å from Zn^{2+} due to interaction with the metal-bound water molecule. Since Cys328 mutation to Ser had previously been shown to strongly influence Zn^{2+} content,[19] it is likely to remain coordinated. Importantly, the classical simulations in both implicit and explicit solvent as well as the QM/MM simulations provide evidence for Cys230 as a potential fourth ligand to Zn^{2+}. In addition, the coordination to Cys230 results in folding of the IDC region and proximity to the DNA binding domain adjacent to the Fe-S cluster of MutY. Indeed, the final structure and overlay with the bacterial structure provides a compelling explanation for how Zn-mediated folding would stabilize this region to potentially mediate more favorable interactions with the DNA substrate. The computational studies along with the similar sequence conservation of Cys230 and Cys318 provided motivation for further evaluation of the impact of mutation of Cys230 on metal content and activity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Metal Analysis of Cys to Ser Variants.

The high sequence similarity of mouse Mutyh and human MUTYH and the higher levels of active fraction of purified mouse Mutyh prompted its use for *in vitro* studies.[19, 49] To test the hypothesis that human MUTYH Cys230 is the fourth Zn^{2+} chelating ligand, we utilized the homology in mouse Mutyh and replaced the corresponding residue, Cys215, alone and together with Cys300 (equivalent to Cys318 in human MUTYH) to Ser and evaluated the impact on levels of coordinated Zn^{2+}. WT mouse Mutyh and the Cys to Ser variants were overexpressed and purified via an optimized protocol that provides for reproducibly high levels of active protein.[19] Overexpression and purification of WT Mutyh, single variants Cys215Ser and Cys300Ser, and the double variant (DV) Cys215Ser/Cys300Ser were analyzed for iron and zinc content using inductively-coupled plasma-mass spectrometry (ICP-MS). WT mouse Mutyh and the single variants displayed similar levels of Zn^{2+} and slight variances in retention of the $[4Fe\text{-}4S]^{2+}$ cluster (Table 1). Notably, the double variant displayed diminished Zn^{2+} loading compared to WT mouse Mutyh or either single variant, suggesting a synergistic chelation of Zn^{2+} by Cys215 and Cys300.

Active Fraction.

The adenine glycosylase activity of Cys215Ser and Cys300Ser, and double variant Cys215Ser/Cys300Ser mouse Mutyh was measured with an 8-oxoG:A-containing 30 bp duplex substrate labeled at the 5'-end of the A-containing strand with $[^{32}P]$-phosphate.[43, 45] Adenine removal was assessed by the extent of strand scission at the abasic site produced by MutY after base quenching via denaturing PAGE. Under multiple-turnover (MTO) conditions ([Mutyh] < [DNA]), a biphasic production curve was observed consisting of a rapid exponential "burst" phase followed by a slow linear steady-state phase due to rate-limiting release of the abasic site-containing product DNA duplex. The amplitude of the burst phase directly relates to the active concentration of the Mutyh.45 The ratio of active enzyme concentration to the enzyme concentration determined at $A280_{nm}$ corresponds to the percent active fraction (Table 1, Figure 4). The mouse Mutyh Cys215Ser/Cys300Ser double variant displayed a substantial drop in active fraction compared to WT or either of the single variants. Notably, the active fraction and levels of Zn^{2+} of the double variant relative to WT directly correlated.

Intrinsic Rate of Glycosidic Bond Cleavage.

Adenine glycosylase assays performed under single-turnover (STO) conditions where the active enzyme concentration is greater than both that of the substrate concentration and K_d allow for isolation of the rate constant k_2 (Figure 4A).[45] Notably, these experiments were performed using the active enzyme concentration to ensure single-turnover conditions to allow for the k_{obs} to k_2 simplification.[12] In these experiments, all of the Cys to Ser variants (single and double) exhibited k_2 values similar to that of WT mouse Mutyh (Table 1, Figure 4). A reduction in the extent of completion was observed with the double variant (Figure 4) and may be an indication of some reduced catalytic fitness. Notably, the correlation of percent active fraction to Zn^{2+} loading indicates that only the Zn^{2+}-loaded Mutyh is active. The lack of sensitivity of the observed rate to altered coordination around the Zn^{2+} ion is consistent with the expectation that the Zn^{2+} ion does not participate directly in base

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

excision catalysis.12 Moreover, this is consistent with our proposed role for the Zn^{2+} linchpin motif in mediating proper interaction with the damaged substrate.

The Zn^{2+} Linchpin Mediates High Affinity for the DNA Substrate.

The correlation of active fraction in mouse Mutyh double variant Cys215Ser/Cys300Ser with the reduced coordination of Zn^{2+} supports the role of the Zinc Linchpin motif in mediating proper engagement of the 8-oxoG:A mismatch required for efficient base excision. We anticipated that this feature would also be reflected in higher affinity of Mutyh forms harboring higher levels of coordinated Zn^{2+}. To test this hypothesis, we designed a fluorescence based *in vitro* assay that utilized a 5' 6-FAM labeled DNA duplex with a central 8-oxoG opposite a noncleavable 2'-deoxyadenosine analog, 2'-deoxy-2'-fluoroadenosine (FA).[50] The competence of WT mouse Mutyh and associated Zn^{2+} ligand Cys to Ser variants to recognize the substrate mispair was gauged by relative apparent K_d values determined in these assays. The robustness of the assay was verified with the WT Mutyh enzyme, which provided K_d values similar to those previously reported using electrophoretic mobility shift assays (Table 1).[51]

The dissociation constants (K_d) were determined for WT mouse Mutyh and the single Cys-to-Ser mouse Mutyh variants (Cys215Ser, Cys300Ser and Cys307Ser), and the double mouse Mutyh variant (Cys215Ser/Cys300Ser) (Table 1, Figures 5 and S7). Based on the total protein concentration, the DNA affinity of each Cys300Ser and Cys215Ser single variant, is similar to WT Mutyh (K_d = 50 ± 8 n*M*). The double variant Cys215Ser/Cys300Ser showed reduced affinity (K_d = 136 ± 14 n*M*). Moreover, mouse Mutyh Cys307Ser (corresponds to human MUTYH Cys325Ser) which has very low levels of Zn^{2+} (~10% of WT), had an even larger 10-fold increase in K_d (560 ± 80 n*M*) relative to WT mouse Mutyh.19 In all cases, the extent of reduced affinity correlates directly with the reduced fraction of Zn^{2+}. In addition, adjusting the enzyme concentration used to take into account percent active fraction results in collapse of the K_d values for all forms to that of WT Mutyh (Figure S7). This underscores that adenine base excision occurs only by the fraction of enzyme with WT-like 8-oxoG:A mismatch affinity. The correlation of substrate affinity, Zn^{2+} loading and active fraction indicates that the differences in binding affinity imparted by the presence or absence of Zn^{2+} are significant to profound alter the ability of Mutyh to excise adenine.

Structural Impact of Zn^{2+} coordination within Mutyh.

The correlation between the amount of coordinated Zn^{2+}, active fraction and high substrate affinity prompted further evaluation of the influence of Zn^{2+} coordination on the overall folding of mouse Mutyh. Circular dichroism (CD) spectroscopy can detect differences in secondary and tertiary structures of proteins[46–47] and therefore was used to reveal potential structural differences between WT Mutyh and associated Zn^{2+} linchpin variants. The CD spectra of all four Zn^{2+} linchpin Cys-to-Ser single variants and the double variant (Cys215Ser/Cys300Ser) in mouse Mutyh exhibited CD spectra similar to the WT (Figure 6 and S8). Interestingly, the alpha helical signal with Cys307Ser and Cys310Ser Mutyh is slightly larger than WT Mutyh, suggesting that there are local structural changes in the absence of Zn^{2+}. Previous ICP-MS analysis demonstrated very low levels of zinc and iron in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cys307Ser and Cys310Ser Mutyh.[19] Previous studies of *E. coli* MutY showed that the protein is capable of refolding to an overall native state despite the lack of a bound Fe-S cluster; however, the Fe-S cluster is required for activity.[9] Notably, the double variant Cys215Ser/Cys300Ser of mouse Mutyh harbored reduced levels of Zn^{2+} but retained the Fe-S cluster cofactor, and exhibited a CD spectrum that has only minor differences from WT Mutyh. The results from the CD experiments are consistent with the absence of major secondary structural changes due to Zn^{2+} coordination. This suggests that the impacts on the functional properties are due to changes in local structure that fold the IDC consistent with the structural models generated herein (Fig. 3 and Fig S4).

Influence of Zn^{2+} Linchpin on Suppressing DNA Mutations.

The ability of human MUTYH and variants to suppress mutations may be conveniently monitored within *E. coli* using the rifampicin resistance assay.[19, 52–53] This assay relies on the accumulation of mutations within the *rpoB* gene encoding the RNA polymerase rpoB subunit which prevents inhibition by the antibiotic rifampicin.[45] Cells expressing WT human MUTYH can complement for the absence of bacterial MutY to efficiently suppress DNA mutations, resulting in low numbers of colonies in the presence of rifampicin. However, expression of MUTYH variants that have compromised activity results in increased mutation frequency and increased cell survival in the presence of rifampicin. The number of rifampicin revertant colonies relative to the total number of cells in an untreated sample is reported as the mutation frequency (f).[52, 54]

In cells expressing human MUTYH single Cys to Ser Zn^{2+} linchpin variants Cys230Ser and Cys318Ser, low mutation frequencies, similar to those for expression of WT human MUTYH, were observed (Table 2, and Figure S9). However, in cells expressing the double variant Cys230Ser/Cys318Ser MUTYH there is a significant increase in mutation frequency (Table 2). Of note, the corresponding double variant in mouse Mutyh (Cys215Ser/Cys300Ser) was found to have significantly reduced levels of Zn^{2+}. Indeed, the increase of the mutation frequency, in conjunction with a decrease in Zn^{2+} content, glycosylase activity and binding affinity of the double Cys to Ser variants (in mouse or human enzyme), provides strong evidence for the role of Cys230 and Cys318 to participate as a Zn^{2+} ligands in human MUTYH.

In analogous experiments with bacteria expressing Cys325Ser or Cys328Ser human MUTYH, exceptionally high mutation frequencies were observed similar to that observed for bacteria lacking MUTYH or MutY. Notably, the corresponding variants in mouse Mutyh (Cys307Ser and Cys310Ser) harbor low levels of Zn^{2+} *and* Fe-S cluster, [19] suggesting that the dramatic increase in mutation frequency may be due to loss of both cofactors. With the human MUTYH double variant Cys230Ser/Cys318Ser, the increased mutation frequency may be more readily attributed to only the loss of Zn^{2+} coordination. Thus, these results further support a critical role of the Zn^{2+} linchpin in the DNA repair activity of MUTYH.

DISCUSSION

In this work, we identify and confirm the complete coordination sphere of the Zn^{2+} ion within the IDC of human MUTYH and illustrate the key role of this Zn^{2+} linchpin motif in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

orchestrating the lesion recognition and damage excision functions of the MUTYH glycosylase. The Zn^{2+} K-edge EXAFS region of the XAS spectrum of Mutyh fit best to a model with four sulfur ligands, consistent with coordination by the previously identified Cys residues, Cys318, Cys325 and Cys328 in the IDC of human MUTYH.[19] Replacement of two of the Cys residues, corresponding to the Cys-X_2-Cys motif (Cys325 and Cys328 in human MUTYH) in the IDC, with Ser were previously shown to result in loss of Zn^{2+} as well as Fe, and significant impairment of repair activity *in vitro* and in cells.[19] In this work, we took advantage of large scale database mining and sequence alignments of MUTYH orthologs to reveal a candidate fourth Cys residue, Cys230 that exhibits a similar degree of conservation as Cys318, a residue implicated as a Zn^{2+} linchpin ligand based its proximity to established ligands, Cys325 and Cys328, located within the IDC of human MUTYH. The newly identified ligand Cys230 is located within the N-terminal domain and is adjacent to the Fe-S cluster, and its associated DNA binding domain. MD simulations of modelled structures in explicit water, followed by QM/MM energy minimization of the Zn^{2+} coordination sphere, showed the feasibility of Cys230 as a ligand to complete the Zn^{2+} coordination sphere of human MUTYH.[19]

Replacement of Cys230, along with Cys318, to Ser in MUTYH and the corresponding residues in mouse Mutyh (Cys215 and Cys300) support their participation in chelation of the Zn^{2+} ion. The double variant of mouse Mutyh, Cys215Ser/Cys300Ser, harbors significantly reduced levels of coordinated Zn^{2+}. The observed levels of Zn^{2+} were not significantly reduced in the single Cys to Ser mutation of either Cys215 or Cys300 likely due to the strong Zn^{2+} chelation afforded by the Cys307-X_2-Cys310 motif of mouse Mutyh. Of particular note, the *in vitro* glycosylase assays show a correlation between the fraction of active enzyme and the levels of bound Zn^{2+} in the double variant of mouse Mutyh, Cys215Ser/Cys300Ser. Similarly, this double variant showed reduced affinity for an 8-oxoG:FA-containing duplex. In addition, the corresponding human MUTYH double variant (Cys230Ser/Cys318Ser) exhibited a significantly compromised ability to suppress DNA mutations in rifampicin resistance assays. Taken all together, this provides compelling evidence that Cys230 and Cys318, along with the previously established Cys325 and Cys328, participate as ligands to the zinc ion in the Zn^{2+} linchpin motif of human MUTYH.

The differential impact of the mutations of the four Cys to Ser variants on levels of coordinated Zn^{2+} and the corresponding functional consequences also reveal insight into role of the zinc linchpin motif in MUTYH-mediated repair. Notably, there is a direct correlation between levels of Zn^{2+} among the entire set of Cys to Ser mutated forms, and respective active enzyme fraction that suggests that only enzymes containing Zn^{2+} are active. This is also consistent with previous work that showed that removal of Zn^{2+} from Mutyh by chelation with 1,10-phenanthroline inactivated the enzyme.19 These results, taken together with the features of base excision catalysis[12] of the bacterial enzyme that lacks the Zn^{2+} linchpin,[19] suggest that Zn^{2+} coordination within the linchpin motif is required for proper engagement of MUTYH on the 8-oxoG:A mismatch to position adenine within the active site for its excision. This idea is further supported by measurements of DNA affinity of the mouse Mutyh double variant Cys215Ser/Cys300Ser (Cys230Ser/Cys318Ser in human MUTYH), and the four single Cys to Ser variants. The forms with reduced levels of Zn^{2+} were able to bind to substrate analog DNA but with reduced affinity; indeed, the magnitude

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of the reduced affinity directly correlates with the extent of Zn^{2+} bound and active fraction. CD studies further demonstrate that loss of Zn^{2+} does not significantly alter overall folding of mouse Mutyh. We also show that reduction of bound Zn^{2+} impacts the ability of MUTYH to mediate suppression of mutations in cells. These results underscore the importance of the Cys-coordinated Zn^{2+} linchpin motif for proper 8-oxoG:A mismatch DNA binding required for proper function of MUTYH in preventing DNA mutations.

In the structural studies of *Gs* MutY, along with the structural models for MUTYH determined herein (Figure 3, Figures S2-S6), the IDC traverses across the DNA major groove, linking the C-terminal 8-oxoG recognition and the N-terminal catalytic domains.[1] Notably, the results herein indicate that the extended IDC in mammalian orthologs and its ability to coordinate Zn^{2+} is required for proper engagement and positioning of the two functional domains. Indeed, the coordination of the Zn^{2+} by Cys230, along with the IDC Cys residues, provides a connection between the DNA binding region of the Fe-S cluster domain in the N-terminal domain and the IDC (Figure 3, and Figure S4). The sequence analysis revealed that many lower eukaryotes harbor only two of the four Cys residues, which suggests that the enzymes from these organisms may lack a coordinated Zn^{2+} ion. A chimeric protein containing the N- and C-terminal domain of *Schizosaccharomyces pombe* (Sp) MutY and the IDC linker of *E. coli* MutY exhibited reduced glycosylase activity and reduced OG:A mismatch affinity.[23] Surprisingly, however, the Sp/Ec chimeric protein exhibited *increased* affinity for undamaged DNA suggesting an active role of the IDC in insuring faithful recognition of substrate over non-substrate DNA. This task is particularly challenging for MutY compared to other glycosylases, in that it must distinguish 8-oxoG:A mispairs over highly abundant and structurally similar T:A base pairs. The nature of the IDC may be tailored for a particular organism to balance an efficient search with high-fidelity repair. Indeed, the presence of the Zinc Linchpin motif in higher eukaryotes suggests the IDC region plays additional roles that may be important for regulating and coordinating the repair activity in ways that are distinctly different or unnecessary in lower eukaryotes or prokaryotes.

The IDC of MUTYH serves as an important locus for interactions with several DNA repair proteins7 and this suggests that Zn^{2+} coordination may also influence these critical protein-protein interactions. For example, human MUTYH Val315 and Glu316 are adjacent to the Zn^{2+} ligand Cys318 and are essential for binding to Hus1, that is part of the Rad9-Rad1-Hus1 damage signaling clamp.[23] Notably, there are several MAP disease variants located within the IDC, such as Cys318Arg,[1] a mutation of one of the Zn^{2+} coordinating ligands, and Gln324His, located just prior to the MUTYH Zn^{2+} linchpin motif's critical C-X_2-C coordination ligands, Cys325 and Cys328.[25, 55–56] Replacement of Cys318 with an Arg residue would be expected to be even more detrimental to Zn^{2+} ion coordination than the Ser replacement studied herein. Gln324His MUTYH has been thought to be a harmless polymorphism and exhibits WT activity *in vitro*; however affinity for Hus1 and 8-oxoG:A repair activity in a cellular context is reduced.[55–56] Moreover, cells expressing Gln324His were found to have a compromised DNA damage response to oxidative stress.[25, 57] The importance of Zn^{2+} coordination on Hus1 binding to MUTYH, as well as the potential perturbations of the Zn^{2+} linchpin due to Hus1 interactions, have yet to be evaluated; however, it is reasonable to consider that alterations of the Zn^{2+} coordination of the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

MUTYH Zn^{2+} linchpin may play a role in the functional impact of MAP variants by altering 8-oxoG:A repair activity *and* coordination of down-stream repair and damage signaling functions.

The database sequence alignments illustrate an evolution of the Zn^{2+} linchpin motif in higher eukaryotes consistent with unique functions in these higher organisms. Our findings that the absence of Zn^{2+} within the IDC reduces MUTYH activity and the known role of MUTYH in preventing mutations associated with oxidative DNA damage suggests that alterations in the zinc linchpin motif under cellular conditions of high oxidative stress may be used as a regulatory mechanism. Indeed, the MUTYH Zn^{2+} linchpin may serve as a ROS sensing motif via reactive Cys thiolates. Previous works have shown that Zn^{2+}-bound Cys in Zn-Cys4 sites are reactive if they are not protected by hydrogen bonds.27, [5859] Oxidation of the Cys ligands in MUTYH may alter metal coordination, or mediate loss of Zn^{2+}, that results in local structural changes that impact MUTYH repair activity and interactions with downstream protein partners. It may seem counterintuitive for high ROS conditions to downregulate an enzyme that acts on ROS-damaged bases; however, when BER is overburdened due to high oxidative stress, reduced MUTYH activity may prevent the accumulation of unrepaired abasic sites that are susceptible to DNA strand breaks. Evidence of a reactive Cys has been observed in the human 8-oxoG glycosylase, OGG1, where the enzyme activity of the cancer associated variant Ser326Cys was found to be altered by ROS treatment.[60,61] Thus, a role of the MUTYH Zn^{2+} linchpin motif as a mediator of cellular damage response may be a key element of the overall role of MUTYH to prevent genomic instability and disease.

We also observed that mutations that alter Zn^{2+} coordination impact the retention of the Fe-S cluster cofactor,[19] suggesting an interplay of these two metal cofactors in MUTYH. In addition, the newly identified Cys ligand, Cys230, is located within the N-terminal domain and lays adjacent to the FeS cluster. Alterations in the Zn^{2+} linchpin of MUTYH may also impact the redox properties of the nearby Fe-S cluster, and therefore the ROS response of the two cofactors are likely intimately related. Eukaryotic polymerases have been shown to possess both Fe-S and Zn^{2+} cofactors, and mutations to Zn^{2+} site Cys ligands were shown to reduce the extent of Fe loading in the Fe-S cluster site.[14] Studies of polymerases also illustrated that the structural role of the Fe-S cluster site can be replaced with Zn^{2+}.[62–63] This suggests an intriguing possibility that the Zn^{2+} site in MUTYH may alternatively coordinate an Fe-S cluster, which could endow additional redox properties to MUTYH by the presence of two adjacent Fe-S cluster cofactors. The potential exciting implications of this idea have prompted ongoing work in our laboratory to explore this possibility.

CONCLUSION

Herein we describe the identification of Cys230 in human MUTYH as the fourth coordinating ligand of the MUTYH Zn^{2+} linchpin motif along with Cys318, Cys325 and Cys328 using large-scale database sequence alignments and computational modeling. Furthermore, metal analysis, along with functional data *in vitro* and in cells provide strong evidence to support synergistic coordination of Cys230 and Cys318 to Zn^{2+}, along with Cys325 and Cys328. In addition, evaluation of the 8-oxoG:A glycosylase activity and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

mismatch DNA affinity of Cys to Ser variants directly correlates with the extent of coordinated Zn^{2+}. MUTYH variants harboring reduced levels of Zn^{2+} within MUTYH also have a compromised ability to suppress DNA mutations demonstrating the requirement of the four Cys-coordinated Zn^{2+} site for MUTYH-mediated DNA repair. This work shows that the Zn^{2+} linchpin motif in MUTYH is required for high substrate affinity and proper placement of the 8-oxoG:A mispairs within the active site to support adenine excision catalysis. The requirement of the Zn^{2+} linchpin in MUTYH and mammalian orthologs for repair suggests a potential role in regulating MUTYH activity in cells and orchestrating downstream responses under conditions of oxidative stress. Moreover, the finely choreographed roles MUTYH plays in preserving the genome may be eroded by disease-associated mutations adjacent to the critical metal binding sites or excessive exposure to environmental toxins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We are grateful to GAANN Fellowship, Humanities Graduate Research Award, and the Schwall Dissertation Fellowship to N.N.N. We are grateful for financial support for this work from the NIH NCI to S.S.D. (Grant CA067985). C.K. was supported by an NIH NIEHS predoctoral fellowship (ES007059). We thank NSF for support of J.R.A. as an REU (CHE 1560479). The Lim laboratory has been supported by grants from Ministry of Science and Technology (MOST) in Taiwan and Academia Sinica. A special thanks to Austin Cole at the ICP-MS facility for help with metal analysis. We would also like to thank Silvia Hilt and John Voss for use of their CD spectrophotometer.

REFERENCES

1. Banda DM, Nuñez NN, Burnside MA, Bradshaw KM and David SS (2017) Free Radical Biology and Medicine 107, 202–215. [PubMed: 28087410]
2. David SS, O'Shea VL and Kundu S (2007) Nature 447 (7147), 941–950. [PubMed: 17581577]
3. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR and Cheadle JR (2002) Nat. Genet 30 (2), 227–232. [PubMed: 11818965]
4. Wallace SS, Murphy DL and Sweasy JB (2012) Cancer Letters 327 (1–2), 73–89. [PubMed: 22252118]
5. Mazzei F, Viel A and Bignami M (2013) Mutat. Res 743–744, 33–43.
6. Out AA, Tops CM, Nielsen M, Weiss MM, van Minderhout IJ, Fokkema IF, Buisine MP, Claes K, Colas C, Fodde R, Fostira F, Franken PF, Gaustadnes M, Heinimann K, Hodgson SV, Hogervorst FB, Holinski-Feder E, Lagerstedt-Robinson K, Olschwang S, van den Ouweland AM, Redeker EJ, Scott RJ, Vankeirsbilck B, Gronlund RV, Wijnen JT, Wikman FP, Aretz S, Sampson JR, Devilee P, den Dunnen JT and Hes FJ (2010) Human Mutation 31 (11), 1205–15. [PubMed: 20725929]
7. Manlove AH; Nuñez NN; David SS The GO Repair Pathway: OGG1 and MUTYH The Base Excision Repair Pathway; Wilson DM; World Scientific: Hackensack, New Jersey, 2017; 63–115.
8. Michaels ML, Pham L, Nghiem Y, Cruz C and Miller JH (1990) Nucleic Acids Res. 18 (13), 3841–5. [PubMed: 2197596]
9. Porello SL, Cannon MJ and David SS (1998) Biochemistry 37 (18), 6465–6475. [PubMed: 9572864]
10. Lukianova OA and David SS (2005) Curr. Opin. Chem. Biol 9 (2), 145–151. [PubMed: 15811798]
11. Fuss JO, Tsai CL, Ishida JP and Tainer JA (2015) Biochimica Et Biophysica acta 1853 (6), 1253–71. [PubMed: 25655665]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

12. Woods RD, O'Shea VL, Chu A, Cao S, Richards JL, Horvath MP and David SS (2016) Nucleic Acids Res 44 (2), 801–810. [PubMed: 26673696]
13. Bartels PL, Zhou A, Arnold AR, Nuñez NN, Crespilho FN, David SS and Barton JK (2017) Langmuir 33 (10), 2523–2530. [PubMed: 28219007]
14. Bartels PL, Stodola JL, Burgers PMJ and Barton JK (2017) J. Am. Chem. Soc 139 (50), 18339–18348. [PubMed: 29166001]
15. O'Brien E, Holt ME, Thompson MK, Salay LE, Ehlinger AC, Chazin WJ and Barton JK (2017) Science 355 (6327).
16. Tse ECM, Zwang TJ and Barton JK (2017) J. Am. Chem. Soc 139 (36), 12784–12792. [PubMed: 28817778]
17. Boon EM, Livingston AL, Chmiel NH, David SS and Barton JK (2003) Proc. Natl. Acad. Sci 100 (22), 12543–12547. [PubMed: 14559969]
18. Grodick MA, Muren NB and Barton JK (2015) Biochemistry 54 (4), 962–73. [PubMed: 25606780]
19. Engstrom LM, Brinkmeyer MK, Ha Y, Raetz AG, Hedman B, Hodgson KO, Solomon EI and David SS (2014) J. Am. Chem. Soc 136 (22), 7829–32. [PubMed: 24841533]
20. Guan Y, Manuel RC, Arvai AS, Parikh SS, Mol CD, Miller JH, Lloyd S and Tainer JA (1998) Nature Structural Biology 5 (12), 1058–64. [PubMed: 9846876]
21. Fromme JC, Banerjee A, Huang SJ and Verdine GL (2004) Nature 427 (6975), 652–6. [PubMed: 14961129]
22. Lee S and Verdine GL (2009) Proc. Natl. Acad. Sci 106 (44), 18497–502. [PubMed: 19841264]
23. Luncsford PJ, Chang D-Y, Shi G, Bernstein J, Madabushi A, Patterson DN, Lu AL and Toth EA (2010) J. Mol. Biol 403 (3), 351–370. [PubMed: 20816984]
24. Parker A, Gu Y, Mahoney W, Lee SH, Singh KK and Lu AL (2001) J. Biol. Chem 276 (8), 5547–55. [PubMed: 11092888]
25. Shi G, Chang DY, Cheng CC, Guan X, Venclovas C and Lu AL (2006) The Biochemical Journal 400 (1), 53–62. [PubMed: 16879101]
26. Lee Y. m. and Lim C (2008) J. Mol. Biol 379 (3), 545–553. [PubMed: 18462757]
27. Lee Y-M and Lim C (2011) J. Am. Chem. Soc 133, 8691–8703. [PubMed: 21574548]
28. Babu CS, Lee Y-M, Dudev T and Lim C (2014) J. Phys. Chem. A 118, 9244–9252. [PubMed: 25116831]
29. Finn RD, Clements J and Eddy SR (2011) Nucleic Acids Res 39 (Web Server issue), W29–W37. [PubMed: 21593126]
30. Pei J, Kim B-H and Grishin NV (2008) Nucleic Acids Res 36 (7), 2295–2300. [PubMed: 18287115]
31. Sali A and Blundell TL (1993) J. Mol. Biol 234, 779–815. [PubMed: 8254673]
32. Allen MP and Tildesley DJ, Computer Simulation of Liquids. Oxford University Press: NY, 1987.
33. Haberthur U and Caflisch A (2008) J. Comput. Chem 29, 701–715. [PubMed: 17918282]
34. MacKerell A, Banavali N and Foloppe N (2000) Biopolymers 56 (4), 257–265. [PubMed: 11754339]
35. MacKerell JAD, Bashford D, Bellott M, Dunbrack R, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WEI, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D and Karplus M (1998) J. Phys. Chem. B 102, 3586–3616. [PubMed: 24889800]
36. Babu CS and Lim C (2006) J. Phys. Chem. A 110, 691–699. [PubMed: 16405342]
37. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM and Karplus M (2009) Journal of computational chemistry 30 (10), 1545–614. [PubMed: 19444816]
38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW and Klein ML (1983) J. Chem. Phys 79, 926–935.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

39. Ryckaert JP, Ciccotti G and Berendsen HJC (1977) J. Comput. Phys 23, 327–341.
40. Sakharov D and Lim C (2005) J. Am. Chem. Soc 127 (13), 4921–4929. [PubMed: 15796557]
41. Sakharov D and Lim C (2009) J. Comp. Chem 30, 191–202. [PubMed: 18566982]
42. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Jr., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J and Fox DJ Gaussian 09, Revision A.02, Gaussian, Inc.: Wallingford CT, 2009.
43. Nunez NN, Majumdar C, Lay KT and David SS (2018) Methods in enzymology 599, 21–68. [PubMed: 29746241]
44. Pope MA and David SS (2005) DNA Repair 4 (1), 91–102. [PubMed: 15533841]
45. Porello SL, Leyes AE and David SS (1998) Biochemistry-Us 37 (42), 14756–14764.
46. Greenfield NJ (2007) Nat. Protocols 1 (6), 2876–2890.
47. Micsonai A, Wien F, Kernya L, Lee Y-H, Goto Y, Réfrégiers M and Kardos J (2015) Proceedings of the National Academy of Sciences 112 (24), E3095–E3103.
48. Majumdar C, Nunez NN, Raetz AG, Khuu C and David SS (2018) Methods in enzymology 599, 69–99. [PubMed: 29746250]
49. Parker AR and Eshleman JR (2003) Cellular and molecular life sciences : CMLS 60 (10), 2064–83. [PubMed: 14618256]
50. Chepanoske CL, Porello SL, Fujiwara T, Sugiyama H and David SS (1999) Nucleic Acids Res 27 (15), 3197–3204. [PubMed: 10454618]
51. Pope MA, Chmiel NH and David SS (2005) DNA Repair 4 (3), 315–325. [PubMed: 15661655]
52. Wolff E, Kim M, Hu K, Yang H and Miller JH (2004) Journal of Bacteriology 186 (9), 2900–2905. [PubMed: 15090533]
53. Kundu S, Brinkmeyer MK, Livingston AL and David SS (2009) DNA Repair (Amst) 8 (12), 1400–10. [PubMed: 19836313]
54. Garibyan L, Huang T, Kim M, Wolff E, Nguyen A, Nguyen T, Diep A, Hu K, Iverson A, Yang H and Miller JH (2003) DNA Repair (Amst) 2 (5), 593–608. [PubMed: 12713816]
55. Brinkmeyer MK and David SS (2015) DNA Repair (Amst) 34, 39–51. [PubMed: 26377631]
56. Raetz AG, Xie YL, Kundu S, Brinkmeyer MK, Chang C and David SS (2012) Carcinogenesis 33 (11), 2301–2309. [PubMed: 22926731]
57. Turco E, Ventura I, Minoprio A, Russo MT, Torreri P, Degan P, Molatore S, Ranzani GN, Bignami M and Mazzei F (2013) Nucleic Acids Res 41 (7), 4093–103. [PubMed: 23460202]
58. Lee Y-M, Duh Y, Wang S-T, Lai MMC, Yuan HS and Lim C (2016) J Am Chem Soc 138 (11), 3856–3862. [PubMed: 26928525]
59. Lee Y-M, Wang Y-T, Duh Y, Yuan HS and Lim C (2013) J Am Chem Soc 135 (38), 14028–14031. [PubMed: 24010488]
60. Morreall J, Limpose K, Sheppard C, Kow YW, Werner E and Doetsch PW (2015) DNA Repair (Amst) 26, 15–22. [PubMed: 25534136]
61. Hill JW and Evans MK (2006) Nucleic Acids Res 34 (5), 1620–32. [PubMed: 16549874]
62. Netz DJ, Stith CM, Stumpfig M, Kopf G, Vogel D, Genau HM, Stodola JL, Lill R, Burgers PM and Pierik AJ (2011) Nat Chem Biol 8 (1), 125–32. [PubMed: 22119860]
63. Klinge S, Núñez-Ramírez R, Llorca O and Pellegrini L (2009) The EMBO journal 28 (13), 1978–1987. [PubMed: 19494830]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

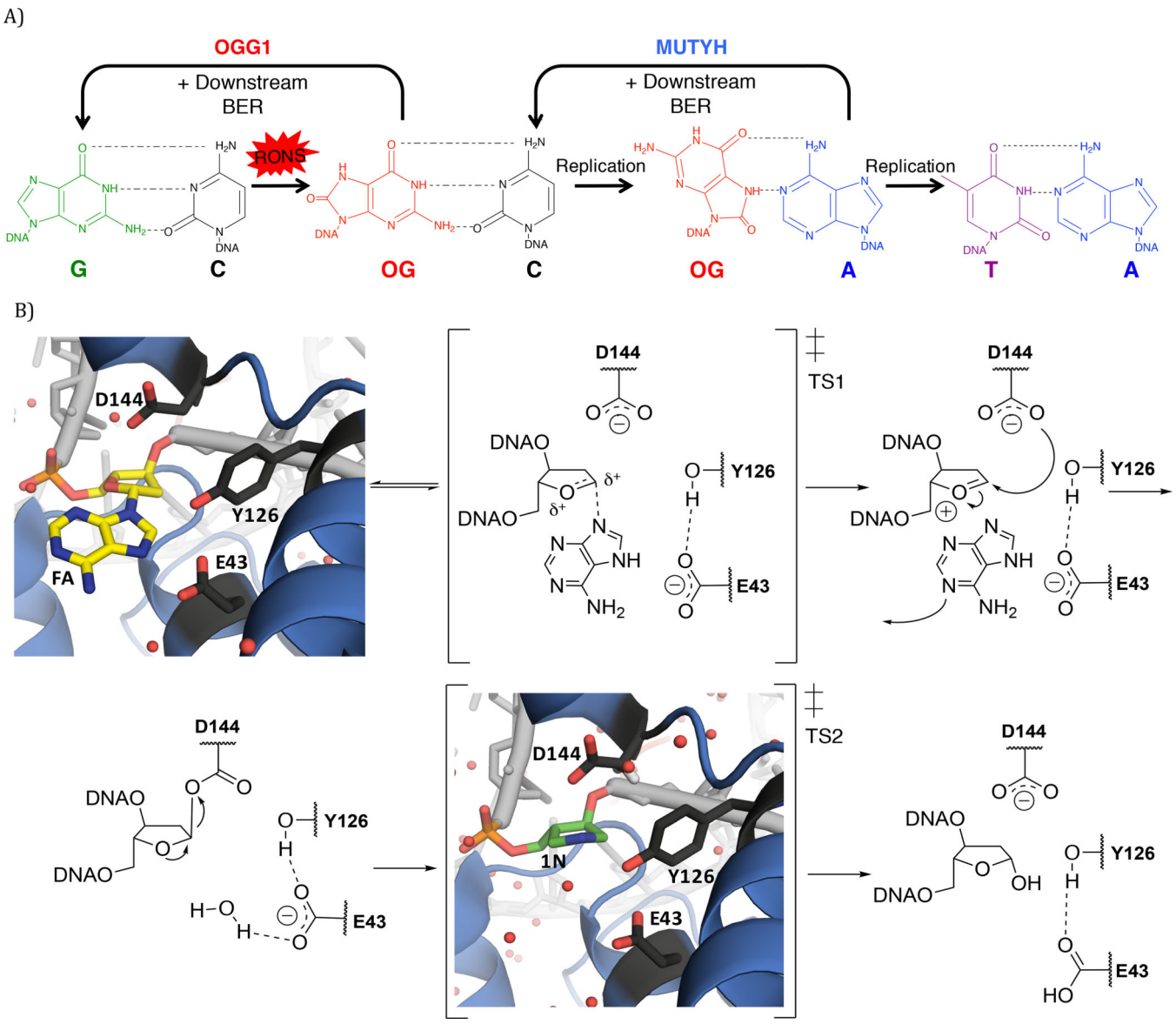


Figure 1.

Proposed mechanism for adenine removal from 8-oxoG:A mismatches by MUTYH and orthologs. A) Endogenous and exogenous sources of reactive oxygen and nitrogen species (RONS) can chemically react with the guanine (G, green) to form 8-oxoG (or OG, red). The presence of 8-oxoG in DNA can result mutations via 8-oxoG:A mismatches. The BER glycosylase OGG1 removes 8-oxoG from 8-oxoG:C pairs, while MUTYH removes the A from 8-oxoG:A mismatches. The G:C base pair is restored by subsequent repair by downstream BER enzymes. B) The mechanism of adenine removal from 8-oxoG:A mismatches by MUTYH is proposed to occur via two S_N1 - like displacement steps resulting in the formation of an abasic (AP) site product. The A nucleotide that is targeted is depicted in the mechanism with the rest of the DNA polymer projecting from the 5' or 3' side indicated as "ODNA". A crystal structure of *Geobacillus stearothermophilus* (Gs) MutY (blue) bound to DNA (light grey) containing a noncleavable substrate base analog arabino

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2’-fluoro-2’-deoxyadenosine, FA (yellow) (PDB ID: 3G0Q) demonstrates the positioning of key residues involved in catalysis. *Gs* MutY bound to DNA containing an azasugar transition state mimic, 1N (green) (PDB ID: 5DPK) also displays these participating amino acids, and position of the proposed water nucleophile.[12] Amino acid residues of MutY that are in close proximity to the FA/1N nucleotides are labeled and colored in dark grey. Water molecules are depicted as red spheres and elements are depicted with oxygen in red, nitrogen in blue, and phosphorous in orange.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

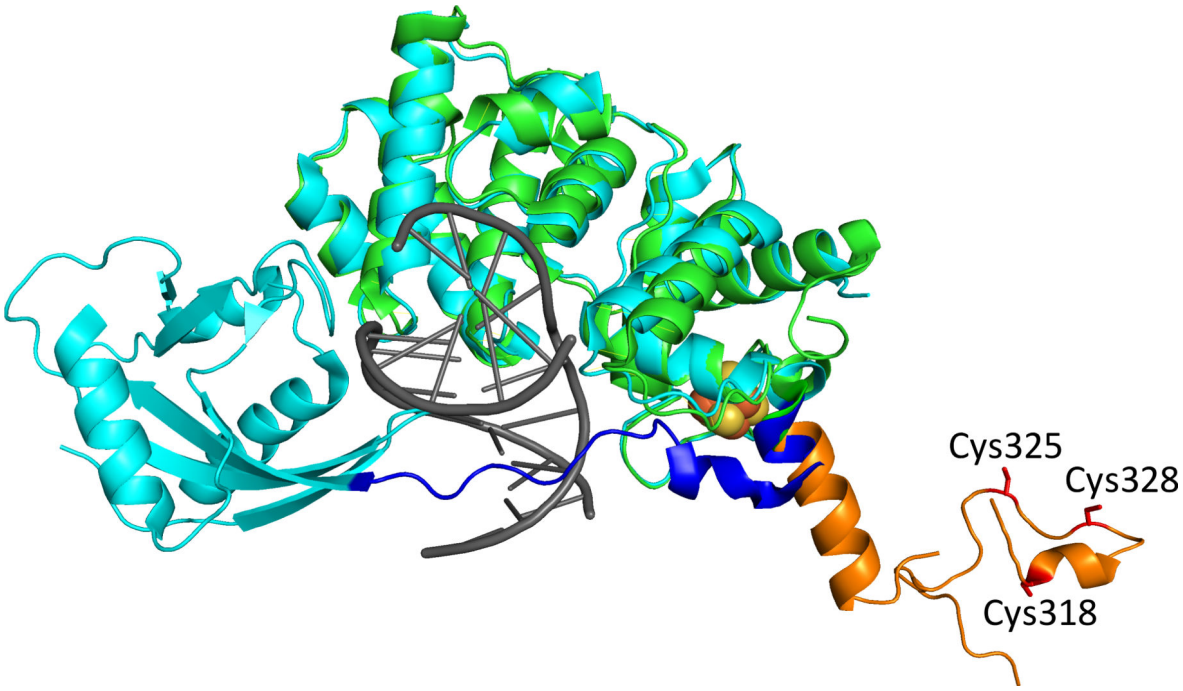


Figure 2.

N-terminal fragment crystal structure of *Homo sapiens* MUTYH (PDB ID: 3N5N, green)[23]. MUTYH residues 293 through 353 correspond to the IDC (orange) aligned to the ortholog structure of *Geobacillus stearothermophilus* MutY (PDB ID: 5DPK, light blue)[12] bound to DNA (grey), demonstrating the significantly shorter IDC (residues 215 through 234, dark blue) found in prokaryotes. Image also depicts established Zn^{2+} ion chelating ligands Cys318, Cys325 and Cys328 (red). The Fe-S cluster is depicted as orange and yellow spheres.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

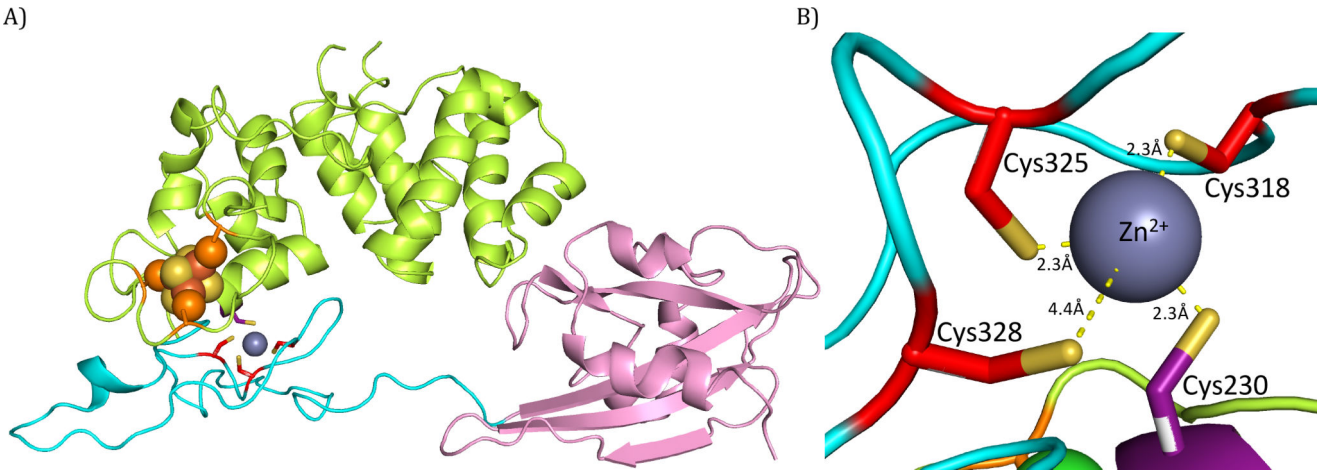


Figure 3.

QM/MM energy-minimized structure of human MUTYH highlighting the four Cys residues interactions with Zn^{2+}. A) A representative structure of human MUTYH from an MD simulation in explicit solvent after QM/MM minimization of the Zn^{2+}-binding site (see Methods). Color coding is as follows: C-terminal, pink; N-terminal, green; IDC, light blue; three previously identified Cys ligands (Cys318, Cys325 and Cys328 in human MUTYH; corresponds to Cys300, Cys307 and Cys310 in mouse Mutyh respectively), red; newly identified fourth Cys ligand (Cys230 in human MUTYH; corresponds to Cys215 in mouse Mutyh), purple; Cys residues coordinating the Fe-S cluster, orange, with the Fe-S cluster designated by orange (Fe) and yellow (S) spheres. B) Close up of MUTYH coordinating Zn^{2+} via four S atoms from Cys230, Cys318, Cys325 and Cys328 (waters removed for clarity). Notably, a single water molecule mediates the interaction between Cys328 and Zn^{2+} (Figure S5). Distances are drawn as yellow dash lines and elements with oxygen in red sticks; nitrogen, in blue; sulfur, in yellow. Cys residues 230, 318, 325 and 328 in human MUTYH correspond to 215, 300, 307 and 310 respectively in mouse Mutyh.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

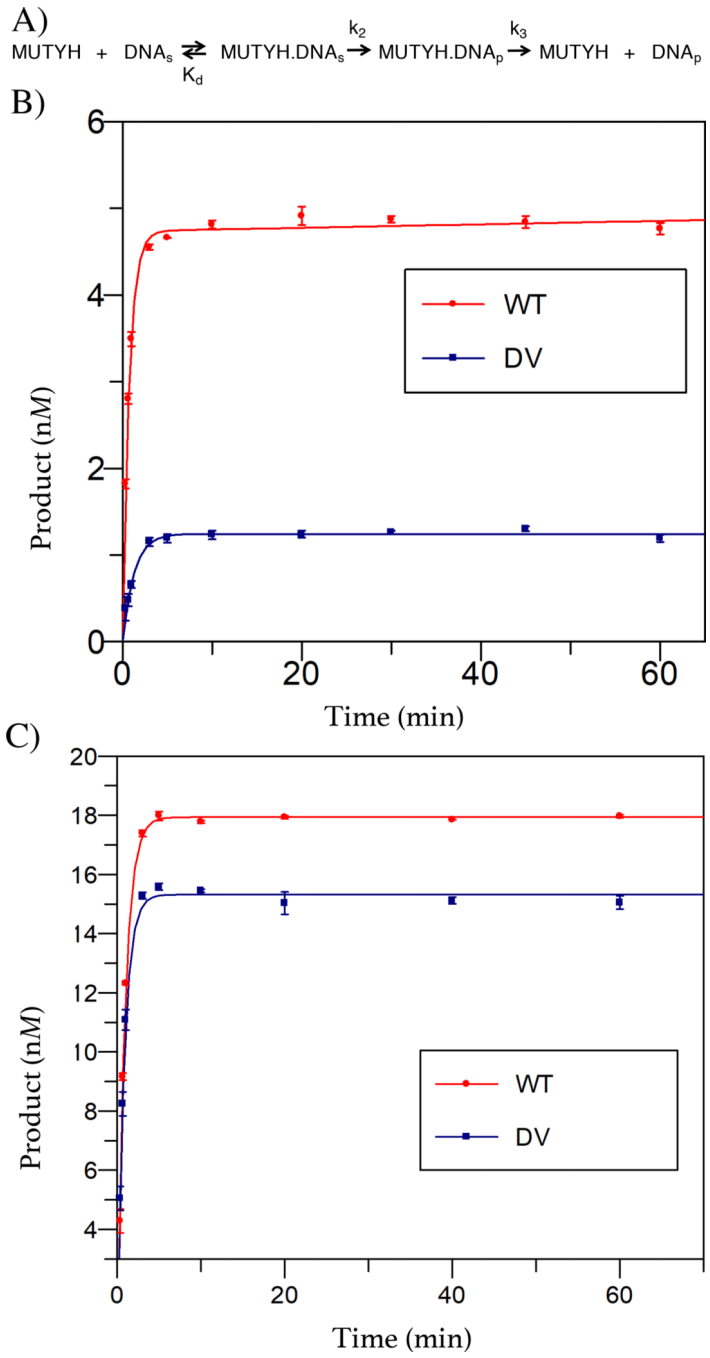


Figure 4.

Adenine glycosylase activity of wild type (WT) and double variant (DV) Cys215Ser/Cys300Ser mouse Mutyh. (A) Minimal kinetic scheme used to evaluate binding (K_d) and subsequent removal of adenine from an 8-oxoG:A base pair-containing duplex substrates. The rate constant k_2, reports on all steps involved in base excision to form an abasic site product. Release of the product from MUTYH is represented by the rate constant k_3. (B) MTO glycosylase assays reveal a reduced active fraction, as evidence by the reduced amplitude of the “burst” phase for the double variant (DV) compared to wild type (WT)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

mouse Mutyh. (C) Graph of WT and DV Mutyh activity under single-turnover conditions ([active enzyme]>>[DNA]) to determine the intrinsic rate of glycosidic bond cleavage, k_2. The data are shown with average and standard error bars from experiments performed with a minimum of three separate sample aliquots.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

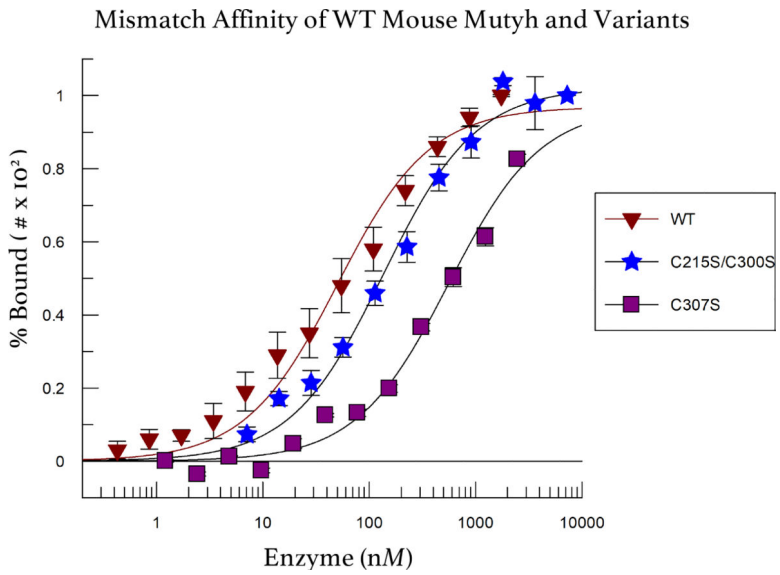


Figure 5.

Mismatch Affinity of WT and Zn^{2+} Linchpin Mouse Mutyh Variants. Fluorescence polarization assays of WT mouse Mutyh and Zn^{2+} linchpin Cys to Ser variants that have reduced Zn^{2+} loading (40%, double variant Cys215Ser/Cys300Ser; 10%, Cys307Ser) show reduced affinity for an 8-oxoG:FA duplex. The K_d values for the double variant Cys215Ser/Cys300Ser and Cys307Ser variant are 136 ± 14 and 560 ± 80 n*M*, respectively, compared to WT Mutyh, 50 ± 8 n*M*. Additional binding data are shown in Figure S7. Cys215, Cys300 and Cys307 in mouse Mutyh correspond to Cys230, Cys318 and Cys325 in human MUTYH.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

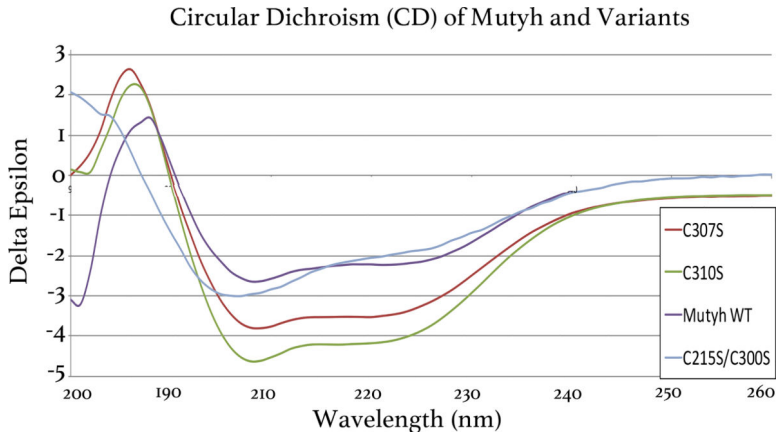


Figure 6.

Circular Dichroism (CD) Spectra of WT MouseMutyh and Cys ligand variants. CD signal is normalized to the protein concentration by conversion of millidegrees (θ) to delta epsilon ($\Delta\varepsilon$, $M^{-1}cm^{-1}$) and are plotted versus wavelength (nm). The CD demonstrates that the dramatic loss in percent activity for Zn^{2+} compromised variants (Cys215, Cys300, Cys307 and Cys310 in mouse Mutyh correspond to Cys230, Cys318, Cys325 and Cys328 in MUTYH, respectively) is not due to a complete loss of protein structure, rather Zn^{2+} coordination plays a local structural role (Figure S8).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:

Metal ion content, *in vitro* enzymatic activity and DNA binding affinity of WT Mutyh and Zn^{2+} ligand Cys to Ser variants.

Mouse Mutyh	Iron / Mutyh[c]	Zinc / Mutyh[c]	%Active fraction[d]	k_2 (min^{-1})[e]	K_d (n*M*)[f]
WT	4.0 ± 0.2	0.85 ± 0.01	42 ± 1	1.1 ± 0.1	50 ± 8
C215S[a]	3.2 ± 0.1	0.77 ± 0.01	43 ± 6	1.2 ± 0.1	63 ± 12
C300S[b]	4.5 ± 0.1	0.81 ± 0.02	44 ± 6	1.3 ± 0.1	74 ± 12
C215S/ C300S	3.2 ± 0.1	0.34 ± 0.01	29 ± 3	1.2 ± 0.1	136 ± 14

[a] Corresponds to Cys230 in human MUTYH.

[b] Corresponds to Cys318 in human MUTYH.

[c] ICP-MS of WT Mutyh and Zn^{2+} ligand Cys to Ser variants expressed as the mole ratio of metal/protein.

[d] The active fraction is determined from the burst amplitudes in adenine glycosylase assays performed under multiple turnover conditions.

[e] The rate constant (k_2) was determined from glycosylase assays performed under single turnover conditions.

[f] DNA affinity was evaluated via fluorescence polarization assays using the enzyme concentration based on $A280_{nm}$. All values in the table represent the average and standard error of at least three separate trials with distinct sample aliquots.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2:

Suppression of mutations in the rifampicin resistance assay of WT MUTYH and Zn^{2+} ligand Cys to Ser variants.[a]

MUTYH (Homo sapiens)	Mutation Frequency (f, X 10^{-8})
WT	0.27 (0.0 – 1.3)
C230S	0.54 (0.0 – 1.6)
C318S	0.64 (0.09 – 3.8)
C230S/C318S	2.1 (0.08 – 14)
C325S*	22 (15–36)
C328S*	23 (15–29)
pMal-c2x*	12 (9–23)

[a]The *rpoB* mutation frequency (f) per cell was calculated by dividing the median number of mutants by the average number of cells in a series of cultures. In parentheses is the 95% confidence limits based on the mean value.

*Previously published numbers for the mutation frequency of Cys325, Cys328 and the empty vector.[19]