Crystal Structure of Maleylacetoacetate Isomerase/Glutathione Transferase Zeta Reveals the Molecular Basis for Its Remarkable Catalytic Promiscuity[†,‡]

Galina Polekhina,[§] Philip G. Board,[||] Anneke C. Blackburn,[||] and Michael W. Parker*[,§]

The Biota Structural Biology Laboratory, St. Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia, and Molecular Genetics Group, Division of Molecular Medicine, John Curtin School of Medical Research, Australian National University, GPO Box 334, Canberra 2601, Australia

Received September 25, 2000; Revised Manuscript Received November 8, 2000

ABSTRACT: Maleylacetoacetate isomerase (MAAI), a key enzyme in the metabolic degradation of phenylalanine and tyrosine, catalyzes the glutathione-dependent isomerization of maleylacetoacetate to fumarylacetoacetate. Deficiencies in enzymes along the degradation pathway lead to serious diseases including phenylketonuria, alkaptonuria, and the fatal disease, hereditary tyrosinemia type I. The structure of MAAI might prove useful in the design of inhibitors that could be used in the clinical management of the latter disease. Here we report the crystal structure of human MAAI at 1.9 Å resolution in complex with glutathione and a sulfate ion which mimics substrate binding. The enzyme has previously been shown to belong to the zeta class of the glutathione *S*-transferase (GST) superfamily based on limited sequence similarity. The structure of MAAI shows that it does adopt the GST canonical fold but with a number of functionally important differences. The structure provides insights into the molecular bases of the remarkable array of different reactions the enzyme is capable of performing including isomerization, oxygenation, dehalogenation, peroxidation, and transferase activity.

The glutathione dependent cis−trans isomerization of maleylacetoacetate (MAA)[1] to fumarylacetoacetate (FAA) (Figure 1) is the penultimate step in the catabolism of phenylalanine and tyrosine. Deficiencies of most enzymes in this pathway lead to a range of disorders including alkaptonuria, phenylketonuria, and several forms of tyrosinemia (*1*). Tyrosinemia type I is the most severe of these disorders and results from deficiency of fumarylacetoacetate hydrolase (FAH), the enzyme in the last step of the pathway (*2*). The accumulation of FAA that occurs in FAH deficiency is thought to cause the liver disease and the high rates of hepatocarcinoma that occur in tyrosinemia type I patients (*3*). While it is thought that MAA may also be toxic, confirmed cases of maleylacetoacetate isomerase (MAAI) deficiency have never been reported. There is one report in an abstract of a patient with a very severe tyrosinemia disorder who had very low MAAI activity (*4*). MAAI activity was also low in the patient's parents, suggesting autosomal recessive inheritance of the condition. Unfortunately, this patient and another with the same disorder passed away

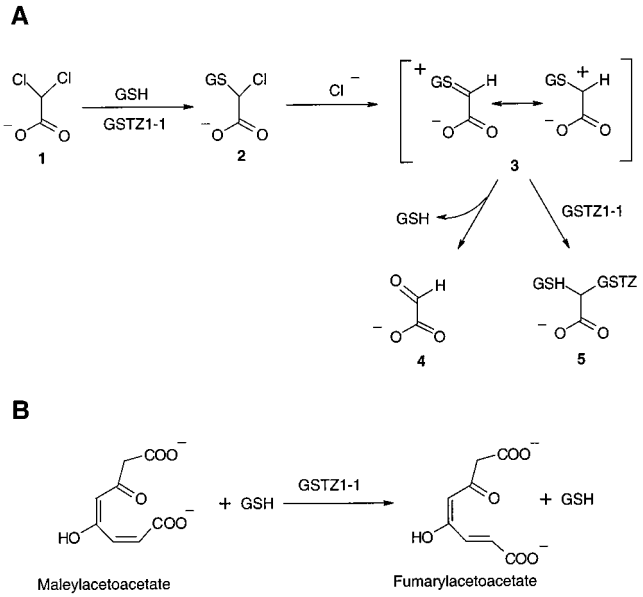

FIGURE 1: Reactions catalyzed by MAAI/GST Z1-1: (a) oxygenation of DCA and (b) isomerization of MAA. In panel a, the numbers refer to dichloroacetic acid (1), *S*-(α-chlorocarboxymethyl) glutathione (2), carbonium-sulfonium intermediate (3), glyoxylate (4), and inactivated MAAI/GST Z1-1 complex (5).

[†] This work was supported by a grant from the Australian Research Council. M.W.P. is an Australian Research Council Senior Research Fellow. This work was also supported by the Australian Synchrotron Research Program, which is funded by the Commonwealth of Australia under the Major National Research Facilities program. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38.

[‡] The crystallographic coordinates have been deposited in the Protein Data Bank under the filename 1FW1.

* To whom correspondence should be addressed. Phone: (0061) 3 9288-2480. Fax: (0061) 3 9416-2676. E-mail: mwp@rubens.its.unimelb.edu.au.

[§] St. Vincent's Institute of Medical Research.

[||] Australian National University.

[1] Abbreviations: APS, Advanced Photon Source; DCA, dichloroacetic acid; DTT, 1,4-dithiothreitol; GSH, glutathione; GST, glutathione *S*-transferase; GST Z1-1, zeta class glutathione *S*-transferase; FAA, fumarylacetoacetate; FAH, fumarylacetoacetate hydrolase; HEPES, (*N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]; HPLC, high-performance liquid chromatography; LinD, lindane degrading reductive dehalogenase; MA, maleylacetone; MAA, maleylacetoacetate; MAAI, maleylacetoacetate isomerase; MAD, multiple anomalous dispersion; SIR, single isomorphous replacement; TCHQ, tetrachlorohydroquinone; TCHQ-D, tetrachlorohydroquinone dehalogenase.

10.1021/bi002249z CCC: $20.00
Published on Web 01/13/2001

shortly after discovery. Inhibitors of MAAI could reduce the formation of FAA and may be useful in the clinical management of hereditary tyrosinemia type I patients (*5*).

Because of technical difficulties in determining its activity, MAAI (EC 5.2.1.2) has been one of the least studied of all the enzymes involved in the catabolism of tyrosine. While most enzymes involved in this pathway are expressed predominantly in the liver, the observation that MAAI is expressed in most tissues led to the suggestion that it may play an additional role in cellular metabolism (*6*). The recent cloning of a cDNA encoding human MAAI revealed that it was identical to the previously described zeta class glutathione transferase, GST Z1-1 (*1*, *7*). The zeta class GSTs are known to be responsible for the metabolism of α-haloacids including dichloroacetic acid (DCA) (Figure 1), a carcinogen in rodents, that is a common contaminant of chlorinated drinking water (*8*−*12*). DCA is also used therapeutically for the treatment of congenital lactic acidosis and homozygous familial hypercholesterolemia (*13*). While GST Z1-1 is thought to be the primary enzyme responsible for the metabolism of DCA, recent studies have shown that GST Z1-1 can be inactivated by DCA (Figure 1) (*14*, *15*).

Little is known about the active-site residues involved in the reactions catalyzed by MAAI/GST Z1-1. The enzyme tetrachlorohydroquinone dehalogenase (TCHQ-D) from *Sphingomonas chlorophenolica* appears to be structurally related to MAAI/GST Z1-1 and catalyses the isomerization of MAA in addition to the dehalogenation of TCHQ (*16*). A conserved N-terminal region cysteine residue appears to be important in both the isomerase and dehalogenase reactions. This implies that a similar residue in MAAI and the zeta class GSTs could be involved in catalysis. A greater understanding of the isomerase and dehalogenase reactions and the mechanism-based inactivation of GST Z1-1 would be gained from knowledge of the structure. In the present study, we have expressed human MAAI/GST Z1-1 and determined its crystal structure at 1.9 Å resolution.

EXPERIMENTAL PROCEDURES

Crystallization and Data Collection. MAAI (polymorphic variant GST Z1a-1a) was expressed and purified as described previously (*15*). The protein was dialyzed against 20 mM HEPES, pH 7.5, 2 mM DTT, and 2 mM GSH and concentrated to about 10 mg mL^{-1} prior to the crystallization. The protein was mixed with an equivolume ratio of a solution containing 100 mM Bicine pH 9.5, 10 mM DTT, 10 mM GSH, 48−52% of saturated ammonium sulfate and 1% *tert*-butyl alcohol and equilibrated over the same solution using the hanging drop vapor diffusion technique at 22 °C. The crystals grew as thin plates up to a size of approximately 0.2 mm × 0.2 mm × 0.05 mm after about one month. They belong to the space group $P42_12$ with cell dimensions $a = b = 100.0$ Å, $c = 56.9$ Å and contain one monomer per asymmetric unit with a solvent content of approximately 50%.

Three different native data sets were used in the process of determining the structure (Table 1). The "native in-house" data were collected from a single frozen crystal at 100 K using a MARResearch imaging (345 mm diameter) plate detector with CuKα X-rays generated by a Rigaku RU-200 rotating anode generator. Glycerol was utilized as a cryoprotectant: 20% glycerol was added to a stabilizing solution of 100 mM bicine, pH 9.5, 52% saturated ammonium sulfate. The "native LNLS" data were collected from a single crystal at room temperature on a MARResearch imaging plate detector using synchrotron radiation at the National Synchrotron Light Laboratory (LNLS), Brazil. The "native APS" data were collected from a single crystal frozen to 100 K on an ADSC Quantum-4 CCD detector using synchrotron radiation at the BioCARS beamline 14-BM-C, Advanced Photon Source, Chicago. It was noted that all crystals displayed strongly anisotropic diffraction patterns that extended along the a^* and b^* axial directions to beyond 2 Å resolution but along c^* to only 2.8 Å resolution. The two heavy atom derivatives, *cis*-$Pt(NH_3)_2Cl_2$ and Pt(ethylenediamine)Cl_2, proved only slightly soluble in artificial mother liquor and hence were prepared as saturated solutions. A data set for the former was collected in-house off a single-crystal frozen to 100 K whereas multiple wavelength anomalous dispersion (MAD) data were collected from the latter derivative on the BioCARS beamline, 14-BM-D, at Chicago. All data were processed using the HKL package (*17*). The relevant data collection statistics are presented in Table 1.

Structure Determination, Model Building, and Refinement. The structure of MAAI was determined by a combination of single isomorphous replacement (SIR) and MAD data. We first attempted to determine the structure by molecular replacement. The native LNLS data set was used in the molecular replacement calculations. A successful solution was obtained using the program MOLREP (*18*) with a search probe consisting of a polyalanine model of the *Lucilia cuprina* GST monomer (*19*) with the variable loops omitted. The solution corresponded to the highest peak in the rotation function (3.41σ with the next highest peak of 3.33σ) and the highest peak in the translation function (5.95σ with the next highest peak of 5.53σ). This solution not only yielded a molecule that packed well within the unit cell but also generated GST dimers, formed through crystallographic 2-fold axes, which closely resembled the quaternary structures previously observed in other GST structures (*20*). However, we were unable to refine the molecular replacement solution successfully. In hindsight, it would appear that the search probe was not sufficiently similar to the structure of MAAI for this approach to yield a high-resolution structure.

The location of heavy atoms sites for the platinum derivatives were determined from Patterson maps. Phases derived from the heavy atom derivatives confirmed the correctness of the molecular replacement solution but were only good enough to build a partial model. Combining the heavy atom phases with the phases derived from the molecular replacement solution did not help the model building efforts. We therefore performed MAD experiments on one of the platinum derivatives (Table 1). MAD phases were calculated using MLPHARE (*21*), treating data at λ_1 as a native data set and at λ_2 and λ_3 as derivatives. The location of platinum sites was determined from the MAD data independently of the SIR data. Both data sets yielded the same single site solution (Table 1). SIGMAA (*21*) was used to combine the SIR and MAD phases and solvent flattening was performed using DM (*21*). The figure-of-merit after density modification was 0.54 for data between 15 and 2.2 Å resolution. Model building into the resultant electron

Table 1: Structure Determination Statistics[a]

Diffraction Data Statistics

data set	resolution (Å)	completeness (%)	I/σ in the last resolution shell	R_{merge}[b] (%)	multiplicity
native LNLS	15−2.3	93.0	3.9	13.7 (44.0)	6.2
native in-house	20−2.6	97.2	4.5	10.4 (25.4)	5.4
cis-$Pt(NH_3)_2Cl_2$	15−2.45	92.8	2.5	13.0 (45.4)	4.0
native APS	30−1.9	98.1	7.4	4.9 (28.6)	6.7

Phasing Statistics − SIR

derivative	soaking time	binding sites	resolution (Å)	R_{iso}[c] (%)	FOM[d] acentric	FOM[d] centric	phasing power[e] acentric	phasing power[e] centric	R_{cullis}[f] acentric	R_{cullis}[f] centric
cis-$Pt(NH_3)Cl_2$	overnight	Met 126	15−3.5	23.4	0.21	0.28	1.53	1.19	0.71	0.63

MAD Data Collection

data set		wavelength (Å)	resolution (Å)	completeness (%)	multiplicity	R_{merge}[b] (%)
Pt-(ethylenediamine)Cl_2	λ_1	1.0718	30−2.2	98.9	8.1	4.9 (14.6)
Pt-(ethylenediamine)Cl_2	λ_2	1.0713	30−2.2	100.0	8.2	5.1 (15.0)
Pt-(ethylenediamine)Cl_2	λ_3	1.0419	30−2.2	98.6	4.2	4.3 (13.4)

Phasing Statistics − MAD as MIR (E_1−native)

derivative	soaking time	binding site	resolution (Å)	$resolution_{ano}$ (Å)	phasing power acentric	phasing power centric	R_{cullis}[f] acentric	R_{cullis}[f] centric	FOM[d] acentric	FOM[d] centric	R_{cullis_ano}[f]
λ_2	overnight	Met 126	30−4.0	30−5.0	0.53	0.33	0.96	0.94	0.31	0.28	0.71
λ_3			30−4.0	30−5.0	0.39	0.25	0.97	0.94			0.77

Final Refinement Statistics

data set	native APS
resolution	30−1.9 Å
R-factor[g]/*R*-free[h]	22.8/27.7%
rmsd bond lengths	0.02 Å
rmsd bond angles	2.0°
rmsd dihedral angles	22.7°
rmsd bonded B's (m−c)	0.95 Å^2
rmsd bonded B's (s−c)	1.60 Å^2

[a] Native LNLS are the native data collected at the National Synchrotron Laboratory, Brazil. Native in-house are the data collected in-house on a Rikagu RU-200 X-ray generator (see text). Native APS are the data collected at the BioCARS facility, Advanced Photon Source, Chicago. Values given in parentheses are for the highest resolution shell (approximately 0.1 Å in width). [b] $R_{merge} = \sum(I - \langle I \rangle)/\sum(I)$, where I is the intensity measurement for a given reflection, $\langle I \rangle$ is the average intensity for multiple measurements of this reflection. [c] $R_{iso} = \sum||F_{PH}| - |F_P||/\sum|F_P|$, where F_{PH} and F_P are the derivative and native structure factor amplitudes, respectively. [d] FOM is the figure-of-merit. [e] Phasing power = root-mean-square of ($|F_H|/E$), where F_H is the calculated heavy-atom structure factor amplitude and E is the residual lack-of-closure error. [f] $R_{cullis} = \sum|||F_{PH}| \pm |F_P|| - |F_H||/\sum||F_{PH}| \pm |F_P||$. [g] *R*-factor $= \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. [h] R_{free} was calculated with 5% of the diffraction data that were selected randomly and not used throughout refinement.

density map was initiated from the partial model derived from the SIR phases. This model included most of the N-terminal domain and the first helix of the C-terminal domain which constituted about 50% of the total scattering matter in the asymmetric unit. Refinement was performed with REFMAC (*22*) making use of the phases from the density modification procedure. Once the C-terminal domain was built the refinement quickly converged. The last few cycles of molecular dynamics refinement were performed with CNS (*23*) and included a bulk solvent correction and restrained refinement of individual *B*-factors. (The R_{free} reference data set was carried over from the REFMAC refinement.)

The final model of MAAI/GST Z1-1 consists of residues 5−212, 1 molecule of GSH, 1 sulfate ion, 1 DTT molecule, and 109 water molecules. The refinement statistics are given in Table 1. The final model is of good quality with 93.9% of residues in the most favorable regions of the Ramachandran plot (*24*). No residues fall in the disallowed regions and only one falls in a generously allowed region: Gln 71 which forms contacts with GSH and is located in a similar region of the Ramachandran plot in other GST structures. The quality of the model is supported by the excellent electron density through out the molecule.

RESULTS

Overall Structure. The molecule adopts the canonical GST fold (Figure 2). There are two domains, an N-terminal thioredoxin-like domain and a C-terminal all α-helical domain. The domains are connected by a linker region between residues 85 and 91. The side chain of Leu 88 from the linker region is wedged between the two domains. The N-terminal domain is composed of a central four-stranded mixed β-sheet (strands are residues 7−10, 32−35, 61−64, and 67−70), which is sandwiched on both sides by alpha-helices. On the solvent exposed side there is helix α2 (residues 47−54) together with some irregular structure, including a large loop comprised of a series of β-turns (residues 38−44) and a turn of 3_{10} helix, connecting helix α2 to the beta-sheet core. Helix α2 and its surrounding structure represent the region of highest temperature factors (between 50 Å^2 and 75 Å^2) in the molecule. An aromatic residue, Phe 50, from helix α2 nestles into a hydrophobic

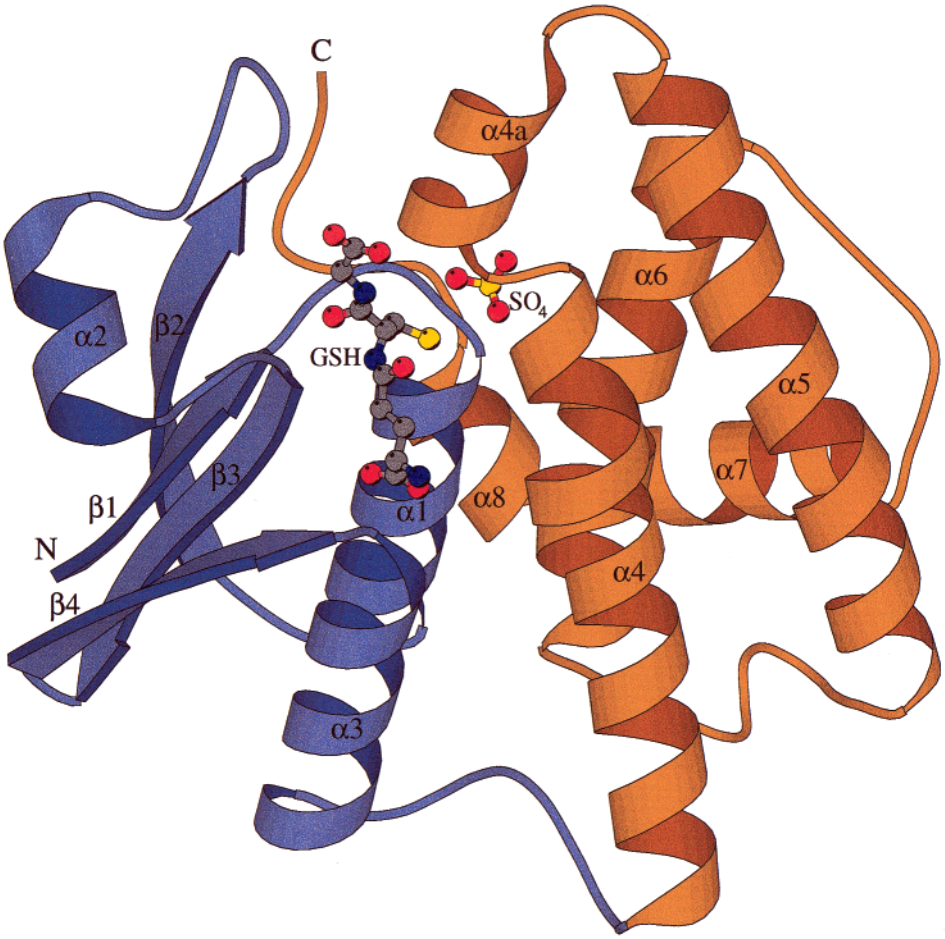

FIGURE 2: Structure of human MAAI/GST Z1-1: a ribbon representation of the monomer indicating the location of secondary structure. GSH and sulfate ions are shown as ball-and-stick and the two domains are shown in different colors. This figure was drawn with BOBSCRIPT (*39*).

pocket formed by residues (Ile 36 and Val 59) of the β-sheet. There are a pair of helices, α1 (residues 14−27) and α3 (residues 71−83), on the opposite side of the β-sheet. The C-terminal domain includes helices α4 (residues 92−110), a turn of 3_{10} helix (residues 111−114), α4a (residues 115 to 123), α5 (residues 123−150), α6 (residues 160−176), and α7 (residues 183−195). The rest of the C-terminus consists of two turns of 3_{10} helix (helix α8; residues 196−200 and 203−207) and a number of sharp bends including a β-turn (residues 207−210) caused by the presence of Pro 204, Pro 208, and Pro 211. Helices α4 and α5 are both over 30 Å long and together with α6 form a tight bundle with approximate parallel/antiparallel alignment of their helical axes. Helix α4 is more accurately described as two helices, α4 and α4a, connected by a turn of 3_{10} helix with a kink at Pro 112. The kink at this residue results in α4a being orientated almost perpendicular to α4. Helices α5 and α6 are connected by a S-shaped bend consisting of a series of β-turns between residues 151 and 156. The axis of helix α7 runs perpendicular to the helical bundle which in turn lies perpendicular to the axis of helix α8. Helix α8 and the C-terminal peptide stretch across the domain interface to pack against the N-terminal domain. Although the last four residues are not visible in the electron density map, it appears likely, based on the positioning of the C-terminus, that they would point away from the molecule into solvent.

The interface between the two domains consists of three regions. At the base, there is a an extensive network of water molecules and a cluster of salt bridges involving residues Glu 77, Glu 81, Arg 96, and Arg 100. The central region is all hydrophobic with contributions from residues Trp 18, the aliphatic moiety of Arg 19, Ile 22, Leu 73, Ile 76, Leu 88, Met 161, Leu 164, Cys 165, Val 167, and Pro 168. The helical towers of the C-terminal domain make contact with the apices of the N-terminal domain through GSH- and sulfate-mediated polar contacts.

GSH and the sulfate ion are located in a very deep crevice (of about 25 Å in depth) between the two domains (Figure 2). The depth of the crevice provides an explanation as to why the enzyme has a relatively weak affinity for GSH purification columns (*7*). The crevice is principally formed by the large loop proceeding helix α2 and helix α4a of the helical tower and these elements effectively bury GSH within the crevice. The DTT molecule is covalently bound to Cys 205. There are three buried charged residues: Arg 19, Arg 21, and Asp 163. All three take part in multiple hydrogen bonding interactions with other polar atoms in the protein core. There are two cis residues in the protein: Pro 60 which is located at the base of the GSH binding site and Pro 86 which is in the domain linker region.

GSH Binding Site. GSH binds to MAAI in an extended conformation and is orientated approximately parallel to the helical axis of α4 with the γ-glutamyl moiety pointing toward the protein core. It forms extensive interactions with the protein including 1 salt bridge and 15 hydrogen-bonding interactions (Figure 3). Every polar atom of the ligand is

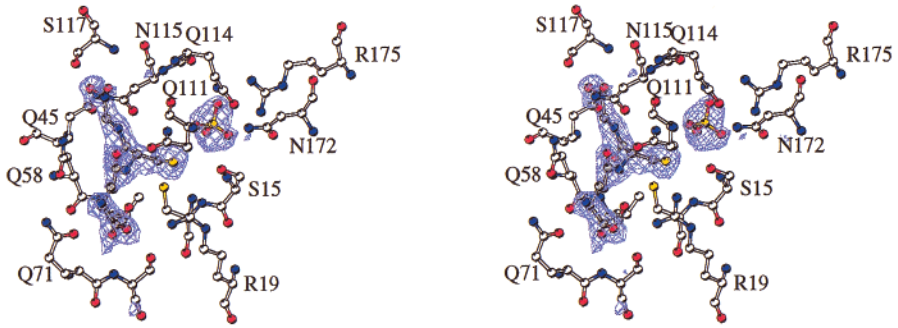

FIGURE 3: Difference Fourier ($F_o - F_c$) electron density map (gray) of the bound ligands at 1.9 Å resolution. The map was calculated from the final model after the bound ligands were omitted from it and a round of positional refinement performed to reduce model bias. The map is contoured at 3σ, with the final model overlayed upon it.

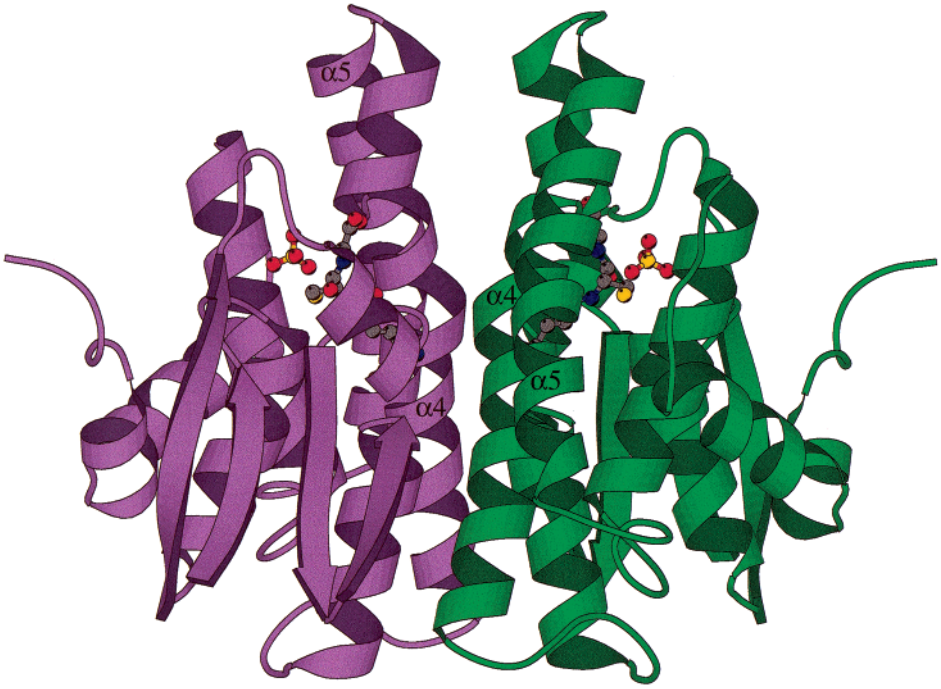

FIGURE 4: Ribbon representation of the dimer with each subunit shown in a different color. This figure was drawn with BOBSCRIPT (*39*).

involved in either direct or water-mediated contacts with the protein. The γ-glutamyl moiety, through its carboxylate oxygens, forms a salt bridge with Arg 19. The backbone of the cysteinyl moiety is involved in an antiparallel sheet interaction with the backbone of Val 59. The amide group of the GSH glycyl moiety is within hydrogen-bonding distance of Asn 115. The carboxylate oxygens of the glycyl moiety form hydrogen bonds with Gln 45 and the main-chain amide of Ser 117. The glycyl moiety sits in a narrow pocket and is shielded from solvent by Phe 46. There are no direct protein contacts to GSH from the neighboring subunit.

The cysteinyl sulfur atom points toward the core of the protein and is located almost directly over the N-terminal end of helix α1. Although close to Cys 16 (2.8 Å), it is not close enough to form a disulfide bridge. The sulfur atom is in hydrogen bonding or van der Waals distance of a number of polar atoms including the amide of Cys 16 (3.5 Å), the amido group of Gln 111 (3.5 Å), and the hydroxyl group of Ser 15 (4.0 Å). It is also close enough (3.0 and 3.4 Å) to interact with two of the oxygen atoms of the bound sulfate ion.

Sulfate Ion in the Active Site. An unexplained tetrahedrally shaped piece of density was observed in a deep cavity at the domain interface (Figure 3). We modeled this density as a sulfate ion due to its presence in the crystallization buffer. The sulfate ion forms numerous contacts with the protein: three of the oxygen atoms are interacting with the side chains of Ser 15, Gln 114, Asn 172, Arg 175, and the main-chain amide of Ser 15. The fourth oxygen atom is hydrogen bonding to two water molecules. The entrance to the sulfate binding site is bordered by a rim of residues with two-thirds of the rim consisting of hydrophobic residues (Tyr 11, Phe 12, Ile 39, Leu 116, Leu 119, and Phe 176) and the rest contributed from the two positively charged residues, Arg 13 and Arg 175.

MAAI Is a Dimer. Previous studies on MAAI have yielded contradictory results with regards to its quaternary structure (*7*, *9*). HPLC gel filtration chromatography indicated the presence of only monomers whereas centrifugation data suggested the enzyme is a dimer. We observe a dimer in the crystals with the dimer axis coincident with a crystallographic 2-fold axis (Figure 4). The intersubunit contacts are extensive (with a total buried surface area of 1538 Å^2/monomer which is within the range of 1350−1700 Å^2 observed in most other GST dimers) and intimate. The interactions are dominated by contributions from the helical towers (α4 and α5) with

contacts extending from the base to the tip of each molecule of the monomer. Interactions at the base consist of a network of water-mediated salt bridges and hydrogen bonds involving Glu 80, Glu 81, Arg 96, and Arg 100. The dimer interface is dominated by van der Waals interactions involving mostly small hydrophobic residues. There are only three aromatic residues in the interface (His 70, Tyr 78, and Trp 130), only two polar residues (Gln 71, Gln 144), which are located close to the periphery of the interaction surface and only a few charged residues (Glu 77, Glu 81, Arg 96, Arg 100, and Asp 104) of which all but Asp 104 are located at the base of the molecule. The only direct polar/charged contact between the two monomers is between Glu 81 and Arg 96 at the base. Other notable interactions include Met 56 which packs into a hydrophobic pocket of the other monomer and aromatic ring stacking of symmetry related Trp 130 at the apex. Given the extensive, intimate contacts between the monomers, it appears likely that the dimer infers considerable stability to the enzyme. The two active sites are separated by a distance of 22 Å, and there is no evidence that the active sites act in a dependent fashion of each other.

DISCUSSION

Mechanism of the Isomerization Activity. MAAI was originally characterized based on its ability to catalyze the isomerization of a cis double bond in maleylacetoacetate (MAA) to the trans configuration (fumarylacetoacetate or FAA) during the catabolism of phenylalanine and tyrosine (Figure 1) (*25*). Although details of the reaction mechanism are unknown it is thought that the catalysis proceeds through a series of steps: in the first step, MAAI catalyses the attack of GSH on the α-carbon of MAA; upon GSH conjugation there would be freedom to rotate about the resulting single bond; following elimination of GSH the product, FAA, would be generated (Figure 1) (*16*).

Other enzymes that utilize GSH in conjugation reactions activate GSH by lowering the pK_a of its thiol (*20*). There are a number of candidates in the crystal structure that might stabilize the thiolate form of GSH. These include Ser 15, Cys 16, Gln 111, and the helix dipole of α1. There is a pronounced electropositive surface surrounding the active site that would attract negatively charged compounds such as MAA toward it. We have modeled MAA into the active site of MAAI and energy minimized the complex (Figure 5). The only assumption in the modeling process was that the acetate carboxylate would most likely be located close to the position of the bound sulfate ion. This orientation ensures that the α carbon is close to the GSH thiol (less than 3.5 Å) so as to ensure a productive complex. The modeling led to a very convincing fit. The acetate carboxyl is within hydrogen bonding distance of the side-chains of Gln 111, Gln 114, and Asn 172 as well as the amide nitrogen of Ser 15. The hydroxyl and keto substituents are within hydrogen bonding distance of Arg 13 and Arg 175 and some of the carbon backbone makes van der Waals interactions with Ile 39 and Leu 116. The other carboxylate could possibly interact with the Arg 13 and hydroxyl of Tyr 11. Thus, the crystal structure provides a molecular basis to explain the specificity of MAAI for the substrate and shows how the active site orientates the substrate for optimal attack by GSH.

Molecular Basis for the Biotransformation of α-Haloacids. On the basis of a number of experimental studies, a reaction mechanism has been proposed to explain the mechanism by which MAAI catalyses the oxidation of a variety of α-haloacids of which DCA has served as a model substrate (Figure 1) (*15*). In the first step, the enzyme catalyses the displacement of one chloride by GSH via a S_N2 reaction to yield *S*-(α-chlorocarboxymethyl)GSH. In the next step, another chloride is lost to yield a carbonium-sulfonium intermediate. The intermediate then decomposes via S_N1-type hydrolysis to form the product, glyoxylate, and release GSH. Thus, GSH is required for the reaction mechanism but is not consumed. DCA is a mechanism-based inactivator of the enzyme whereby it is proposed that an alternative pathway exists in the last step so that the carbonium-sulfonium intermediate covalently attaches itself to the enzyme (Figure 1) (*15*).

The crystal structure of MAAI provides a number of insights into the catalytic mechanism of biotransformation. First, the presence of the carboxylate group in the substrate appears essential (*8*). For example, α-haloacid esters and α-haloacetamides are not substrates. Model building suggests that DCA fits best into the active site by orientating the molecule so that the carboxylate moiety forms a salt bridge with Arg 175, hence mimicing the binding of the sulfate ion in the active site (Figure 5). The interactions observed between the sulfate ion and surrounding residues are compensated for in the DCA model: the main-chain amide of Ser 15 and the side chain of Gln 114 interact with the DCA carboxylate, Asn 172 interacts with Gln 114 and Arg 175 and the side chain of Ser 15 interacts with a buried water molecule. The salt bridge with Arg 175 optimally orientates DCA for GSH attack on the α-carbon and displacement of one of the chlorides. Thus, the structure suggests the carboxylate of the haloacid is essential for correct orientation of the substrate in the active site. Second, *β*-haloacids are not substrates for MAAI (*8*). The crystal structure suggests that the binding site cannot correctly orientate the *β*-carbon for GSH attack. Third, GSH cannot be replaced by other sulfur-containing compounds such as L-methionine, *N*-acetyl-L-cysteine or DTT (*8*). The crystal structure demonstrates that the positioning of the thiol sulfur with respect to the α-carbon of the substrate is critical and this can only be achieved by appropriate binding in the GSH-binding site which is highly specific for GSH. Fourth, the proposed reaction mechanism predicts a carbonium-sulfonium ion intermediate. There are a number of polar groups nearby that could stabilize the intermediate, particularly the hydroxyl of Ser 15.

Implications for Reductive Dehalogenation Mechanisms of Related Enzymes. It has recently been demonstrated that MAAI shares a close relationship with the enzyme tetrachlorohydroquinone dehalogenase (TCHQ-D) (*16*). The latter enzyme catalyzes the replacement of chlorine atoms on tetrachlorohydroquinone and tricholohydroquinone with hydrogen atoms during the biodegradation of pentachlorophenol. TCHQ-D was shown to catalyze the isomerization of maleylacetone (an analogue of MAA) at a comparable rate to the MAAI enzyme. Furthermore, there are sequence similarities in the active site regions of both enzymes (overall identity is 15%). It has been argued that TCQH-D recently evolved from MAAI because the latter enzyme is involved

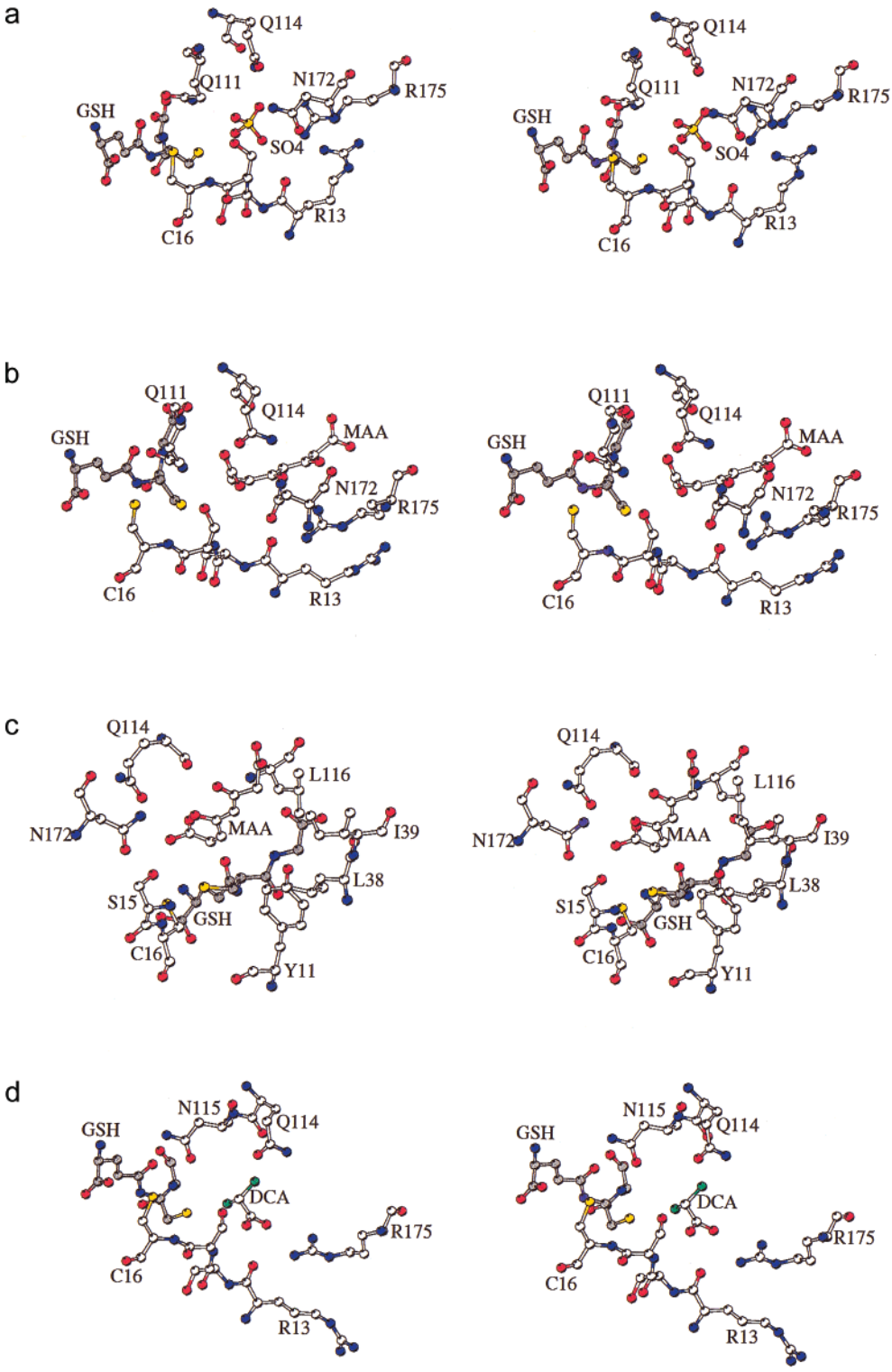

FIGURE 5: Stereoviews of the active site of MAAI complex models. (a) Sulfate ion. (b, c) MAA viewed from two different orientations. (d) DCA. The MAA and DCA models were constructed by manually docking the ligands into the active site using the location of the sulfate ion as a guide. Both models were energy minimized after removal of the sulfate ions and nearby water molecules.

in a presumably ancient pathway of aromatic amino acid catabolism whereas TCHQ-D catalyzes a very specialized reaction on a modern day pollutant [pentachlorophenol was first introduced as a wood preservative in 1936 (*26*)]. Another related enzyme is the lindane degrading reductive dehalogenase (LinD or 2,5-dichlorohydroquinone dehalogenase) which catalyzes a similar reaction to TCHQ-D and has low but detectable sequence identity (<20%) with both TCHQ-D and MAAI (*28*).

On the basis of sequence alignments, a consensus motif was discovered that could be used to identify MAAI/zeta class GSTs, TCHQ-D, and related enzymes (*16*). The motif, LYSYWR/LSSCSXR/KVRIAL, is located in the N-terminal domain of MAAI (residues 8−24) and consists of the

C-terminal half of strand β1, all but the last turn of helix α1 and the β-turn that connects the two structural elements. All three structural regions fulfill important roles: the helix forms part of the core of the domain and contributes to the domain interface; the β-strand packs against the core of the domain and Tyr 9 contributes to the hydrophobic pocket to which helix α2 packs; the β-turn forms the base of the active site and contributes residues involved in binding of GSH and the sulfate ion. Important residues from the motif include Tyr 11 which packs against the large loop and likely orientates it for optimal involvement in GSH binding and catalysis, Phe 12 and Arg 21 which interact with the C-terminal peptide region, Arg 13 which forms part of the positive lip at the active site entrance, Ser 15 and Cys 16 which interact with GSH and the sulfate ion, and Arg 19 which lies at the base of the GSH-binding site and forms a salt bridge with the GSH γ-glutamyl caboxylate.

There are two other regions of high similarity in the published sequence alignments: residues 106−116 and 175−184. Part of each region contributes to the active site in MAAI. Gln 111 interacts with the sulfate ion. Its replacement by Arg in LinD would hence be conservative but its replacement by Gly in TCHQ-D would result in a loss of one of the sulfate ion ligands. Gln 114 also interacts with the sulfate ion and is replaced by Ala in TCHQ-D and His in LinD. Asn 115 interact with GSH and is replaced by Arg in TCHQ-D and Lys in LinD. Residues 114 and 115 would not interact with sulfate or GSH in TCHQ-D and LinD any longer because there is a 6−10 residue insertion before residue 114 in these enzymes. Leu 116, Arg 175, and Phe 176 surround the entrance to the active site and Arg 175 appears to be the residue responsible for anchoring α-haloacids into the active site. These residues are replaced by Arg, Asn, and Phe, respectively, in TCHQ-D and by Ile, Gly, and Met in LinD. The presence of Arg at residue 116 would introduce an extra positive charge to the active site entrance. The replacement of Arg 175 by any residue would be expected to effect the efficiency of reaction with carboxylate substrates such as MAA and α-haloacids. Indeed TCHQ-D has been found to be less efficient in the isomerase reaction (*16*). The other region that provides major contributions to the active site is the large loop after strand β2 in MAAI: this region bears little sequence similarity with LinD and TCHQ-D. Overall, the comparative structure-based sequence analysis suggests the GSH binding sites of MAAI, LinD, and TCHQ-D are quite similar, but the adjacent active site is not, and in particular, it appears that TCHQ-D has lost the sulfate-binding site.

In TCHQ-D, Ser 11 has been hypothesized to stabilize the GSH thiolate and Cys 13 has been shown to act as the nucleophile. The equivalent residues in MAAI are Ser 14 and Cys 16. Ser 14 is located in the β-turn between the first strand and helix in MAAI. It forms hydrogen-bonding interactions with the main-chain carbonyl of Phe 11 and hence seems to play an important structural role. If it did stabilize the GSH thiolate there must be significant conformational changes. It is possible that MAAI is complexed to the protonated form of GSH in the crystal structure and conformational changes do occur on deprotonation (see below). Cys 16 is in the active site but is mostly obscured by the bound GSH. In order for it to act as a nucleophile GSH would have to be displaced.

Preliminary data suggests MAAI does not possess TCHQ dehalogenase activity but TCHQ-D has both isomerization and α-haloacid oxygenation activities supporting the hypothesis that TCHQ-D evolved from MAAI (*16*). However, kinetic studies suggest the catalytic efficiencies are low in both cases (*16*) and this would tally with our suggestion that TCHQ-D might not have the sulfate binding pocket to anchor the carboxylate group of MAA or DCA. Cys 13 of TCHQ-D (Cys 16 in MAAI) has been shown to be essential for the isomerization of MA (an analogue of MAA), oxygenation of DCA, and dehalogenation of its natural substrate (*16*). In addition, this cysteine was shown not to be involved in GSH binding or ionization. These data, together with the crystal structure of MAAI support, the suggestion that an early step in the mechanism of each reaction would involve nucleophilic attack of GSH followed in a subsequent step by nucleophilic attack of an intermediate by the cysteine. Surprisingly, the equivalent cysteine in MAAI is not essential for reactivity (P. G. Board, unpublished results).

MAAI as a Glutathione Transferase. The crystal structure of MAAI reveals it belongs to the GST superfamily despite exhibiting either little or no activity with standard substrates and the lack of any significant sequence identity (less than 25%) with other members for which there are crystal structures. Like other GSTs, it assembles as a dimer with an N-terminal thioredoxin-like fold and a C-terminal all α-helical domain. The most significant differences between MAAI and GSTs occur in the helix α2 and surrounding regions, large distortions in the apex region of the helical towers, and the presence of a small, polar active site in MAAI compared with the much larger, hydrophobic active sites typical of other members of the GST family (*20*).

There are many structural features of MAAI found in other GST classes, although it does not appear that MAAI resembles more closely any one class in particular. MAAI has a truncated C-terminus like the beta (*28*), delta (*19*), mu (*29*), phi (*30*), pi (*31*), and sigma (*32*) class enzymes. The predominantly hydrophobic dimer interface is found in most GSTs except the beta (*28*) and sigma (*32*) class enzymes. MAAI lacks the typical V-shaped dimer interface and hence resembles the theta (*33*) and beta (*28*) classes in this regard. An intersubunit "lock-and-key" hydrophobic interaction has been described in the mammalian GST structures (*20*). This involves an aromatic residue termed the "key", from the loop preceding strand β3 and a hydrophobic "lock" from helices α4 and α5 of the other monomer. MAAI uses Met 56 as a "key" but the pocket is located closer to the apex of the helical towers compared to the mammalian GSTs. The anchoring of helix α2 to the β-sheet core of the N-terminal domain, via an aromatic residue nestling into a hydrophobic pocket, has previously been observed in many of the nonmammalian GSTs (*19*, *28*, *30*). The large loop preceding helix α2 is reminiscent of the mu loop previously seen in the rat mu class structure (*29*) and the β-hairpin seen in the beta class structure (*28*).

The almost buried active site is reminiscent of the human theta class structure but in the latter case the active site is covered by a C-terminal helical extension (*33*). The MAAI residues involved in binding GSH are nearly all conserved or conservatively substituted in the GSTs (*20*). One exception is the contact with the amino group of the γ-glutamyl moiety which is normally fulfilled by an acidic residue in GSTs but

replaced by a water molecule in MAAI. More will be said about the GSH thiol interaction in the next section. There are no direct contacts between GSH and the neighboring monomer in MAAI like the phi (*30*) and delta (*19*) class enzymes. GSH contacts with the C-terminal domain have been observed previously in the beta (*28*) and theta (*33*) enzymes. The presence of a cysteine residue in the active site has previously been observed in the beta (*28*) and omega (*34*) classes, but in both these cases, GSH was found covalently bound to the cysteine. With respect to the beta class sequence, the insertion of an extra residue at position 15 (Ser 15) in MAAI is sufficient to alter the position and conformation of Cys 16 so that it cannot get close enough to the GSH thiol to form a covalent bond. Cys 32 of the ω enzyme superimposes closest to Ser 14 of MAAI; however, the latter points in a completely different direction so that it cannot interact with GSH. An active site sulfate ion has also been observed in the human theta class enzyme where it plays a role in the sulfatase activity of that enzyme (*33*). There is no obvious evolutionary relationship between the MAAI and theta sulfate sites as they are more than 4 Å apart, the theta site is occluded by helix α4a in MAAI and different residues are involved in binding to the sulfate.

Evolutionary Origins of MAAI. MAAI/GST Z1-1 appears very likely to represent an ancient class of GST given its distribution in a wide range of organisms including bacteria, fungi, nematodes, plants, and animals (*7, 16*). It also seems likely that it is most closely related to the theta class of enzymes that have similar broad distributions and in which some of the members such as dichloromethane dehalogenase (*35*) have very specific activities like MAAI. Both MAAI and dichloromethane dehalogenase, as well as some mammalian theta class GSTs, can oxygenate dichloromethane (*36*, *37*) providing another link between the two classes. A hallmark feature of the theta class GSTs is the presence of a catalytic serine residue that stabilizes the thiolate form of GSH. In the theta (*33*) and delta (*19*) class GSTs, the serine residue (Ser 9 and Ser 11, respectively) is in hydrogen-bonding distance of the GSH thiol, and in the case of the omega class enzyme, the serine is replaced by cysteine that forms a covalent linkage with GSH through the thiol (*34*). Superposition of the MAAI crystal structure on the delta, omega, and theta crystal structures shows that the Ser/Cys residue superimposes on Ser 14 of MAAI. However, the Ser 14 hydroxyl points away from the active site and forms an interaction with the main-chain atom of a nearby residue in the β-turn. If Ser 14 is a catalytic residue in MAAI, then the structure must under go some conformational changes so that Ser 14 can interact with GSH. It is possible that we have crystallized the protonated form of GSH and that deprotonation would be accompanied by the required conformational changes. There is a precedent for this scenario in the pi class GST, where protonation causes a loss of direct interaction of the catalytic tyrosine with GSH and the thiol sulfur points in a different direction to that seen in other GSH complexes (*38*).

Concluding Remarks. The extraordinary catalytic versatility of MAAI is, in part, explained by the crystal structure of the enzyme. Interestingly, the active site is a lot smaller and more polar than most other GSTs and these properties make the enzyme much more selective in the substrates that it recognizes. The other major distinguishing features of the active site compared to other GSTs is its high electropositivity and the location of Arg 175 which appears to play an essential anchoring role for fixing substrates in an optimal orientation for attack by GSH.

ACKNOWLEDGMENT

We thank Dr. Igor Polikarpov and Jose Brandao for help with data collection at the National Synchrotron Light Laboratory (LNLS) in Brazil. The experiment was supported with funds from FAPESP. We thank Harry Tong and other staff at BioCARS for their help with data collection during our visit to APS.

REFERENCES

1. Fernández-Cañón, J. M., and Peñalva, M. A. (1998) *J. Biol. Chem. 273*, 329−337.
2. Poudrier, J., Lettre, F., Scriver, C. R., Larochelle, J., and Tanguay, R. M. (1998) *Mol. Genet. Metab. 64*, 119−125.
3. Russo, P., and O'Regan, S. (1990) *Am. J. Hum. Genet. 47*, 317−324.
4. Berger, R., Michals, K., Galbraeth, J., and Matalon, R. (1988) *Pediat. Res. 23*, 328.
5. Lindstedt, S., Holme, E., Lock, E. A., Hjalmarson, O., and Standvik, B. (1992) *Lancet 340*, 813−817.
6. Fernández-Cañón, J. M., Hejna, J., Reifsteck, C., Olson, S., and Grompe, M. (1999) *Genomics 58*, 263−269.
7. Board, P. G., Baker, R. T., Chelvanayagam, G., and Jermiin, L. S. (1997) *Biochem. J. 328*, 929−935.
8. Tong, Z., Board, P. G., and Anders, M. W. (1998) *Chem. Res. Toxicol. 11*, 1332−1338.
9. Tong, Z., Board, P. G., and Anders, M. W. (1998) *Biochem. J. 331*, 371−374.
10. DeAngelo, A. B., Daniel, F. B., Most, B. M., and Olson, G. R. (1996) *Toxicology 114*, 207−221.
11. Bull, R. J., Sanchez, I. M., Nelson, M. A., Larson, J. L., and Lansing, A. J. (1990) *Toxicology 63*, 341−359.
12. Weisel, C. P., Kim, H., Haltmeier, P., and Klotz, J. B. (1999) *Environ. Health Perspect. 107*, 103−110.
13. Stacpoole, P. W., Henderson, G. N., Yan, Z., and James, M. O. (1998) *Environ. Health Perspect. 106*, 989−994.
14. Anderson, W. B., Board, P. G., Gargano, B., and Anders, M. W. (1999) *Chem. Res. Toxicol. 12*, 1144−1149.
15. Tzeng, H. F., Blackburn, A. C., Board, P. G., and Anders, M. W. (2000) *Chem. Res. Toxicol. 13*, 231−236.
16. Anandarajah, K., Kiefer, P. M., Jr., Donohoe, B. S., and Copley, S. D. (2000) *Biochemistry 39*, 5303−5311.
17. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
18. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr. 30*, 1022−1025.
19. Wilce, M. C. J., Board, P. G., Feil, S. C., and Parker, M. W. (1995) *EMBO J. 14*, 2133-2143.
20. Armstrong, R. N. (1997) *Chem. Res. Toxicol. 10*, 2−18.
21. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr., Sect. D 50*, 760−763.
22. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr., Sect. D 53*, 240−255.
23. Brünger, 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., Sect. D 54*, 905−921.
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
25. Knox, W. E., and Edwards, S. W. (1955) *J. Biol. Chem. 216*, 489−498.
26. Cline, R. E., Hill, R. H., Phillips, D. L., and Needham, L. L. (1989) *Arch. Environ. Contam. Toxicol. 18*, 475−481.
27. Miyauchi, K., Suh, S. K., Nagata, Y., and Takagi, M. (1998) *J. Bacteriol. 180*, 1354−1359.

28. Rossjohn, J., Polekhina, G., Feil, S. C., Allocati, N., Masulli, M., Di Ilio, C., and Parker, M. W. (1998) *Structure 6*, 721−734.
29. Ji, X., Zhang, P., Armstrong, R. N., and Gilliland, G. L. (1992) *Biochemistry 31*, 10169−10184.
30. Reinemer, P., Prade, L., Hof, P., Neuefeind, T., Huber, R., Zettl, R., Palme, K., Schell, J., Koelln, I., Bartunik, H. D., and Bieseler, B. (1996) *J. Mol. Biol. 255*, 289−309.
31. Reinemer, P., Dirr, H. W., Ladenstein, R., Huber, R., Lo Bello, M., Federici, G., and Parker, M. W. (1992) *J. Mol. Biol. 227*, 214−226.
32. Ji, X., van Rosenvinge, E. C., Johnson, W. W., Tomarev, S. I., Piatigorsky, J., Armstrong, R. N., and Gilliland, G. L. (1995) *Biochemistry 34*, 5317−5328.
33. Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelvanayagam, G., Tan, K.-L., Board, P. G., and Parker, M. W. (1998) *Structure 6*, 309−322.
34. Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F., and Pandit, J. (2000) *J. Biol. Chem. 275*, 24798−24806.
35. La Roche, S. D., and Leisinger, T. (1990) *J. Bacteriol. 172*, 164−171.
36. Hashmi, M., Dechert, S., Dekant, W., and Anders, M. W. (1994) *Chem. Res. Toxicol. 7*, 291−296.
37. Blocki, F. A., Logan, M. S., Baoli, C., and Wackett, L. P. (1994) *J. Biol. Chem. 269*, 8826−8830.
38. Oakley, A. J., Lo Bello, M., Battistoni, A., Ricci, G., Rossjohn, J., Villar, H. O., and Parker, M. W. (1997) *J. Mol. Biol. 274*, 84−100.
39. Esnouf, R. M. (1999) *Acta Crystallogr., Sect. D 55*, 938−940.

BI002249Z