Determination of the Structure of *Escherichia coli* Glyoxalase I Suggests a Structural Basis for Differential Metal Activation†

Molly Min He,‡ Susan L. Clugston,§ John F. Honek,§ and Brian W. Matthews*,‡

Howard Hughes Medical Institute, Institute of Molecular Biology, Department of Physics, 1229 University of Oregon, Eugene, Oregon 97403-1229, and Department of Chemistry, University of Waterloo, Waterloo, Ontario N2L 3G1 Canada

Received April 14, 2000; Revised Manuscript Received May 19, 2000

ABSTRACT: The metalloenzyme glyoxalase I (GlxI) converts the nonenzymatically produced hemimercaptal of cytotoxic methylglyoxal and glutathione to nontoxic *S*-D-lactoylglutathione. Human GlxI, for which the structure is known, is active in the presence of Zn^{2+}. Unexpectedly, the *Escherichia coli* enzyme is inactive in the presence of Zn^{2+} and is maximally active with Ni^{2+}. To understand this difference in metal activation and also to obtain a representative of the bacterial enzymes, the structure of *E. coli* Ni^{2+}-GlxI has been determined. Structures have also been determined for the apo enzyme as well as complexes with Co^{2+}, Cd^{2+}, and Zn^{2+}. It is found that each of the protein−metal complexes that is catalytically active has octahedral geometry. This includes the complexes of the *E. coli* enzyme with Ni^{2+}, Co^{2+}, and Cd^{2+}, as well as the structures reported for the human Zn^{2+} enzyme. Conversely, the complex of the *E. coli* enzyme with Zn^{2+} has trigonal bipyramidal coordination and is inactive. This mode of coordination includes four protein ligands plus a single water molecule. In contrast, the coordination in the active forms of the enzyme includes two water molecules bound to the metal ion, suggesting that this may be a key feature of the catalytic mechanism. A comparison of the human and *E. coli* enzymes suggests that there are differences between the active sites that might be exploited for therapeutic use.

Although metal ions play an important role in many enzyme-catalyzed reactions (*1*), the respective roles that are played by the metal ions themselves and the surrounding ligands are often poorly understood (*2*−*8*). A case in point is glyoxalase I, which, as explained below, has unexpected variability in its response to the binding of different metals, depending on the source of the enzyme.

Glyoxalase I (GlxI; *S*-D-lactoylglutathione lyase, EC 4.4.1.5),[1] the first of two enzymes in the pathway to convert cytotoxic α-keto aldehydes, such as methylglyoxal, into nontoxic α-hydroxycarboxylic acids, requires a metal ion for catalytic activity (*9*). This enzyme system is important in that an increase in methylglyoxal can produce toxic effects by reacting with DNA, RNA, and proteins (*10*−*13*). This toxicity can be exploited by inhibitors designed to target GlxI and has been utilized in the design of anticancer and antimalarial agents (*13*−*15*). Similarly, specific inhibitors for the bacterial GlxI may be effective antibiotics (*10*).

In *Homo sapiens*, *Saccharomyces cerevisiae*, and *Pseudomonas putida* GlxI, the essential metal has been shown to be zinc (*16*−*18*). In contrast, however, GlxI from *E. coli* is completely inactive in the presence of this ion. Maximal activity is found in the presence of Ni^{2+}, and reduced activity is found with Co^{2+}, Cd^{2+}, and Mn^{2+} (*19*, *20*). This was especially puzzling given the relatively high amino acid sequence homology between the enzymes.

The structure of human GlxI has been reported (*21*) and has indicated it to be a member of the so-called $\beta\alpha\beta\beta\beta$ superfamily of metalloproteins (*22*). It is the only eukaryotic protein from the family whose structure has been determined and is somewhat remote from the known bacterial members of this family. To obtain information on a representative bacterial enzyme, the structure of *E. coli* GlxI has been determined. Also, to ameliorate our knowledge of protein−nickel interactions, an area of current interest in metalloenzymology (*23*−*25*), and to try to understand why the human and *E. coli* enzymes have very different metal requirements, the structure of the *E. coli* enzyme was determined in the apo form as well as in complex with nickel, cobalt, cadmium, and zinc. The results suggest that it is the geometry of metal coordination, rather than the metal ion itself, that correlates with catalytic activity.

MATERIALS AND METHODS

Materials. HEPES was purchased from BDH (Toronto, ON), and seleno-*L*-methionine (SeMet) was purchased from Sigma Chemical Co. All other reagents used in the growth and purification of the enzyme were described previously (*19*). Ampicillin and carbenicillin were used at a concentration of 50 μg/mL in Luria-Bertani broth and 30 μg/mL in minimal media M9 (*26*).

† This work was supported by NSERC Canada (J.F.H) and NIH Grant GM20066 (B.W.M). S.L.C. was supported by a postgraduate scholarship from NSERC.

* To whom correspondence should be addressed: Institute of Molecular Biology, 1229 University of Oregon, Eugene, OR 97403. Telephone: (541) 346-2572. Fax: (541) 346-5870. E-mail: brian@uoxray.uoregon.edu.

‡ University of Oregon.

§ University of Waterloo.

[1] Abbreviations: GlxI, glyoxalase I; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; MOPS, 4-morpholine-propanesulfonic acid; SeMet, seleno-*L*-methionine.

10.1021/bi000856g CCC: $19.00
Published on Web 07/01/2000

Table 1: Data Collection and Phasing Statistics

data set	Ni^{2+}	Co^{2+}	Cd^{2+}	Zn^{2+}	apo	Ni^{2+} (SeMet)[a]
space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_12_12_1$
cell dimensions						
a (Å)	46.05	46.09	45.93	46.24	46.28	80.46
b (Å)	56.48	56.57	56.39	57.17	57.20	85.63
c (Å)	46.71	46.75	46.79	46.99	46.99	122.96
β (deg)	95.4	95.5	95.2	95.4	95.2	
resolution range (Å)	20−1.5	20−1.9	20−1.9	20−1.8	20−1.7	20−2.5
measured reflections	116 872	116 932	47 320	134 520	81 372	103 752
unique reflections	38 619	17 383	19 871	22 460	29 051	32 488
completeness (%)[b]	98.9 (98.0)	90.0 (88.7)	90.3 (90.0)	98.3 (98.2)	97.8 (93.1)	94.2 (91.0)
R_{sym} (%)[c]	6.9 (28.9)	5.1 (11.0)	10.2 (27.6)	4.6 (5.2)	5.1 (24.1)	4.8 (26.2)
$<I/\sigma(I)>$	12.8 (1.9)	33.2 (6.1)	8.7 (1.5)	38.9 (13.6)	11.5 (2.2)	19.2 (3.4)
phasing power[d]						3.2
R_{cullis}[e]						0.57

[a] The data statistics for Ni^{2+} SeMet are calculated from the data set collected at a remote wavelength ($\lambda = 0.9665$ Å). [b] Numbers in parentheses are for the highest resolution bins. [c] $R_{sym}(I) = (\Sigma_{hkl}\Sigma_i \mid I(i) - <I(hkl)> \mid)/\Sigma_{hkl}<I(hkl)>$. [d] The phasing statistics are calculated from the data set collected at the peak wavelength ($\lambda = 0.9792$ Å). The phasing power, calculated for the acentric reflections, is the root-mean-square value of ($|F_H|/|F_{PH} - |F_P + F_H|\ |$). [e] $R_{cullis} = \Sigma\ |F_{PH} - |F_P + F_H|\ | / \Sigma\ |F_{PH} - F_P|$ and is calculated for the acentric reflections.

Protein Production and Purification. Native *E. coli* MG1655/pGL10 was grown, GlxI was expressed, and the protein was purified as described (*19*). For incorporation of SeMet, M9 (supplemented with 0.4% glucose, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.001% uracil, and 5 μM $NiCl_2$) was inoculated with *E. coli* MG1655/pGL10 grown in LB_{Amp}. SeMet (0.3 mM) and additional $NiCl_2$ (7.5 μM) was added to the culture when the OD_{600} reached ∼0.5. Following 30 min additional growth at 37 °C, GlxI synthesis was induced for 8 h with IPTG (0.5 mM). The cells were harvested and the protein was purified in the same manner as described for the native enzyme (*19*). SeMet incorporation was monitored by electrospray mass spectrometry, provided by the Biomedical Mass Spectrometry Laboratory, University of Waterloo.

Enzyme Preparation for Crystallization. The precipitated protein following isoelectric focusing has the highest purity and was utilized for all crystallographic work. The protein (in 50 mM MOPS, pH 7.0, Chelex treated, Na^+ form) was concentrated using an Amicon Centricon concentrator (YM10; Millipore, Bedford, MA), and the buffer was changed to 50 mM HEPES, pH 7.0 (Chelex treated). For the various metal-substituted forms of the enzyme, 2−2.5 mol equiv of metal ($NiCl_2$, $ZnCl_2$, $CdCl_2$, $CoCl_2$, or $MnCl_2$) were added to the dimeric apoenzyme, prior to concentration. The enzyme was concentrated to between 12 and 37 mg/mL. Concentration was determined utilizing the Bradford assay (*27*) with bovine albumin as the standard, and enzyme activities were performed as previously described (*19*). Protein samples were frozen in liquid nitrogen and stored at −80 °C in eppendorf tubes. Tubes were soaked in 10% nitric acid for ∼30 min to remove any extraneous metals followed by liberal rinsing with Chelex-treated water.

Crystallization of Ni^{2+}-, Co^{2+}-, Cd^{2+}-, Zn^{2+}-, Apo-, and SeMet-Ni^{2+}-GlxI. The various metal-bound forms of native GlxI were readily crystallized by vapor diffusion in hanging drops with PEG 1000 and PEG 8000 as a precipitant. The protein solution (5 μL of 12−37 mg/mL) was mixed with an equal volume of well liquor containing 5−10% PEG 1000 and 5−10% PEG 8000. After approximately one week at room temperature, crystals of dimensions ∼0.8 × 0.3 × 0.2 mm appeared. Although the apo-GlxI protein gave crystals under similar conditions, they appeared to be less stable. More stable crystals with better morphology were obtained by placing the trays at 4 °C for 2 weeks or longer. The crystals of SeMet-Ni^{2+}-GlxI could only be obtained when additional buffer (50 mM HEPES, pH 7.0) was added to the hanging drops containing ∼11 mg/mL protein, and they differed in morphology from the native GlxI crystals. In addition, these crystals diffracted very poorly until subjected to one or more cycles of crystal annealing (*28*, *29*). The crystals were flash-cooled in a stream of cold nitrogen and then transferred to mother liquor with 30% PEG 400 at room temperature. After a 1-min incubation the crystals were flash-cooled again in the presence of 30% PEG 400.

Data Collection and Structure Determination. The diffraction data for native Ni^{2+}- and apo-GlxI were collected on beamline 9-1 at the Stanford Synchrotron Radiation Laboratory. The data for the crystals of Co^{2+}-, Cd^{2+}-, and Zn^{2+}-GlxI were recorded under cryo conditions on a Rigaku R-AXIS imaging plate mounted on a rotating anode source. All native GlxI data sets were integrated and scaled using the Mosflm and Scala package (*31*). Four data sets of SeMet-Ni^{2+}-GlxI crystals were collected at different wavelengths, maximizing the K-edge anomalous dispersion effects of selenium, on beamline 5-2 at the Advanced Light Source (Berkeley, CA). These data sets were integrated and scaled using HKL2000 (*32*). Data collection statistics can be seen in Table 1. The space group of native GlxI crystals was determined to be $P2_1$ with two molecules per asymmetric unit. In contrast, the SeMet-Ni^{2+}-GlxI crystals grew in space group $P2_12_12_1$ with six molecules per asymmetric unit.

Attempts to determine the structure of the native *E. coli* enzyme in space group $P2_1$ using molecular replacement based on the human enzyme (*21*) were unsuccessful. This led to the use of the selenomethionine form which, as noted, crystallized in a different space group. The positions of the 12 Se sites were determined by Shake 'n Bake (*33*). The program SHARP (*34*) was used to refine the positions of the Se sites and to calculate the experimental electron density map. The phasing statistics can be seen in Table 1.

Model Building, Refinement, and Analysis. The model of SeMet-Ni^{2+}-GlxI was built with the program O (*35*, *36*). The noncrystallographic symmetry operators relating the three dimers in the asymmetric unit were determined using O and refined with CNS_Solve (*37*). Refinement of the structure

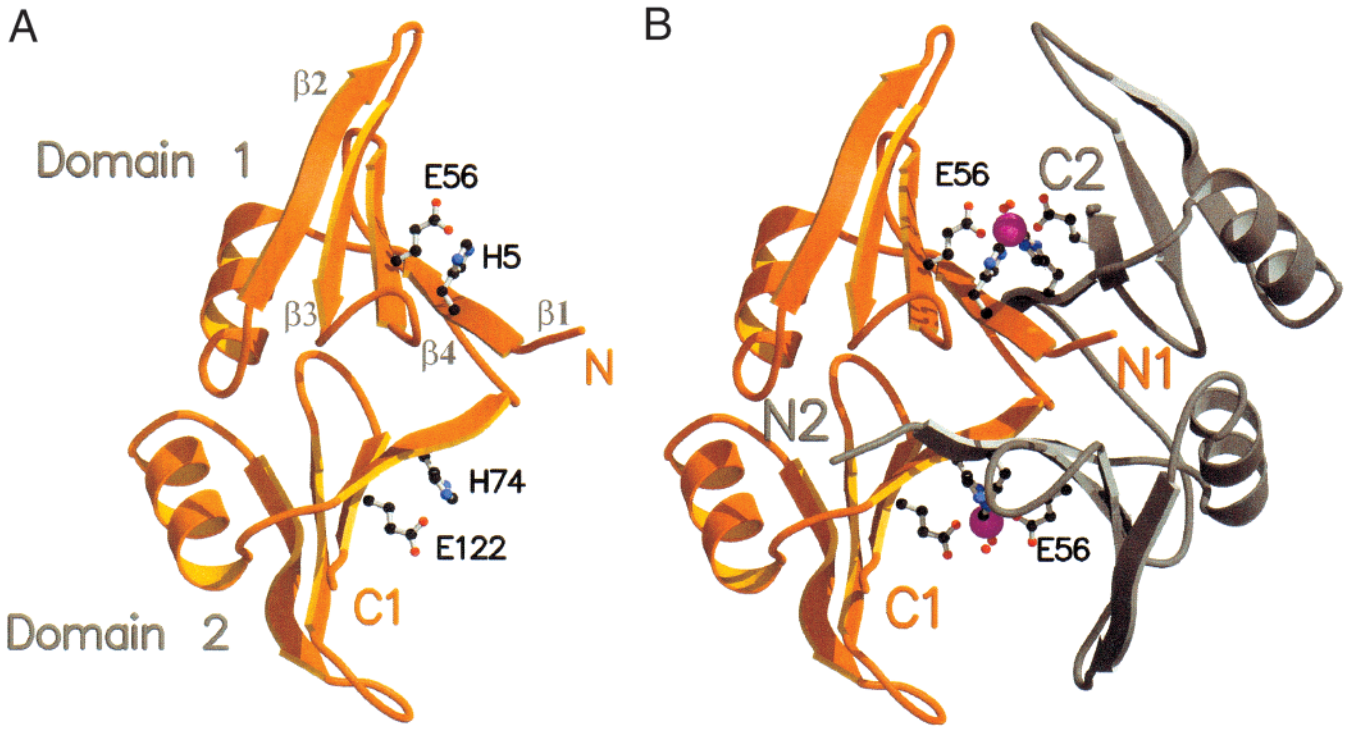

FIGURE 1: Schematic representation of *E. coli* GlxI. (A) Monomer; (b) dimer. The two monomers in the dimer are colored orange and gray, respectively. The nickel (magenta) and its coordinating protein and water ligands are shown in a ball-and-stick representation. The active site is situated within a curved β-sheet, which is formed on dimerization. Prepared using MOLSCRIPT and Render3D (*57*, *58*).

was first carried out using CNS_Solve (*37*). This involved a combination of simulated annealing (*38*) interspersed with manual rebuilding into maps averaged using RAVE (*39*). Further refinement was carried out using TNT (*40*). Five percent of the total reflections from thin shells were excluded from the refinement process and used to calculate *R*-free (*41*). Strict noncrystallographic symmetry was maintained throughout, until *R* had dropped to 25% and *R*-free had dropped to 30%.

At this stage, the model was used in a molecular replacement search (*42*) leading to the successful determination of the GlxI structure in the $P2_1$ crystal form. This not only made it possible to refine (*40*) the structures of the various metal-substituted forms of the enzyme but also to take advantage of the higher resolution data obtainable for these crystals. The structure validation program PROCHECK (*43*) shows that the final models for all structures have reasonable geometry, with good root-mean-square (RMS) deviation from target values (Table 2; *44*).

The presence and absence of the metal-bound water was confirmed by the calculation of omit-maps. Overlay of the structures was done by using program EDPDB on the basis of least-squares fitting (*45*).

Electrostatic potentials were calculated and the molecular surfaces were displayed by using the program GRASP (*46*). The metal ions in both *E. coli* and *H. sapiens* GlxI proteins and the inhibitor ligand in the *H. sapiens* GlxI were excluded from these calculations.

Table 2: Refinement Statistics

protein[a]	Ni^{2+}	Co^{2+}	Cd^{2+}	Zn^{2+}	apo
resolution (Å)	20−1.5	20−1.9	20−1.9	20−1.8	20−1.7
R_{factor} (%)	20.5	16.9	17.7	18.6	18.8
R_{free} (%)	27.2	25.4	28.1	26.0	25.5
Δ_{bonds} (Å)[b]	0.01	0.005	0.007	0.007	0.006
Δ_{angles} (deg)[b]	1.987	1.439	1.580	1.603	1.483
average *B* value (Å^2)					
protein	19.5	22.0	18.2	18.5	23.5
metal	17.6	25.6	15.1	18.0	
overall	18.5	23.8	16.7	18.3	
no. of water molecules	266	263	273	235	288
residues in disallowed regions of Ramachandran plot (%)[c]	0.0	0.0	0.0	0.0	0.0

[a] Values are for the dimer in the asymmetric unit. [b] Discrepancies from ideal bond lengths and angles as defined by Engh and Huber (*44*). [c] Using a stringent boundary as defined by Kleywegt and Jones (*39*).

RESULTS

Overall Structure. The 135-amino acid native *E. coli* GlxI structure is composed of two identical subunits and has an overall fold similar to that of the 184-amino acid *H. sapiens* enzyme (*21*). Each subunit is itself made up of two domains (residues 3−60 and 72−126) that are linked by an intervening segment of 12 residues (Figure 1A). Each of these domains is composed of a $\beta\alpha\beta\beta\beta$ motif forming a mixed β-sheet (Figure 1A). The sequence of the strands within the sheet is $\beta_1\beta_4\beta_3\beta_2$. Within their respective "core" regions, the two domains have quite similar structures (e.g., the C^α atoms of residues 3−36 and 72−97 superimpose within 0.3 Å). The solvent-exposed loops connecting strands β_2 and β_3 (residues 36−38 and 104−107) and β_3 and β_4 (residues 47−51 and 115−119), however, have very different conformations (Figure 2A).

In the dimer, the β_1 strand of the first domain from subunit A associates in an antiparallel fashion with the β_1 strand from the second domain of subunit B. This leads to a curved 8-stranded β-sheet (Figure 1B). A reciprocal interaction

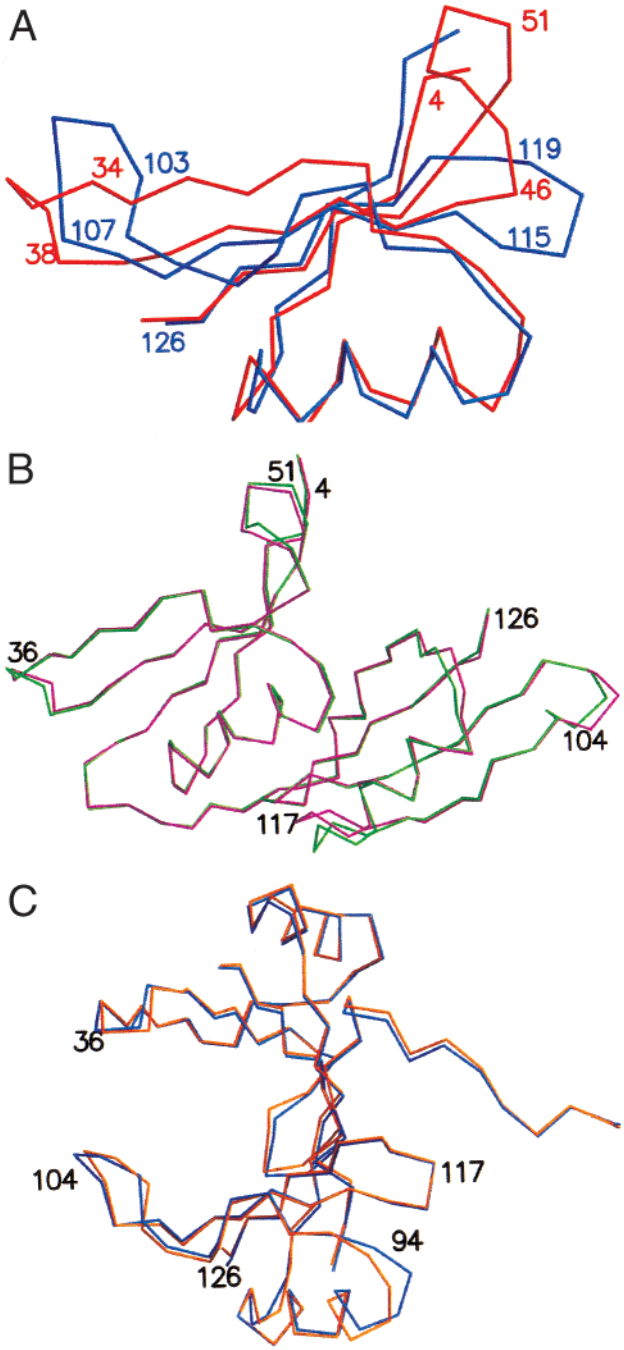

FIGURE 2: (A) Superposition of domain 1 (red) of the Ni^{2+}-GlxI structure on domain 2 (blue). (B) Superposition of monomer 1 (magenta) on monomer 2 (green). (C) Superposition of the respective domains that constitute the active site (see text).

between the second domain of subunit A and the first domain of subunit B leads to a second curved β-sheet. The two metal-binding sites per dimer are located within these curved β-sheets.

The two monomers are very close in structure (Figure 2B), their C^{α} atoms being superimposable within 0.47 Å. To compare the conformation in the vicinity of the two active sites, the first domain from the first monomer plus the second domain from the second monomer was overlaid on to the second domain from the first monomer plus the first domain from the second monomer (Figure 2C). In this case, the RMS deviation of the C^{α} atoms is only about 0.54 Å with the largest difference in the loop regions from residues Gln91−Gly94 and Val103−Thr108. The C^{α} atoms in these two loops have B factors ranging from 40 to 70 $Å^2$, significantly higher than the average value of 19.5 $Å^2$ for the protein as a whole (Figure 3). Higher backbone B factors were also found for the same two loops in SeMet-Ni^{2+}-GlxI, whose structure was solved in space group $P2_12_12_1$ (data not shown). The high B factors of the two loops in both crystal packing environments indicate their flexibility. The B factors of residues 91−94 in subunit A are about 25 $Å^2$ higher than those in subunit B for all the GlxI structures solved in space group $P2_1$. This is consistent with the observation that residues 91−94 in subunit B are involved in crystal contacts, while the same residues in subunit A are exposed to the solvent. Although the loop formed by residues 103−108, which is "above" the active site, is very mobile, the backbone in the vicinity of the metal-binding residues is among the most rigid part of the entire protein (Figure 3).

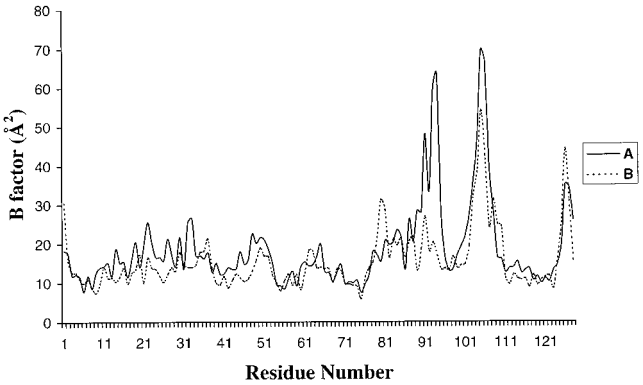

FIGURE 3: B factors of the C^{α} atoms in subunit A (solid line) and subunit B (dotted line) of Ni^{2+}-GlxI.

Toward the carboxy-terminus of the molecule the electron density could be followed clearly until Glu126, at which point it became much weaker. It could, however, still be followed until it entered the active site of an adjacent molecule in the crystal. At this point, the density became much stronger and could be modeled as Gly134−Asn135, the two residues at the carboxy-terminus of the protein. The intervening electron density between Glu126 and Gly134 is long enough to account for the intervening seven residues, but was not sufficiently well-defined to permit the inclusion of these residues in the final model.

Metal Coordination. Although biochemical data have suggested that only one metal ion per dimer is required to fully activate the protein (*19*), two such ions are seen at essentially equivalent positions (Figure 1B). In the native *E. coli* Ni^{2+}-GlxI structure, four protein residues (His5 and Glu56 from one monomer and His74 and Glu122 from the other) and two water molecules are coordinated to the nickel (Figure 4A). Very similar ligation is seen in the two active sites.

The coordination around the nickel ion is octahedral with close to ideal geometry (Figure 5, Tables 3 and 4). The metal lies in the plane defined by N^{ϵ} of the two histidine residues and the two water molecules. Monodentate coordination is seen for each of the two axial carboxylates. Substitution of different metal ions in the active site of GlxI does not change the overall structure of the protein. The Co^{2+}, Cd^{2+}, and Zn^{2+} enzymes have a C^{α} RMS deviation not exceeding 0.3 Å when compared to Ni^{2+}-GlxI. The changes that do occur are

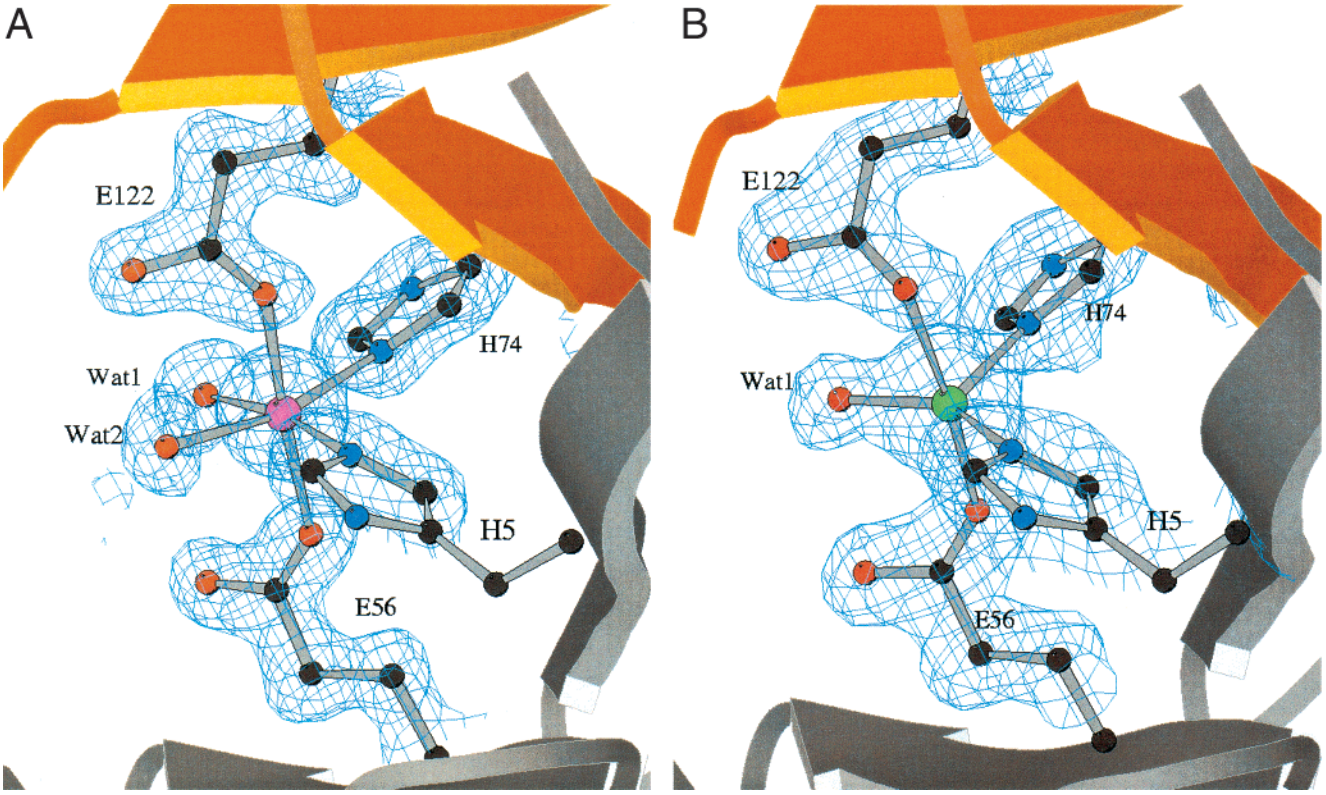

FIGURE 4: Active sites of (A) Ni^{2+}- and (B) Zn^{2+}-GlxI. The density shown is from a map calculated with amplitudes ($2F_o-F_c$) where F_o and F_c are, respectively, the observed structure amplitudes and those calculated from the final refined structure. The resolution is 1.5 Å for Ni^{2+}-GlxI and 1.8 Å for Zn^{2+}-GlxI. The octahedral arrangement of the nickel ligands and the trigonal bipyramidal coordination of the zinc ligands can be clearly seen. No distances or angle restrains were applied to the coordination geometry during refinement. Figures were prepared using Bobscript (*59*).

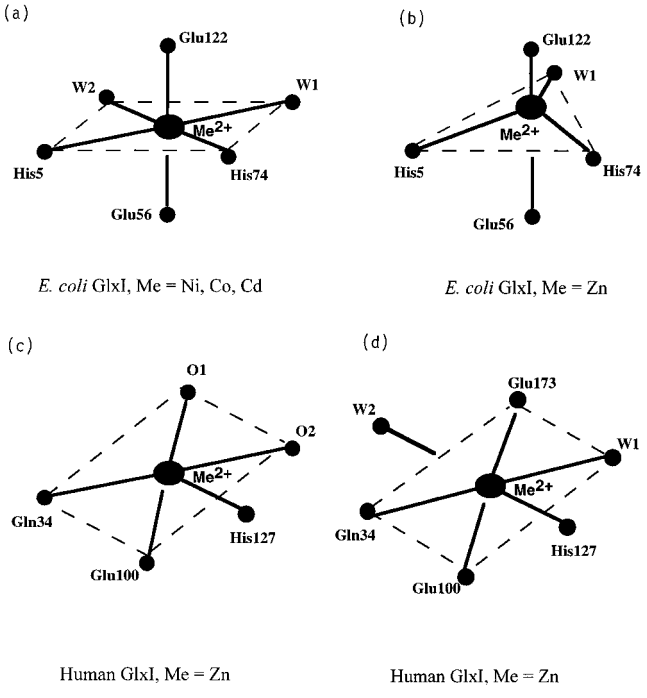

FIGURE 5: Metal coordination in different GlxI structures. (a) Complexes of *E. coli* GlxI with Ni^{2+}, Co^{2+}, and Cd^{2+}. The two water ligands are labeled W1 and W2. (b) Complex of *E. coli* GlxI with Zn^{2+}. (c) Complex of human GlxI with *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione. The two oxygen ligands from the hydroxamate inhibitor are labeled O1 and O2. (d) Complex of human GlxI with *S*-*p*-nitrobenzyloxycarbonylglutathione (coordinates taken from ref *51*).

Table 3: Coordination Distances for Metal Ligands in *E. coli* Glyoxalase I Complexes

	Ligand distance (Å)							
	Ni^{2+}		Co^{2+}		Cd^{2+}		Zn^{2+}	
ligand	site 1	site 2	site 1	site 2	site 1	site 2	site 1	site 2
His5 N^{ϵ}	2.2	2.1	2.3	2.2	2.4	2.3	2.2	2.1
W1	2.1	2.1	2.3	2.2	2.4	2.4	2.0	2.1
His74 N^{ϵ}	2.3	2.3	2.4	2.3	2.6	2.5	2.2	2.2
Glu56 O^{ϵ}	2.1	2.1	2.1	2.1	2.4	2.3	2.1	2.1
Glu122 O^{ϵ}	2.1	2.1	2.2	2.4	2.8	2.4	2.1	2.4
W2	2.2	2.2	2.4	2.4	2.4	2.6		

localized to the immediate vicinity of the metal sites. As with the native nickel enzyme, the enzymatically active Co^{2+} and Cd^{2+} derivatives also exhibit octahedral coordination. The coordination distances are somewhat longer (Table 3), consistent with the increased radii of Co^{2+} and Cd^{2+} as compared to Ni^{2+} (*47*). In contrast, the inactive Zn^{2+} derivative displays trigonal bipyramidal coordination with the same four protein ligands plus a single water molecule (Figures 4B and 5B). The geometries of the zinc ions in the two active sites are slightly different. In one case, the zinc-bound water superimposes on one of the two water ligands in the nickel complex and has close-to-perfect trigonal bipyramidal geometry. In the other active site, the water does not superimpose on either of the water ligands of the Ni^{2+}-structure resulting in distorted trigonal bipyramidal coordination.

The overall structure of apo-GlxI is very similar to that of the metal-bound forms. The RMS discrepancy of the C^{α} atoms relative to the nickel enzyme is 0.32 Å. This is

Table 4: Coordination Geometry for Metal Ligands in *E. coli* Glyoxalase I Complexes

	bond angles, octahedral (deg)							bond angles, trigonal bipyramidal (deg)		
		Ni^{2+}		Co^{2+}		Cd^{2+}			Zn^{2+}	
atoms	ideal	site 1	site 2	site 1	site 2	site 1	site 2	ideal	site 1	site 2
O^{ϵ}122-Me-N^{ϵ}74	90	81	82	84	90	87	90	90	90	92
N^{ϵ}5-Me-O^{ϵ}122	90	99	95	91	79	79	95	90	88	88
O^{ϵ}56-Me-N^{ϵ}5	90	79	82	81	85	85	84	90	83	88
W1-Me-O^{ϵ}56	90	88	88	85	87	80	81	90	87	93
W1-Me-O^{ϵ}122	90	93	96	103	108	104	100	90	95	86
O^{ϵ}122-Me-O^{ϵ}56	180	179	176	172	163	164	179	180	169	173
W1-Me-N^{ϵ}5	180	163	165	161	163	161	161	120	143	131
N^{ϵ}74-Me-W1	90	91	92	90	91	92	89	120	101	112
N^{ϵ}74-Me-O^{ϵ}56	90	99	96	98	96	93	91	90	99	98
N^{ϵ}74-Me-N^{ϵ}5	90	100	99	105	104	102	102	120	116	116
W2-Me-N^{ϵ}5	90	92	93	93	90	93	95			
W2-Me-O^{ϵ}56	90	97	98	101	101	104	103			
W1-Me-W2	90	79	79	78	77	78	78			
W2-Me-O^{ϵ}122	90	83	84	79	77	79	76			
W2-Me-N^{ϵ}74	180	161	162	155	158	157	160			

consistent with the circular dichroism spectrum, which does not suggest any significant structural change upon addition of Ni^{2+} ions to the apoenzyme (*19*). There are some changes in the immediate vicinity of the metal sites. In the apoenzyme, the metal ligands are somewhat more loosely packed, and the two glutamates are directed away from the vacant sites.

DISCUSSION

On the basis of sequence similarity with the *H. sapiens*, *S. cerevisiae*, and *P. putida* GlxI enzymes, all of which require Zn^{2+} (*16*, *18*), *E. coli* GlxI might have been expected to be a zinc metalloenzyme. The *E. coli* enzyme does, in fact, bind Zn^{2+} (as shown here) but is inactive in the presence of this ion. It does, however, display activity with Ni^{2+}, Co^{2+}, Cd^{2+}, and Mn^{2+} (*19*).

Comparison with Human GlxI and Other Members of the βαββββ Superfamily. The primary sequence of *H. sapiens* GlxI contains 184 amino acids (*17*), the N-terminal Met being lost in the active enzyme. Subsequent references to *H. sapiens* GlxI residue numbers refer to the intact 184 amino acid sequence. Following the recent determination of the *H. sapiens* GlxI crystal structure (*21*), a novel protein *βαββββ* superfamily was identified (*22*). Despite very low sequence homology and no similarity in enzymatic function, *H. sapiens* GlxI has been shown to be structurally related to the bacterial bleomycin resistance protein (BRP) from *Streptoalloteichus hindustanus* (*48*) and 2,3-dihydroxybiphenyl 1,2-dioxygenase from *Burkholderia cepacia* (*49*). The fosfomycin resistance protein (FosA), a Mn^{2+}-dependent glutathione-*S*-transferase, has also been suggested to be a member of this superfamily (*50*). A detailed analysis of the structural similarities and postulated evolution of this family of enzymes has been given by Bergdoll et al. (*22*). With the exception of FosA, which has not been structurally characterized, all members of the *βαββββ* family have overall structures that are much more similar to that of *E. coli* GlxI than to that of human GlxI. Therefore, the determination of the structure of the *E. coli* enzyme provides a more representative member of the family as a whole.

As expected, the overall structure of *E. coli* GlxI is similar to that of its human counterpart (*21*). There are, however, several substantive differences. A 29-amino acid N-terminal arm which wraps around the adjacent subunit in the *H. sapiens* protein is absent in the *E. coli* enzyme as well as in other bacterial members of the *βαββββ* family. This extended arm may help to stabilize the dimeric form of the human enzyme, although its absence in the *E. coli* enzyme shows that such an arm is not essential for dimer formation. Since this arm lies mostly on the surface of the enzyme its absence is not expected to substantially perturb the rest of the structure.

Another difference is that *E. coli* GlxI has a significantly larger active site. In the human enzyme, the active site consists of a largely hydrophobic pocket of volume about 70 $Å^3$. In contrast, the active site in the *E. coli* enzyme consists of a deep solvent channel about 10−15 Å in diameter (Figure 6). This increase in solvent accessibility is predominantly due to the deletion of a 15-amino acid segment that would lie approximately between residues 51−52 in the bacterial protein. This deletion eliminates an α-helix and a loop that lies along the side of the active site in the human enzyme. As a result, the active site of *E. coli* GlxI may have a very different response to inhibitors and substrates. These are being explored by studies of mutant enzymes and inhibitor complexes. Two smaller regions (residues 107−111 and 120−123 in *H. sapiens* GlxI) not present in the *E. coli* enzyme correspond to small loops lying between the two *βαββββ* motifs. These differences are mirrored within the other known superfamily members and suggest the *E. coli* enzyme is more similar to the other known members of the superfamily than is the *H. sapiens* GlxI. The presumed evolution of this family of enzymes has been discussed by Bergdoll et al. (*22*).

The four C-terminal residues of *E. coli* GlxI (Gly132-Asn135) can be proteolytically removed without affecting catalytic activity (*19*). In the electron density map, the residues following Glu126 cannot be traced accurately due to weak electron density, with the exception of the final two residues (Gly134−Asn135), which bind within the active site of an adjacent protein molecule in the crystal. The mobility of the C-terminal region explains its susceptibility to proteolytic degradation. Analysis of the B factors of the *E. coli* GlxI structure (Figure 3) reveals two additional regions with higher than average mobility, namely, residues 91−94 and 104−106. The second region lies "above" the active site,

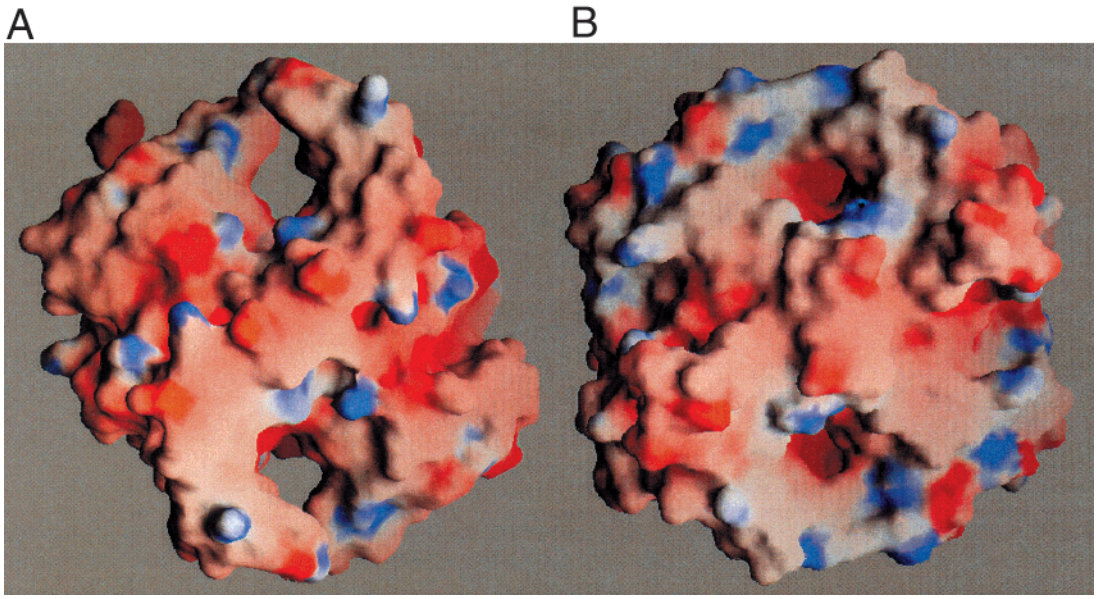

FIGURE 6: Molecular surfaces of the *E. coli* (A) and *H. sapiens* (B) GlxI proteins. Electrostatic potentials were calculated using GRASP (*46*) including all atoms in the refined model. Positive potential (>15 mV) is colored blue, neutral potential (0 mV) is colored gray, and negative potential (<−15 mV) is colored red. Molecules are shown in approximately the same alignment as in Figure 1B.

suggesting a possible involvement in substrate binding. There is some evidence for this in the recent structure determination of the *H. sapiens* GlxI with a proposed transition state analogue (*51*), which indicates that C^{α} of Lys157 moves 3 Å toward the active site. This residue is located within a loop that corresponds to Pro102−Val103−Lys104−Gly105−Gly106−Thr107−Thr108−Val109 in the *E. coli* enzyme. Thus, Lys104 in the *E. coli* enzyme may play a role similar to Lys157 in the human enzyme.

Metal Binding Sites. The structures presented here show that each of the active forms of *E. coli* GlxI has an octahedral metal coordination, with two participating water molecules. Electron paramagnetic resonance indicates that the Mn^{2+}-substituted enzyme, which is active, is also octahedrally coordinated (Markham, G. D., Clugston, S. L. and Honek, J. F., unpublished results). However, X-ray analysis of the enzyme cocrystallized or soaked in up to 5 mM Mn^{2+} showed weak electron density at the metal binding sites, suggesting that the affinity for this ion in the crystal is low. Direct comparison with the metal coordination of the active Zn^{2+} *H. sapiens* enzyme is difficult as the structure lacking an inhibitor has not yet been reported. However, the available inhibitor complexes indicate that Zn^{2+} is coordinated by four protein ligands (Gln34, Glu100, His127, Glu173) plus one or possibly two water molecules with square pyramidal or octahedral geometry. Liganding with the inhibitor substitutes for the binding of the second water molecule in some cases (*21*, *51*, *52*).

In the case of the *E. coli* enzyme, the two residues at the carboxy-terminus of one molecule bind within the active site of an adjacent molecule in the crystal. The location and mode of binding is similar to that of benzyl-GSH cocrystallized in the active site of the *H. sapiens* GlxI (Figure 7). In particular, O11 and N1 of benzyl-GSH form hydrogen bonds with $N^{\delta 2}$ and $O^{\delta 1}$ of Asn104 that mimic those between O and $N^{\delta 2}$ of Asn135 and Asn60 in the *E. coli* GlxI enzyme. The presence of these two amino acids in the active site does not, however, appear to contribute to or to directly influence the ligation at the metal site.

Even though human and *E. coli* GlxI have homologous amino acid sequences and similar three-dimensional structures, they do not have identical metal ligands. As shown in Figure 5, Gln34 of the human enzyme is replaced with a histidine in *E. coli* GlxI. In the absence of structural information, this variability might have suggested that this is a nonessential amino acid. Even with the knowledge of the respective structures, it might suggest that the presence of the glutamine would correlate with a requirement for zinc while the histidine ligand would correlate with the use of nickel. This, however, cannot be the case since the *S. cerevisiae* and *P. putida* GlxI enzymes are active with zinc yet use histidine as a metal ligand (*20*). The key factor responsible for activity, therefore, seems to be geometry.

Previous studies of the zinc endopeptidase astacin have revealed a relationship between the geometry around the metal center and enzymatic activity of various metal-substituted forms of the enzyme (*53*). Trigonal bipyramidal ligation was observed for each of the three active forms of the enzyme, Zn^{2+}, Cu^{2+}, and Co^{2+}, whereas octahedral coordination occurred for the Ni^{2+} enzyme and tetrahedral for Hg^{2+}-astacin, both of which were inactive. The results suggested that the correct orientation of metal-liganded water molecules was necessary for catalytic activity. In the case of asticin, Zn^{2+} and Co^{2+} give the appropriate orientation, while Ni^{2+} does not. The structural analysis of *E. coli* GlxI displays a similar correlation between geometry and activity except that in this case Ni^{2+}, Co^{2+}, and Cd^{2+} appear to have the correct orientation of bound water molecules, while Zn^{2+} does not.

Reaction Mechanism. On the basis of solvent isotope exchange studies on the *S. cerevisiae* enzyme, the overall mechanism of GlxI has been proposed to involve proton abstraction, formation of an enediol(ate), with reprotonation at the adjacent carbon (*54*). A specific mechanism has been proposed based on the recent crystallographic studies of *H. sapiens* GlxI in the presence of a putative transition state analogue (*51*). In this mechanism, it is proposed that the six-coordinate zinc, including two water ligands, becomes five-

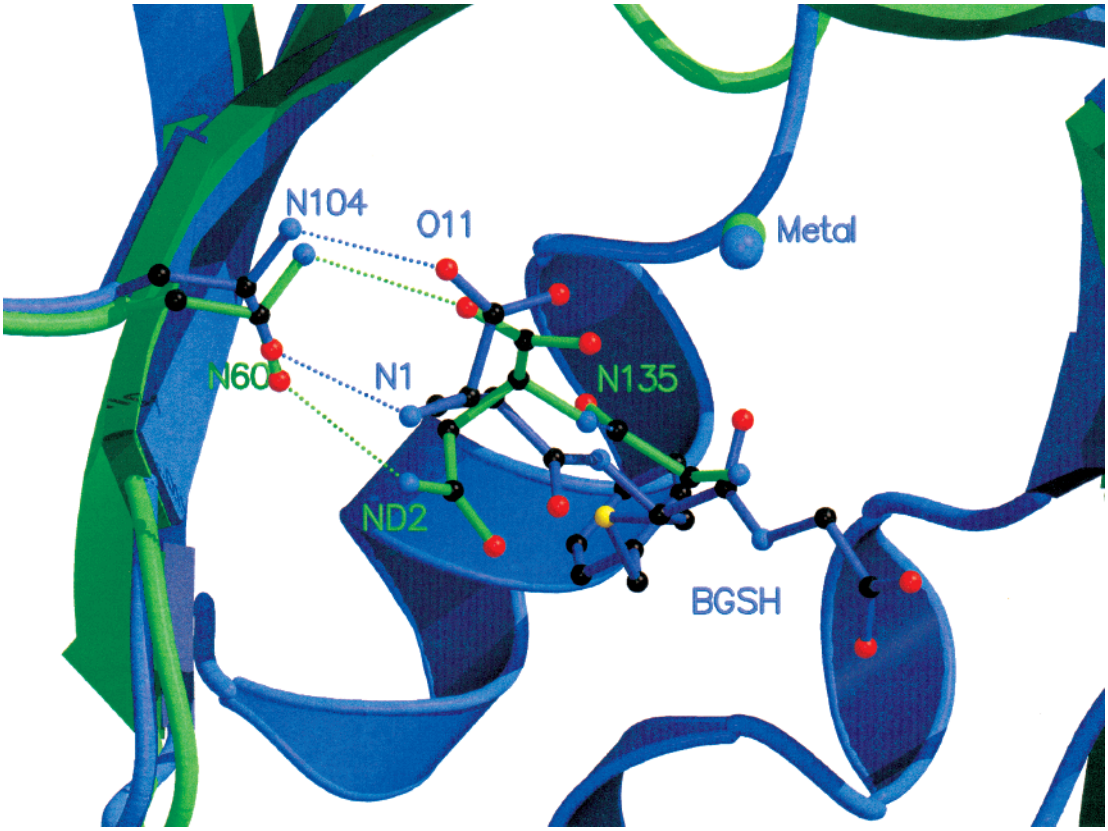

FIGURE 7: Superposition of the active site of *E. coli* GlxI (green) on that of *H. sapiens* GlxI (blue). Benzyl-GSH, which is bound to the human enzyme, is shown in blue with its hydrogen bonds to Asn shown as dotted lines. Gly134 and Asn135, which bind within the active site of *E. coli* enzyme and make analogous hydrogen bonds to Asn60, are shown in green.

coordinate upon substrate binding. The two oxygen atoms of the substrate (seen with the enediol inhibitor) replace the two water ligands. The binding of the substrate is also proposed to displace the zinc ligand Glu173, which is then free to act as the catalytic base. On the basis of the present knowledge of the crystal structure of the *E. coli* enzyme, a related mechanism could be proposed, i.e., it could be envisaged that Glu122 in the *E. coli* enzyme, the counterpart of Glu173, would be displaced from the metal on substrate binding and act as the catalytic base. There are, however, a number of uncertainties. For example, while an incoming substrate might be envisaged to displace a glutamate from zinc, it would be less likely to do so for nickel (*53*).

In principle, the following mechanisms might be considered for *E. coli* GlxI, based on our current structural knowledge: (i) The substrate (or transition state) might displace a single water molecule plus Glu122, such that a six-coordinate metal is still present and allows for the production of the general catalytic base, Glu122, analogous to that proposed for the *H. sapiens* enzyme. (ii) On the basis of the octahedral geometry for all the activating metals for *E. coli*, a mechanism that utilizes the two water molecules to polarize and possibly serve in the proton abstraction could be proposed. Previous work on the *H. sapiens* enzyme has indicated the importance of two water molecules directly interacting with the substrate (*55*, *56*). (iii) Both water molecules might be displaced from the metal and an additional unspecified specific or general base might participate. However, the only putative general base present in the active site is Glu49, which is approximately 8.0 Å from the metal center. In the absence of fairly substantial structural changes, the involvement of Glu49 would seem doubtful.

The most remarkable feature of the present study is the correlation between catalytic activity and the coordination at the metal binding site. Each of the protein−metal complexes that has octahedral (6-fold) geometry yields an active enzyme. This includes the complexes of the *E. coli* enzyme with Ni^{2+}, Co^{2+}, and Cd^{2+}, as well as the structures reported for the human Zn^{2+} enzyme. Conversely, the complex of the *E. coli* enzyme with Zn^{2+} has trigonal bipyramidal (5-fold) coordination and is inactive. It would seem clear that octahedral geometry is a prerequisite for activity. Why is this the case? Two possible reasons can be suggested. First, it may be that a rather special geometry is required to have a glutamic acid (e.g., Glu173 in the human enzyme) act as both a metal ligand and as a catalytic base (assuming it does, in fact, fulfill both roles). Second, it may be that the mechanism of action requires that two water molecules be bound to the metal ion either in the resting enzyme or at some time during catalysis. Six-fold coordination would allow four protein ligands plus two such water molecules, but this would not be the case for 5-fold metal ligation (Figure 5).

ACKNOWLEDGMENT

M. He thanks Drs. Martin Sagermann, Blaine Mooers, and Mike Quillin for help with all aspects of the crystallography and Dr. Allen Orville for help with the figures. We are also grateful to Dr. Elisabeth Daub for her advice on the SeMet incorporation.

REFERENCES

1. Michl, J., Ed. (1996) Bioinorganic Enzymology (Holm, R. H., Solomon, E. I., Eds.) Chemical Reviews, Vol. 96, American Chemical Society, Washington, DC.
2. DeGrado, W. F., and Summa, C. M. (1999) *Annu. Rev. Biochem. 68*, 779−819.
3. Hunt, J. A., Ahmed, M., and Fierke, C. A. (1999) *Biochemistry 38*, 9054−9062.
4. Karlin, S., Zhu, Z.-Y., and Karlin, K. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 14225−14230.
5. Rulísek, L., and Vondrásek, J. (1998) *J. Inorg. Biochem. 71*, 115−127.
6. Vallee, B. L., and Auld, D. S. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 220−224.
7. Volbeda, A., Fontecilla-Camps, J. C., and Frey, M. (1996) *Curr. Opin. Struct. Biol. 6*, 804−812.
8. Wisz, M. S., Garrett, C. Z., and Hellinga, H. W. (1998) *Biochemistry 37*, 8269−8277.
9. Vander Jagt, D. L. (1989) In *Coenzymes and Cofactors VIII: Glutathione Part A* (Dolphin, D., Poulson, R., Avramovic, O., Eds.), pp 597−641, John Wiley and Sons Inc., New York.
10. Ferguson, G. P., Tötemeyer, S., MacLean, M. J., and Booth, I. R. (1998) *Arch. Microbiol. 170*, 209−219.
11. Inoue, Y., and Kimura, A. (1995) In *Advances in Microbial Physiology* (Poole, R. K., Ed.), Vol. 37, pp 177−227, Academic Press Ltd., New York.
12. Oya, T., Hattori, N., Mizuno, Y., Miyata, S., Maeda, S., Osawa, T., and Uchida, K. (1999) *J. Biol. Chem. 274*, 18492−18502.
13. Thornalley, P. J. (1998) *Chem.-Biol. Interact. 111−112*, 137−151.
14. Barnard, J. F., Vander Jagt, D. L., and Honek, J. F. (1994) *Biochim. Biophys. Acta 1208*, 127−135.
15. Kavarana, M. J., Kovaleva, E. G., Creighton, D. J., Wollman, M. B., and Eiseman, J. L. (1999) *J. Med. Chem. 42*, 221−228.
16. Aronsson, A. C., Marmstål, E., and Mannervik, B. (1978) *Biochem. Biophys. Res. Commun. 81*, 1235−1240.
17. Ridderström, M., and Mannervik, B. (1996) *Biochem. J. 314*, 463−467.
18. Saint-Jean, A. P., Phillips, K. R., Creighton, D. J., and Stone, M. J. (1998) *Biochemistry 37*, 10345−10353.
19. Clugston, S. L., Barnard, J. F. J., Kinach, R., Miedema, D., Ruman, R., Daub, E., and Honek, J. F. (1998) *Biochemistry 37*, 8754−8763.
20. Clugston, S. L., Daub, E., Kinach, R., Miedema, D., Barnard, J. F. J., and Honek, J. F. (1997) *Gene 186*, 103−111.
21. Cameron, A. D., Olin, B., Ridderström, M., Mannervik, B., and Jones, T. A. (1997) *EMBO J. 16*, 3386−3395.
22. Bergdoll, M., Eltis, L. D., Cameron, A. D., Dumas, P., and Bolin, J. T. (1998) *Protein Sci. 7*, 1661−1670.
23. Hausinger, R. P. (1987) *Microbiol. Rev. 51*, 22−42.
24. Maroney, M. J. (1999) *Curr. Opin. Chem. Biol. 3*, 188−199.
25. Watt, R. K., and Ludden, P. W. (1999) *Cell. Mol. Life Sci. 56*, 604−625.
26. Davis, R. W., Botstein, D., and Roth, J. R. (1980) In *Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
27. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
28. Harp, J. M., Timm, D. E., and Bunick, G. J. (1998) *Acta Crystallogr. D54*, 622−628.
29. Harp, J. M., Hanson, B. L., Timm, D. E., and Bunick, G. J. (1999) *Acta Crystallogr. D55*, 1329−1334.
30. Leslie, A. G. W. (1999) Mosflm 6.0 Users' manual 6.0 edit. MRC Laboratory of Molecular Biology, Cambridge.
31. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
32. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
33. Hauptman, H. A. (1997) *Methods Enzymol. 277*, 3−13.
34. de La Fortelle, E., and Bricogne, G. (1997) *Methods Enzymol. 276*, 472−494.
35. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. O. (1991) *Acta Crystallogr. A47*, 110−119.
36. Jones, T. A., and Kjeldgaard, M. O. (1997) *Methods Enzymol. 277*, 173−208.
37. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
38. Brunger, A. T., Krukowski, A., and Erickson, J. W. (1990) *Acta Crystallogr. A46*, 585−593.
39. Kleywegt, G., and Jones, T. (1994) In *From First Map to Final Model* (Bailey, S., Hubbard, R., Waller, D., Eds.), pp 59−66, EPSRC, Daresbury Laboratory, Daresbury, UK.
40. Tronrud, D. E. (1997) *Methods Enzymol. 277*, 306−319.
41. Brunger, A. T. (1992) *Nature 355*, 472−475.
42. Kissinger, C. R., and Gehlhaar, D. K. (1997) EPMR Users' manual. Agouron Pharmaceuticals, Inc., San Diego.
43. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
44. Engh, R. A., and Huber, R. (1992) *Acta Crystallogr. A47*, 392−400.
45. Zhang, X. J., and Matthews, B. M. (1995) *J. Appl. Crystallogr. 28*, 624−630.
46. Nicholls, A., and Honig, B. (1991) *J. Comput. Chem 12*, 435−445.
47. Pauling, L. (1960) *The Nature of the Chemical Bond*, 3rd ed., Cornell University Press, Ithaca, NY.
48. Dumas, P., Bergdoll, M., Cagnon, C., and Masson, J.-M. (1994) *EMBO J. 13*, 2483−2492.
49. Han, S., Eltis, L. D., Timmis, K. N., Muchmore, S. W., and Bolin, J. T. (1995) *Science 270*, 976−980.
50. Bernat, B. A., Laughlin, L. T., and Armstrong, R. N. (1997) *Biochemistry 36*, 3050−3055.
51. Cameron, A. D., Ridderström, M., Olin, B., Kavarana, M. J., Creighton, D. J., and Mannervik, B. (1999) *Biochemistry 38*, 13480−13490.
52. Ridderström, M., Cameron, A. D., Jones, T. A., and Mannervik, B. (1998) *J. Biol. Chem. 273*, 21623−21628.
53. Gomis-Rüth, F.-X., Grams, F., Yiallouros, I., Nar, H., Küsthardt, U., Zwilling, R., Bode, W., and Stöcker, W. (1994) *J. Biol. Chem. 269*, 17111−17117.
54. Hall, S. S., Doweyko, A. M., and Jordan, F. (1976) *J. Am. Chem. Soc. 98*, 7460−7461.
55. Sellin, S., Eriksson, L. E. G., and Mannervik, B. (1982) *Biochemistry 21*, 4850−4857.
56. Sellin, S., Rosevear, P. R., Mannervik, B., and Mildvan, A. S. (1982) *J. Biol. Chem. 257*, 10023−10029.
57. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
58. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
59. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−134.

BI000856G