Supporting Information for

“A Zinc Linchpin Motif in the MUTYH Glycosylase Interdomain Connector is Required for Efficient Repair of DNA Damage”

Lisa M. Engstrom†, Megan K. Brinkmeyer†, Yang Ha‡§, Alan G. Raetz†, Britt Hedman§, Keith O. Hodgson‡§, Edward I. Solomon†, Sheila S. David†*

†Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, California, 95616
‡Department of Chemistry, Stanford University, Stanford, California, 94305
§ Stanford Synchrotron Radiation Lightsource, SLAC, Stanford University, Menlo Park, California 94025

Index of Supporting Information
Materials and Methods ........................................................................................................S2
Zn K-edge XAS data ........................................................................................................S8
Additional Metal Ion analysis ..........................................................................................S11
Representative Mutyh activity .........................................................................................S12
Rifampicin resistance assay data .....................................................................................S13
References ......................................................................................................................S13
Materials and Methods:

Murine Mutyh overexpression and purification. Several constructs were made of the WT and MutyhΔ333-515 for metal analysis, adenine glycosylase assays, and XAS experiments. A pQE30 vector containing the Mutyh gene that expresses Mutyh with an N-terminal hexa-His tag was provided by Prof. J. Miller (UCLA). The Mutyh gene was removed from this vector using restriction enzyme digestion using BamHI, separated by agarose gel electrophoresis, and purified (QIAQuick). SalI and NotI restriction sites were added to the 5’- and 3’-ends, respectively, using PCR amplification and the following primers:

5’-CGCGTCGACTCAGCCAGGCCAAGCCTTCCAGC and
5’-CGTGGCGGCCGCTCACTGGGTAGTACTGTTGGG.

Subsequently, the Mutyh gene was double digested using SalI and NotI, and then cloned into the pET28a vector. Correct ligation of the Mutyh gene into the pET28a vector was confirmed by sequencing. The truncated protein (MutyhΔ333-515) was constructed by introduction of a stop codon at position 332 in the Mutyh gene using site-directed mutagenesis (Stratagene, QuikChange) and confirmed by DNA sequencing. All Mutyh variants were overexpressed in the pET28a vector.

Mutyh was overexpressed in BL21(DE3) cells. Protein expression was induced by the addition of 1 mM IPTG and incubated at 30 °C for 6 hours. The cells were harvested by centrifugation (8000 rpm, 10 minutes, 4 °C) and the pellet resuspended in 50 mL of ice cold Buffer I (20 mM sodium phosphate buffer pH 7.6, 10 % glycerol) supplemented with 1 mM PMSF, and stored at -80 °C. The cells were thawed and lysed by sonication (6 minutes, 30 second cycles) and centrifuged (8,000 rpm, 10 minutes). Imidazole and NaCl was added to the supernatant to bring the final concentration to 1 M NaCl and 20 mM imidazole. The supernatant was combined and batch bound to Ni-NTA slurry (Qiagen, 3 mL/50 mL supernatant) by gentle rocking for 1 hour at 4 °C. The protein-bound resin was poured into 3 empty PD-10 columns (GE Healthcare). After the lysate had completed its flow-through, each column was washed with 10 mL Ni Wash Buffer (20 mM sodium phosphate buffer pH 7.6, 1 M NaCl, 20 mM imidazole, 10% glycerol). Protein was eluted by addition of 2-5 mL of Ni Elute Buffer (20 mM sodium phosphate buffer pH 7.6, 300 mM NaCl, 500 mM imidazole, 10% glycerol). The elute was buffer exchanged using murine Mutyh concentration buffer (20 mM sodium phosphate pH 7.6,
200 mM NaCl, 1 mM EDTA, 10% glycerol) and transferred to a conical tube. To cleave the His-tag from the protein, 2 U thrombin/1 mg protein was added to the protein and incubated at 4 °C for 16 - 20 hours. Following incubation, the thrombin reaction was inhibited by the addition of 1 mM PMSF, incubated at 4 °C for two hours. Ni-NTA resin, pre-rinsed in Buffer I, was added to the digested protein and allowed to batch bind for 1 hour, 4 °C. The protein-bound resin was poured into an empty PD-10 column and the flow-through collected, followed by 2mL Ni Wash Buffer. The column was eluted with 3mL Ni Elute Buffer. The combined flow-through and wash was diluted 8-fold with Buffer IIA (20 mM sodium phosphate buffer pH 7.6, 1 mM EDTA) and syringe filtered with a 0.2 µm filter. The protein was further purified using an AKTA FPLC (GE Healthcare) and a heparin column (GE Healthcare). The protein was eluted using a NaCl gradient by increasing the amount of Buffer IIB (20 mM sodium phosphate pH 7.6, 1 M NaCl, 1 mM EDTA) from 0-100% over 20 column volumes. Eluent fractions were combined, buffer exchanged using murine Mutyh concentration buffer, and concentrated using an Amicon Ultrafiltration Concentrator with 10,000 MWCO (Millipore). The protein was diluted with glycerol (40% final concentration) and stored at -80 °C. Fractions were run on an SDS-PAGE to ascertain purity and found to be greater than 95% pure. The concentration of total protein was determined using an estimated ε\textsubscript{280} = 85,370 M\textsuperscript{-1}cm\textsuperscript{-1} (ExPASy), and the concentration of the iron sulfur cluster was estimated using ε\textsubscript{410} = 17,000 M\textsuperscript{-1}cm\textsuperscript{-1}.

The purification of Mutyh\Delta333-515 in the pQE30 vector was modified from a previously reported procedure.\textsuperscript{2} Cells were grown and centrifuged as previously described. The cells were thawed and lysed by sonication (6 minutes, 30 second cycles) and centrifuged (10,000 rpm, 5 minutes). The supernatant was removed and the pellet resuspended in 30 mL Buffer I supplemented with PMSF, repeating sonication and centrifugation. Purification via Ni-NTA resin was carried out as described above. Following elution from the Ni-NTA column, the protein was further purified using an AKTA FPLC (GE Healthcare) and a cation exchange column as described above. The concentration of total protein was determined using an estimated ε\textsubscript{280} = 57,910 M\textsuperscript{-1}cm\textsuperscript{-1}.

Trace Metal Analysis by ICP-MS. Clear polypropylene was used whenever possible as well as certified metal-free autoclaved pipette tips and centrifuge tubes. MilliQ water used in sample preparation was first passed through a Chelex column (BioRad) to remove any trace elements.
Samples were prepared by addition of 10 - 40 µL enzyme and diluting to a total volume of 2 - 3 mL with MilliQ water to afford a final concentration of approximately 100 ppb Fe. Method blanks were prepared by addition of 10 – 40 µL Mutyh Concentration Buffer and diluting to a total volume of 2 - 3 mL with water. Samples and method blanks were submitted in triplicate to the Interdisciplinary Center for Inductively-Coupled Plasma Mass Spectrometry at the University of California, Davis.

**DNA substrates.** DNA oligonucleotides containing normal phosphoramidites were synthesized by Integrated DNA Technologies. Oligonucleotides containing OG were obtained from the Cores Facility at the University of Utah Medical Center. All oligonucleotides were HPLC purified prior to use on a Beckman Gold Nouveau system using ion-exchange chromatography with a Waters AP1DEAE 8HR column. The following duplex was used in γ-³²P labeled kinetics assays: 5’-CTG TAA CGG GAG CT 3’ / 5’-CTA GTG GCT CCA TGA TCG-3’, where X = OG.

**Glycosylase assays.** Adenine glycosylase assays were performed as previously described. Two total protein concentration was determined by the Bradford method. Active site titrations (AST) and multiple-turnover (MTO) experiments were performed using 20 nM DNA and 2, 4, and 8 nM Mutyh in 20 mM Tris-HCl, pH 7.6, 1 mM EDTA, 0.1 mg/mL BSA, and 30 mM NaCl. Single-turnover (STO) experiments were carried out in order to separately determine the rate of N-glycosidic bond cleavage. In these experiments the enzyme was in excess over the concentration of DNA, such that [DNA] = 20 nM and [E] = 100 nM. In all experiments, reactions were incubated at 37 °C. Aliquots were removed at designated time points and quenched in a final concentration of 0.2 M NaOH followed by heating at 90 °C for 10 minutes. An equal volume of formamide loading dye was added and the aliquots were incubated an additional 10 minutes at 90 °C. The samples were resolved on a 15% (19:1) denaturing polyacrylamide gel in 1X TBE buffer at 1500 V for 2 hours and the resulting gel visualized using autoradiography by subsequent overnight exposure to a phosphor storage screen. The resulting scan was quantified using ImageQuaNT (v.5.2) data was fit using the appropriate rate equations in GraFit (v.5). A representative plot of product formation as a function of time for the reaction of C300S Mutyh using both MTO and STO conditions is shown in Figure S4.
Appropriate fitting of the data allowed for determination of the active site concentration and relevant rate constants as described previously for *E.coli* MutY.\(^3\) The following scheme was used to determine the kinetic parameters for the various Mutyh enzymes:

\[
\text{Mutyh} + (\text{DNA})_n \xrightarrow{k_1} \text{Mutyh\textbullet}(\text{DNA})_n \xrightarrow{k_2} \text{Mutyh\textbullet}(\text{DNA})_p \xrightarrow{k_3} \text{Mutyh} + (\text{DNA})_p
\]

All enzyme concentrations listed are the active enzyme concentration, unless otherwise specified.

**Rifampicin resistance assay.** The general procedure was similar to a previously described procedure from this laboratory.\(^4\) The *pMal-c2x* expression vector and the *pMalMUTYH WT, C318S, C325S, C328S, and C318S/C325S/C328S* plasmids were individually transformed into the *E. coli* strain GT100 *mutY*::mini-Tn10 *mutM*. A minimum of 16 independent overnight cultures were grown in LB media containing 100 µg/ml ampicillin and 15 µg/ml tetracycline from the transformed colonies. To determine the number of viable cells, a 10\(^{-7}\) fold dilution was plated on LB agar plates containing the same antibiotics, incubated at 37 ºC overnight, and the resulting colonies were counted. Cells from the same culture were plated onto LB agar plates containing the appropriate antibiotics and supplemented with 100 µg/ml rifampicin. The median number of colonies from the rifampicin-containing plates was divided by the average value for the viable cells to calculate the mutation frequency (\(f\)).

**XAS Data Collection.** XAS data were measured at the 20-pole wiggler Beam Line 7-3 at Stanford Synchrotron Radiation Lightsource (SSRL) with storage ring parameters of 3 GeV and 80–100 mA. A Si(220) double-crystal monochromator was used for energy selection, and a Rh-coated mirror upstream of the monochromator was used for harmonic rejection and collimation. Bacterial overexpression of Mutyh\(\Delta333-515\) afforded significantly higher amounts of protein compared to WT Mutyh and provided concentrations needed for XAS experiments. A Mutyh\(\Delta333-515\) solution (100 µL total volume, 800 µM [Zn]\(_T\), 40% glycerol, in sodium phosphate (20 mM) buffer, pH 7.6, containing 200 mM NaCl, 1 mM EDTA) was transferred into a Lucite XAS cell with 37 µm Kapton tape windows, flash frozen in an isopentane/liquid
nitrogen slush bath and stored in liquid nitrogen prior to data collection. The sample was maintained at 10 K during the data collection, using an Oxford Instruments CF1208 liquid helium continuous-flow cryostat. A Canberra 30-element solid-state Ge detector array was used to record Zn Kα fluorescence data. Through the use of Soller slits and a Cu filter placed between the sample cryostat and the detector window, signal intensity from inelastic scattering and Zn Kβ fluorescence was substantially diminished. Internal energy calibration was performed by simultaneous measurement of the absorption of a Zn foil placed between two ionization chambers filled with Ar located after the sample. The first inflection point of the foil XAS edge was assigned to 9660.70 eV. No signs of photoreduction of the metal site were observed during the scans. A total of 8 scans were measured for the Zn Mutyh (800 µM) sample.

XAS Data Analysis. After inspection of raw data and averaging, the average data file was normalized, using the program PYSPLINE, by fitting a second-order polynomial to the pre-edge region and subtracting from the entire data range with control points, followed by fitting a four-region spline function over the post-edge region. The data were normalized to an edge jump of 1.0 at 9675 eV. By means of the least-squares fitting program OPT, a component of the EXAFSPAK suite of software, EXAFS data over the range $k = 2$−$13 \text{ Å}^{-1}$ were fit using initial ab initio theoretical phase and amplitude functions calculated from FEFF 7.0 on the basis of the protein crystal structure 1JZQ (for Zn-4Cys model) and 2CO8 (for Zn-3Cys1His model) as starting models. Atomic coordinates from the crystal structure model were adjusted as necessary as fits were further refined. During fit optimization, the inter-atomic distance between the absorbing and backscattering atom (R) and the mean-square thermal and static deviation in R ($\sigma^2$) were varied for all components. The threshold energy ($\Delta E_0$) was allowed to vary for each fit but constrained to the same value for all components. The amplitude reduction factor ($S_0^2$) was maintained at a value of 1.0 throughout analysis.

XAS Results. XAS is an element specific technique that probes the local environment of a metal ion at an atomic-level resolution. For XAS data collected at the metal K-edge, the X-ray Absorption Near Edge Structure (XANES) region of the spectrum is sensitive to the effective charge on the metal ion and the coordination geometry, while the Extended X-ray Absorption Fine Structure (EXAFS) region provides geometric information on how many, and what types of
atoms are at what distance from the center atom. The normalized Zn K-edge XAS spectrum in Figure S1 showed that the site contains a Zn(II) center. The redox state was confirmed by both the edge energy as well as the fact that there is no pre-edge feature. The data were fit with all possible ligand combinations: ZnS$_4$, ZnS$_3$(O/N), ZnS$_2$(O/N)$_2$, ZnS (O/N)$_3$, ZnS(O/N)$_4$ (Note that N and O are not distinguishable in EXAFS, and N was used in the fitting). The ZnS$_4$ model gave the best fit to the experimental data (Figure S2, red and Table S1). The 4th ligand may be N/O, but in order to get a good fit, the addition of two N/O was required. The ZnS$_3$(O/N)$_2$ model (Figure S3) gives a slightly better fit to the data, especially in the low $k$ region (Figure S2, green and Table S1). The Zn-O/N bond lengths are longer (2.08 Å) than average Zn-O/N bond length (~2.0 Å). This reflects the effect of having three strong thiolate donors will weaken the bonding interactions of uncharged N/O ligands.
Figure S1. Normalized Zn K-edge XAS spectrum of Zn site in MutylΔ333-515
Figure S2. EXAFS and Fourier transform of MutyhΔ333-515 data (black) with fits for the Zn-4Cys model (red) and Zn-3Cys2N/O model (green).
Figure S3. Zn-3Cys2N/O Model used to fit the EXAFS.

Table S1

|                | Zn(Cys)₄ (Red) | Zn(Cys)₃(O/N)₂ (Green) |
|----------------|----------------|------------------------|
| Coordination number (N), interatomic distance (R, Å), mean-square thermal and static disorder in distance (σ², Å²), and EXAFS threshold energy adjustment from 9675 eV (∆E₀, eV) were varied in the fits. Estimated errors are ± 0.02 Å in R, ± 0.0001 Å² in σ², and ± 20% in N. The goodness of fit, F, is defined as $F = [\sum k^6(\chi_{\text{exper.}} - \chi_{\text{calc.}})^2 / \sum k^6(\chi_{\text{exper.}})^2]^{0.5}$ |
Table S2. Metal ion content of WT and mutant Mutyh proteins determined by ICP-MS.\(^a\)

| Metal | Mutyh  | Mutyh  | Mutyh  | C300S\(^c\) | C307S\(^c\) | C310S\(^c\) | D296A\(^c\) | E298A\(^c\) | E299A\(^c\) | MUTYH\(^d\) |
|-------|--------|--------|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|       |        | Δ333-515\(^b\) | Δ333-515\(^c\) |             |             |             |             |             |             |             |
| Fe    | 2.5 ± 0.3 | 2.8 ± 0.3 | 2.3 ± 0.2 | 2.6 ± 0.2 | 0.29 ± 0.01 | 0.84 ± 0.19 | 1.7 ± 0.4 | 1.7 ± 0.4 | 1.9 ± 0.2 | 0.11 ± 0.01 |
| Zn    | 0.6 ± 0.1 | 0.4 ± 0.1 | 0.7 ± 0.1 | 0.61 ± 0.02 | 0.081 ± 0.002 | 0.085 ± 0.016 | 0.60 ± 0.1 | 0.58 ± 0.1 | 0.65 ± 0.06 | 0.25 ± 0.04 |
| Cu\(^e\) | 0.0016 | 0.018 | 0.0072 | 0.018 | 0.0048 | 0.0043 | 0.0065 | 0.0068 | 0.18 |
| Mn\(^e\) | < DL\(^f\) | 0.0069 | 0.0004 | 0.0011 | -0.0005 | 0.0003 | 0.0001 | 0.0002 | 0.0069 |
| Ni\(^e\) | 0.0052 | 0.022 | 0.044 | 0.029 | 0.035 | -0.0012 | -0.005 | -0.005 | 0.022 |

\(^a\) Values are expressed as molar ratio of metal:protein. The ratio of metal ion to protein was determined by dividing the concentration of metal ions by the concentration of the total protein (determined by \(\varepsilon_{280} = 85,370\) M\(^{-1}\)cm\(^{-1}\) for WT Mutyh and \(\varepsilon_{280} = 57,910\) M\(^{-1}\)cm\(^{-1}\) for MutyhΔ333-515). \(^b\) Mutyh was overexpressed in \(E.\ coli\) using a pQE30 vector containing a non-cleavable N-terminal hexa-His affinity tag. \(^c\) Mutyh was overexpressed in \(E.\ coli\) using a pET28a vector with the N-terminal hexa-His affinity tag cleaved during purification. \(^d\) MUTYH (human) was overexpressed in baculovirus-infected insect cells. \(^e\) The standard deviation for the low amounts of Cu, Mn, and Ni is ~ 25% of the average value. \(^f\) The amount of metal ion present was below instrument detection levels.
Figure S4. Representative production curves of adenine glycosylase activity of C300S Mutyh under MTO conditions (grey squares) and STO conditions (blue circles). 20 nM DNA containing a central OG:A base pair was incubated with 2 nM Mutyh under MTO conditions and with 40 nM Mutyh under STO conditions. Reactions incubated in 20 mM Tris-HCl pH 7.6, 30 mM NaCl, 1 mM EDTA, 1 mg/mL BSA, at 37 °C. Fitting of data from 3-5 separate experiments provided rate constants listed in Table 2.
Table S3: Rifampicin resistance assay for WT MUTYH and the Cys to Ser modified MUTYH enzymes.\(^a\)

| Plasmid                  | Mutation frequency \((f, \times 10^6)b\) | \(f/f_{(WT)}\) |
|--------------------------|-----------------------------------------|-----------------|
| GT100 muty-mutm- cells   | 22 (17-41)                              | 12              |
| pMal-c2x                 | 12 (9-23)                                | 6               |
| pMal\textit{MUTYH} WT   | 1.9 (1.3-2.1)                           | 1               |
| pMal\textit{MUTYH} C318S| 2.6 (1.3-3.3)                            | 1.4             |
| pMal\textit{MUTYH} C325S| 22 (15-36)                              | 12              |
| pMal\textit{MUTYH} C328S| 23 (15-29)                              | 12              |
| pMal\textit{MUTYH} C318S\(\text{C325S C328S}\) | 35 (29-53)                              | 19              |

\(^a\)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. Note: C318, C325, and C328 human MUTYH correspond to C300, C307, and C310 mouse Mutyh, respectively. \(^b\)95% confidence limits based on the mean value are listed in parentheses.

(1) Yang, H.; Clendenin, W. M.; Wong, D.; Demple, B.; Slupska, M. M.; Chiang, J.-H.; Miller, J. H. \textit{Nucleic Acids Res.} \textbf{2001}, \textit{29}, 743.
(2) Pope, M. A.; David, S. S. \textit{DNA Repair} \textbf{2005}, \textit{4}, 91.
(3) Porello, S. L.; Leyes, A. E.; David, S. S. \textit{Biochemistry} \textbf{1998}, \textit{37}, 14756.
(4) Kundu, S.; Brinkmeyer, M. K.; Livingston, A. L.; David, S. S. \textit{DNA Repair (Amst)} \textbf{2009}, \textit{8}, 1400.
(5) Kundu, S.; Brinkmeyer, M. K.; Eigenheer, R. A.; David, S. S. \textit{DNA Repair (Amst)} \textbf{2010}. 

S13