Site-Directed Mutagenesis of the Cysteine Ligands to the [4Fe−4S] Cluster of *Escherichia coli* MutY†

Marie-Pierre Golinelli, Nikolas H. Chmiel, and Sheila S. David*

Department of Chemistry, University of Utah, 315 S. 1400 E., Salt Lake City, Utah 84112-0850

Received September 24, 1998; Revised Manuscript Received March 19, 1999

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 [Michaels et al. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7022−7025]. MutY prevents DNA mutations resulting from the misincorporation of A opposite OG by using N-glycosylase activity to remove the adenine base. An interesting feature of MutY is that it contains a $[4Fe-4S]^{2+}$ cluster that has been shown to play an important role in substrate recognition [Porello, S. L., Cannon, M. J., David, S. S. (1998) *Biochemistry 37*, 6465−6475]. Herein, 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, and alanine. The extent to which the various mutations reduce the levels of protein overexpression suggests that coordination of the $[4Fe-4S]^{2+}$ cluster provides stability to MutY in vivo. The ability of the mutated enzymes to bind to a substrate analogue DNA duplex and their in vivo activity were evaluated. Remarkably, the effects are both substitution and position dependent. For example, replacement of cysteine 199 with histidine provides a mutated enzyme that is expressed at high levels and exhibits DNA binding and in vivo activity similar to the WT enzyme. These results suggest that histidine coordination to the iron−sulfur cluster may be accommodated at this position in MutY. In contrast, replacement of cysteine 192 with histidine results in less efficient DNA binding and in vivo activity compared to the WT enzyme without affecting levels of overexpression. The results from the site-directed mutagenesis suggest that the structural properties of the iron−sulfur cluster coordination domain are important for both substrate DNA recognition and the in vivo activity of MutY.

DNA is subject to a variety of chemical modifications resulting from hydrolytic reactions, the action of alkylating agents, radiation, and oxidative stress (*1*). Moreover, errors in DNA replication can lead to base pair mismatches resulting in permanent alterations of the genome's informational content. 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 illustrating the importance of DNA repair in the prevention of diseases such as cancer (*2*). For example, a direct link between defective DNA mismatch repair and hereditary nonpolyposis colon cancer has been established (*3*, *4*).

The base excision repair (BER)[1] pathway is primarily responsible for the repair of damage to heterocyclic DNA bases (*5*, *6*). DNA glycosylases are the marquee players in the BER pathway by virtue of their ability to recognize damaged or mismatched bases and catalyze *N*-glycosidic bond cleavage to effect base release. Subsequent action of apurinic apyrimidinic (AP) endonucleases, DNA polymerase, and DNA ligase restores the correct, undamaged nucleotide (*6*). In *Escherichia coli*, two BER glycosylases, MutM (also called FPG) and MutY, play a coordinated role in preventing mutations caused by oxidative damage to guanine in DNA (Figure 1) (*7*). MutM removes the oxidatively damaged residue, 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG), from DNA (*8*). If the OG lesion escapes MutM, miscoding by OG during DNA replication may result in the formation of an OG:A mismatch (*9*, *10*). If the OG:A base pair is left unrepaired, replication of the "A" strand secures a permanent G-to-T transversion mutation. MutY circumvents such DNA mutations by removing adenine bases from OG:A base pairs (*11*, *12*). Restoration of the G:C base pair is achieved by the action of an AP endonuclease, DNA polymerase, DNA ligase, and MutM (*13*). MutY is also active in the removal of adenine from G:A base pairs within DNA (*14*); however,

†This work was supported by the University of Utah Research Foundation and the Chemistry Department of the University of Utah. S.S.D. is an A.P. Sloan Research Fellow (1998−2000). DNA sequencing was performed by the DNA Sequencing Facility at the Huntsman Cancer Institute, University of Utah, supported in part by NCI grant #5P30CA42014.

* To whom correspondence should be addressed. Tel: 801-585-9718. Fax: 801-581-8433. E-mail: david@chemistry.chem.utah.edu.

[1] Abbreviations: AP, apurinic apyrimidinic; BER, base excision repair; BSA, bovine serum albumin; CD, circular dichroism; DTT, dithiothreitol; *E. coli*, *Escherichia coli*; EDTA, ethylenediaminetetraacetic acid; endo III, endonuclease III; F, 2′-deoxyformycin A; Fd, ferredoxin; FCL, $[4Fe-4S]^{2+}$ cluster loop; HiPIP, high-potential [4Fe−4S] cluster-containing protein; IPTG, isopropyl β-D thiogalactoside; LB, Luria Bertani broth; OG, 7,8-dihydro-8-oxo-2′-deoxyguanosine; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; TBE, tris-borate-EDTA buffer; Tris, tris(hydroxymethyl)aminomethane; WT, wild type; *C. pasterianum*, *Clostridium pasterianum*; *D. gigas*, *Desulforibrio gigas*; *C. vinosum*, *Chromatium vinosum.*

10.1021/bi982300n CCC: $18.00
Published on Web 06/01/1999

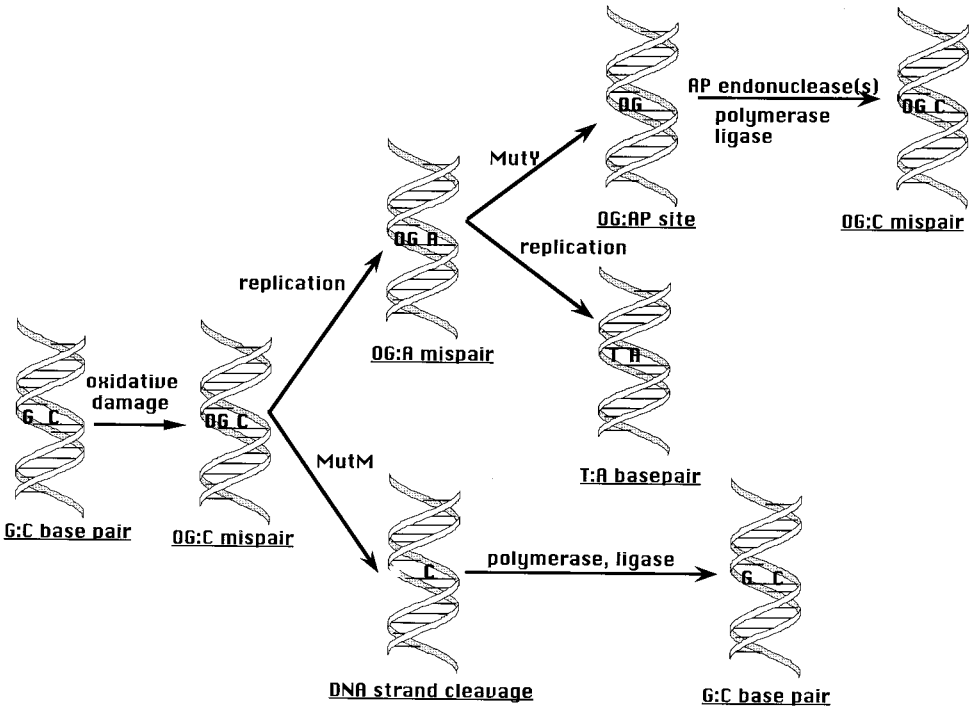

FIGURE 1: Repair system for 7,8-dihydro-8-oxo-2′-deoxyguanosine involving MutM (FPG) and MutY. Adapted from ref 6 and 7.

the rate of adenine removal from OG:A substrates is significantly faster than from G:A substrates (*15*).

MutY belongs to a superfamily of BER enzymes which recognize a wide variety of damaged or inappropriate bases. Structural studies of two superfamily members, *E. coli* endonuclease III (endo III, Figure 2) and AlkA, indicate that these two enzymes have a similar overall three-dimensional fold and contain a helix−hairpin−helix (HhH) motif which is highly conserved within the superfamily (*16*−*19*). *E. coli* MutY exhibits particularly high sequence homology to endo III. Both enzymes are involved in the repair of oxidative DNA damage; however, in contrast to MutY, endo III displays glycosylase activity for DNA containing damaged pyrimidines such as ring-saturated and ring-fragmented thymine derivatives (*20*−*22*). A unique feature of both *E. coli* MutY and *E. coli* endo III is that they have been shown to contain a $[4Fe-4S]^{2+}$ cluster (*23*, *24*). The spacing of cysteine ligands in endo III is unique from that found in other [4Fe−4S] proteins (*25*), indicating that this enzyme is representative of a distinct class of Fe−S proteins. A number of glycosylase enzymes in the BER superfamily, including *Micrococcus luteus* UV endonuclease and a G:T-specific thymine glycosylase from *Methanobacterium thermoformicicum*, contain the same conserved pattern of cysteine residue spacing, suggesting that they also contain a similarly coordinated $[4Fe-4S]^{2+}$ cluster (*6*).

The crystal structure of *E. coli* endo III (*17*, *26*) (Figure 2) shows that four cysteines, arranged with contiguous spacing (Cys187-X_6−Cys194-X_2−Cys197-X_5−Cys203) near the C-terminus, coordinate the cluster to form a distinct metal cluster coordination domain. On the basis of limited protease digestion (*27*, *28*) and sequence homology to endo III (*29*), MutY may be divided into two domains: an N-terminal domain (Met 1-Lys 225) that possesses high sequence homology to endo III (66.3% similar and 23.8% identical in a 181 amino acid span) and a C-terminal extension domain (Gln 226-Val 350). On the basis of the homology to endo III, cysteine residues 192, 199, 202, and 208 were proposed to be the ligands to the [4Fe−4S] cluster in MutY (*29*). The recently reported X-ray structure of MutY's N-terminal domain confirms this cluster coordination (*30*).

The presence of a $[4Fe-4S]^{2+}$ cluster in several DNA repair enzymes is unusual since iron−sulfur clusters are more commonly found in proteins involved in electron transfer reactions (*31*). However, many nonredox roles for Fe−S clusters have been uncovered, suggesting that they are versatile cofactors (*32*, *33*). In endo III, the observation that the cluster is resistant to oxidation or reduction has been used as an argument against its involvement in electron transfer (*23*). The X-ray crystal structure of endo III (*26*) reveals the presence of common amino acids used in DNA phosphate recognition (lysine and arginine) within a solvent-exposed loop between the first two cysteine ligands to the Fe−S cluster. This region, termed the iron−sulfur cluster loop (FCL) motif, appears nicely poised for interactions with DNA (*17*). Indeed, replacement of a lysine residue by glutamic acid in the FCL of endo III resulted in a >100-fold increase of K_m (*17*). These results led to the conclusion that the cluster in endo III may act indirectly in DNA recognition by stabilizing the binding motif or more directly as a structural motif for interaction with DNA (*17*). Recent studies from our laboratory involving the refolding of MutY showed that, in vitro, assembly of the cluster is not required for initiation of global folding of the enzyme, and the cluster does not contribute to the overall thermal stability of the protein (*24*). However, the presence of the cluster is required for MutY's ability to bind to substrate DNA. Not surprisingly, efficient substrate recognition is needed to support MutY's adenine glycosylase activity, and therefore MutY lacking the cluster is also completely inactive in vitro.

Cysteine is the most common ligand to iron−sulfur clusters in proteins. However, chemical models have also indicated that a wide variety of ligands including phenolates, alkoxides, carboxylates, and imidazolates can coordinate iron

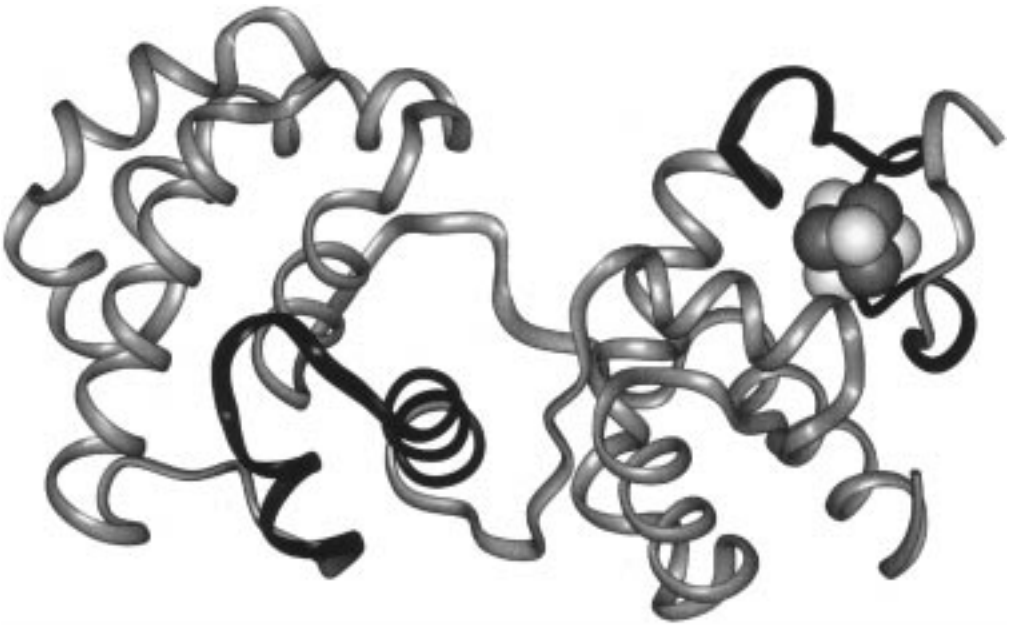

FIGURE 2: Three-dimensional structure of *E. coli* endonuclease III from the X-ray crystallographic data (*17*). The structure is shown in a ribbon representation, with the [4Fe−4S] cluster shown using a space-filling representation. The helix−hairpin−helix motif and the iron−sulfur coordination domain (containing all four cysteines ligating the cluster) are shaded in black. The implicated active site lies within the cleft between these two sequence motifs. The figure was generated using the Brookhaven PDB file ABK2.

in iron−sulfur compounds (*34*−*37*). In contrast, the natural occurrence of [4Fe−4S] clusters coordinated by amino acids other than cysteine is exceedingly rare. Documentated cases include *Desulfovibrio gigas* Ni−Fe hydrogenase (3 Cys, 1 His) (*38*), the Fe-only hydrogenase from *Clostridium pasteurianum* (3 Cys, 1 His) (*39*), and a ferredoxin from the hyperthermophile *Pyrococcus furiosus* (3 Cys, 1 Asp) (*40*). Several modified Fe−S proteins have been produced using site-directed mutagenesis that contain a noncysteinyl ligand coordinated to an iron−sulfur cluster (*41*). Indeed, the substitution of cysteine by serine in iron−sulfur proteins has been commonly used for the identification of ligands (*41, 42*). In some cases, a modified [4Fe−4S] cluster utilizing coordination by the introduced ligand has been produced (*43, 44*). However, due to the fact that protein-bound iron−sulfur clusters usually provide stability to the protein, noncysteinyl ligands may significantly destabilize the iron−sulfur cluster. Typically, mutagenesis of the cysteine ligands results in loss of the cluster or modification of its nuclearity ([3Fe−4S]) (*44*−*46*). Furthermore, the less stable nature of noncysteinyl coordination often makes isolation of these modified proteins difficult (*41*). This has been illustrated in *Chromatium vinosum* HiPIP where individual Cys-to-Ser substitution of the four cysteines coordinating the [4Fe−4S] cluster allowed for isolation and purification of only the C77S mutated enzyme (*47*).

In the present work, site-directed mutagenesis has been used to individually alter MutY's [4Fe−4S] cluster cysteine ligands to serine, histidine, or alanine. The effects of these modifications on the stability of the enzyme, its ability to bind DNA, and its in vivo activity have been determined. These results suggest that the [4Fe−4S] cluster provides stability to the MutY protein in vivo. Furthermore, alteration of the [4Fe−4S] cluster coordination is found to affect both the DNA binding properties and in vivo activity of MutY. Interestingly, serine, histidine, and alanine replacements of the cysteine ligands produce contrasting effects on the stability and activity of the protein, suggesting that histidine and serine in some positions may serve as ligands to the iron−sulfur cluster in MutY.

MATERIALS AND METHODS

General Methods, Bacterial Strains, Materials, and Instrumentation. Common *E. coli* strains JM101 and JM109 were used in this work (*48*). In JM101 *mutY* $^{-}$ the chromosomal copy of the mutY gene was disrupted by transduction of *mutY*::mini-tn *10* into JM101. Strain CC104 is a derivative of the strain P90C containing a *lac*Z mutation at residue 461 of β-galactosidase (*49*)**.** *E. coli* strain GT100 is *ara*, Δ(*gpt-lac*)*5*, *rpsL*, [F′*lacI*^{q}*L8*, *proA*$^{+}$*B*$^{+}$] (*7*). These strains are more fully described by Miller (*48*). The *E. coli* strains CC104 *mutM*::mini-kan *mutY*::mini-tn *10*, GT100 *mutY*::mini-tn *10 mutM*$^{-}$, and JM101 *mutY*::mini-tn *10* were generous gifts from Dr. M. L. Michaels (Amgen). The plasmid containing the *mutY* gene, pKKYEco, was kindly provided by Drs. M. L. Michaels and J. H. Miller. All common DNA manipulations were performed using standard protocols (*50*). Enzymes were purchased from Boehringer Mannheim. All 2′-cyanoethyl phosphoramidites were purchased from ABI, except the 7,8-dihydro-8-oxo-2′-deoxyguanosine phosphoramidite which was purchased from Glen Research. The 2′-deoxyformycin A phosphoramidite was synthesized as described previously (*51*). DNA oligonucleotides were synthesized by standard phosphoramidite chemistry on an Applied Biosystems DNA synthesizer (model 392) as per the manufacturer's protocol. Oligonucleotides for MutY assays were purified and handled as described previously (*52*). 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 purchased from New England Biolabs in the presence of γ-^{32}P-ATP from Amersham Life Sciences. Labeled oligonucleotides were

purified using the Nensorb 20 purification system (DuPont-NEN). UV−visible absorption 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. Site-directed mutagenesis was performed by a modification (*53, 54*) of a method (*55*) which uses two successive rounds of polymerase chain reaction (PCR) to first introduce the desired mutation and then amplify the entire gene. The pKKYEco plasmid, in which a 2 kb DNA fragment containing the *E. coli* MutY gene was cloned between the *Eco*RI (5′-end) and *Pst*I (3′-end) restriction sites of the expression vector pKK223-3 (Pharmacia), was used to prepare a new plasmid, pKKYEco2. This new construction, used as template for PCR reactions, differs from the parent plasmid (pKKYEco) in that the *Pst*I site is located directly after the Stop codon of *mutY*. The appropriate oligonucleotide primers (Supporting Information) were used to introduce the desired mutation.

The PCR reactions were performed in a GeneAmp PCR system 2400 (Perkin-Elmer) in a 100 μL reaction volume. For the first PCR, 10 ng of template, 75 pmol of primer 01 (5′-ctg aaa tga gct gtt gac-3′) and the primer containing the desired mutation (Supporting Information), 20 nmol of each dNTP, and 2.5 units of *Pwo* DNA polymerase were used in a buffer containing 10 μL of 10X *Pwo* buffer. The settings of the thermocycler were 94 °C for 1 min, 45 °C for 1 min, and 72 °C for 1 min, over 30 cycles. The second PCR reaction is performed under the same conditions as the first except that 75 pmol of primer 03 (5′-ggc tga aaa tct tct ctc-3′) and 10 μL of the first PCR reaction are used as the primers. The PCR products were analyzed on a high melting point agarose gel and purified by phenol extraction. The product of the second PCR was digested by *Pst*I and *Eco*RI enzymes and purified on a low melting point agarose gel. The resulting fragment of DNA was ligated with T4 DNA ligase between the *Eco*RI and *Pst*I sites of the pKK223-3 plasmid. The modified plasmid was transformed into *E. coli* strain JM109 and plated on LB agar plates containing 100 mg/L ampicillin. Subsequently, an inoculated 5 mL LB culture containing 100 mg/L ampicillin was grown, and the plasmid was purified from the cells with the Wizard*Plus* MiniPrep kit (Promega). The sequence was confirmed using automated fluorescent sequencing (*56*) on an ABI Prism 377 automated DNA sequencer (PE Applied Biosystems) at the Huntsman Cancer Institute sequencing facility (University of Utah).

Protein Labeling by ^{35}S-Cysteine. *E. coli* strain JM101 *mutY*::mini-tn *10* harboring the pKKYEco2 plasmid containing *mutY* or a modified *mutY* gene was grown in LB medium containing 100 mg/L ampicillin and 15 mg/L tetracycline. When the measured optical density at 600 nm reached 0.9, 1 mL of culture was induced with 1 mM isopropyl β-D thiogalactoside (IPTG). After 15 min of induction, approximately 5 μCi of ^{35}S-cysteine (Amersham) was added, and the cells were grown for three additional hours. The cells were concentrated by centrifugation, and the protein products were resolved by electrophoresis on a 12% SDS polyacrylamide gel. The dried gel was exposed to a storage phosphor screen for a minimum of 8 h, and the resulting storage phosphor autoradiogram was quantified using ImageQuaNT software (Molecular Dynamics).

Preparation of Cellular Lysates for DNA Binding Assays. *E. coli* strain JM101 *mutY*::mini-tn *10* harboring the appropriate plasmid (pKK223−3, pKKYEco2, or a modified pKKYEco2) was grown overnight in LB medium containing 100 mg/L ampicillin and 15 mg/L tetracycline. Overnight cultures (1.5 mL) were used to inoculate 150 mL of LB containing ampicillin and tetracycline. The culture was grown at 37 °C until an optical density at 600 nm of approximately 0.9 was reached. At this point, the culture was induced with 1 mM IPTG and grown at 37 °C for three additional hours. Cells were concentrated by centrifugation and resuspended in 1 mL of buffer A (50 mM Tris-HCl, pH 8.0, 2 mM EDTA, 5 mM DTT, 250 mM NaCl, 5% glycerol). The resuspended cells were sonicated and then centrifuged at 14 000 rpm at 4 °C for 15 min. The cellular lysate was recovered as the supernatant and the protein concentration determined by Bradford assay (Biorad) (*57*).

Gel Retardation Assay with Cellular Lysates. Gel retardation assays (*52, 58*) were performed using the following duplex: d(5′-CGATCATGGAGCCAC(OG)AGCTCCCGTTACAG-3′)•d(3′-GCTAGTACCTCGGTGFTCGAGGGCAATGTC-5′). The OG-strand was 5′-^{32}P-end labeled prior to annealing to the complementary strand. Reactions (20 μL) containing 500 pM duplex, 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% glycerol, and 0.1 mg/mL BSA were incubated at 37 °C with various amounts of cellular lysate. A range of cellular lysate concentrations was used by dilution with a dilution buffer (20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 20% glycerol). After 30 min of incubation, nondenaturing loading dye was added, and the samples were electrophoresed on a 6% nondenaturing polyacrylamide gel (17 cm × 14 cm × 0.3 cm) with 0.5X TBE buffer at 200 V for 15 min and at 100 V for 2 h at 4 °C. The dried gel was exposed to a storage phosphor screen for 8−10 h. The storage phosphor autoradiogram was quantified using ImageQuaNT software (MD).

The percent bound DNA versus log[protein] was plotted, and the concentration required to give 50% bound DNA (C_{50}) was determined. At least three separate measurements using freshly prepared cellular lysates from independent cultures were averaged for the reported C_{50} values.

In Vivo Activity. The pKK223-3 or pKKYEco2 plasmid containing the *mutY* (or a mutated *mutY* gene) was transformed into *E. coli* strain CC104 (*12, 49*) *mutM*::mini-kan *mutY*::mini-tn *10* (*7*). A minimum of 16 independent overnight cultures were grown in LB medium containing 100 mg/L ampicillin and 50 mg/L kanamycin. To determine the number of viable cells, a 10^7-fold dilution was plated on LB agar plates containing ampicillin (100 mg/L) and kanamycin (50 mg/L). A variable volume was also plated on two different types of agar plates to test for mutator frequency (*48*). One set of LB plates containing ampicillin (100 mg/L), kanamycin (50 mg/L), and rifampicin (100 mg/L) was used to measure rifampicin revertants (Rif^r). The other set consisted of minimal lactose plates containing ampicillin (100 mg/L) and kanamycin (50 mg/L) to determine Lac$^+$ revertants. After overnight incubation at 30 °C, Rif^r colonies were counted. Lac$^+$ colonies were counted after 2 days of incubation at 30 °C. The experiments with GT100 *mutM*$^-$ *mutY*::mini-tn *10* were similar except that the plates did not contain kanamycin and only the Rif^r colonies were measured.

MutY Homology Model. The sequences of *E. coli* MutY and endo III were aligned as reported previously by Michaels et al. (*29*) using the Homology module of Insight II (MSI, v. 97.0) on a Silicon Graphics O_2 workstation. Coordinates for the endo III X-ray crystal structure (*17*) were obtained from the Brookhaven Protein Data Bank, file accession number 2ABK. Using these coordinates, homologous residues of MutY were mapped onto the structure of endo III. Loop regions were generated for nonhomologous sections of MutY and spliced between homologous portions. The $[4Fe-4S]^{2+}$ coordinates for endo III were used to map the cluster to the model of MutY. A "bump" check routine was used to verify that no areas of significant van der Waals overlap among side chains were present. Hydrophobicity analysis of the MutY model was accomplished using the Engleman−Steitz algorithm.

RESULTS

Choice of Mutations. In *E. coli* endo III, cysteines 187, 194, 197, and 203 are ligands of the [4Fe−4S] cluster (*23*, *26*). On the basis of the sequence alignments between *E. coli* endo III and *E. coli* MutY, cysteines 192, 199, 202, and 208 were proposed to be the ligands of the cluster in MutY (*29*). Recently, this proposal was confirmed by the X-ray structural characterization of the catalytic domain of MutY (*30*). To probe the role of the iron−sulfur cluster in MutY, we individually altered each of the four cysteine ligands to histidine, serine, and alanine. The replacement of a single cysteine with alanine removes a possible ligand to the cluster at this position. However, cysteine replacement by serine or histidine will alter the type of ligand available for coordination to the cluster.

Levels of Overexpression of Mutated Enzymes. In numerous examples, replacement of cysteine ligands to an iron−sulfur cluster using site-directed mutagenesis has resulted in a significant reduction of the amount of protein produced (*41*). In cases where the WT gene is well-expressed, such a decrease in the amount of polypeptide chain is unlikely to result from a modification of a transcription−translation series of events. Rather, the lower level of expression of the modified gene is more likely related to instability of the mutated protein. This is particularly true with mutations that affect assembly of metal sites since the correct insertion of an Fe−S cluster is often a determining factor in the folding of an Fe−S protein (*41*). If the mutated enzyme does not retain the cluster or contains a malformed cluster, it may be improperly folded and therefore more susceptible to degradation by cellular proteases.

To study the involvement of each of these four cysteines in the in vivo stabilization of the protein, the amount of polypeptide chain produced during overexpression of each modified MutY protein was determined. The different mutated plasmids were prepared using a PCR-based method (Supporting Information), and the enzymes were overexpressed in *E. coli* under conditions similar to the WT enzyme. Proteins produced during the overexpression were radioactively labeled by incorporation of ^{35}S-cysteine after induction with IPTG. The proteins in the crude extracts were then analyzed by SDS−PAGE, and the amount of MutY polypeptide chain was quantified using storage phosphor autoradiography. To eliminate interference with the WT enzyme, the overexpression of the mutated enzymes was performed in an *E. coli* strain lacking the chromosomal copy of WT MutY. Using this procedure, the levels of overexpression of each mutated form relative to the WT enzyme were determined (Table 1).

Individual replacements of the cysteines at positions 192, 199, 202, and 208 by serine dramatically lowers the overexpression of the enzyme. Indeed, the maximum overexpression observed for the serine substitutions (C192S, C199S)[2,3] is less than 15% of WT MutY. The overexpression of the mutated forms containing histidine is significantly higher on average. This is especially striking for the C192H and C199H forms, which are overexpressed at levels similar to the WT enzyme. In contrast, when one of these two cysteines (192, 199) is replaced by an alanine, no polypeptide chain was detected in the ^{35}S-labeling experiment. The observation that individual alterations at all four cysteine positions affects overexpression levels is consistent with their involvement in the coordination of the cluster in the WT form of MutY. Furthermore, these results suggest that coordination of the cluster is involved in stabilizing MutY in vivo.

The effects of these mutations are also highly dependent on the position of the modified residue. The instability of the C192A and C199A enzymes indicates that a ligand at these positions stabilizes the protein. The removal of a ligand may produce a modified protein with a [3Fe−4S] cluster, a [4Fe−4S] cluster with an alternative ligand, or no cluster. It is possible that none of these forms are able to sufficiently sustain MutY's proper folding to prevent its degradation by cellular proteases. The higher overexpression of the histidine and serine mutated forms at positions 192 and 199 compared to the analogous alanine mutations suggests that histidine and serine may coordinate the cluster. In contrast, the mutation of cysteine 202 or 208 to serine or histidine dramatically reduces the amount of protein that is observed. This indicates that the presence of cysteine residues at positions 202 and 208 is required for efficient protein overexpression. However, histidine replacement is again favored over serine as in positions 192 and 199, since quantifiable amounts of the C202H (5%) and C208H (20%) forms relative to the WT form are observed, while no polypeptide chain could be observed for the corresponding serine mutated enzymes.

It is anticipated that the low overexpression of many of the mutated enzymes will severely hamper their purification. Specifically, alterations or loss of the cluster could occur during the process of purification. Indeed, attempts to purify some of the mutated forms resulted in loss of the cluster. Thus, to obtain information on all of the mutated forms, even those that are highly destabilized, we have taken advantage of the ability of MutY present in cellular lysates to bind with high affinity to a substrate analogue. Moreover, the activity of the mutated forms can be evaluated in vivo using MutY's ability to prevent G:C → T:A transversion mutations.

DNA Binding Experiments with Cellular Lysates. The WT form of MutY specifically recognizes OG:A mismatches and

[2] An example of the abbreviation used to represent mutated MutY enzymes is as follows: C192S, cysteine 192 replaced by a serine.

[3] The X-ray crystal structure of the catalytic domain of MutY was published (*30*) during the time that this manuscript was undergoing review; however, the crystallographic coordinates are not yet available.

Table 1: Stability, DNA Binding, and in Vivo Activity Data Obtained on the Cluster Ligand Mutated Forms of MutY

				in vivo activity (CC104)[b]		in vivo activity (GT100)[b]	
protein	relative overexpression (%)[a]	maximal % bound DNA	C_{50} (ng) in cell extracts	no. of colonies (Rif^r, Lac$^+$)	% activity[c]	no. of colonies (Rif^r)	% activity
control	0[d]	0		1780 ± 580 2814 ± 880	0	980 ± 160	0
WT	100	90	115 ± 30	7 ± 5 24 ± 19	100	2 ± 1	100
C192S	12	0		21 ± 6 29 ± 12	100	9 ± 2	99
C192H	78	9		435 ± 120 607 ± 151	78	973 ± 205	0
C192A	0[d]	9		29 ± 11 126 ± 108	98	255 ± 74	76
C199S	12	75	490 ± 300	19 ± 9 9 ± 4	100	3 ± 1	100
C199H	99	85	180 ± 100	10 ± 5 14 ± 1	100	1 ± 1	100
C199A	0[d]	5		1 ± 1 11 ± 2	100	16 ± 3	98
C192A−C199A	0[d]	0		1654 ± 242 3174 ± 846	4	1011 ± 310	0
C202S	0[d]	0		18 ± 2 60 ± 28	99	704 ± 74	28
C202H	5	0		327 ± 110 870 ± 170	83	762 ± 175	23
C208S	0[d]	5		10 ± 3 34 ± 18	100	740 ± 56	24
C208H	20	10		34 ± 22 22 ± 5	100	608 ± 213	38

[a] This is expressed as a percentage compared to the amount of the WT form. The amount of polypeptide chain from a given experiment is determined by the percent of MutY polypeptide chain compared to the rest of the proteins produced during the overexpression. The error in these values is estimated to be 10−20%. [b] An in vivo activity of 100% is given to the cells transformed by pKKYEco2 (WT form of MutY) and 0% for those transformed by the plasmid lacking the *mutY* gene (pKK223-3). The values for Rif^r and Lac$^+$ reversions were determined by averaging at least three separate experiments. Both CC104 and GT100 strains used in these experiments are *mutY*$^-$ *mutM*$^-$. [c] This represents an average of the percent activity obtained from Rif^r and Lac$^+$ assay. [d] A percent polypeptide chain of "0" is reported when no band corresponding to MutY is detected on the SDS polyacrylamide gel using ^{35}S-Cys labeling of the overexpressed proteins.

catalyzes the hydrolysis of the *N*-glycosidic bond of the misincorporated 2′-deoxyadenosine (*13*). Previous work in our laboratory (*52*) has shown that an OG:F (F = 2′-deoxyformycin A) mismatch within a DNA duplex is recognized with high affinity ($K_d \sim 500$ pM) by MutY. However, the nucleotide F contains a *C*-glycosidic linkage which provides resistance to the glycosylase activity of MutY. Thus, oligonucleotide duplexes containing an OG:F mismatch are effective substrate analogues that allow an evaluation of specific binding by MutY without the complications of the enzymatic reaction.

A nondenaturing gel retardation method (*58*) was used to monitor specific complex formation between the OG:F duplex and enzyme present in cellular lysates prepared from the overexpression of MutY (WT or mutated form) in a *mutY*$^-$ *E. coli* strain. To reduce protein degradation, cellular lysates were prepared immediately prior to performing the gel retardation assay. Using cellular lysates from the overexpression of WT MutY, only one retarded band was observed (Figure 3B). This band co-migrates with the retarded band observed when purified WT MutY is used (Figure 3C) and is representative of the MutY/DNA substrate analogue complex (*52*). In the analogous experiment with cells transformed with the plasmid pKK223-3, in which no copy of *mutY* is present, no retarded band at the position corresponding to the MutY−DNA complex was observed (Figure 3A); however, when an extremely high concentration of crude extract was used (Figure 3A, lane 1), a more highly retarded band was observed with a relative mobility consistent with results obtained using purified *E. coli* MutM (data not shown). It is likely that this band comes from MutM present in the cellular lysates, since the chromosomal *mutM* gene is not disrupted in this *E. coli* strain. In experiments using cellular lysates containing overexpressed MutY, this band was not observed (Figure 3B) as a consequence of MutY's higher affinity for the OG:F duplex and higher concentration relative to MutM. Importantly, these results show that we can specifically monitor binding of MutY to the OG:F duplex in cellular lysates.

The formation of the MutY−DNA complex was monitored as a function of the total protein concentration in the cellular lysates to determine a C_{50} value which represents the amount of total protein (in ng) necessary to obtain 50% bound DNA (Figure 3D). When WT MutY is overexpressed, a C_{50} of 115 ± 30 ng was determined. Among all the different mutated forms, high affinity binding was only observed with the C199S and C199H forms (Table 1). Indeed, the C_{50} for the cellular lysates containing C199H (180 ± 100 ng) and C199S (490 ± 300 ng) are in the same range as the value obtained with the extracts containing WT MutY. The average C_{50} value for C199S is slightly higher than for the WT or the C199H form, possibly a consequence of the lower overexpression of this mutated form (Table 1). Surprisingly, in similar experiments using the corresponding serine and histidine mutations at position 192, significant binding to the substrate analogue duplex is not observed even though the overexpression of these forms is similar to those at position 199.

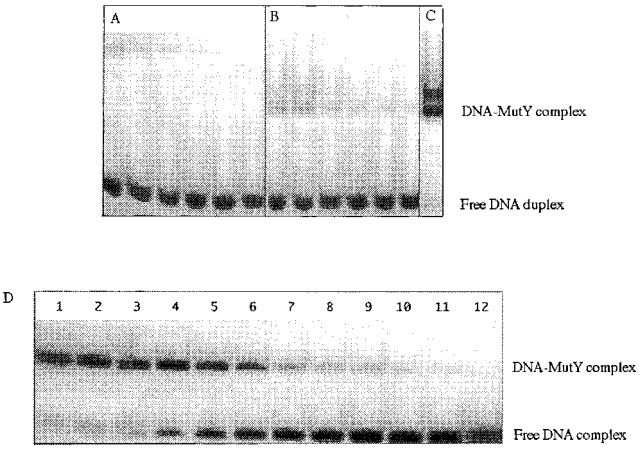

FIGURE 3: Storage phosphor autoradiogram of gel retardation assay to detect binding of MutY in cellular lysates to a substrate analogue 30-mer duplex containing a central OG:F base pair. Lanes in panel A contain cellular extract from *E. coli* JM101 *mutY*::mini-tet strain transformed by pKK223-3 (expression plasmid lacking *mutY* gene). Lanes in panel B contain crude extract from the same strain containing the plasmid pKKYEco2 (contains WT MutY). Lanes in panel C contain purified WT MutY. The amount of crude extract (in ng total protein) used in panels A and B is 75, 38, 19, 9, 5, and 2 ng (left to right). Panel D is a representative gel retardation assay used to quantitate the DNA binding of WT MutY in cellular lysates. The amount of cellular lysate (in ng total protein) used was as follows: 2500 (lane 1), 1250 (lane 2), 625 (lane 3), 312 (lane 4), 156 (lane 5), 78 (lane 6), 39 (lane 7), 20 (lane 8), 10 (lane 9), 5 (lane 10), 2 (lane 11), 1 (lane 12). The OG-strand of the OG:F duplex (500 pM) was 5′-end labeled with ^{32}P. The C_{50} obtained for this particular experiment was 134 ng. Average values from at least three separate determinations are listed in Table 1.

Using extremely high concentrations of cellular lysates, a small amount of MutY−DNA complex (<10%) with the C192H, C192A, C199A, C208S, and C208H forms was observed. The C192H form is overexpressed to high levels, and therefore the low amount of binding suggests that alteration of the cysteine ligand to histidine at this position severely compromises the DNA binding affinity of the enzyme. In contrast, the C192A, C199A, C208S, and C208H forms were overexpressed at low levels. Indeed, the ^{35}S-labeled polypeptide chain for the C192A, C199A, and C208S forms could not be detected by SDS−PAGE. The observation of a DNA-MutY complex suggests that a small amount of these mutated forms are present in the cellular extracts. Interestingly, with the C192A−C199A form, no polypeptide chain was observed by SDS−PAGE and no DNA binding could be observed in the gel retardation assay even at extremely high concentrations of cellular extract (Table 1). The double mutation ensures that the polypeptide chain cannot coordinate an iron−sulfur cluster like the WT enzyme. Taken together, these results suggest that the small amount of the C192A, C199A, C208S, C208H enzymes that was present contains a cluster of some form (a [3Fe−4S] or a [4Fe−4S]) in order to account for the ability to bind the substrate analogue OG:F duplex.

In Vivo Activity. Miller et al. have shown that in the absence of MutY, cells exhibit an increase in G:C → T:A transversion mutations (*12*). In the *E. coli* strain CC104, the *lac*Z gene is inactivated by mutation of Glu 461 (codon GAG) to Gln (codon GCG), and the cells can revert back to the Lac$^+$ phenotype only via a G:C → T:A transversion (*49, 59*). As a consequence, the absence of MutY activity results

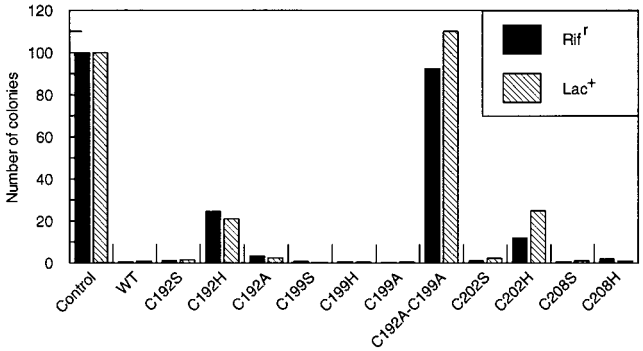

FIGURE 4: Number of Rif^r and Lac$^+$ colonies obtained with the in vivo experiment performed with *E. coli* strain CC104 *mutY*::mini-tn *10 mutM*::mini-kan transformed with pKKYEco2 plasmid (WT or modified *mutY* genes). The results are normalized with a value of 100 attributed to the result with *E. coli* transformed with pKK223-3 plasmid.

in a high level of reversion to the Lac$^+$ phenotype, enabling the *E. coli* to grow on lactose as the only carbon source. In addition, the in vivo mutation frequency can also be determined by using the ability of rifampicin to block bacterial RNA polymerases (*60*). In the absence of MutY activity, a high level of DNA mutations creates modifications in the rifampicin binding site of RNA polymerase, thus providing the *E. coli* with resistance to rifampicin. Miller et al. have also shown that if only the chromosomal *mutY* gene is disrupted, the mutation rate, as judged by both the generation of colonies able to grow on rifampicin containing plates (rifampicin revertants, Rif^r) and the rate of G:C → T:A tranversions in a Lac$^+$ reversion assay, is low due to the presence of MutM (*7, 12, 61*). Indeed, when a CC104 strain where *mutY* and *mutM* are disrupted was used, the mutation rate was significantly higher (approximately 1790 Rif^r and 2814 Lac$^+$ colonies for 10^8 cells compared to 21 and 59, respectively, in the CC104 *mutY*$^-$ strain, Table 1). Expression of WT MutY by transformation of the CC104 *mutY*$^-$ *mutM*$^-$ *E. coli* strain with pKKYEco2 significantly decreased the rate of G:C → T:A transversion mutations (7 Rif^r and 26 Lac$^+$ colonies for 10^8 cells). Thus, the presence of MutY in these cells was able to significantly reduce the appearance of mutations. It is important to note that in these experiments, MutY was constitutively expressed due to the fact that the *E. coli* strain CC104 is *lac*I$^-$ and consequently the *tac* promotor of pKKYEco2 was not repressed.

The in vivo activity of the different mutated forms of MutY is shown in Figure 4 and Table 1. Surprisingly, all of the serine and the two single alanine mutated forms (C192A, C199A) exhibited an in vivo activity comparable to the WT form, even though minimal overexpression was observed. In contrast, the mutated forms containing histidine exhibited a lower average in vivo activity, despite displaying more efficient overexpression than the serine counterparts. This is most striking for the C192H form which has a relatively low in vivo activity compared to the other mutated forms. The in vivo activity is completely abolished only with the C192A−C199A form. In this case, overexpression and DNA binding were not observed. The high level of in vivo activity for the C192A, C199A, C202S, and C208S forms suggests that only a small amount of active enzyme is needed for complete complementation and that expression of the WT enzyme provides much more enzyme than is needed. This

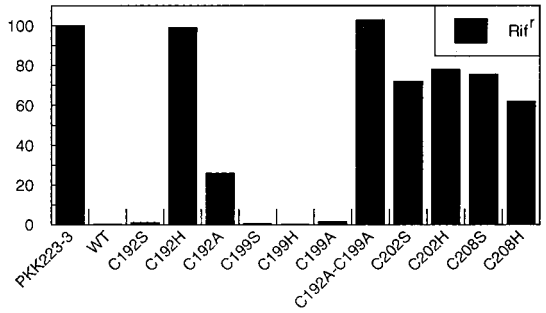

FIGURE 5: Number of Rif^r and Lac$^+$ colonies obtained with the in vivo experiment performed with *E. coli* strain GT100 *mutY*::mini-tn *10 mutM*::mini-kan containing the pKKYEco2 plasmid (WT or modified *mutY* gene). The results are normalized with a value of 100 attributed to the result with *E. coli* transformed by the corresponding plasmid (pKK223-3) lacking *mutY*.

hypothesis is further supported by the fact that the reversion frequency of the *E. coli* strains CC104 and CC104 *mutY*$^-$ transformed with pKKYEco are similar despite the larger cellular concentration of MutY in the latter (*7*). This feature of the assay may be masking some of the differences in the activity of the mutated forms.

To differentiate between the active forms, similar experiments were performed using the *lac*I^q *E. coli* strain GT100 (*7*), in which the *tac* promoter of pKK223-3 is strongly repressed in the absence of IPTG. This strain does not have the mutation in the *lac*Z gene and therefore only Rif^r can be used to determine the mutation frequency. Nonetheless, our results in CC104 (Table 1) indicate that the variation in G:C → T:A transversions (Lac$^+$ assay) is proportional to the variation in overall mutation rate (Rif^r assay). Using this strain, full complementation was observed with *E. coli* expressing the WT enzyme; however, differences in the mutation rates were observed for mutated forms with low overexpression (Figure 5 and Table 1). For example, the histidine and serine mutated forms at positions 202 and 208 are significantly less active than the WT, consistent with the lower levels of expression of these mutated forms. However, even using this *lac*I^q strain, a small amount of protein is required for full complementation as illustrated by the fact that C199A was able to provide complete complementation even though its presence was not detected in the overexpression experiments. This indicates that the low in vivo activity of some mutated forms may not only be a consequence of low amounts of the protein but also be due to an intrinsic lower activity. Indeed, an interesting trend is observed at position 192, where relative levels of overexpression (C192H > C192S > C192A) do not correlate with the relative activity (C192S > C192A > C192H). This suggests that the low in vivo activity of C192H is a consequence of modification of the properties of the protein resulting from the mutation, rather than the amount of protein present.

DISCUSSION

On the basis of the homology to *E. coli* endo III, the cysteines in *E. coli* MutY at positions 192, 199, 202, and 208 were proposed to be the ligands to the [4Fe−4S] cluster (*29*), and this has been recently confirmed by the X-ray structural characterization of the catalytic domain of *E. coli* MutY (*30*). In the present work, our results demonstrate that substitution of one of these four cysteines by serine severely compromises the expression of the MutY polypeptide chain. This is most likely due to a reduction in the amount of completely folded protein due to cluster destabilization. This assertion is further supported by the observed differences in overexpression when these cysteines are substituted by a different type of potential ligand (histidine) or with an amino acid containing a side chain that cannot coordinate a metal ion (alanine). At all four positions, the relative overexpression indicates that replacing the cysteine ligands with histidine is considerably less destabilizing than the analogous serine substitution. Substitution of cysteine by alanine at positions 192 and 199 does not provide a ligand to the cluster, and no polypeptide chain is observed with either alanine substitution. Thus, these results suggest that both serine and histidine may serve as ligands to the cluster to account for the higher overexpression. Taken together, these results are consistent with a requirement for the presence of the [4Fe−4S] cluster in order to establish a stable, folded conformation of MutY in vivo.

Recent work in our laboratory (*24*) has shown that in vitro MutY can be refolded with and without reassembly of the [4Fe−4S] cluster as judged by the lack of major differences in the CD spectrum. These results suggest that [4Fe−4S] cluster assembly is not required in initiating enzyme folding or in stabilizing MutY's overall structure in vitro. Such results may seem to be contradictory to those reported herein based on the overexpression of cluster ligand mutated forms. However, this likely reflects differences in the processes of MutY folding in vitro and in vivo. Indeed, the assembly and insertion of Fe−S clusters into proteins in vivo is not well-understood. In vivo, the absence or alteration of the cluster may not affect the overall structure, but may result in a local structural modification that increases the protein's susceptibility to proteolytic degradation. Furthermore, the in vitro refolding was monitored by CD spectroscopy which is not a sensitive technique for detecting minor changes in protein structure. Consequently, the resistance to proteolytic degradation may be a more sensitive assay for the complete folding of MutY. Previous work with native MutY has shown that limited proteolytic digestion provides two stable fragments (*27*, *28*) and therefore, it will be interesting to determine if MutY enzyme forms lacking the cluster are more susceptible to proteolysis in vitro.

Significantly, though MutY can be refolded without assembly of the [4Fe−4S] cluster, this apo-enzyme form does not exhibit adenine glycosylase activity or the ability to bind to a substrate analogue DNA duplex (*24*). On the basis of these previous results, it is likely that a mutated MutY enzyme that does not contain the iron−sulfur cluster will be unable to bind the substrate analogue OG:F duplex. Thus, the observation of in vivo activity and DNA binding with some of the mutated enzymes is suggestive of the presence of an Fe−S cluster in these modified enzymes. Most notably, the in vivo activity and DNA binding properties of the C199H and C199S enzymes are comparable to those of the WT enzyme, which suggests that these forms contain the [4Fe−4S] cluster.

The C199A form exhibits activity in the in vivo assay, suggesting that a small amount of cluster-containing protein is present in the cells. However, the replacement of cysteine with alanine requires a major modification of the cluster

nuclearity or coordination. In this case, the absence of measurable overexpression and minimal DNA binding indicates that this form is highly unstable. This is in direct contrast with the results obtained with the C199H and C199S enzymes. Indeed, the results with C199H and C199S are consistent with a significantly different coordination which provides stability when the cells are lysed for the DNA binding experiments. This suggests that the [4Fe−4S] cluster present is likely coordinated by the new ligand that has been introduced. Histidine coordination to a [4Fe−4S] cluster has been observed in the crystallographic characterization of *D. gigas* Ni−Fe hydrogenase (*38*) and *C. pasterianum* Fe-hydrogenase (*39*). Serine coordination to a [4Fe−4S] cluster has been implicated in the C184S mutated form of the β subunit of *E. coli* nitrate reductase (*62*) and in the C148S and C151S forms of the FrdB subunit of *E. coli* fumarate reductase (*44*). In addition, NMR studies of the C77S mutated form of *C. vinosum* HiPIP has confirmed serinate coordination to the [4Fe−4S] center (*47, 63*).

Based on the differences in side chain length, the serine mutation might be expected to minimally perturb the overall structure compared to the incorporation of histidine. However, the high levels of expression for the C199H and C192H forms suggest that histidine coordination may be more easily accommodated by MutY's [4Fe−4S] cluster. In inorganic model compounds, Fe(II/III)−OR distances (1.9 Å) are considerably shorter than the corresponding Fe(II/III)−SR distances (2.3 Å) (*36, 64*). In the case of histidine, the side chain is longer than that in cysteine; however, Fe−N(His) distances are expected to be significantly shorter (approximately 2.0−2.1 Å) than Fe−S(Cys) distances (*36, 65*). In examples with serine incorporation into Fe−S proteins, the observed Fe−O(Ser) distances are variable (1.9−2.3 Å), suggesting that the distance adopted in proteins is dependent on the protein matrix (*66, 67*). Thus, incorporation of both serine and histidine may involve movement of the side chain or polypeptide chain to accommodate cluster coordination or the coordination of the cluster may not be optimal. Both of these conditions may affect the stability as well as activity of the mutated enzyme.

By taking advantage of the high sequence and structural homology of MutY and endo III around the [4Fe−4S] cluster, a model of this region in MutY was generated (Figure 6).[3] In endo III, the hydrophilic FCL motif (Thr 186 to Cys 194) has been implicated in DNA binding (*17*). The corresponding FCL region in MutY (Ile 191 to Cys 199) is also highly hydrophilic, containing positively charged residues commonly used in DNA recognition (Arg 194, Lys 196, Lys 198) (Figure 6). The ability to accommodate substitutions of Cys 192 and Cys 199 in MutY is likely a consequence of the flexibility and hydrophilicity of the polypeptide loop between these residues. In contrast, Cys 202 and Cys 208 are located in a hydrophobic portion of the protein, and mutations in hydrophobic segments are generally more destabilizing than in hydrophilic regions (*68*). The differences in stability of mutated forms depending on the position of the modified cysteine may also be a consequence of a difference in the function of these four cysteines in stabilizing the structure of the protein. In endo III, the four corresponding cysteines participate in hydrogen bonding with the amide backbone or an arginine side chain. The loss of similar types of hydrogen bonding at the corresponding cysteines in MutY

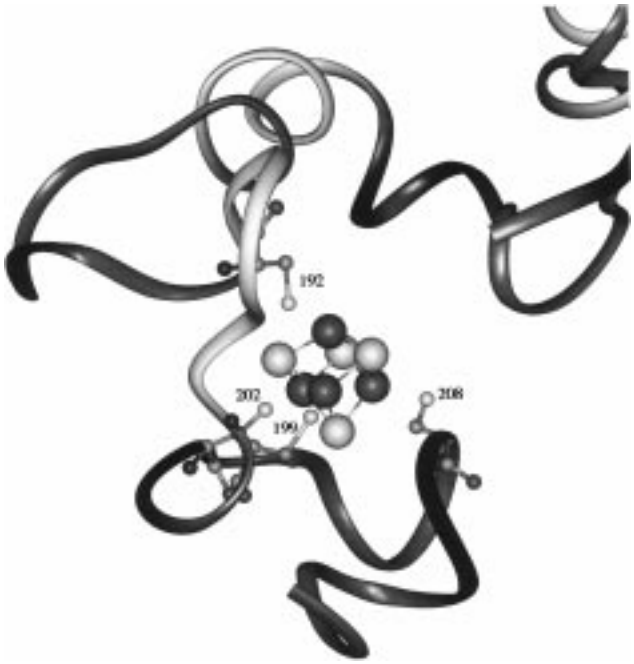

FIGURE 6: Homology model of P26 domain of MutY based on endo III crystal structure (*17*) showing the [4Fe−4S] cluster coordination domain containing the four cysteine ligands (192, 199, 202, and 208). The polypeptide backbone is represented by a ribbon with the side chains of the cysteine ligands shown. The [4Fe−4S] cluster is shown in a ball-and-stick representation. The polypeptide chain has been shaded to indicate hydrophobicity versus hydrophilicity. Darkly shaded regions are hydrophobic, while lightly colored regions are hydrophilic. Neutral regions are shaded in gray.

by the mutations may destabilize the protein. Moreover, in MutY, cysteines 202 and 208 are located near the C-terminal extension domain, and replacement of these cysteines may increase the flexibility of the C-terminal end of the protein, consequently decreasing its stability.

The most dramatic effects on the DNA binding and in vivo activity are observed in the substitutions at position 192. The replacement of cysteine with histidine and serine at position 192 results in similar levels of overexpressed protein as the corresponding replacements at position 199. This suggests that these forms may contain a [4Fe−4S] cluster coordinated by the introduced amino acid to account for the increased stability of these forms relative to the C192A form. However, the C192H and C192S proteins do not exhibit high affinity for the OG:F substrate analogue duplex as was observed with the corresponding C199S/H enzymes. In the in vivo assay using the GT100 strain, both C192S and C192A exhibit near WT activity, while C192H has no activity. The inactivity in the in vivo assay and the inability of C192H enzyme to bind the OG:F duplex may be due to the lack of the [4Fe−4S] cluster. However, this explanation seems inconsistent with its high relative overexpression. A more likely explanation is that the reduced DNA binding ability of C192H is a direct result of histidine coordination which may be unfavorable to the activity relative to the cluster coordination present in the C192S and C192A enzymes.

Unusual coordination of the [4Fe−4S] cluster may alter the redox properties of the cluster (*35, 47, 44, 63*), and this could account for the lower activity of the C192H enzyme. However, the inability to reduce or oxidize the $[4Fe-4S]^{2+}$ cluster of endo III suggests that this class of DNA repair

enzymes is not involved in a redox reaction (*23*). A more likely explanation for the reduced activity of the C192H form is that this mutation has resulted in a local modification of the protein's conformation that reduces efficient substrate recognition. In the endo III structure (*17*, *26*), the Sγ atom of Cys 187 (equivalent to 192 in MutY) forms a hydrogen bond with the Nη2 atom of Arg 147 (equivalent to Arg 147 in MutY). Assuming that similar interactions may be present in MutY, mutation of Cys 192 to histidine would result in loss of this hydrogen bond and potentially modify the conformation of MutY's FCL. Conformational changes in the polypeptide chain proximal to position 192 may also be required in order to accommodate the histidine ligation. Alterations of the conformation of MutY's FCL, which contains residues that may be involved in DNA recognition, may explain the decreased ability of the C192H form to interact with DNA. This suggests that the $[4Fe-4S]^{2+}$ cluster in MutY may provide the scaffolding for proper positioning of the amino acids for interactions with DNA in a manner analogous to that previously proposed for endo III (*17*).

To establish whether histidine or serine is coordinated to the [4Fe−4S] cluster, purification of the C192H, C192S, C199H, and C199S MutY enzymes is in progress. If purification of the holoproteins is possible, spectroscopic studies will be useful in determining the coordination of the [4Fe−4S] cluster and the effect of such atypical coordination on the properties of a [4Fe−4S] cluster. If unique spectroscopic signatures of these modified MutY forms are identified, this may also provide spectroscopic handles for determining the effects of substrate and substrate analogue binding on the properties of MutY's [4Fe−4S] coordination domain. A thorough characterization of the DNA binding and enzymatic properties of some of these modified forms is also in progress. This will provide further insight into the specific roles the [4Fe−4S] cluster and the FCL motif play in detecting OG:A mismatches within DNA.

ACKNOWLEDGMENT

We are appreciative of generous sharing of *E. coli* strains and plasmids by Dr. Mark Michaels, as well as helpful and useful comments concerning this work. The F-containing oligonucleotide was synthesized by Dr. Heiko Kuhn.

SUPPORTING INFORMATION AVAILABLE

A table of primers used in the PCR reactions to prepare the mutated forms of MutY. This material is available free of charge via the Internet at http://pubs.acs.org.

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. Modrich, P. (1994) *Science 266*, 1959−1960.
4. Modrich, P. and Lahue, R. (1996) *Annu. Rev. Biochem. 65*, 101−133.
5. Krokan, H. E., Standal, R., and Slupphaug, G. (1997) *Biochem. J. 325*, 1−16.
6. David, S. S., and Williams, S. D. (1998) *Chem. Rev. 98*, 1221−1261.
7. Michaels, M. L., Cruz, C., Grollman, A. P., and Miller, J. H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7022−7025.
8. Tchou, J., and Grollman, A. P. (1993) *Mutat. Res. 299*, 277−287.
9. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature 349*, 431−434.
10. Wood, M. L., Esteve, A., Morningstar, M. L., Kuziemko, G. M., and Essigmann, J. M. (1992) *Nucleic Acids Res. 20*, 6023−6032.
11. Michaels, M. L., Tchou, J., Grollman, A. P., and Miller, J. H. (1992) *Biochemistry 31*, 10964−10968.
12. Nghiem, Y., Cabrera, M., Cupples, C. G., and Miller, J. H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2709−2713.
13. Michaels, M. L., and Miller, J. H. (1992) *J. Bacteriol. 174*, 6321−6325.
14. Tsai-Wu, J.-J., Liu, H.-F., and Lu, A.-L. (1992) *Proc. Nat. Acad. Sci. U.S.A. 89*, 8779−8783.
15. Porello, S. L.; Leyes, A. E., and David, S. S. (1998) *Biochemistry 37*, 14756−14764.
16. Nash, H. M., Bruner, S. D., Schärer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S., and Verdine, G. L. (1996) *Curr. Biol. 6*, 1230−1233.
17. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., and Tainer, J. A. (1995) *EMBO J. 14*, 4108−4120.
18. Labahn, J., Schärer, O. D., Long, A., Ezaz-Nikpay, K., Verdine, G. L., and Ellenberger, T. E. (1996) *Cell 86*, 321−329.
19. Yamagata, Y., Kato, M., Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., Yasumura, K., Tomita, K.-I., Ihara, K., Fujii, Y., Nakabeppu, Y., Sekiguchi, M., and Fujii, S. (1996) *Cell 86*, 311−319.
20. Katcher, H. L., and Wallace, S. S. (1983) *Biochemistry 22*, 4072−4081.
21. Hatahet, Z., Kow, Y. W., Purmal, A., Cunningham, R. P., and Wallace, S. S. (1994) *J. Biol. Chem. 269*, 18814−18820.
22. Breimer, L. H., and Lindahl, T. (1984) *J. Biol. Chem. 259*, 5543−5548.
23. Cunningham, R. P., Ashara, 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.
24. Porello, S. L., Cannon, M. J., and David, S. S. (1998) *Biochemistry 37*, 6465−6475.
25. Fukuyama, K., Nakahara, Y., Tsukihara, T., Katsube, Y., Hase, T., and Matsubara, H. (1988) *J. Mol. Biol. 199*, 183−193.
26. Kuo, C.-F., McRee, D. E., Fisher, C. L., O'Handley, S. F., Cunningham, R. P., and Tainer, J. A. (1992) *Science 258*, 434−440.
27. Gogos, A., Cillo, J., Clarke, N. D., and Lu, A.-L. (1996) *Biochemistry 35*, 16665−16671.
28. Manuel, R. C., Czerwinski, E. W., and Lloyd, R. S. (1996) *J. Biol. Chem. 271*, 16218−16226.
29. Michaels, M. L., Pham, L., Nghiem, Y., Cruz, C., and Miller, J. H. (1990) *Nucleic Acids Res. 18*, 3841−3845.
30. Guan, Y., Manuel, R. C., Arvai, A. S., Parikh, S. S., Mol, C. D., Miller, J. H., Lloyd, R. S., and Tainer, J. S. (1998) *Nat. Struct. Biol. 5*, 1058−1063.
31. Johnson, M. K. (1994) in *Encyclopedia of Inorganic Chemistry* King, R. B., Ed. Vol. 4, John Wiley & Sons, New York.
32. Flint, D. H., and Allen, R. M. (1996) *Chem. Rev. 96*, 2315−2334.
33. Beinert, H., Holm, R. H., and Münck, E. (1997) *Science 277*, 653−659.
34. Holm, R. H., Ciurli, S., and Weigal, J. A. (1990) in *Progress in Inorganic Chemistry* (Lippard, S. J., Ed.) Vol. 38, John Wiley & Sons, Inc., New York.
35. Beardwood, P., and Gibson, J. F. (1992) *J. Chem. Soc., Dalton Trans.* 2457−2466.
36. Salifoglou, A., Simopoulos, A., Kostikas, A., Dunham, R. W., Kanatzidis, M. G., and Coucouvanis, D. (1988) *Inorg. Chem. 27*, 3394−3406.
37. Cleland, W. E., Holtman, D. A., Sabat, M., Ibers, J. A., DeFotis, G. C., and Averill, B. A. (1984) *J. Am. Chem. Soc. 105*, 6021−6031.
38. Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1995) *Nature 373*, 580−587.
39. Peters, J. W., Lanzilotta, W. N., Lemon, B. J., and Seefeldt, L. C. (1998) *Science 282*, 1853−1858.

40. 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.
41. Moulis, J.-M., Davasse, V., Golinelli, M.-P., Meyer, J., and Quinkal, I. (1996) *JBIC 1*, 2−14.
42. Crouse, B. R., Sellers, V. M., Finnegan, M. G., Dailey, H. A., and Johnson, M. K. (1996) *Biochemistry 35*, 16222−16229.
43. Xia, B., Cheng, H., Bandarian, V., Reed, G. H., and Markley, J. L. (1996) *Biochemistry 35*, 9488−9495.
44. Kowal, A. T., Werth, M. T., Manodori, A., Cecchini, G., Schröder, I., Gunsalus, R. P., and Johnson, M. K. (1995) *Biochemistry 34*, 12284−12293.
45. Mehari, T., Qiao, F., Scott, M. P., Nellis, D. F., Zhao, J., Bryant, D. A., and Goldbeck, J. H. (1995) *J. Biol. Chem. 270*, 28108−28117.
46. Shen, B., Jollie, D. R., Diller, T. C., Stout, C. D., Stephens, P. J., Burgess, B. K. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10064−10068.
47. Agarwal, A., Li, D., and Cowan, J. A. (1996) *J. Am. Chem. Soc. 118*, 927−928.
48. Miller, J. H. (1992) *A short course in bacterial genetics*, Cold Spring Harbor, New York.
49. Cupples, C. G., and Miller, J. H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5345−5349.
50. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning, A laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
51. Kuhn, H., Smith, D. P., and David, S. S. (1995) *J. Org. Chem. 60*, 7094−7095.
52. Porello, S. L., Williams, S. D., Kuhn, H., Michaels, M. L., and David, S. S. (1996) *J. Am. Chem. Soc. 118*, 10684−10692.
53. Kammann, M., Laufs, J., Schell, J., and Gronenborn, B. (1989) *Nucleic Acids Res. 17*, 5404.
54. Meyer, J., Fujinaga, J., Gaillard, J., and Lutz, M. (1994) *Biochemistry 33*, 13642−13650.
55. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res. 16*, 7351−7365.
56. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. D., and Hood, L. E. (1986) *Nature 321*, 674−679.
57. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
58. Carey, J. (1991) *Methods Enzymol. 208*, 103−117.
59. Miller, J. H., and Michaels, M. L. (1996) *Gene 179*, 129−132.
60. Wehrli, W., Knüsel, F., Schmid, K., Steheln, M. (1968) *Proc. Natl. Acad. Sci. U.S.A. 61*, 667−673.
61. Cabrera, M., Nghiem, Y., and Miller, J. H. (1988) *J. Bacteriol. 170*, 5405−5407.
62. Augier, V., Guigliarelli, B., Asso, M., Bertrand, P., Frixon, C., Giordano, G., Chippaux, M., and Blasco, F. (1993) *Biochemistry 32*, 2013−2023.
63. Babini, E., Bertini, I., Borsari, M., Capozzi, F., Dikiy, A., Eltis, L. D., and Luchinat, C. (1996) *J. Am. Chem. Soc. 118*, 75−80.
64. Cleland, W. E., and Averill, B. A. (1984) *Inorg. Chem. 23*, 4192−4197.
65. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure 4*, 567−579.
66. Xiao, Z., Lavery, M. J., Ayhan, M., Scrofani, S. D. B., Wilce, M. C. J., Guss, J. M., Tregloan, P. A., George, G. N., and Wedd, A. G. (1998) *J. Am. Chem. Soc. 120*, 4135−4150.
67. Bentrop, D., Bertini, I., Capozzi, F., Dikiy, A., Eltis, L., and Luchinat, C. (1996) *Biochemistry 35*, 5928−5936.
68. Bordo, D., and Argos, P. (1991) *J. Mol. Biol 217*, 721−729.

BI982300N