# Noncysteinyl Coordination to the [4Fe-4S]$^{2+}$ Cluster of the DNA Repair Adenine Glycosylase MutY Introduced via Site-Directed Mutagenesis. Structural Characterization of an Unusual Histidinyl-Coordinated Cluster†,‡

Troy E. Messick,§ Nikolas H. Chmiel,|| Marie-Pierre Golinelli,||,⊥ Michael R. Langer,|| Leemor Joshua-Tor,*,§ and Sheila S. David*,||

*Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, and W. M. Keck Structural Biology Laboratory, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724*

*Received November 8, 2001; Revised Manuscript Received January 16, 2002*

ABSTRACT: The *Escherichia coli* DNA repair enzyme MutY plays an important role in the recognition and repair of 7,8-dihydro-8-oxo-2′-deoxyguanosine−2′-deoxyadenosine (OG·A) mismatches in DNA. MutY prevents DNA mutations caused by the misincorporation of A opposite OG by catalyzing the deglycosylation of the aberrant adenine. MutY is representative of a unique subfamily of DNA repair enzymes that also contain a [4Fe-4S]$^{2+}$ cluster, which has been implicated in substrate recognition. Previously, we have used site-directed mutagenesis to individually replace the cysteine ligands to the [4Fe-4S]$^{2+}$ cluster of *E. coli* MutY with serine, histidine, or alanine. These experiments suggested that histidine coordination to the iron−sulfur cluster may be accommodated in MutY at position 199. Purification and enzymatic analysis of C199H and C199S forms indicated that these forms behave nearly identical to the WT enzyme. Furthermore, introduction of the C199H mutation in a truncated form of MutY (C199HT) allowed for crystallization and structural characterization of the modified [4Fe-4S] cluster coordination. The C199HT structure showed that histidine coordinated to the iron cluster although comparison to the structure of the WT truncated enzyme indicated that the occupancy of iron at the modified position had been reduced to 60%. Electron paramagnetic resonance (EPR) spectroscopy on samples of C199HT indicates that a significant percentage (15−30%) of iron clusters were of the [3Fe-4S]$^{1+}$ form. Oxidation of the C199HT enzyme with ferricyanide increases the amount of the 3Fe cluster by approximately 2-fold. Detailed kinetic analysis on samples containing a mixture of [3Fe-4S]$^{1+}$ and [4Fe-4S]$^{2+}$ forms indicated that the reactivity of the [3Fe-4S]$^{1+}$ C199HT enzyme does not differ significantly from that of the WT truncated enzyme. The relative resistance of the [4Fe-4S]$^{2+}$ cluster toward oxidation, as well as the retention of activity of the [3Fe-4S]$^{1+}$ form, may be an important aspect of the role of MutY in repair of DNA damage resulting from oxidative stress.

DNA is subject to a variety of chemical modifications resulting from hydrolytic reactions, the action of alkylating agents, radiation, and oxidative stress. Such modifications can cause disruptions in DNA synthesis and produce permanent alterations in the genome (*1*). Fortunately, elaborate repair pathways exist in all organisms to protect them from the potential deleterious and mutagenic effects of DNA damage and mismatches (*2*). There are an increasing number of examples showing that such pathways may be involved in the prevention of diseases such as cancer (*2*, *3*). The base excision repair (BER)[1] pathway is primarily responsible for the repair of damage to heterocyclic DNA bases (*4*, *5*). In this pathway, the marquee players are the DNA glycosylases

† This work was supported by the National Institutes of Health (CA 67985), the University of Utah Research Foundation, and the Department of Chemistry of the University of Utah. S.S.D. is an A. P. Sloan Research Fellow (1998−2002). DNA sequencing was performed by the DNA sequencing facility at the University of Utah Medical School, which is supported in part by NCI Grant 5P30CA43014. The NSLS at BNL is funded by the U.S. Department of Energy Office of Basic Energy Sciences, Division of Material Sciences and Division of Chemical Sciences under Contract DE-AC02-98CH10886.

‡ This paper is dedicated to the memory of Barbara Burgess. Her beautiful work on Fe-S proteins has been an inspiration to us.

* To whom correspondence and reprint requests should be addressed. S.S.D.: tel, (801) 585-9718; fax, (801) 587-9657; e-mail, david@chemistry.chem.utah.edu. L.J.: tel, (516) 367-8821; fax, (516) 367-8873; e-mail, leemor@cshl.org.

§ Cold Spring Harbor Laboratory.

|| University of Utah.

⊥ Present address: Laboratoire d'Enzymologie et Biochimie Structurales, Centre National de la Recherche Scientifique, 1 Avenue de la Terrasse, 91190 Gif-sur-Yvette, France.

[1] Abbreviations: Abs, absorbance; AP, apurinic−apyrimidinic; bp, base pair; BER, base excision repair; BSA, bovine serum albumin; C199H, MutY with a histidine at position 199 in place of cysteine; C199HT, Stop 225 form containing the C199H mutation; C199S, MutY with a serine at position 199 in place of the cysteine; DTT, dithiothreitol; *E. coli*, *Escherichia coli*; EDTA, ethylenediaminetetraacetic acid; endo III, *E. coli* endonuclease III; F, 2′-deoxyformycin A; FCL, iron−sulfur cluster loop (residues Ile 191−Cys 199 in *E. coli* MutY); IPTG, isopropyl β-D-thiogalactoside; LB, Luria−Bertani broth; MTO, multiple turnover; nt, nucleotide; OG, 7,8-dihydro-8-oxo-2′-deoxyguanosine; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; PMSF, phenylmethanesulfonyl fluoride; ROS, reactive oxygen species; SDS, sodium dodecyl sulfate; STO, single turnover; Stop 225, truncated form of *E. coli* MutY containing residues Met 1−Lys 225; TBE, Tris−borate−EDTA buffer; Tris, tris(hydroxymethyl)aminomethane; WT, wild-type MutY.

10.1021/bi012035x CCC: $22.00 
Published on Web 03/03/2002

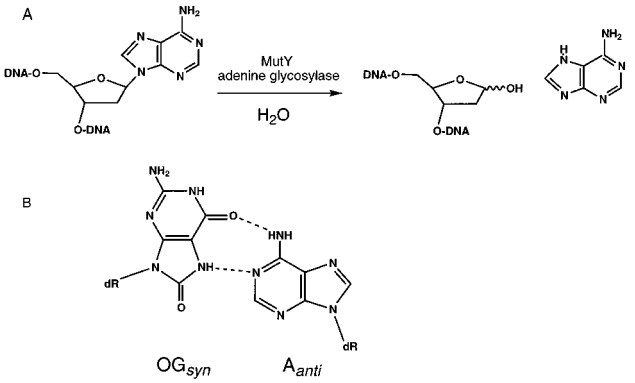


FIGURE 1: (A) Adenine glycosylase activity of MutY. The reaction of MutY with OG·A and G·A mismatches results in hydrolysis of the *N*-glycosyl bond to produce an abasic site-containing DNA product and a free adenine base. (B) Conformation of an OG·A mismatch in B-form DNA based on NMR and X-ray crystallography (*56*, *57*).

that recognize mismatched or damaged bases and catalyze N-glycosidic bond cleavage (Figure 1A). The subsequent action of apurinic−apyrimidinic (AP) endonucleases and phosphodiesterases provides the appropriate 3′-hydroxyl and 5′-phosphate ends for reparation of the DNA by DNA polymerase and DNA ligase (*5*).

The oxidation of 2′-deoxyguanosine (G) to 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG) is one of the most common types of damage to DNA in cells (*6*). When OG is present in DNA, it can mispair with A (Figure 1B) during a replication event and introduce a permanent G-to-T transversion mutation (*7*). In *Escherichia coli*, the "GO" repair pathway is dedicated to the prevention of mutations caused by oxidative damage to guanine (*8*). This pathway relies on the action of three proteins: MutT, MutM, and MutY. The MutT protein catalyzes the hydrolysis of d(OGTP) to remove it from the dNTP pool and prevent its incorporation in DNA (*9*). Both MutM and MutY are DNA glycosylases; MutM (also called Fpg) removes OG from DNA (*10*) while MutY removes A when mispaired with an OG (Figure 1B) (*11*). Although an OG·A mismatch (Figure 1B) is the preferred substrate (*12*), MutY is also able to remove A mispaired with G or C (*13*, *14*).

*E. coli* MutY exhibits particularly high sequence homology to *E. coli* endonuclease III (endo III) (*15*). Endo III is a glycosylase specific for the removal of oxidatively damaged pyrimidines in DNA (*4*). A unique feature of these two enzymes is that they contain an $[4Fe\text{-}4S]^{2+}$ cluster (*16*, *17*). The $[4Fe\text{-}4S]^{2+}$ cluster forms a cuboidal structure where the iron and sulfide ions are arranged in two overlapping tetrahedrons. The four iron atoms of the cluster are coordinated via a conserved and unique spacing of cysteine thiolates, Cys-$X_6$-Cys-$X_2$-Cys-$X_5$-Cys. The same pattern of cysteine residues is also found in some other BER glycosylases that include *Micrococcus luteus* UV endonuclease, a pyrimidine dimer glycosylase (*18*); *Methanobacterium thermoautotrophicum* MIG, a glycosylase specific for removal of uracil and thymine mispaired with G (*19*); and *Thermatoga maritima* MpgII, a glycosylase that targets alkylated bases (*20*). These enzymes presumably contain a similarly coordinated $[4Fe\text{-}4S]^{2+}$ as that found in MutY and endo III.

The presence of a $[4Fe\text{-}4S]^{2+}$ cluster in several DNA repair enzymes is unusual since iron−sulfur clusters are commonly found in proteins involved in electron-transfer reactions (*21*). However, there are many nonredox roles for Fe-S clusters that have been uncovered, suggesting that they are versatile cofactors (*22*, *23*). A relatively new role for iron−sulfur clusters that is of particular biological interest involves their participation as sensors of Fe levels or reactive oxygen species (ROS) in regulation of gene expression or enzymatic activity (*22*, *24*−*27*). For example, the Fe-S cluster in the transcription factor FNR (for fumarate nitrate reduction) appears to be involved in oxygen sensing (*24*). FNR controls the transcription of over 100 genes to facilitate adaptation of *E. coli* to growth in $O_2$ limiting conditions. The reaction of oxygen with the $[4Fe\text{-}4S]^{2+}$ cluster results in conversion of this cluster to a $[2Fe\text{-}2S]^{2+}$ cluster that affects the dimerization state, DNA binding affinity, and transcriptional activation capabilities of FNR (*24*).

Modification of the ligation to an Fe-S cluster via site-directed mutagenesis is often used to investigate the functional properties of an Fe-S cluster as well as determine the spectroscopic and electrochemical consequences of noncysteinyl ligation (*28*). However, attempts to express proteins containing [4Fe-4S] clusters with altered ligation are often hampered by the unstable nature of the mutated protein (*28*). Previously, we individually replaced the cysteine residues (Cys 192, Cys 199, Cys 202, and Cys 208) involved in the coordination of the $[4Fe\text{-}4S]^{2+}$ cluster of MutY with histidine, serine, or alanine (*29*). Only the C192H/S and C199H/S polypeptide chains were expressed at significant levels. In particular, the incorporation of histidine at these positions provided relative levels of overexpression of the mutated protein that rivaled that of the WT protein. On the basis of the higher level of expression due to the presence of a potential ligand (histidine and serine versus alanine), we proposed that the serine or histidine introduced in position 192 or 199 may serve as a ligand to the $[4Fe\text{-}4S]^{2+}$ cluster.

The ability of mutated MutY proteins to suppress DNA mutations relative to the WT enzyme provides a useful method for evaluating the consequences of the cysteine ligand replacements on the in vivo activity of MutY (*29*). Such analyses of mutation frequencies revealed that though the C199H and C192H polypeptides were overproduced at similar levels, only the C199H form had in vivo activity similar to the WT form. This result suggests that the structural properties of the Fe-S cluster coordination domain are important in the in vivo functioning of MutY. This is also consistent with the observation (*17*) that the amino acid loop containing Cys 192 and Cys 199, referred to as the iron−sulfur cluster loop (FCL), is solvent exposed and highly hydrophilic, containing a number of positively charged amino acids often used in DNA recognition (Arg 194, Lys 196, Lys 198). Indeed, recently, our laboratory showed that replacement of each of these positively charged amino acids within the FCL with alanine greatly affects substrate recognition (*30*). Moreover, K198A MutY was shown to have a significantly reduced ability to remove A from G·A mispairs. This suggests that the FCL is intimately involved in both damage recognition and base removal.

In the present work, we have overproduced and purified the C199H and C199S enzymes. The C199H mutation was also introduced in a truncated form of MutY (Stop 225)

which contains the first 225 residues of MutY yet retains catalytic activity. Henceforth, this doubly mutated form will be referred to as C199HT. The kinetic properties of the C199H, C199S, and C199HT enzymes were similar to those of the WT and Stop 225 enzymes, indicating that the mutations minimally disturb the catalytic properties of the enzymes. Structural characterization of C199HT and comparison to Stop 225 showed that the occupancy of the Fe coordinated by His 199 is 60% compared to the cysteine 199 coordinated Fe in Stop 225. Thus, samples of C199HT contain both $[4Fe\text{-}4S]^{2+}$ with a histidinyl coordination and $[3Fe\text{-}4S]^{1+/0}$ clusters. EPR analysis of C199HT and samples of C199HT oxidized with ferricyanide or reduced with dithionite confirmed the presence of $[3Fe\text{-}4S]^{1+}$ along with $[4Fe\text{-}4S]^{2+}$ clusters in C199HT. Kinetic analysis of C199HT samples containing significant amounts of the $[3Fe\text{-}4S]^{1+}$ form indicates behavior that is quite similar to that of Stop 225, suggesting that the presence of either a $[4Fe\text{-}4S]^{2+}$ with a coordinated histidine or a $[3Fe\text{-}4S]^{1+}$ cluster is sufficient to support efficient adenine glycosylase activity of MutY.

Table 1: Data Collection and Refinement Statistics

| | C199HT | Stop 225 |
|---|---|---|
| data reduction | | |
| space group | *C*2 | *C*2 |
| unit cell dimensions | 83.547, 49.899, 71.007, $\beta = 122.564$ | 82.877, 49.425, 70.278, $\beta = 122.785$ |
| resolution limits (highest shell) (Å) | 50.0−1.70 (1.76−1.70) | 40.0−1.74 (1.80−1.74) |
| wavelength (Å) | 1.100 | 1.100 |
| no. of measured reflections | 117089 (10899) | 132299 (12220) |
| no. of unique reflections | 27245 (2691) | 24736 (2445) |
| completeness (%) | 99.9 (99.4) | 100 (99.9) |
| $\langle I/\sigma(I)\rangle$ | 17.6 (2.6) | 29.3 (6.4) |
| $R_{sym}$ | 0.086 (0.505) | 0.058 (0.282) |
| refinement | | |
| reflections used | 50.0−1.70 (all data) | |
| *R*-factor (%) | 19.1 | |
| *R*-free (%) | 20.8 | |
| bond length rmsd (Å) | 0.006 | |
| bond angle rmsd (deg) | 1.1 | |
| Ramachandran plot core/allowed (%) | 91.6/8.4 | |

## MATERIALS AND METHODS

*General Methods, Bacterial Strains, Materials, and Instrumentation.* The *E. coli* laboratory strain JM101 was used in this work (*31*). In JM101*mutY*−, the chromosomal copy of the *mutY* gene was disrupted by transduction of *mutY*:: mini-tn*l0* into JM101. The plasmid containing the *mutY* gene, pKKYEco, was kindly provided by M. L. Michaels and J. H. Miller. All common DNA manipulations were performed using standard protocols (*32*). All 2′-deoxycyanoethyl phosphoramidites were purchased from ABI, except 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG) phosphoramidite, which was purchased from Glen Research. DNA oligonucleotides were synthesized by standard phosphoramidite chemistry on an Applied Biosystems 392 DNA synthesizer as per the manufacturer's protocol. Oligonucleotides for MutY assays were purified and handled as described previously (*12*, *33*). The oligonucleotides used for PCR reactions were purified using oligonucleotide purification cartridges (Perkin-Elmer). All buffers and other reagents used were purchased from Fisher, Sigma, or USB. 5′-End labeling was performed with T4 polynucleotide kinase (New England Biolabs) using [$\gamma$-$^{32}$P]ATP from Amersham Pharmacia Biotech, Inc. Labeled oligonucleotides were purified using ProbeQuant G-50 Micro Columns (Amersham Pharmacia Biotech, Inc.). Other enzymes were purchased from Roche. UV−visible spectroscopy was performed on a Hewlett-Packard 8452A diode array spectrophotometer. Storage phosphor autoradiography was performed using a Molecular Dynamics STORM 840 phosphorimager.

*Site-Directed Mutagenesis.* The cloning of the genes containing a cysteine mutation introduced in the full-length MutY was described previously (*29*). The mutated forms C199HT and C199ST were made by introduction of the stop codon at position 226 (*34*) into the C199H and C199S gene-containing plasmids as described previously for the Stop 225 mutated form (*29*).

*Purification of the Mutated Forms. E. coli* culture and protein purification were performed in a manner similar to that previously described (*34*) with the exception that the pellet resulting from the ammonium sulfate precipitation was dissolved in 30−50 mL of deoxygenated buffer A (20 mM sodium phosphate, pH 7.5, 1 mM EDTA, 1 mM DTT, 5% glycerol) containing 100 mM NaCl. MutY was then purified to homogeneity by column chromatography on a BioLogic chromatography system (Bio-Rad) using deoxygenated buffers. For the purification of C199S, the DNA and protein precipitations were performed in a glovebag under argon, and the protein was eluted from the chromatography columns into flasks flushed with argon. In addition, cation exchange was the only chromatographic purification that was employed. The C199HT and Stop 225 proteins used for crystallographic experiments were purified a second time on a heparin (Bio-Rad) column and eluted with a gradient of KCl in a buffer containing 20 mM HEPES/KOH, pH 7.5, 1 mM DTT, and 5% glycerol and concentrated via ultrafiltration to approximately 80 $\mu$M by UV−visible absorption at 410 nm ($\epsilon = 17000\ M^{-1}\ cm^{-1}$). Active site concentration determinations of preparations of MutY and relevant mutated forms were performed as described previously (*12*).

*Crystallization.* Crystals of C199HT and of Stop 225 were obtained using the hanging drop diffusion method from equal volumes of 60 $\mu$M protein solution mixed with a well solution containing 1.6−1.8 M ammonium sulfate and 100 mM Tris-HCl, pH 8.5. Crystals grew at 16 °C as thin plates to an optimum size of 0.3 mm × 0.2 mm × <0.01 mm after 3−5 days. Crystals were flash frozen after being soaked in a cryoprotectant solution containing the mother liquor and 20% glycerol for 3−5 min.

*Data Collection and Structure Determination.* Data for both C199HT and Stop 225 crystals were collected at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory (BNL) at beamline X26C, using a 30 cm MAR image plate and an incidental beam wavelength of 1.100 Å. Data were processed using the HKL package (*35*). Data collection statistics are listed in Table 1. The structure of C199HT was solved by isomorphous replacement (difference Fourier) using the coordinates from the Stop 225 structure (1MUY), removing the iron−sulfur cluster, and replacing C199 with glycine (*17*), since the crystals were of the same crystal form.

A difference Fourier $F_o - F_c$ electron density map clearly indicated altered density around the iron cluster and the histidine residue (data not shown). In addition, an $F_o$(Stop225) $- F_o$(C199HT) electron density map was calculated to minimize model bias (Figure 4). The iron cluster and histidine were added to the model, and refinement was carried out using CNS (version 1.0) (*36*) and SHELX(*37*). Refinement cycles were alternated with manual rebuilding with the program O (*38*). A total of 232 water and 3 glycerol molecules as well as a sulfate ion were added during later stages of the refinement. Anisotropic *B*-factors were refined for all atoms of the iron−sulfur cluster. In addition, the occupancy of the histidine-coordinated iron was refined and examined by careful inspection of difference electron density maps. The model was refined to an *R*-factor of 19.1% and an *R*-free of 20.8% at 1.7 Å resolution. Refinement statistics are shown in Table 1. The PDB coordinates were deposited in PDB (PDB code 1KQJ).

*DNA Duplexes Used for $K_d$ and Kinetic Experiments.* The following DNA duplexes were used: d(5′-CGATCATGG-AGCCAC(X)AGCTCCCGTTACAG-3′) and d(3′-GCTAG-TACCTCGGTG(Y)TCGAGGGCAATGTC-5′) with X = 2′-deoxyguanosine (G) or 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG) and Y = 2′-deoxyadenosine (A), 2′-deoxyformycin A (F), or 2′-deoxycytidine (C). The OG (or G) strand ($K_d$ experiments) or A strand (glycosylase assays) was 5′-$^{32}$P-end labeled with T4 polynucleotide kinase (New England Biolabs). In the case of glycosylase assays, 3−5% of the end-labeled A-containing strand was added to the unlabeled A-containing strand. For both experiments, the complementary strand was then added in 10− 20% excess, and strands were annealed in a buffer composed of 20 mM Tris-HCl (pH 7.6), 10 mM EDTA, and 150 mM NaCl.

*Gel Retardation Assay.* Gel retardation assays (*39*) were performed as described previously (*12*, *40*). Reactions containing <10 pM (estimated on the basis of the upper limit of isolation after labeling) appropriate duplex in a buffer containing 20 mM Tris-HCl, pH 7.6, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% glycerol, and 0.1 mg/mL BSA were incubated with various amounts of MutY at 25 °C. After incubation for 30 min, the reaction mixtures were loaded on a 6% nondenaturing polyacrylamide gel and electrophoresed in 0.5× TBE (Tris−borate−EDTA) for at least 2 h. The dried gel was exposed to a storage phosphor screen for 8−12 h. The storage phosphor autoradiogram was quantified using ImageQuant version 4.2a software (Molecular Dynamics). The percent bound DNA was plotted as a function of protein concentration, and an equilibrium dissociation constant ($K_d$) was determined by fitting the data with the equation for one-site ligand binding using the program GraFit 4.0 (Erithacus). The reported $K_d$ value represents an average of at least four separate experiments and has been corrected on the basis of the active site concentration determined for the relevant MutY or mutated MutY preparation used.

*Adenine Glycosylase Activity.* The adenine glycosylase activity was monitored under conditions of single and multiple turnover as described in detail previously (*12*, *34*). Briefly, the reaction solution contained 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, and 30 mM NaCl. The multiple-turnover experiments were performed with OG·A substrates to determine the active site concentration of MutY preparations and with both substrates to determine values for $k_3$. Single-turnover experiments were used to determine values for $k_2$. The multiple-turnover experiments were performed in buffer solutions containing 20 nM OG·A or G·A duplex DNA and the appropriate amount of active MutY necessary to provide a burst amplitude corresponding to 5−10% product formation. The single-turnover experiments were performed by incubation of 20 nM duplex DNA with 30 nM active protein.[2] For all kinetic experiments, the DNA solution was equilibrated at 37 °C, and the protein was added; subsequently, aliquots were removed at time points ranging from 15 s to 60 min, quenched by addition of 0.1 M NaOH, heated at 95 °C for 5 min, and then analyzed by denaturing PAGE (15% 19:1 acrylamide−bisacrylamide). The gels were exposed to a storage phosphor screen for at least 6 h. The resulting storage phosphor autoradiogram was quantified using ImageQuant V4.2a. The data were fit using GraFit 4.0 with the relevant equations (*12*, *34*) to extract values for the $k_2$ and $k_3$ rate constants. In the multiple-turnover kinetics with OG·A-containing DNA, the amplitude of the burst phase was used to calculate the active site concentration of each enzyme preparation. The percent active enzyme concentrations (based on the Bradford determination of total protein) for the enzyme preparations used in this work were as follows: C199H, 37%; C199S, 17%; C199HT, 41%, 76%, and 39%. These percent active site concentrations are consistent with the range of values routinely observed with WT and Stop 225 enzymes.

*Electron Paramagnetic Resonance Spectroscopy.* Low-temperature EPR spectra were recorded on a Bruker ESP-300E spectrometer equipped with an Oxford Instruments ESR-900E helium flow cryostat (Utah State University). Samples (36−80 $\mu$M) of C199HT and Stop 225 of approximately 200 $\mu$L volume in sodium phosphate buffer (20 mM, pH 7.5) containing 500 mM NaCl and 5% glycerol were transferred to EPR tubes and frozen in liquid nitrogen. C199HT oxidation was performed by addition of a 10-fold excess of freshly made potassium ferricyanide to the protein sample and incubation at room temperature for 30 min. Ferricyanide was removed by several cycles of concentration and dilution via ultrafiltration. Data were collected at a microwave frequency of 9.6 GHz and compared with spectra of a 100 $\mu$M Cu-EDTA standard run under nonsaturating conditions.

## RESULTS

*Expression and Purification of the Mutated Forms.* Previously (*29*), we have individually mutated the cysteines at positions 192, 199, 202, and 208, which are involved in the coordination of the $[4Fe\text{-}4S]^{2+}$ cluster of *E. coli* MutY (*17*). By radiolabeling the polypeptide chain with $^{35}$S during overexpression and analyzing the soluble cellular extracts by SDS−PAGE, we observed that polypeptides containing a mutation at position 202 or 208 were produced at very

[2] It is generally thought that single-turnover experiments should have enzyme in great excess (3−5-fold) over substrate. With MutY, we have found, under our assay conditions, that when the active enzyme concentration is greater than a 2-fold excess over DNA, the observed rate constants for adenine removal are diminished. This may be due to aggregation of MutY or substrate inhibition. However, the single-turnover experiments are completely valid when enzyme is only slightly in excess of substrate, and our rate constant analysis requires only the [E] > $K_d$. See Porello et al. (*12*) for more details.

Table 2: Rate Constants for Adenine Glycosylase Activity and Dissociation Constants ($K_d$s) for C199H, C199S, and C199HT Forms Compared to WT and Stop 225 MutY

| | OG·A[a] | | G·A[a] | | OG·F[a] |
|---|---|---|---|---|---|
| enzyme | $k_2$ ($min^{-1}$)[b] | $k_3$ ($min^{-1}$)[c] | $k_2$ ($min^{-1}$)[b] | $k_3$ ($min^{-1}$)[c] | $K_d$ (nM)[d] |
| WT[e] | >10 | 0.004 ± 0.002 | 1.6 ± 0.2 | 0.03 ± 0.01 | 0.28 ± 0.02 |
| C199H | >10 | 0.002 ± 0.001 | 1.9 ± 0.2 | 0.04 ± 0.01 | 0.08 ± 0.03 |
| C199S | >10 | 0.001 ± 0.001 | 1.8 ± 0.2 | 0.07 ± 0.02 | 0.23 ± 0.03 |
| C199HT | 0.43 ± 0.05 | 0.02 ± 0.01 | 0.26 ± 0.07 | 0.03 ± 0.01 | nd[f] |
| Stop 225[g] | 0.4 ± 0.1 | 0.010 ± 0.006 | 0.16 ± 0.02 | 0.03 ± 0.01 | 80 ± 30 |

[a] Designates central base pairs of the 30 bp duplex used in this work, where OG = 7,8-dihydro-8-oxo-2′-deoxyguanosine and F = 2′-deoxyformycin A. [b] The data are from single-turnover experiments using 30 nM MutY (active site concentration) and 20 nM DNA at 37 °C. These data represent averages from several different preparations of enzyme. The error reported is the standard deviation of the average. [c] These rate constants were determined from multiple-turnover experiments using 20 nM DNA and between 1 and 2 nM MutY (active site concentration) at 37 °C. These data represent averages from several different preparations of enzyme. The error reported is the standard deviation of the average. [d] $K_d$ values were determined using gel retardation methods with <10 pM DNA and variable amounts of MutY incubated at 25 °C. Each value is an average of at least four separate determinations. [e] WT data previously reported in refs *40* and *41*. [f] Not determined. [g] Stop 225 data previously reported in ref *34*.

low levels while the forms containing the corresponding mutations at position 192 or 199 were overexpressed to a greater extent. In particular, C192H and C199H forms were produced at 78% and 100% relative to the WT enzyme. When a serine was introduced at either position, the relative overexpression was reduced to approximately 10% of the WT enzyme; however, virtually no polypeptide chain could be detected with the analogous alanine substitution. On the basis of these results, we had proposed that histidine and serine may serve as ligands to the $[4Fe\text{-}4S]^{2+}$ cluster to stabilize the folded protein, thereby reducing degradation by cellular proteases. The effects of alterations of the cysteine ligands on the in vivo activity of MutY were assessed by determining the ability of modified forms to suppress DNA mutations compared to the WT enzyme (*29*). In this in vivo assay, only the C199H and C199S forms exhibited activity comparable to that of the WT enzyme. In contrast, the C192H form exhibited virtually *no activity* in the in vivo assay, even though the polypeptide chain was well expressed.

To more fully investigate functional and structural properties of these modified MutY enzymes, several of the proteins were overproduced in *E. coli*, and purification was attempted. To avoid oxidative degradation of the cluster, all buffers were deaerated, and column fractions were collected under argon. Even with these precautions to avoid oxidative degradation of the cluster, only the C199H and C199HT enzymes were amenable to homogeneous purification. The Fe-S cluster of the C199S form was found to be particularly sensitive to oxidative degradation, and therefore, a streamlined purification involving only the cation-exchange column was employed.

*Adenine Glycosylase Activity.* To investigate the effects of the ligand replacements on the ability of MutY to remove adenine from OG·A and G·A mismatches, kinetic experiments on the C199H and C199S forms were performed. Previously (*12*), we have shown that, under conditions where the DNA concentration was greater than the concentration of MutY, the reaction of WT and Stop 225 MutY with both OG·A and G·A substrates was characterized by biphasic kinetics, displaying an exponential burst of product formation followed by a linear steady-state phase. This type of kinetic behavior suggests that the product release step is rate-limiting. On the basis of the results with the WT enzyme, a minimal kinetic mechanism was proposed (*12*) as shown in Scheme 1.

Scheme 1

$$\text{MutY} + (\text{DNA})_S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{MutY}\cdot(\text{DNA})_S \xrightarrow{k_2} \text{MutY}\cdot(\text{DNA})_P \xrightarrow{k_3} (\text{DNA})_P + \text{MutY}$$

Using single-turnover experiments (STO) where [MutY] > [DNA], the step involving glycosidic bond cleavage (characterized by $k_2$) can be measured without complications associated with the product release step ($k_3$). With both WT (*12*) and Stop 225 MutY (*34*), the rate for adenine removal ($k_2$) with an OG·A-containing duplex is faster than with the analogous G·A-containing duplex. In contrast, MutY is able to turn over more rapidly with G·A as shown by the reverse trend for $k_3$ with G·A- versus OG·A-containing duplexes. Under conditions of multiple turnover ([DNA] > [MutY]), both the C199H and C199S enzymes exhibited kinetic behavior similar to that of the WT enzyme, indicating that a similar kinetic scheme could be applied to both enzymes. Using single-turnover kinetic experiments similar to those used with the WT enzyme, the kinetic rate constants $k_2$ for the C199H and C199S enzymes with a 30 base pair duplex containing either a central OG·A or G·A mismatch were determined (Table 2). In the case of the OG·A substrate, the rate of the reaction with both enzymes was too fast to be measured at 37 °C using our manual method; however, the lower limit was found to be the same as determined for the WT enzyme. In the case of the G·A substrate, the kinetic rate constants ($k_2$) for C199H and C199S were found to be similar to those for the WT enzyme. The turnover rates as determined by $k_3$ with both OG·A and G·A substrates for C199H and C199S were nearly identical to those determined for the WT enzyme. These results show that substitution of Cys 199 with histidine or serine minimally perturbs the adenine glycosylase activity of MutY.

*Dissociation Constant ($K_d$) Determination.* MutY specifically recognizes OG·A mismatches and catalyzes the hydrolysis of the N-glycosidic bond of the misincorporated 2′-deoxyadenosine (*11*). Previous work in our laboratory (*40*, *41*) has shown that duplexes containing a central OG·F (F = 2′-deoxyformycin A) bp bind with high affinity to WT MutY ($K_d$ = 280 ± 20 pM). The F nucleotide contains a C-glycosidic linkage that provides resistance to the glycosylase activity of MutY and therefore eliminates complications associated with enzymatic turnover when evaluating

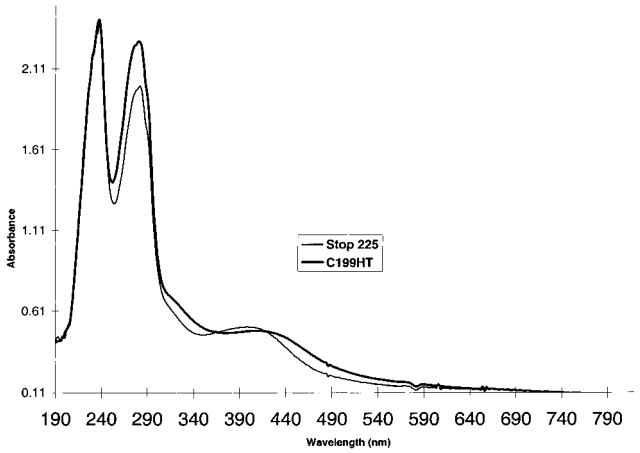


FIGURE 2: UV−visible absorption spectra of C199HT (thick line) and Stop 225 (thin line).

FIGURE 3: Overall superposition of wild-type Stop 225 in blue and C199HT in red. The two crystal structures overlap with an rmsd of the Cα atoms of 0.279 Å (for 225 atoms). The iron−sulfur cluster, green and yellow, respectively, is located in the C-terminal half of the structure. This figure as well as Figure 4 was prepared with Bobscript (*58*, *59*) and Raster3D (*60*, *61*).

substrate binding. To determine the binding characteristics of C199H and C199S enzymes, a similar nondenaturing gel retardation method was used to detect the presence of a specific MutY−DNA complex and determine the relevant equilibrium dissociation constants. Using this method, the $K_d$ values determined for the C199H and C199S forms with the OG·F-containing duplex were 80 ± 30 and 230 ± 30 pM, respectively (Table 2). These values are quite similar to that previously determined for the WT enzyme (280 ± 20 pM), revealing that both mutant forms exhibit substrate recognition properties similar to that of the WT enzyme. Interestingly, the C199H enzyme exhibits a slightly higher affinity for the OG·F duplex compared to the WT enzyme. This may be due to the higher positive charge on the cluster with a histidine substituting for cysteine and is consistent with the involvement of this region in the DNA recognition and repair properties of MutY (*29*, *30*).

*Preparation of C199HT.* To characterize the effect of the cysteine substitutions at position 199 on the structure of the enzyme, the C199S and C199H mutations were introduced in a truncated form of *E. coli* MutY, Stop 225 (*34*), which contains the first 225 amino acids that constitute the catalytic domain of MutY. The Stop 225 enzyme has been crystallized and the structure determined previously (*17*). Though we were unable to purify C199S Stop 225 (C199ST), C199H Stop 225 (C199HT) was purified to homogeneity. The kinetic parameters ($k_2$, $k_3$) for C199HT for both G·A and OG·A substrates are similar to those obtained with the truncated form Stop 225 (Table 2). Indeed, the C199H mutation seems to have similar effects when introduced in the full-size MutY or in the truncated form Stop 225. These results indicate that the C199HT is probably a valid representation of the N-terminal domain of C199H.

*UV−Visible Absorption Spectroscopy of C199H and C199HT.* WT and Stop 225 MutY exhibit a broad absorption feature in the UV−visible spectrum centered around 410 nm due to cysteinyl thiolate-to-Fe(III) charge transfer. The corresponding spectrum of the C199H or C199HT has a distinct red shift in this absorption feature of approximately 10−20 nm (Figure 2), which is consistent with a modification of the cluster environment. Due to the similarity of absorption characteristics between $[4Fe\text{-}4S]^{2+}$ and $[3Fe\text{-}4S]^{1+}$ clusters (*42*), the origin of the red shift is not likely to be attributed to the presence of a $[3Fe\text{-}4S]^{1+}$ cluster. The red shift is most likely due to the replacement of one of the coordinated cysteines with histidine, which would be expected to lower the energy of the cysteinate-to-Fe(III) charge-transfer transitions. The similarity in the UV−visible spectrum suggests that the C199H and C199HT enzymes contain a similarly coordinated $[4Fe\text{-}4S]^{2+}$ cluster.

Anaerobic addition of excess dithionite (10−100×) to samples of C199H at pH 7.6 and 8.8 did not result in any change in the UV−visible spectrum, suggesting that the $[4Fe\text{-}4S]^{2+}$ cluster of C199H is resistant to reduction, similar to the cluster in the WT enzyme.

*Crystal Structure of C199HT.* The overall structure of the C199HT mutant is very similar to that of Stop 225 (Figure 3). Briefly, the protein can be divided into two similarly sized lobes separated by a deep cleft: the N-terminal region (residues 20−130) contains a continuous six-helix bundle and the C-terminal region (residues 131−225) contains the $[4Fe\text{-}4S]^{2+}$ cluster. Superimposing the C199HT and WT Stop 225 structures on each other results in a root-mean-square deviation (rmsd) of Cα atoms of 0.279 Å, mostly due to a slightly increased (0.780°) bend between the two lobes in the C199H mutant. This could be due to slight packing differences between the two crystals.

An omit map lacking the iron−sulfur cluster and the side chain atoms of residue 199 shows clear density for the histidine coordinated to the cluster (not shown). The interatomic distance between the Nδ atom of histidine and its liganded iron is 2.0 Å. This value is in good agreement with the average distance of 2.13 Å between Nδ or Nϵ and iron in 2402 cases reported in the metalloprotein database (as of December 2000) (http://metallo.scripps.edu).

To visualize differences between C199HT and Stop 225 unbiased by the model, we collected data from Stop 225 crystals under conditions similar to those of C199HT and subtracted the observed structure factors (Stop 225 minus C199HT) to generate an $F_o - F_o$ electron density map (Figure 4). Positive electron density, which would correspond to density from Stop 225 that does not appear in C199HT, is shown in red; negative electron density, which corresponds

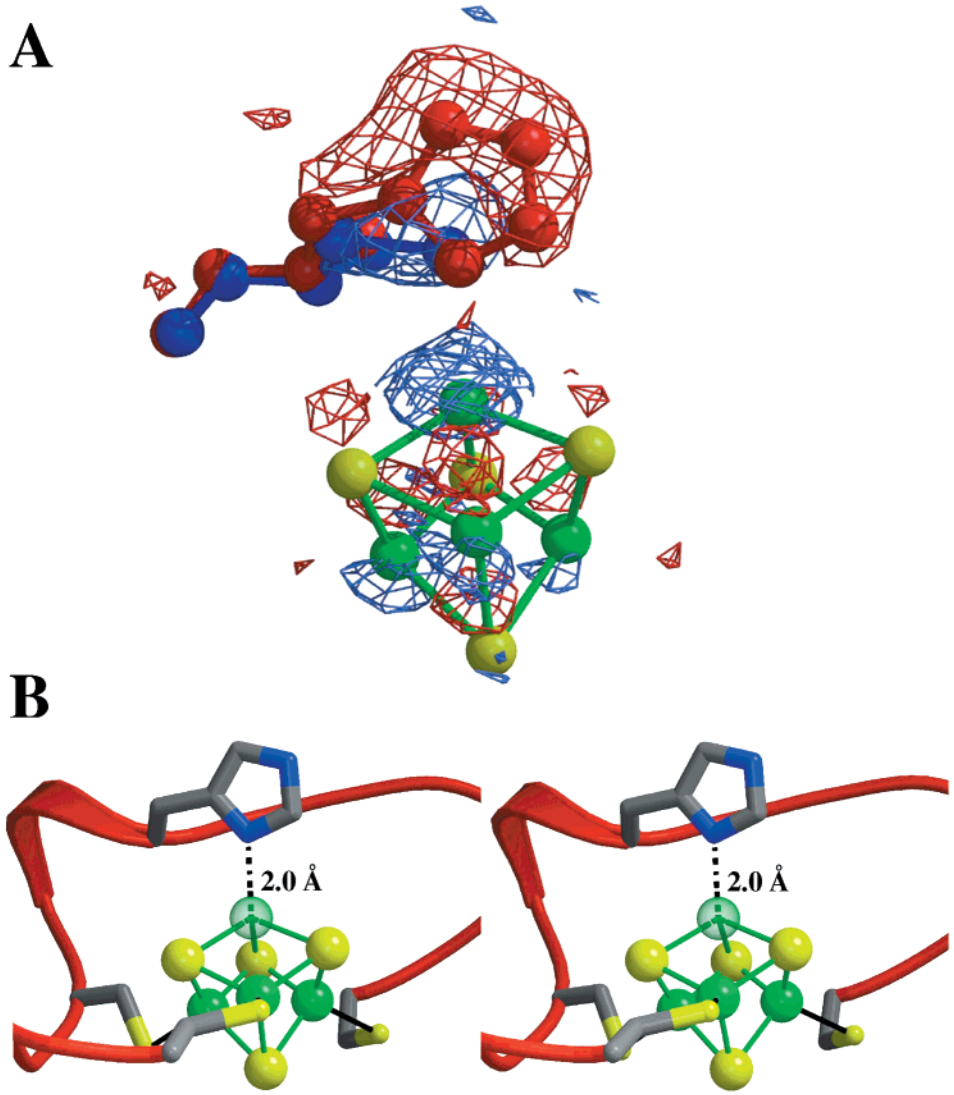


FIGURE 4: Iron−sulfur cluster region. (A) $F_o - F_o$ omit map of the region around the iron−sulfur cluster: in red, electron density seen in the C199HT data but not in the WT Stop 225; in blue, electron density in the WT Stop 225 data but not in C199HT. (B) Stereo ribbon diagram of residues 192−208 of both structures showing the intermolecular distance of 2.0 Å between the Nδ of H199 and Fe atoms. The occupancy of the Fe atom ligating to the histidine is 60% that of WT Stop 225.

to density from C199HT, but not in Stop 225, is shown in blue. One can clearly see the density around the imidazole ring of the histidine for C199H. The difference map shows that the occupancy of that Fe is reduced to around 60% in C199HT compared to Stop 225. Thus, the iron coordinated by the histidine in C199HT is more labile compared to the corresponding iron in Stop 225; no significant differences in the rest of the molecule can be seen. In particular, there is no significant difference in the Fe cluster loop (FCL) motif. The overall similarity in the structures is consistent with the similarity in the activity of C199HT and Stop 225.

*Electron Paramagnetic Resonance Spectroscopy.* The observation of reduced occupancy of the histidine-coordinated iron in the X-ray structure of C199HT indicated the presence of a three-iron cluster which may also be observed by EPR spectroscopy (i.e., $S = {}^1/_2$ [3Fe-4S]$^{1+}$). Notably, both [4Fe-4S]$^{2+}$ ($S = 0$) and [3Fe-4S]$^0$ ($S = 0$) clusters may also be present but would be EPR silent. Indeed, EPR spectra (Figure 5) of samples of C199HT at liquid helium temperatures (9.5 K) exhibited resonances near $g \sim 2$, characteristic of an $S = {}^1/_2$ [3Fe-4S]$^{1+}$ cluster (*42*). Spectra were taken of samples from three different preparations of C199HT, as well as samples which had been incubated with a 10-fold excess of ferricyanide. All of the samples had spectra that displayed features characteristic of [3Fe-4S]$^{1+}$ clusters. Upon integration relative to a Cu-EDTA standard, the concentration of the $S = {}^1/_2$ [3Fe-4S]$^{1+}$ cluster was determined to be 15−30% of the total sample (Table 3). This amount is significantly higher than what is routinely observed with Stop 225 where the observed signal corresponds to approximately 5% of the cluster concentration. Treatment of C199H or C199HT samples with excess dithionite resulted in complete loss of an EPR signal. This is consistent with this signal arising from a [3Fe-4S]$^{1+}$ cluster which should be readily reduced to the EPR-silent [3Fe-4S]$^0$ cluster. The fact that no signal is observed for a reduced [4Fe-4S]$^{1+}$ cluster under these conditions is consistent with the lack of change in the UV−visible spectrum under similar conditions, suggesting that the histidine-coordinated [4Fe-4S]$^{2+}$ cluster is resistant to reduction with dithionite. This is similar to what we have observed with the WT enzyme; however, this is somewhat surprising, since the introduction of the histidine ligand might have been expected to produce a cluster that was more amenable to reduction.

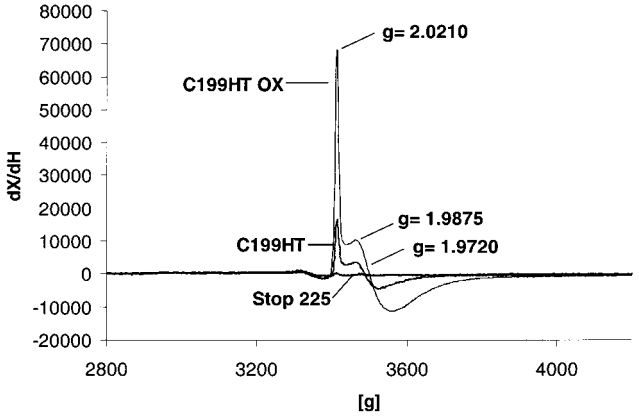


FIGURE 5: EPR spectra of 36 $\mu$M C199HT$_{OX}$ (thin black line), 47 $\mu$M C199HT (thick gray line), and 80 $\mu$M Stop 225 (thick black line). Samples (200 $\mu$L volume) were in sodium phosphate buffer (20 mM, pH 7.5) containing 500 mM NaCl and 5% glycerol. Spectra were taken at 9.5 K and 5 mW power with a modulation amplitude of 13 G and a sweep width of 1400 G centered at 3500 G. The total sweep time was 84 s with a time constant of 82 ms. C199HT and C199HT$_{OX}$ *g*-values marked on the spectrum are 2.020 at 3414 G and 1.9904 at 3467 G. Integration of the spectra yields approximately 68%, 32%, and 5% [3Fe4S]$^{1+}$ (relative to total cluster concentration) for C199HT$_{OX}$, C199HT, and Stop 225, respectively.

Table 3: EPR Quantitation of [3Fe-4S]$^{1+}$ Cluster in C199HT and Stop 225 Enzymes

| enzyme | [cluster][a] ($\mu$M) | [total protein][b] ($\mu$M) | [3Fe-4S][c] ($\mu$M) | % [3Fe-4S] based on total protein | % [3Fe-4S] based on UV−vis | % activity[d] |
|---|---|---|---|---|---|---|
| C199HT(1) | 50 | 65 | 10 | 16 | 21 | 41 |
| C199HT(2) | 13 | 15 | 2.8 | 18 | 21 | 76 |
| C199HT(3) | 47 | 75 | 15 | 20 | 32 | 39 |
| C199HT(3)$_{OX}$[e] | 36 | 58 | 25 | 43 | 68 | 42 |
| Stop 225 | 80 | 85 | 4 | 4.9 | 4.7 | 46 |

[a] This indicates the concentration of a particular sample labeled 1, 2, or 3 based on absorbance at 410 nM using $\epsilon$ = 17000 M$^{-1}$ cm$^{-1}$. This relates to the total "cluster" concentration of the sample. [b] This indicates the concentration of a sample designated 1, 2, or 3 based on the total protein concentration as determined using the Bradford assay with BSA as a standard. [c] The concentration of the [3Fe-4S]$^{1+}$ cluster based on integration of the EPR signal relative to a Cu(EDTA) standard. [d] We routinely express percent activity in relation to the Bradford concentration. [e] OX = designates sample oxidized with a 10-fold excess of ferricyanide.

Integration of the EPR signal for the oxidized protein revealed an increase in the amount of [3Fe-4S]$^{1+}$ cluster in the sample (approximately 2-fold). This is consistent with other examples of modified cluster coordination in which ferricyanide oxidation of a [4Fe-4S]$^{2+}$ protein has resulted in formation of the [3Fe-4S]$^{1+}$ form (*43*, *44*). Thus, this suggests that loss of the histidine-coordinated iron may be mediated, at least in part, by oxidative reactions.

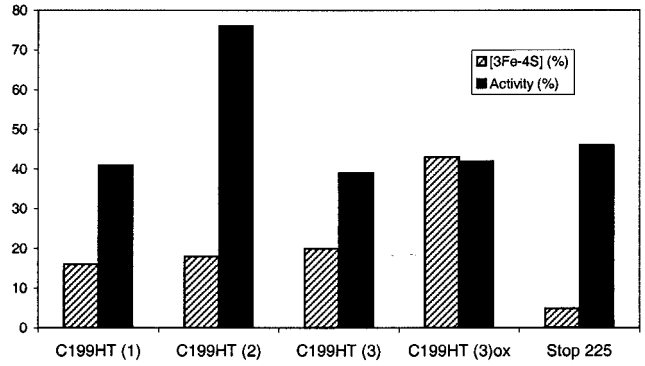


FIGURE 6: Bar graph illustrating the amount of [3Fe-4S]$^{1+}$ cluster present in various samples relative to the percent active site concentration determined from the adenine glycosylase assays with an OG·A duplex. Note that there does not appear to be a direct correlation.

*Kinetics of C199HT Samples Containing [3Fe-4Fe]$^{1+}$ Clusters.* To determine what effect the [3Fe-4Fe]$^{1+}$ cluster has on enzymatic activity, a kinetic analysis was performed similar to that with full-length MutY and C199H/S mutants. Several protein preparations of C199HT were used, and the kinetic data and activity were compared to the values reported for Stop 225 (*34*). One C199HT sample was incubated with excess ferricyanide to determine if this would alter the activity of MutY. The results of the adenine glycosylase assays are shown in Table 4. EPR measurements were taken of similar preparations and integrated to find the percent of [3Fe-4S]$^{1+}$ cluster present in the sample, and these data are shown in Table 3. These data taken together indicate that the amount of [3Fe-4S]$^{1+}$ in the sample does not correlate with the enzymatic activity. Indeed, the active protein concentration determined from an active site titration using an OG·A substrate does not correlate with the amount of 3Fe cluster (Figure 6). The different C199HT preparations (Tables 3 and 4, designated 1, 2, and 3) have similar percentages of [3Fe-4S]$^{1+}$ cluster, yet have very different percent active protein concentrations. Moreover, the active site concentrations of the C199HT samples are in the same range as Stop 225, which has only a small amount of [3Fe-

Table 4: Comparison of Adenine Glycosylase Activity of C199HT and Stop 225

| enzyme | OG·A mispair[a] $k_2$ (min$^{-1}$)[b] | OG·A mispair[a] $k_3$ (min$^{-1}$)[c] | G·A mispair[a] $k_2$ (min$^{-1}$)[b] | G·A mispair[a] $k_3$ (min$^{-1}$)[c] | % active[d] |
|---|---|---|---|---|---|
| C199HT(1) | 0.48 ± 0.08 | 0.012 ± 0.006 | 0.25 ± 0.04 | 0.03 ± 0.01 | 41 |
| C199HT(2) | 0.52 ± 0.02 | 0.018 ± 0.007 | 0.34 ± 0.06 | 0.05 ± 0.03 | 76 |
| C199HT(3) | 0.38 ± 0.02 | 0.015 ± 0.005 | 0.20 ± 0.04 | 0.020 ± 0.004 | 39 |
| C199HT(3)$_{OX}$[e] | 0.38 ± 0.03 | 0.019 ± 0.009 | 0.24 ± 0.04 | 0.016 ± 0.006 | 42 |
| Stop 225 | 0.29 ± 0.08 | 0.015 ± 0.004 | 0.17 ± 0.03 | 0.026 ± 0.007 | 46 |

[a] Designates central base pair of the 30 bp duplex used in this work, where OG = 7,8-dihydro-8-oxo-2′-deoxyguanosine. [b] The data are from single-turnover experiments using 30 nM C199HT or Stop 225 (active site concentration) and 20 nM DNA at 37 °C. These data were obtained from three to four kinetic experiments on a specific preparation of enzyme designated 1, 2, or 3. The error reported is the standard deviation of the average. [c] These rate constants were determined from multiple-turnover experiments using 20 nM DNA and between 1 and 2 nM C199HT or Stop 225 (active site concentration) at 37 °C. These data were obtained from three to four kinetic experiments on a specific preparation designated 1, 2, or 3. The error reported is the standard deviation of the average. [d] Active site concentrations are determined from multiple-turnover experiments with an OG·A substrate as described previously (*12*, *34*). The percent active concentration is determined relative to the Bradford determined concentration for total protein concentration. [e] OX indicates samples oxidized with a 10-fold excess of ferricyanide.

4S]$^{1+}$cluster. Oxidation with ferricyanide increased the amount of [3Fe-4S]$^{1+}$ but did not affect the amount of active protein in the sample. These results suggest that the [3Fe-4S]$^{1+}$ cluster is active and that the observation of less than 100% active site concentration samples of MutY may be more reasonably attributed to the presence of improperly folded protein, or apoprotein, than the presence of a [3Fe-4S]$^{1+}$ cluster. Detailed kinetic analyses of the various C199HT samples indicate that both of the rates including the chemistry ($k_2$) and product release ($k_3$) are similar to Stop 225 for all of the C199HT protein preparations tested. However, it is possible that in these experiments the rate constants, which have been determined at relatively high enzyme concentrations, may be dominated by the [4Fe-4S]$^{2+}$ cluster-containing MutY in the sample. However, the fact that the kinetic rate constants were unaffected even when greater than 50% of the sample contains a [3Fe-4S]$^{1+}$ cluster, as in C199HT$_{ox}$ sample, suggests that the [3Fe-4S]$^{1+}$ cluster-containing forms are not inactive. Indeed, the similar overall enzymatic activity suggests that the [3Fe-4S]$^{1+}$ cluster-containing MutY may participate in the glycosylase activity as efficiently as the WT protein.

## DISCUSSION

While rare, noncysteinyl coordination to iron−sulfur clusters in proteins has been observed in nature and is likely important in the catalytic activity of the protein. Examples of naturally occurring noncysteinyl coordination in [4Fe-4S] clusters include *Desulfovibrio gigas* Ni−Fe hydrogenase (3 Cys, 1 His) (*45*), the Fe-only hydrogenase from *Clostridium pasteurianum* (3 Cys, 1 His) (*46*), and a ferredoxin from the hyperthermophile *Pyrococcus furiosus* ferredoxin (3 Cys, 1 Asp) (*47*). The presence of histidine may alter the midpoint potentials of proteins involved in redox chemistry. For example, Rieske proteins containing [2Fe-2S] clusters with histidine ligands have higher midpoint potentials than [2Fe-2S] clusters with exclusively cysteinyl ligation (*48*). In contrast, the histidine-coordinated [4Fe-4S]$^{2+/1+}$ observed in the Fe-only hydrogenase from *C. pasteurianum* does not exhibit a dramatically altered redox potential (*49*). This has led to the suggestion, based on its location at the protein surface, that this unusual cluster may be important for mediating interactions with its relevant redox partners (*46*). Introduction of histidine ligands to a [4Fe-4S] cluster via site-directed mutagenesis has not been previously observed. Thus, this work represents the first structural and biochemical characterization of a mutated [4Fe-4S] protein containing histidinyl ligation.

Previous work in our laboratory suggested that both Cys 192 and Cys 199 in MutY may be individually replaced with histidine and serine to produce a [4Fe-4S] cluster-containing protein. While mutations at position 192 yielded protein that was not stable enough to survive purification, mutant enzymes C199S, C199H, and C199HT were amenable to overexpression and purification. Kinetic analysis of these mutants reveals that catalytic activity is not significantly modified by the mutations at position 199, yielding rate constants for adenine removal ($k_2$, $k_3$) similar to that of WT (C199S, C199H) or Stop 225 (C199HT). In the case of C199H, a slightly increased affinity for a substrate analogue was observed, possibly due to the increased positive charge around the cluster and FCL relative to the cysteine-containing counterparts. This is consistent with participation of this region in DNA binding. Additionally, the crystal structure of C199HT was solved to 1.7 Å resolution. The structure reveals a global fold similar to that of Stop 225. C199 in the Stop 225 structure is the most exposed to the outer solvent of all of the iron cluster ligating cysteines. The distance between C199 and four well-ordered water molecules averages about 4.8 Å. The ability to accommodate histidine ligation appears to be achieved by shifting the waters away from the cluster. In contrast, cysteine 192 is surrounded by other residues that may help to orient the cysteine toward the cluster. In particular, one of the nitrogens (NH2) of the guanidinium group of arginine 147 is within hydrogen-bonding distance (3.2 Å) to the S$\gamma$ of cysteine 192. Thus, the tight molecular environment and introduction of electrostatic repulsion around position 192 might destabilize the histidine coordination at this position, thus yielding an unstable protein that is not readily purified.

Histidine 199 is ligated to the iron−sulfur cluster with an interatomic distance of 2.0 Å between the Fe and N$\delta$. This distance is comparable to the two other histidine-ligated iron−sulfur cluster structures from the *D. gigas* Ni−Fe (*45*) and the *C. pasteurianum* Fe-only hydrogenases (*46*). Previously, we have shown that the Fe cluster loop (FCL) (residues 191−199) of MutY is important for recognition of its damaged DNA substrate. C192 and C199 can be thought of as anchors, tethering the two ends of this highly hydrophilic loop. The results in this report show that histidine at position 199 can also anchor the FCL, without diminishing its catalytic proficiency. The ability to accommodate a cysteine or a histidine at position 199 in MutY may be a reflection of the unique properties of the iron−sulfur cluster in this protein. The importance of the protein matrix in stabilizing alternative cluster ligation is also supported by the observation that small peptide ferredoxin maquettes did not appear to support ligation by histidine or aspartate (*50*). The flexible nature of the FCL in MutY which permits incorporation of a histidine anchor may also be an important aspect of the FCL's involvement in the recognition of DNA damage. In the reported structures of DNA repair enzymes bound to their DNA substrate, significant alterations in the protein and DNA structure have been observed (*4*).

An interesting feature of the X-ray structure is the reduced occupancy (60%) for the iron coordinated to His 199, suggesting the presence of some [3Fe-4S]$^{1+/0}$ cluster in the sample. This was confirmed by EPR studies on samples of C199HT, where a signal ($g \sim 2.01$) was seen which integrated to 15−30% of the total sample, depending on the enzyme preparation used. This indicates that the Fe coordinated to the histidine is labile. In fact, formation of [3Fe-4S]$^{1+}$ clusters has been more commonly observed when replacing ligating amino acids via site-directed mutagenesis than coordination of the introduced ligand to a [4Fe-4S] cluster. For example, Adams and colleagues (*51*) showed that when the unique aspartate ligand (Asp 14) was changed to a histidine in the ferredoxin from *P. furiosus*, the [3Fe-4S]$^{1+/0}$ cluster form predominates. Ferricyanide oxidation of C199HT increased the amount of the [3Fe-4S]$^{1+}$ cluster approximately 2-fold. Surprisingly, however, the kinetic properties of enzyme samples containing significant amounts of the [3Fe-4S]$^{1+}$ cluster are essentially the same as for WT Stop 225, which contains predominantly [4Fe-4S]$^{2+}$ clusters.

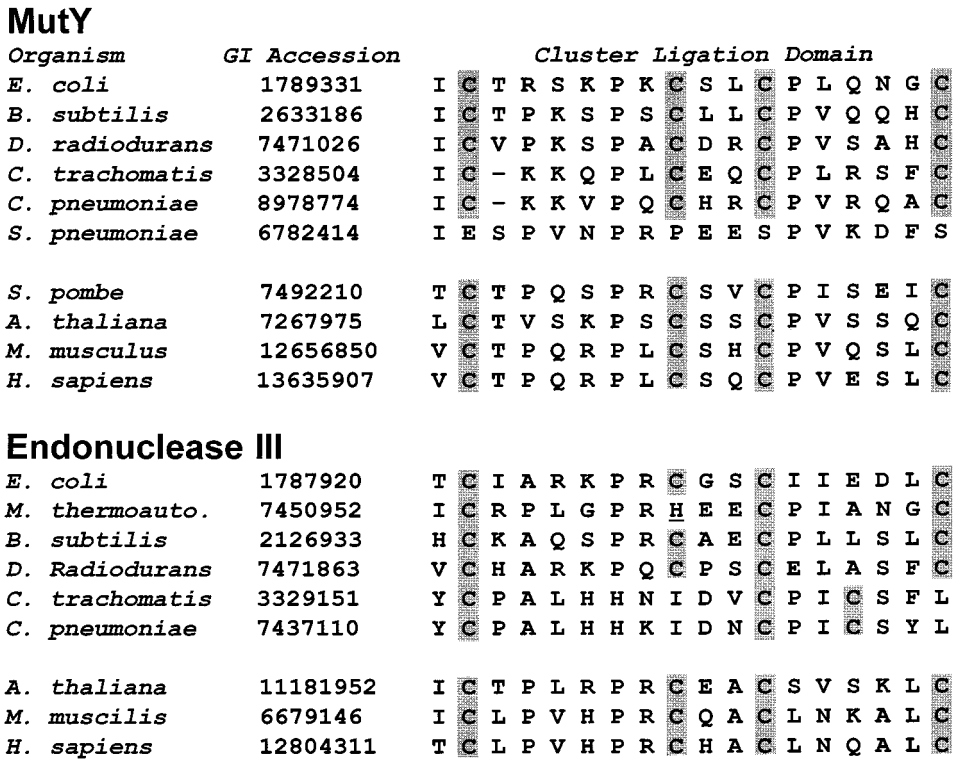


FIGURE 7: Alignment of the FCL region of MutY and endo III enzymes from several organisms.

This suggests that the presence of the $[3Fe\text{-}4S]^{1+}$ cluster does not significantly diminish the ability of the enzyme to catalyze removal of adenine from G·A and OG·A mismatches. This result also provides a possible rationale for our previous observations that mutation of C199 to histidine, serine, or *alanine* did not decrease the efficiency of MutY to prevent DNA mutations in an in vivo assay (*29*). This suggests that the in vivo activity detected for C199A MutY *may possibly* be due to the activity of the 3Fe form. However, it should be noted that while substitutions at position 199 are well tolerated for in vivo activity which requires little active protein, only the presence of a suitable ligand (i.e., histidine or cysteine) provides WT levels of overexpression. Thus, the presence of a completely coordinated [4Fe-4S] cluster in MutY appears to be necessary for overexpression and may be a requirement for the stability of the protein within the cell.

Recently, Camba and Armstrong (*44*) proposed that the conversion of a $[4Fe\text{-}4S]^{2+}$ to a $[3Fe\text{-}4S]^{+}$ cluster may serve as a way to protect enzyme activity in the presence of a highly oxidizing environment. In their studies with *C. pasteurianum* ferredoxin they showed that while oxidation to the 3Fe form was relatively easy, the resulting 3Fe form was stable and resistant to further oxidation. The $[4Fe\text{-}4S]^{2+}$ cluster in this ferredoxin could also be reconstituted in the presence of iron and a reducing environment. Due to the role of MutY in the repair of oxidatively damaged DNA, the enzyme itself might have to function in oxidizing environments. While we have shown that apoMutY is not active (*52*), the cluster conversion reaction could be used as a way for the enzyme to resist cluster loss (and activity loss) entirely. The small reorganizational energy needed for the conversion would allow for the maintenance of the same cuboidal structure as is present in $[4Fe\text{-}4S]^{2+}$ clusters. On the basis of the critical importance of the FCL motif of MutY in recognition of DNA damage, this would allow MutY to take a one-electron oxidation "hit", yet still maintain the same protein scaffold present in the $[4Fe\text{-}4S]^{2+}$ form.

Jung et al. (*53*) recently reported that the [4Fe-4S] cluster in C42D *Azotobacter vinlandii* FdI contains a coordinated aspartate ligand, as demonstrated by X-ray crystallography, and is resistant to conversion to a 3Fe form. The authors suggest that the presence of a noncysteinyl ligand alone is not sufficient for conversion to a 3Fe cluster but rather that the flexibility of the protein backbone and the extent of hydrophobic packing and electrostatic interactions around the cluster are the primary determinants in the 4Fe to 3Fe conversion reaction. Indeed, in *A. vinlandii* FdI, the aspartate is snuggly secured by van der Waals contacts with three hydrophobic residues, while in examples of readily converted proteins the introduced aspartate is flanked by one or two glycine residues. In the case of MutY, the FCL appears flexible enough to accommodate histidine at position C199H, and this may also facilitate loss of the coordinated iron in the C199HT form. Indeed, the structure reveals that the FCL is on the surface of the protein and available to solvent. Moreover, the main chain of His 199 is right on the surface, and there are no strong stabilizing interactions for the His residue as there are for the Asp in the *A. vinlandii* FdI structure.

There are many bacterial and eukaryotic homologues (>50) to *E. coli* MutY and *E. coli* endo III that have been identified. The presence and spacing of the cysteine ligands are almost exclusively conserved within this group. However, there are examples of alterations of the second cysteine position in the Nth/MutY family of genes corresponding to Cys 199 in *E. coli* MutY (Figure 7). The wild-type Nth/EndoIII gene from the recently sequenced genome of *M. thermoautotrophicum* (PIR accession number B69202) has a histidine at the second cysteine position. This suggests that

the cysteine-to-histidine substitution may be a viable, active, natural alternative to the cysteine at this position. Isoleucine is another substitution found at the second position in Nth/EndoIII genes from the *Chlamydia* genera, suggesting that a [3Fe-4S] cluster may be present naturally. In the case of MutY enzymes, the presence of four cysteines is conserved except in three putative bacterial MutY proteins from *Streptococcus pneumoniae*, *Streptococcus pyogenes*, and *Treponema pallidum* which lack all four cysteine ligands (*54*). This suggests that these putative MutY enzymes may lack the $[4Fe\text{-}4S]^{2+}$ cluster entirely. Interestingly, in *S. pneumoniae*, hyper-recombination was found to be dependent on the presence of this atypical MutY homologue (*54*). Thus, the properties of damage recognition and repair by these unusual MutY homologues may involve significantly different mechanisms than their *E. coli* counterpart.

## ACKNOWLEDGMENT

We are extremely appreciative of Professor Rick Holz and Dr. David Bienvenue (Utah State University) in providing access to the EPR spectrometer and help in obtaining EPR spectra.

## REFERENCES

1. Lindahl, T. (1993) *Nature 362*, 709−715.
2. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, Washington, DC.
3. Peltomaki, P. (2001) *Hum. Mol. Genet. 10*, 735−740.
4. David, S. S., and Williams, S. D. (1998) *Chem. Rev. 98*, 1221−1261.
5. Schärer, O. D., and Jiricny, J. (2001) *BioEssays 23*, 270−281.
6. Shigenaga, M. K., Park, J.-W., Cundy, K. C., Gimeno, C. J., and Ames, B. N. (1990) *Methods Enzymol. 186*, 521−530.
7. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature 349*, 431−434.
8. Michaels, M. L., and Miller, J. H. (1992) *J. Bacteriol. 174*, 6321−6325.
9. Maki, H., and Sekiguchi, M. (1992) *Nature 355*, 273−275.
10. Tchou, J., Kasai, H., Shibutani, S., Chung, M.-H., Laval, J., Grollman, A. P., and Nishimura, S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4690−4694.
11. Michaels, M. L., Tchou, J., Grollman, A. P., and Miller, J. H. (1992) *Biochemistry 31*, 10964−10968.
12. Porello, S. L., Leyes, A. E., and David, S. S. (1998) *Biochemistry 37*, 14756−14764.
13. Radicella, J. P., Clark, E. A., and Fox, M. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9674−9678.
14. Tsai-Wu, J.-J., Liu, H.-F., and Lu, A.-L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8779−8783.
15. Michaels, M. L., Pham, L., Nghiem, Y., Cruz, C., and Miller, J. H. (1990) *Nucleic Acids Res. 18*, 3841−3845.
16. Cunningham, R. P., Asahara, H., Bank, J. F., Scholes, C. P., Salerno, J. C., Surerus, K., Münck, E., McCracken, J., Peisach, J., and Emptage, M. H. (1989) *Biochemistry 28*, 4450−4455.
17. Guan, Y., Manuel, R. C., Arvai, A. S., Parikh, S. S., Mol, C. D., Miller, J. H., Lloyd, R. S., and Tainer, J. A. (1998) *Nat. Struct. Biol. 5*, 1058−1064.
18. Piersen, C. E., Prince, M. A., Augustine, M. L., Dodson, M. L., and Lloyd, R. S. (1995) *J. Biol. Chem. 270*, 23475−23484.
19. Begley, T. J., and Cunningham, R. P. (1999) *Protein Eng. 12*, 333−340.
20. Begley, T. J., Haas, B. J., Noel, J., Shekhtman, A., Williams, W. A., and Cunningham, R. P. (1999) *Curr. Biol. 9*, 653−656.
21. Johnson, M. K. (1994) *Encylcopedia of Inorganic Chemistry*, Wiley, New York.
22. Flint, D. H., and Allen, R. M. (1996) *Chem. Rev. 96*, 2315−2334.
23. Beinert, H., Holm, R. H., and Münck, E. (1997) *Science 277*, 653−659.
24. Kiley, P. J., and Beinert, H. (1998) *FEMS Microbiol. Rev. 22*, 341−352.
25. Ding, H., and Demple, B. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 8445−8449.
26. Pomposiello, P. J., and Demple, B. (2001) *Trends Biotechnol. 19*, 109−114.
27. Rouault, T. A., and Klausner, R. D. (1996) *J. Biol. Inorg. Chem. 1*, 494−499.
28. Moulis, J.-M., Devasse, V., Golinelli, M.-P., Meyer, J., and Quinkal, I. (1996) *J. Bioinorg. Chem. 1*, 2−14.
29. Golinelli, M.-P., Chmiel, N. H., and David, S. S. (1999) *Biochemistry 38*, 6997−7007.
30. Chepanoske, C. L., Golinelli, M. P., Williams, S. D., and David, S. S. (2000) *Arch. Biochem. Biophys. 380*, 11−19.
31. Miller, J. H. (1992) *A short course in bacterial genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
32. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning, A laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
33. Leipold, M. D., Muller, J. G., Burrows, C. J., and David, S. S. (2000) *Biochemistry 39*, 14984−14992.
34. Chmiel, N. H., Golinelli, M.-P., Francis, A. W., and David, S. S. (2001) *Nucleic Acids Res. 29*, 553−564.
35. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
36. Brünger, A. T., et al. (1998) *Acta Crystallogr. D54*, 905−921.
37. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol. 277*, 319−343.
38. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
39. Carey, J. (1991) *Methods Enzymol. 208*, 103−117.
40. Chepanoske, C. L., Porello, S. P., Fujiwara, T., Sugiyama, H., and David, S. S. (1999) *Nucleic Acids Res. 27*, 3197−3204.
41. Porello, S. L., Williams, S. D., Kuhn, H., Michaels, M. L., and David, S. S. (1996) *J. Am. Chem. Soc. 118*, 10684−10692.
42. Johnson, M. K., Duderstadt, R. E., and Dunn, E. C. (1999) *Adv. Inorg. Chem. 47*, 1−82.
43. Johnson, M. K., Spiro, T. G., and Mortenson, L. E. (1982) *J. Biol. Chem. 257*, 2447−2452.
44. Camba, R., and Armstrong, F. A. (2000) *Biochemistry 39*, 10587−10598.
45. Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1995) *Nature 373*, 580−587.
46. Peters, J. W., Lanzilotta, W. N., Lemon, B. J., and Seefeldt, L. C. (1998) *Science 282*, 1853−1858.
47. Calzolai, L., Gorst, C. M., Bren, K. L., Zhou, Z.-H., Adams, M. W. W., and La Mar, G. N. (1997) *J. Am. Chem. Soc. 119*, 9341−9350.
48. Stephens, P. J., Jollie, D. R., and Warshel, A. (1996) *Chem. Rev. 96*, 2491−2513.
49. Adams, M. W. W. (1987) *J. Biol. Chem. 262*, 15054−15061.
50. Mulholland, S. E., Gibney, B. R., Rabanal, F., and Dutton, P. L. (1998) *J. Am. Chem. Soc. 120*, 10296−10302.
51. Duderstadt, R. E., Staples, C. R., Brereton, P. S., Adams, M. W. W., and Johnson, M. K. (1999) *Biochemistry 38*, 10585−10593.
52. Porello, S. L., Cannon, M. J., and David, S. S. (1998) *Biochemistry 37*, 6465−6475.
53. Jung, Y. S., Bonagura, C. A., Tilley, G. J., Gao-Sheridan, H. S., Armstrong, F. A., Stout, C. D., and Burgess, B. K. (2000) *J. Biol. Chem. 275*, 36974−36983.
54. Samrakandi, M. M., and Pasta, F. (2000) *J. Bacteriol. 182*, 3353−3360.
55. Bradford, M. M. (1976) *Anal. Chem. 72*, 248−254.

56. McAuley-Hecht, K. E., Leonard, G. A., Gibson, N. J., Thomson, J. B., Watson, W. P., Hunter, W. N., and Brown, T. (1994) *Biochemistry 33*, 10266−10270.
57. Kouchakdijian, M., Bodepudi, V., Shibutani, S., Eisenberg, M., Johnson, F., Grollman, A. P., and Patel, D. J. (1991) *Biochemistry 30*, 1403−1412.
58. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
59. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−134.
60. Bacon, D. J., and Anderson, W. F. (1988) *J. Mol. Graphics 6*, 219−220.
61. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.

BI012035X