Crystal Structure of *Tritrichomonas foetus* Inosine-5′-monophosphate Dehydrogenase and the Enzyme−Product Complex†

Frank G. Whitby,[‡,§,||] Hartmut Luecke,[||,⊥,▽] Peter Kuhn,[⊥] John R. Somoza,[@] Jorge A. Huete-Perez,[‡,#] John D. Phillips,[+] Christopher P. Hill,[∞] Robert J. Fletterick,[‡,@] and Ching Chung Wang*[,‡]

Department of Pharmaceutical Chemistry, School of Pharmacy, 926 Medical Sciences Building, University of California, San Francisco, California 94143-0446, Stanford Synchrotron Radiation Laboratory, Stanford, California 94309, Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143-0448, Department of Medicine, University of Utah School of Medicine, Salt Lake City, Utah 84132, and Department of Biochemistry, University of Utah School of Medicine, 50 North Medical Drive, Salt Lake City, Utah 84132

Received April 16, 1997; Revised Manuscript Received June 3, 1997[⊗]

ABSTRACT: Inosine-5′-monophosphate dehydrogenase (IMPDH) is an attractive drug target for the control of parasitic infections. The enzyme catalyzes the oxidation of inosine monophosphate (IMP) to xanthosine monophosphate (XMP), the committed step in *de novo* guanosine monophosphate (GMP) biosynthesis. We have determined the crystal structures of IMPDH from the protozoan parasite *Tritrichomonas foetus* in the apo form at 2.3 Å resolution and the enzyme-XMP complex at 2.6 Å resolution. Each monomer of this tetrameric enzyme is comprised of two domains, the largest of which includes an eight-stranded parallel β/α-barrel that contains the enzyme active site at the C termini of the barrel β-strands. A second domain, comprised of residues 102−220, is disordered in the crystal. IMPDH is expected to be active as a tetramer, since the active site cavity is formed by strands from adjacent subunits. An intrasubunit disulfide bond, seen in the crystal structure, may stabilize the protein in a less active form, as high concentrations of reducing agent have been shown to increase enzyme activity. Disorder at the active site suggests that a high degree of flexibility may be inherent in the catalytic function of IMPDH. Unlike IMPDH from other species, the *T. foetus* enzyme has a single arginine that is largely responsible for coordinating the substrate phosphate in the active site. This structural uniqueness may facilitate structure-based identification and design of compounds that specifically inhibit the parasite enzyme.

Tritrichomonas foetus is an anaerobic protozoan parasite that infects the urogenital tract of cows, causing bovine trichomoniasis which can result in sterility and abortion (*1, 2*). As with most parasitic infections, there is currently no effective treatment for tritrichomonal infection. Protozoan parasites, however, rely on metabolic pathways that differ significantly from those of their mammalian hosts (*3*), suggesting that inhibition of crucial parasite enzymes could be useful in developing selective chemotherapeutic agents. *T. foetus*, which is incapable of *de novo* synthesis of purine nucleotides (*4*), preferentially salvages hypoxanthine from its host and converts it to inosine 5′-monophosphate (IMP)[1] via the enzyme hypoxanthine-guanine-xanthine phosphoribosyltransferase (HGXPRTase), itself an attractive drug target (*5*). IMP is then converted to xanthosine-5′-monophosphate (XMP) by the enzyme inosine-5′-monophosphate dehydrogenase (IMPDH, EC 1.1.1.205) (*6*) which is the rate-limiting enzyme in guanine nucleotide synthesis (*6*). The *in vitro* growth of *T. foetus* has been shown to be arrested by the uncompetitive inhibitor mycophenolic acid (MPA), which is a potent inhibitor of mammalian and bacterial IMPDH (*7, 8*). This inhibition in *T. foetus* was reversed by supplementing the growth medium with guanine or guanosine, suggesting the validity of IMPDH as a target for antitrichomonal chemotherapy (*9*).

A single-copy gene for *T. foetus* IMPDH encodes a protein of 503 amino acids (*10*) that shares about 25−30% amino acid sequence identities with known bacterial (*11*), parasite (*12*), and mammalian (*13*−*15*) forms of the enzyme, including the two human isoforms (*16*). IMPDHs from all of these sources are homotetramers of ∼55 kDa subunits that share a similar mechanism of catalysis. Most other known nicotinamide adenine dinucleotide (NAD)-dependent dehydrogenases undergo a conformational change following cofactor binding, thereby promoting substrate binding and catalysis (*17*). IMPDH from all species studied to date, however, has an ordered bi-bi kinetic mechanism where IMP

† This work was supported in part by NIH Grants AI-19391 to C.C.W. and GM56445 to H.L. F.G.W. was supported in part by a grant from the Lucille P. Markey Charitable Trust.

* Corresponding author.
‡ School of Pharmacy, University of California.
§ Present address: Department of Biochemistry, University of Utah School of Medicine, 50 N. Medical Dr., Salt Lake City, UT 84132. E-mail: frankw@solitude.med.utah.edu.
|| F.G.W. and H.L. made equal contributions to this publication.
⊥ Stanford Synchrotron Radiation Laboratory.
▽ Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697.
@ Department of Biochemistry and Biophysics, University of California.
# Department of Pathology, 4150 Clement St., University of California, San Francisco, CA 94121.
+ Department of Medicine, University of Utah School of Medicine.
∞ Department of Biochemistry, University of Utah School of Medicine.

⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1997.

[1] Abbreviations: IMPDH, inosine-5′-monophosphate dehydrogenase; IMP, inosine monophosphate; XMP, xanthosine monophosphate; GMP, guanosine monophosphate; HGXPRTase, hypoxanthine−guanine−xanthine phosphoribosyltransferase; MPA, mycophenolic acid; NAD, nicotinamide adenine dinucleotide; MIR, multiple isomorphous replacement; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; DTT, dithiothreitol; DTNB, 5,5′-dithiobis(2-nitrobenzoate); FMN, flavin mononucleotide; PCMBS, *p*-(chloromercuri)benzenesulfonate; BME, β-mercaptoethanol.

S0006-2960(97)00885-4 CCC: $14.00

Table 1: Data Collection and Heavy-Atom Statistics

	native	K_2PtCl_4	thimerosal	XMP
diffraction limit (Å)	2.3	3.0	3.0	2.6
completeness (%)[a]	97.8 (3.7)	99.5 (5.2)	95.5 (4.2)	94.3 (4.7)
R_{merge}[b] (%)	5.7 (25.0)	11.1 (19.7)	6.5 (19.7)	10.4 (26.8)
R_{iso}[c] (18−3.0 Å) (%)		23.1	21.9	
phasing power[d] (18−3.0 Å)		1.57	1.55	
R_{cullis}[e] (18−3.0 Å) (%)		0.581	0.601	
R_{kraut}[f] (18−3.0 Å) (%)		0.065	0.087	
number of heavy-atom sites		2	2	

[a] Overall completeness redundancy of data = (observed reflections/unique reflections). [b] $R_{merge} = \sum|I-\langle I\rangle|/I$ (R_{merge} in the highest-resolution shell). [c] $R_{iso} = \sum|F_{PH}-F_P|/\sum F_P$. [d] Phasing power = root mean square (rms) of f_H/E, where f_H = calculated heavy-atom structure-factor amplitude and E = lack of closure ($\sum||F_{PH}\pm F_P|-f_H|$), where F_{PH} = structure-factor amplitude of the derivative crystal and F_P = structure-factor amplitude of the native crystal. [e] $R_{cullis} = \sum|||F_{PH}|(obs)\pm|F_P|(obs)|-|F_H|(calc)|/\sum||F_{PH}|(obs)\pm|F_P|(obs)|$. [f] $R_{kraut} = \sum||F_{PH}|(obs)-|F_{PH}|(calc)|/\sum|F_{PH}|(obs)$.

binds before NAD and NADH is released before XMP. Interestingly, IMPDHs from other species, including plant, bacterial, and mammalian enzymes, require K^+ for catalysis, while the *T. foetus* enzyme does not (*6*).

Hydrolytic dehalogenation of 2-F-IMP and 2-Cl-IMP and their subsequent conversion to XMP by human IMPDH in the absence of NAD indicate that the reaction mechanism proceeds via a nucleophilic attack at the C2-position of IMP, thus implying the formation of a tetrahedral intermediate (*18, 19*). Our previous studies demonstrated the formation of a covalent intermediate between IMP and the Cys-319 of *T. foetus* IMPDH during the enzyme-catalyzed reaction (*20*). In addition, Link and Straub (*21*) identified a covalent complex of IMP and the conserved Cys-331 of human IMPDH, thus suggesting that these cysteine residues are the nucleophiles in the active sites of IMPDHs that catalyze the hydride transfer from the C2-position of IMP to NAD. Link and Straub (*21*) also demonstrated that MPA binding to human IMPDH traps and stabilizes this covalent intermediate. This observation was supported by the studies of Fleming et al. (*22*), which indicated that MPA inhibition of hamster IMPDH occurred after IMP and NAD binding, hydride transfer, and NADH release, apparently by trapping a transient covalent product of the hydride transfer reaction before a hydrolysis step preceding XMP release. These findings agree well with our previous postulation that MPA appears to inhibit completion of the IMPDH enzymatic cycle at a stage following hydride transfer and NADH release by partially occupying the NADH binding site (*18*). Recent elucidation of the crystal structure of hamster IMPDH in complex with IMP and MPA demonstrated this mode of inhibition and revealed a covalent adduct between Cys-331 and the C2-position of the XMP precursor reaction intermediate (*23*). *T. foetus* IMPDH is, however, about 1000-fold more weakly inhibited by MPA than is the mammalian enzyme (*6*). MPA inhibits mammalian enzymes with a K_i of 10−37 nM (*16*) and *T. foetus* IMPDH with a K_i of 7−9 μM (*6, 18*). This difference in inhibition by MPA suggests that the cofactor binding site of *T. foetus* IMPDH differs from that of mammalian enzymes and that these differences could be exploited in designing a specific inhibitor by modification of MPA.

We report here the X-ray crystal structures of *T. foetus* IMPDH in the apo form and in complex with XMP. The structure is characterized by a high degree of disorder in specific parts of the molecule, suggesting that *T. foetus* IMPDH has unusual flexibility. In contrast to that of IMPDHs from other species, the binding of the substrate phosphate moiety involves an arginine residue. This distinctive means of substrate binding can guide the design of specific inhibitors targeted against the parasite enzyme and may serve as a model for IMPDH from other protozoan parasites.

EXPERIMENTAL PROCEDURES

The expression, purification, and preliminary crystallographic analysis of *T. foetus* IMPDH have been described elsewhere (*6, 10, 20, 24*). Briefly, the protein was overexpressed in *Escherichia coli* and purified by anion exchange chromatography. Crystals were grown at room temperature in 2.2 M ammonium sulfate from protein initially in the wide concentration range of 5−97 mg/mL by hanging, sitting drop and liquid diffusion methods. The size of the crystals increased with increasing initial protein concentration, the largest crystals being almost 2 mm in each dimension. The space group is $P432$ ($a = b = c = 157.25$ Å, $\alpha = \beta = \gamma = 90°$) with one molecule in the asymmetric unit (*24*). High-quality XMP complex crystals were obtained both by soaking crystals in mother liquor containing XMP at 50 mM and by cocrystallization with 30−50 mM XMP in the crystallization buffer. Crystals soaked in and grown in XMP at concentrations below 2 mM did not show substantial occupancy of the active site.

All data used in this study were collected at 4 °C on a Mar image plate detector at beamline 7-1 of the Stanford Synchrotron Radiation Laboratory (SSRL). As a consequence of the high space group symmetry, data were collected very rapidly with a high level of redundancy. Typically, a complete data set was collected from one crystal as 20 contiguous, 10 s, 1° oscillation images. Data were indexed and integrated with the program DENZO (*25*) and were scaled and reduced using ROTAVATA and AGROVATA (*26*).

Initial phases were determined at 3.0 Å resolution for the apoenzyme crystal by the MIR method, with derivatives prepared by soaking. Heavy-atom sites for a thimerosal derivative were identified using the programs of the XtalView package (*27*) and SHELXS-90 (*28*). XtalView was used for heavy-atom refinement, rapid difference Fourier analysis of putative derivatives, and calculation of initial SIR and MIR maps which showed clearly the molecular packing arrangement and solvent boundaries. Final heavy-atom analysis was carried out using the PHASES package (*29*). Of the several derivatives identified, one tetrachloroplatinate and one thimerosal derivative were determined to be of the highest quality and were used in phasing (Table 1).

A few helices were apparent in the initial 3.0 Å MIR map, although it was difficult to trace extensive regions of the density. The map quality was dramatically improved by solvent flattening using PHASES (*29, 30*), and the resulting map showed clearly the eight-stranded β/α-barrel and several other regions of recognizable secondary structure. An initial model consisting of 273 alanine residues was built in several isolated segments using the program O (*31*). Rigid-body refinement, simulated annealing, and positional and *B* factor refinement were then performed using X-PLOR (*32*). Model

Table 2: Refinement Statistics for *T. foetus* IMPDH

	apo	XMP complex
resolution range (Å)	8.0−2.3	8.0−2.6
number of protein atoms	3211	3124
number of solvent atoms	57	51
R-factor (%)	21.8 (20.6 with a cutoff $I/\sigma I$ of >2) (using 28 851 reflections)	20.3 (19.0 with a cutoff $I/\sigma I$ of >2) (using 19 545 reflections)
free *R*-factor (%)	26.5 (25.7 with a cutoff $I/\sigma I$ of >2) (for 2945 reflections)	26.4 (24.7 with a cutoff $I/\sigma I$ of >2) (for 1919 reflections)
rms deviations from ideal geometry		
bond lengths (Å)	0.011	0.012
bond angles (deg)	1.593	1.663
mean temperature factors (Å^2)		
main chain	25.33	23.20
side chains	29.38	26.75
water molecules	35.80	30.30
XMP		22.38

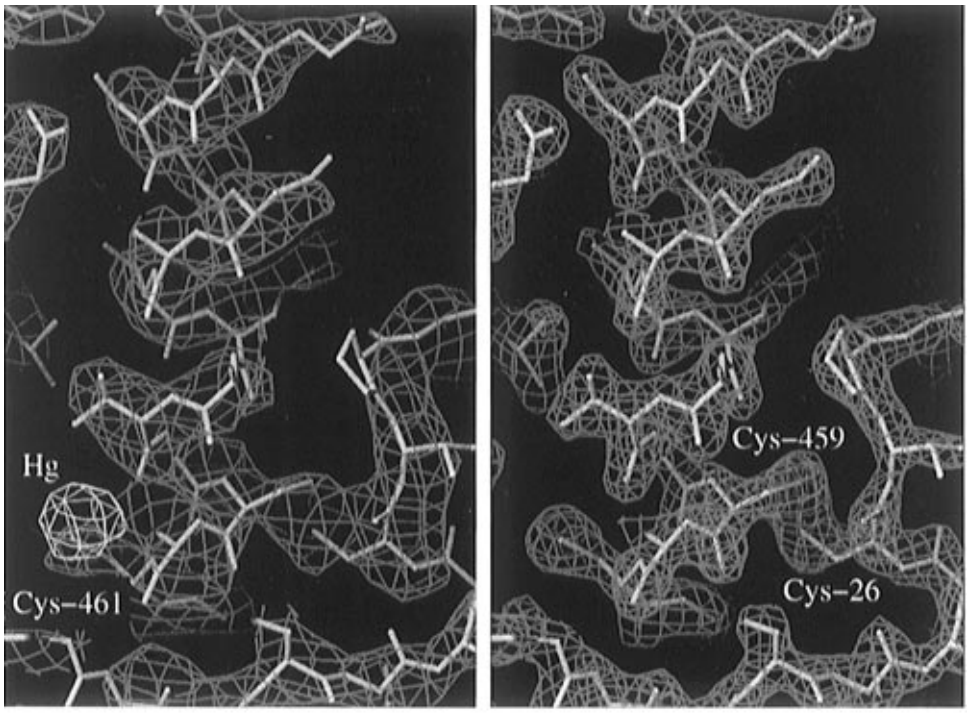


FIGURE 1: Electron density maps. (a) Solvent-flattened MIR map (blue) calculated at 3.0 Å resolution and contoured at 1.5 σ and mercury $F_o - F_c$ map (yellow) contoured at 4.0 σ showing the covalent interaction between Cys-461 and thimerosal that was used in phasing. Cys-26 and Cys-459 were not reactive with thimerosal or PCMBS. The refined XMP complex model is shown in white. (b) The same region showing a $2F_o - F_c$ map of the refined XMP complex (blue) calculated at 2.6 Å resolution and contoured at 1.8 σ. An omit map ($F_o - F_c$), contoured in pink at 2.0 σ, shows the disulfide linkage between Cys-26 and Cys-459. All atoms of residues 26 and 459 were omitted from the model prior to simulated annealing refinement and map calculation with X-PLOR (*32*) [figure produced using O (*31*)].

phases were combined with initial MIR phases using the sigma (Read) procedure as implemented in X-PLOR, followed by solvent flattening in PHASES. The model was manually re-fitted, and the additional density was fitted with more alanine residues. Two additional rounds of manual rebuilding of the polyalanine model followed by refinement, phase combination, and solvent flattening were performed. To ensure that solvent flattening had not removed ordered regions of the protein density, unflattened MIR maps were compared with the refined partial model, but it was clear that some portions of the molecule were disordered.

At this stage, the density was well-defined for many side chains and sequence assignments were made. Cysteine and methionine side chains, which were targeted by thimerosal and tetrachloroplatinate derivatives, respectively, provided useful guides for model building. Water molecules were included in the model where a strong spherical difference density corresponded to chemically reasonable interactions with ordered protein residues. Water molecules were deleted if their *B* factor was refined to greater than 50 Å^2. Refinement statistics are reported in Table 2. The resulting apoenzyme model includes residues 2−101, 222−313, 328−412, and 432−483 and 51 water molecules.

During the course of model building, the published *T. foetus* amino acid sequence (*10*) did not correspond well to the electron density in the area of residues 269−275. Thus, this region of the gene encoding IMPDH was resequenced, and it was found that errors in the initial nucleotide sequence determination had resulted in the mistaken assignment of the amino acid sequence for residues 269−276 and residue 289. The region starting with residue 269 had previously been assigned the sequence SAENYNRL and has now been

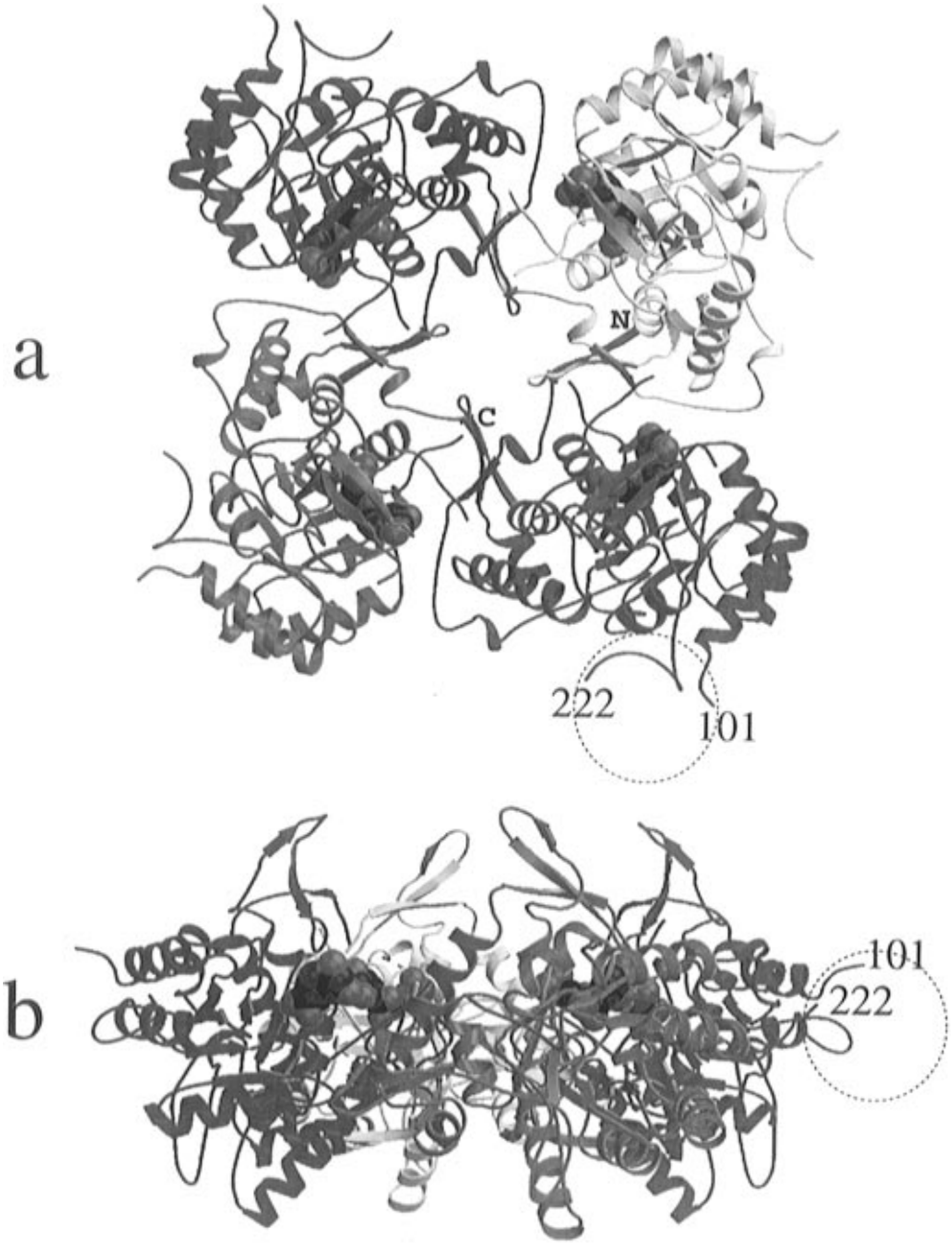


FIGURE 2: Ribbon representations of IMPDH. The tetramer shown parallel (a) and perpendicular (b) to the 4-fold axis. Each subunit is shown in a different color, and a space filling model of XMP is shown in the active site. The N- and C-terminal residues of the model are indicated for the subunit drawn in purple. The subunit interface is formed primarily by hydrophilic contacts and contains a β-sheet interaction between single β-strands from each monomer. A dashed circle represents an approximate location of the second domain (between residues 101 and 222) for which electron density is not observed [figure produced using MOLSCRIPT (*44*) and RASTER3D (*45*)].

correctly identified as WQKITIGW. Residue 289 was corrected from a proline to an alanine. Corrections to the gene sequence have been submitted to GenBank, and the accession number for the sequence will remain L18917.

The XMP complex data were phased with the refined apoenzyme model, and the substrate was located by difference Fourier analysis. The substrate was included in the model along with 48 water molecules and refined at 2.6 Å resolution. Refinement statistics are reported in Table 2. The resulting enzyme complex model includes residues 2−101, 222−313, 325−412, and 432−483. Atomic coordinates have been deposited in the Brookhaven Protein Data Bank (1708475).

For determination of free sulfhydryls, purified protein suitable for crystallization was exhaustively dialyzed against a buffer containing 10 mM Tris-HCl (pH 7.4) and 10% glycerol at 4 °C. Five nanomoles of the protein was diluted into 0.5 mL of dialysis buffer, and the spectrophotometer was blanked with this solution. Fresh DTNB was added to the protein to a final concentration of 1 mM from a 10 mM stock in water. The change in absorbance at 412 nm was measured with a Cary model 219 spectrophotometer (Varian Instruments, Palo Alto, CA). The reaction was complete in 45 min and was repeated three times.

Equilibrium sedimentation analysis was performed using a Beckman XL-A centrifuge. The sample buffer contained 100 mM Tris-HCl (pH 7.4), 100 mM KCl, and 20 mM BME. Three protein concentrations (0.8, 2.0, and 6.0 μM) were monitored at wavelengths of 280 and 230 nm. The samples were spun at 8000 rpm until they reached equilibrium (46

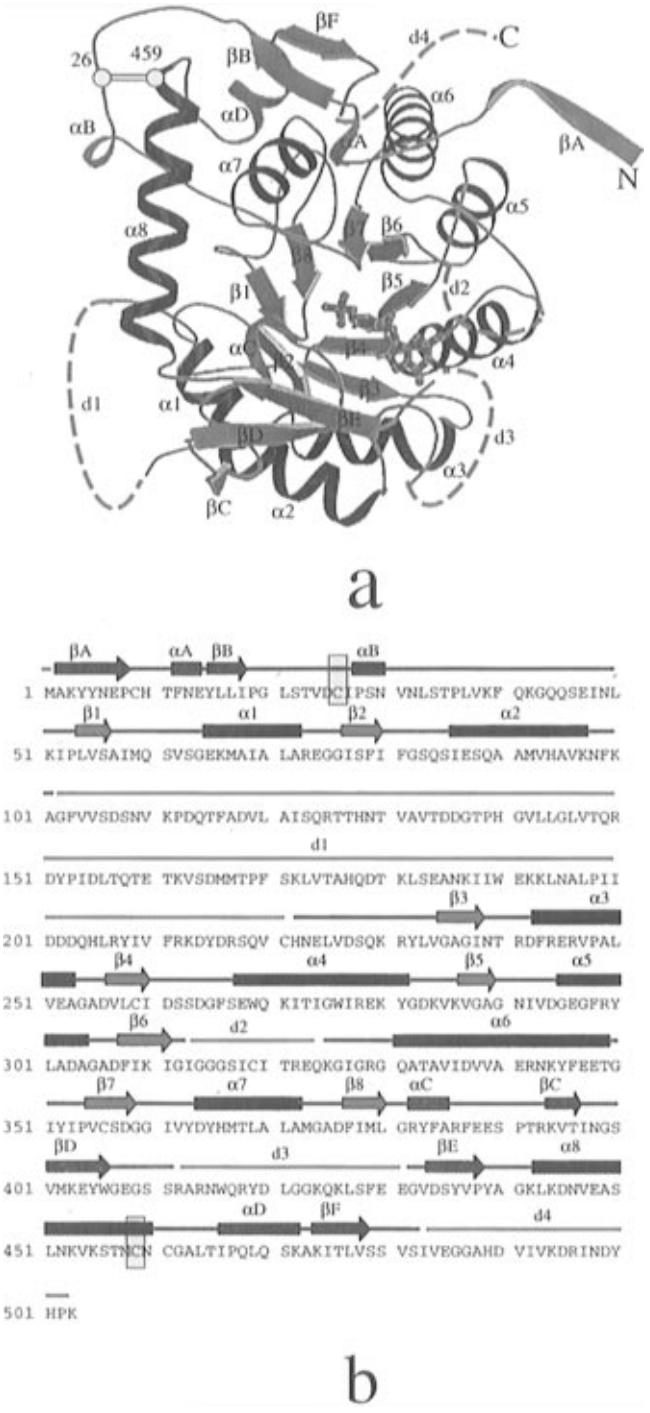


FIGURE 3: Secondary structure elements of *T. foetus* IMPDH. (a) A ribbon diagram viewed approximately along the axis of the parallel α/*β*-barrel, showing XMP bound at the C-terminal ends of the barrel *β*-strands. The strands of the barrel are shown in red, helices of the barrel in blue, and all other main chain residues in green. XMP is shown as a ball-and-stick model. Light blue dashed lines represent the four disordered regions of the protein (labeled d1−d4) but are not meant to imply any structural knowledge of these regions [figure produced using MOLSCRIPT (*44*)]. (b) A diagram showing the correspondence of secondary structure elements to the amino acid sequence of *T. foetus* IMPDH.

h). At each speed, 80 scans were recorded with a step size of 0.003 mm and averaged. These data were analyzed using the Microcal Origin software, version 1.8 that was provided with the instrument and a calculated value for V_{bar} of 0.738. A plot of the apparent molecular weight vs absorbance was generated from the equilibrium gradient at each speed (*33*). Attempts to fit the data as a single species, varying the

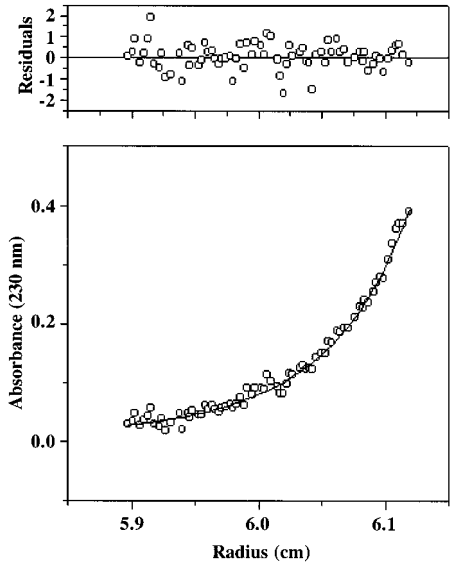


FIGURE 4: Analysis of *T. foetus* IMPDH by sedimentation equilibrium. Data are shown for a protein sample at equilibrium and scanned at 230 nm. The data were fit to a self-associating tetramer−octamer model. The upper panel shows the residual differences between the experimental observations and the best fit curve as a function of radial position.

apparent molecular weight, resulted in highly nonrandom residuals, suggesting that it was necessary to model the system as an associating equilibrium. The best global fit over all concentration ranges provided a goodness of fit of 0.3628 and was achieved by modeling the system as a tetramer−octamer association (*34*) with a dissociation constant of 1.4 ± 0.5 μM.

RESULTS AND DISCUSSION

The structure of the apoenzyme was determined by the multiple-isomorphous replacement (MIR) method (Table 1) and refined at 2.3 Å resolution (Table 2). The crystallographic *R*-factor for the current model is 21.8% (*R*-free = 26.5%) for all reflections between 8.0 and 2.3 Å resolution (Table 2). The model has rms deviations from ideality of 0.011 Å in bond lengths and 1.593° in angles. The model includes 57 water molecules and 329 amino acid residues out of 503 residues in the protein. Despite the high degree of disorder of several parts of the molecule, the quality of the electron density maps attests to the correctness of the structure determination (Figure 1). Furthermore, inspection of the electron density maps indicated that the original sequence determination included several errors that were later confirmed by resequencing. Analysis of the model with PROCHECK (*35*) indicated good stereochemistry (90.8% of the residues are in most favorable regions, 8.8% of residues are in additionally allowed regions). Diffraction data were also collected from IMPDH crystals that were soaked in XMP. This product complex crystal structure, which includes 332 residues and 51 water molecules, has been refined at 2.6 Å resolution to a crystallographic *R*-factor of 20.3% (*R*-free = 26.4%) for all reflections between 8.0 and 2.6 Å resolution (Figure 2).

T. foetus IMPDH is a two-domain protein of 503 amino acid residues. The principal domain includes a parallel eight-stranded *β*/α-barrel and is formed by residues 2−101 and 221−503 (Figure 3). The smaller domain, residues 102−220, is not visible in the electron density maps, presumably due to disorder. An obvious void is available in the crystal

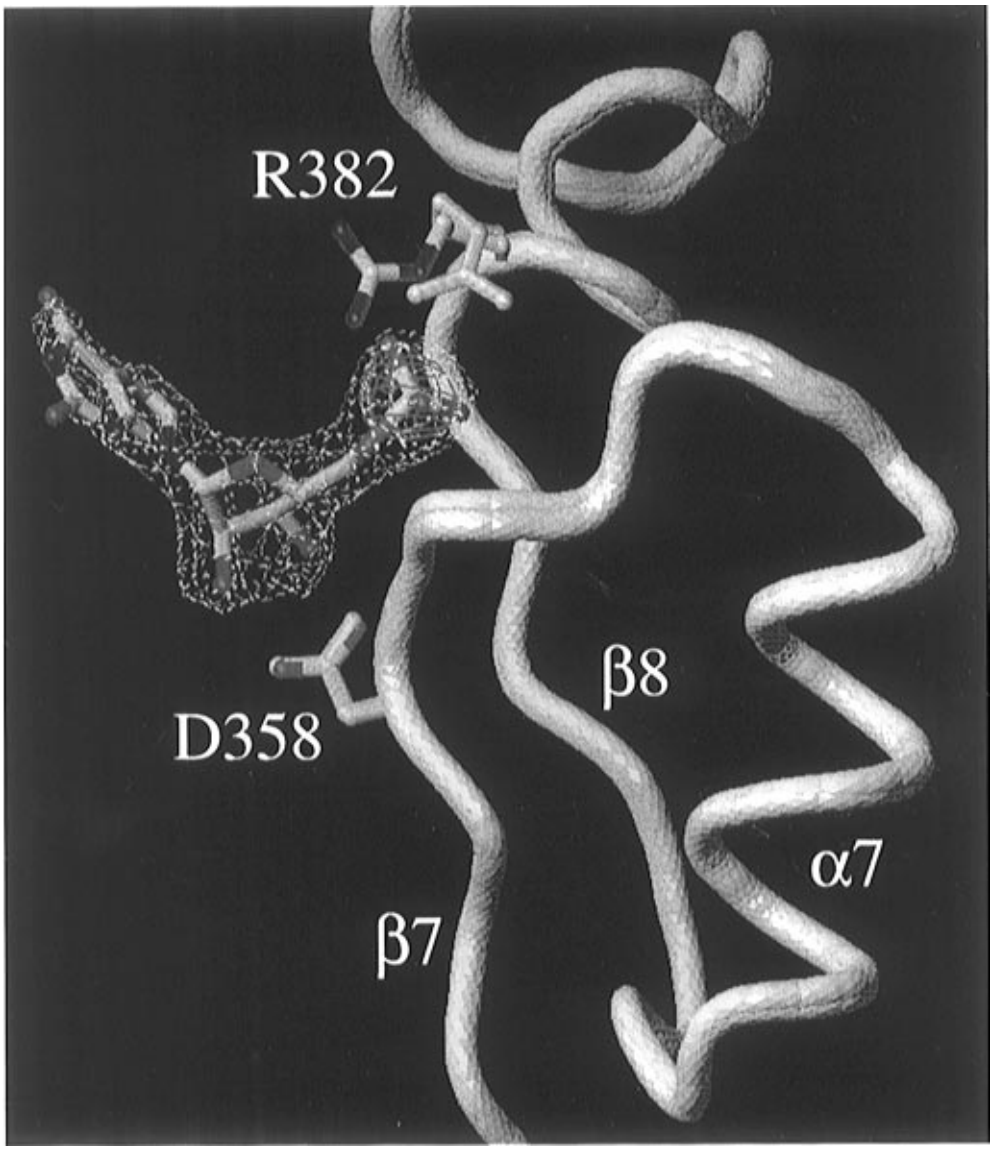
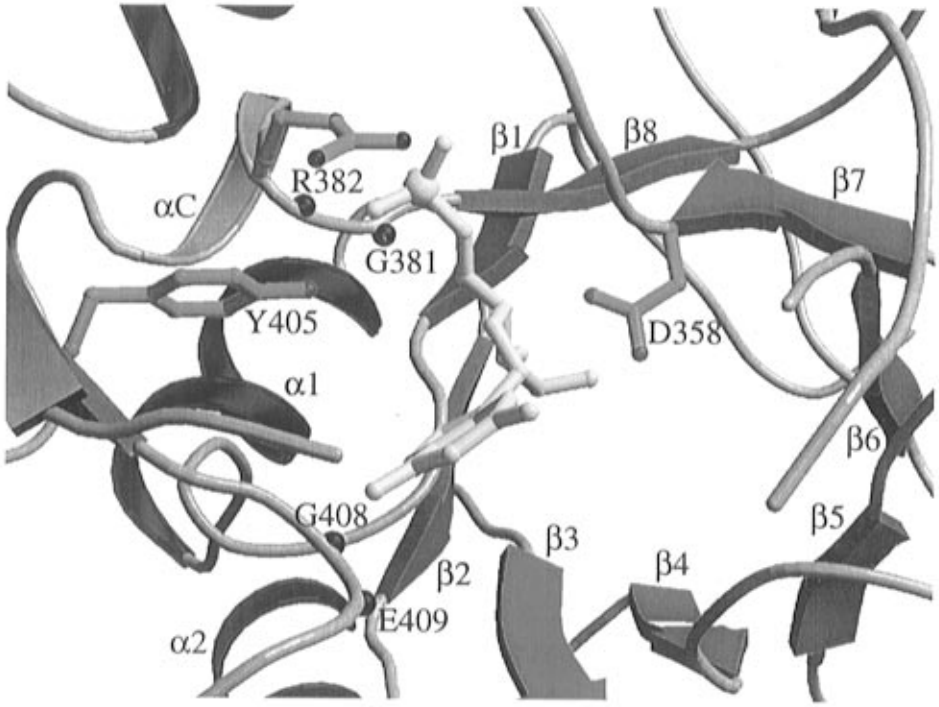


FIGURE 5: Nucleotide binding site of *T. foetus* IMPDH. The positions of the main chain atoms of the apo structure (white ribbon) and the XMP complex model (yellow ribbon) are nearly identical. The refined model of XMP and a white omit map ($F_o - F_c$) calculated at 2.6 Å resolution and contoured at 4.5σ shows the position of XMP in the active site. The pink map is a selenium difference map ($F_o - F_c$) contoured at 5.0σ showing that a sulfate ion occupies the substrate phosphate binding site in the apo structure. The R382 and D358 side chains are shown for the apo (white) and XMP complex (colored) structures [figure produced using O (*31*)].

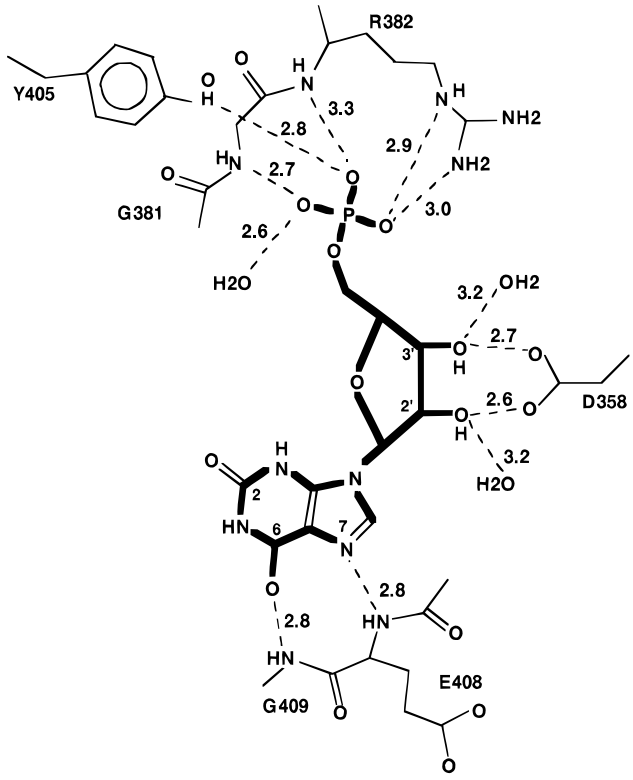
lattice to accommodate a domain of this size. Analysis by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) and electrospray mass spectrometry demonstrates that the crystals contain intact IMPDH with no evidence of proteolysis (data not shown). This second domain is also highly disordered in hamster IMPDH crystals (*23*). The function of the second domain is not known and its sequence is not apparently similar to those of other proteins. Other disordered segments of the IMPDH structure include three flexible loops formed by residues 314−324 (between strand 6 and helix 6 of the barrel), 413−431 (following helix 8 of the barrel), and 484−503 (the C terminus). The apparent disorder of the glycine-rich loop (residues 314−324) that contains the active site nucleophile, Cys-319, suggests that a high degree of flexibility may be needed for the catalytic cycle.

As is commonly seen in β/α-barrel structures, the IMPDH active site is located at the C-terminal end of the β−strands (hereafter termed the C-terminal end of the barrel) (Figure 2). The structure of the barrel domain strongly resembles that of hamster IMPDH (*23*) and spinach glycolate oxidase (*36*, *37*). Furthermore, IMPDH bears strong sequence similarity to GMP reductase and the glycolate oxidase family of proteins, including tryptophan synthase, indole-glycerol-phosphate synthase, flavocytochrome b_2, lactate oxidase, and trimethylamine dehydrogenase (*38*, *39*). Sequence alignment indicates that GMP reductase also resembles the IMPDH barrel domain, but lacks the disordered second domain.

Dynamic light scattering indicated that purified *T. foetus* IMPDH is monodisperse in solution with an apparent molecular mass of 262 kDa (data not shown), in reasonable agreement with the mass of 222 kDa expected for a tetramer. However, *T. foetus* IMPDH has been reported to elute from a gel filtration column as two species of 270 and 340 kDa under low-salt conditions (*6*), indicating that the protein may form higher-order aggregates such as hexamers or octamers under certain conditions. Equilibrium sedimentation experiments provided further evidence of higher-level association (Figure 4), suggesting that dimerization of *T. foetus* IMPDH



a



b

FIGURE 6: Active site region of *T. foetus* IMPDH. (a) A ribbon diagram of the XMP complex model showing amino acid side chains and main chain amides that hydrogen bond with XMP. The numbering of the secondary structure elements corresponds to Figure 3 [figure produced using MOLSCRIPT (*44*)]. (b) A diagrammatic representation of hydrogen bonding between XMP and residues in the active site.

tetramers for forming 444 kDa octamers has an equilibrium dissociation constant of 1−2 μM.

In the crystal, two tetrameric rings pack face to face, forming an octameric assembly of approximately the same dimensions as glycolate oxidase which has also been reported to form tetramers and octamers in solution (*37*). Hamster IMPDH crystals contain a closely similar C_4 tetramer (*23*), although no evidence of octamer formation has been reported. Formation of the *T. foetus* IMPDH cyclically symmetric (C_4) tetramer from four monomer subunits apparently buries a

T.FOETUS (1708475)	353	IPVCS	D	GGIVYDYHMTLALAMGADFIML	GR	YFARFEESPTRKVTINGSVMKE	Y	WGEGS	410
L.ESCULENTUM(GOX) (1063400)	200	IPVFL	D	GGVRRGTDVFKALALGASGIFI	GR	PVVFSLAAEGEAGVKKVLQMLR	D	EFELT	257
SPINACH(GOX) (65974)	280	VPVFL	D	GGVRRGTDVFKALALGAAGVFI	GR	PVVFSLAAEGEAGVKKVLQMMR	D	EFELT	337
PUMPKIN(GOX) (217909)	280	VPVFL	D	GGVRRGTDVFKALALGASGIFI	GR	PVVFSLAAEGEAGVRKVLQMLR	D	EFELI	337
A.VIRIDANS(LOX) (1478355)	291	VPIVF	D	SGVRRGEHVAKALASGADVVAL	GR	PVLFGLALGGWQGAYSVLD...	Y	FQKDL	344
HAMSTER (304517)	359	VPVIA	D	GGIQNVGHIAKALALGASTVMM	GS	LLAATTEAPGEYFFSDGIRLKK	Y	RGMGS	416
HUMAN-II (602458)	359	VPVIA	D	GGIQNVGHIAKALALGASTVMM	GS	LLAATTEAPGEYFFSDGIRLKK	Y	RGMGS	416
HUMAN-I (307067)	359	VPIIA	D	GGIQTVGHVVKALALGASTVMM	GS	LLAATTEAPGEYFFSDGVRLKK	Y	RGMGS	416
M.MUSCULUS (425158)	359	VPVIA	D	GGIQNVGHIAKALALGASTVMM	GS	LLAATTEAPGEYFFSDGIRLKK	Y	RGMGS	416
L.DONOVANI (159361)	355	VPCTA	D	GGLRQVGDICKALAIGANCAML	GG	MLSGTTETPGEYFFKGGVRLKV	Y	RGMGS	412
T.BRUCEI (162136)	353	VPCVA	D	GGLRNVGDVCKALAVGANVAML	GS	MIAGTSETPGEYFFKDGMRLKG	Y	RGMGS	410
E.COLI K12 (146275)	333	IPVIA	D	GGIRFSGDIAKAIAAGASAVMV	GS	MLAGTEESPGEIELYQGRSYKS	Y	RGMGS	390
D.MELANOGASTER (387594)	378	VPVIA	D	GGIQSIGHIVKAIALGASAVMM	GS	LLAGTSEAPGEYFFSDGVRLKK	Y	RGMGS	435
H.INFLUENZAE (1222143)	334	IPVIA	D	GGIRFSGDIAKAIAAGASCVMV	GS	MFAGTEEAPGEIELYQGRAFKS	Y	RGMGS	391
B.SUBTILIS (39959)	336	KTIIA	D	GGIKFSGDITKALAAGGHAVML	GS	LLAGTSESPGETEIYQGRRFKV	Y	RGMGS	393
S.PYOGENES (924848)	338	KTIIA	D	GGIKYSGDIVKALAAGGNAVML	GS	MFAGTDEAPGETEIYQGRKFKT	Y	RGMGS	395
C.GRISEUS (124426)	359	VPVIA	D	GGIQNVGHIAKALALGASTVMM	GS	LLAATTEAPGEYFFSDGIRLKK	Y	RGMGS	416
A.CALCOACETICUS (400057)	331	IPLIA	D	GGIPFSGDMAKAIGAGASTIMV	GS	LLAGTEEAPGEVEFFQGRYYKA	Y	RGMGS	388
B.ARABIDOPSIS (1100063)	350	IPVIA	D	GGISNSGHIVKALVLGASTVMM	GS	FLAGSTEAPGGYEYTNGKRIKK	Y	RGMGS	407
M.LEPRAE (466944)	369	VPVIA	D	GGLQYSGDIAKALAAGASTTML	GS	LLAGTAEAPGELIFVNGKQFKS	Y	RGMGS	426
P.FALCIPARUM (*)	392	TKTIA	D	GGIKNSGNIVKALSLGADFVML	GN	LLAATEESCSEYYFENNVRLKI	Y	RGMGS	449
B.BURGDORFERI (532792)	257	ICIIA	D	GGIRFSGDVVKAIAAGADSVMI	GN	LFAGTKESPSEEIIYNGKKFKS	Y	VGMGS	314
P.FURIOSUS (595287)	329	LYVIA	D	GGIKYSGDIVKAIAAGADAVML	GN	LLAGTKEAPGKEVIINGRKYKQ	Y	RGMGS	386
P.CARINII (1272244)	291	VPTIA	D	GGIENIGHITKALALGASAVMM	GN	LLAGTTESPGQYYYRDGQRLKS	Y	RGMGS	348
S.CEREVISIAE (1363745)	363	IPCMA	D	GGVQNIGHITKALALGSSTVMM	GG	MLAGTTESPGEYFYQDGKRLKA	Y	RGMGS	420
HUMAN(GMPR) (544455)	214	GHIIS	D	GGCTCPGDVAKAFGAGADFVML	GG	MFSGHTECAGEVFERNGRKLKL	F	YGMSS	271
E.COLI(GMPR) (78478)	217	GMIVS	D	GGCTTPGDVAKAF.ARADFVML	GG	MLAGHEESGGRIVEENGEKFML	F	YGMSS	270
A.LUMBRICOIDES(GMPR) (159661)	217	GHVMS	D	GGCTNPGDVAKAFGGGADFVMI	GG	LLAGHDQCGGEVVEKDGKKYKL	F	YGMSS	274

FIGURE 7: Multiple sequence alignment of the phosphate-binding site region of proteins having sequence homology with respect to *T. foetus* IMPDH. The name of the enzyme encoded by the sequence is indicated in parentheses, where GOX = glycolate oxidase, LOX = lactate oxidase, GMPR = GMP reductase, and no designation indicates IMPDH. Residues that coordinate the nucleotide phosphate (Gly-381, Arg-382, and Tyr-405) and ribose (Asp-358) in *T. foetus* IMPDH and the homologous residues of the other sequences are boxed. GenBank accession numbers are indicated in parentheses (*G. Prosise, H., Luecke, and A. A. James, unpublished).

total of 8898 Å^2 of molecular surface area (2224 Å^2 per monomer−monomer interface). The subunit interface involves primarily hydrophilic interactions where the N-terminal β-strand (βA in figure 3) forms a three-stranded antiparallel β-sheet with two strands (βB and βF) of the adjacent subunit (Figure 2). Formation of the octameric complex seen in the crystal buries a total of 20 803 Å^2 of molecular surface area.

The relatively extended N-terminal 50 residues are located on the surface of the barrel and in the subunit interface. This region is linked to helix 8 of the barrel (α8 in Figure 3) through a disulfide bond between Cys-26 and Cys-459. The IMPDH enzymatic activity is typically assayed in a buffer containing 1 mM dithiothreitol (DTT) (*6, 20*). The specific activity of *T. foetus* IMPDH increased more than 10-fold, however, when 20 mM BME was added to the assay buffer (F. G. Whitby, J. A. Huete-Perez, and C. C. Wang, unpublished results). This suggests that reduction of the disulfide linkage may allow the protein to adopt a more active conformation, perhaps by alteration of the nearby subunit interface. Quantification of the free sulfhydryls with 5,5′-dithiobis(2-nitrobenzoate) (DTNB) (*40, 41*) indicated that, of the eight cysteine residues, there are 5.1 ± 0.3 mol of free sulfhydryls per mole of monomer, suggesting that there are six free sulfhydryls per monomer and that the remaining two cysteine residues are occupied in a disulfide linkage. In contrast, all eight of the cysteines of human IMPDH are reported to be in the reduced state (*42*).

Isomorphous crystals were grown from solutions in which the usual 2.2 M ammonium sulfate precipitant was replaced with ammonium selenate. Subsequent difference electron density maps revealed a strong peak near the C-terminal end of barrel strands 7 and 8 (Figure 5), at the site equivalent to the phosphate position of the flavin mononucleotide (FMN) cofactor in the glycolate oxidase structure (*36, 37*). This indicates that a sulfate ion occupies the substrate phosphate binding site in the apo-IMPDH crystal.

The XMP molecule was clearly defined in apo-IMPDH-phased difference maps (Figure 5), with the substrate at the base of a broad active site cavity and the phosphate group of the nucleotide occupying the same position as the bound sulfate ion of the apo structure. Although there is ample space surrounding the substrate for cofactor binding, it is not clear how much of this space is occupied by the flexible active site loops. The apo- and XMP-bound IMPDH structures exhibit comparable levels of disorder.

Crystal packing interactions may hold the active site in a more open conformation than is found in solution, thereby preventing the substrate from ordering the active site. Residues 391−403 form an extended loop at one edge of the active site and protrude approximately 15 Å from the otherwise relatively flat tetramer to form hydrogen bonding interactions with subunits from the other C_4 tetramer of the octameric assembly (Figure 2a). Whereas these interactions may occur in the solution octamer, in the absence of these crystal lattice contacts, this loop might fold down toward the 4-fold axis, partially covering and possibly ordering residues in the active site. Evidence for a conformational change in this loop upon substrate binding is provided by limited proteolysis studies on hamster IMPDH in which the equivalent loop is protected in the presence of substrate, product, and the inhibitor MPA (*43*). On the other hand, the crystal structure of hamster IMPDH in complex with substrate and MPA also has an extended conformation for the loop equivalent to residues 391−403 of the *T. foetus* enzyme and in general displays a similar degree of disorder throughout. This is despite different packing interactions in the hamster IMPDH crystal. Thus, crystal packing

interactions do not appear to fully explain the high degree of disorder in these crystals and the open conformation of the 391−403 loop.

The active site geometry of *T. foetus* IMPDH is, overall, very similar to that of hamster IMPDH (Figure 6). Asp-358 and two water molecules that hydrogen bond to the ribose hydroxyl groups have exact counterparts in hamster IMPDH (Asp-364) (*23*). This aspartate is conserved in all species of IMPDH, in GMP reductase, and in several other members of the glycolate oxidase family (Figure 7) (*39*). The main chain amide nitrogens of Glu-408 and Gly-409 correspond to those of Met-414 and Gly-415 in the hamster enzyme. In *T. foetus* IMPDH, the phosphate oxygens make one hydrogen bond to each of the Tyr-405 hydroxyls, Gly-381 amide nitrogen, and a water molecule. Comparable interactions are seen in the hamster enzyme. The most distinctive residue involved in substrate binding by *T. foetus* IMPDH is Arg-382, which forms three hydrogen bonds to the substrate phosphate through the main chain amide and the guanidinium. The phosphate binding site of spinach glycolate oxidase is structurally very similar, with the amide nitrogen and guanidinium group of an arginine residue hydrogen bonding to the phosphate oxygens (*37*, *38*). Amino acid sequence alignment indicates that coordination of the phosphate group by Arg-382 in the *T. foetus* enzyme is unique among the IMPDHs (*23*). This suggests that arginine-specific reagents or substrate analogs might serve as selective *T. foetus* IMPDH inhibitors.

ACKNOWLEDGMENT

We thank Bill Weis, Anand Kolatkar, and Henry Bellamy for assistance with data collection and use of computer equipment, Ehmke Pohl for help with the use of SHELXS, and Joseph Sin and James McCloskey for mass spectroscopy analysis.

REFERENCES

1. Speer, C. A., and White, M. W. (1991) *Large Anim. Vet. 46*, 18−20.
2. Fitzgerald, P. R. (1986) *Veterinary Clinic of North America: Food Animal Practice*, W. B. Saunders, Philadelphia.
3. Wang, C. C. (1984) *J. Med. Chem. 27*, 1−9.
4. Wang, C. C., Verham, R., Rice, A., and Tzeng, S.-F. (1983) *Mol. Biochem. Parasitol. 8*, 325−337.
5. Somoza, J. R., Chin, M. S., Focia, P. J., Wang, C. C., and Fletterick, R. J. (1996) *Biochemistry 35*, 7032−7040.
6. Verham, R., Meek, T. D., Hedstrom, L., and Wang, C. C. (1987) *Mol. Biochem. Parasitol. 24*, 1−12.
7. Carter, J. B., Franklin, T. J., Jones, D. F., Leonard, B. J., Mills, J. D., Turner, R. W., and Turner, W. B. (1969) *Nature 223*, 848−850.
8. Franklin, T. J., and Cooke, J. M. (1969) *Biochem. J. 113*, 515−524.
9. Wang, C. C., Verham, R., Cheng, H.-W., Rice, A., and Wang, A. L. (1984) *Biochem. Pharmacol. 33*, 1328−1329.
10. Beck, J. T., Zhao, S., and Wang, C. C. (1994) *Exp. Parasitol. 78*, 101−112.
11. Gilbert, H. J., Lowe, C. R., and Drabble, W. T. (1979) *Biochem. J. 183*, 481−494.
12. Wilson, K., Collart, F. R., Huberman, E., Stringer, J. R., and Ullman, B. (1991) *J. Biol. Chem. 266*, 1665−1671.
13. Collart, F. R., and Huberman, E. (1988) *J. Biol. Chem. 263*, 15769−15772.
14. Natsumeda, Y., Ohno, S., Kawasaki, H., Konno, Y., Weber, G., and Suzuki, K. (1990) *J. Biol. Chem. 265*, 5292−5295.
15. Tiedeman, A. A., and Smith, J. M. (1991) *J. Biochem. 97*, 289−293.
16. Carr, S. F., Papp, E., Wu, J. C., and Natsumeda, Y. (1993) *J. Biol. Chem. 268*, 27286−27290.
17. Cantor, C., and Schimmel, P. R. (1980) *The Conformation of Biological Macromolecules*, W. H. Freeman & Co., San Francisco.
18. Hedstrom, L., and Wang, C. C. (1990) *Biochemistry 29*, 849−854.
19. Antonino, L. C., and Wu, J. C. (1994) *Biochemistry 33*, 1753−1759.
20. Huete-Perez, J., Wu, J. C., Whitby, F. G., and Wang, C. C. (1995) *Biochemistry 34*, 13889−13894.
21. Link, J. O., and Straub, K. (1996) *J. Am. Chem. Soc. 118*, 2091−2092.
22. Fleming, M. A., Chambers, S. P., Connelly, P. R., Nimmersgen, E., Fox, T., Bruzzese, F. J., Hoe, F. J., Fulghem, J. R., Livingston, D. j., Stuver, C. M., Sintchak, M. D., Wilson, K. P., and Thomson, J. A. (1996) *Biochemistry 35*, 6990−6997.
23. Sintchak, M. D., Fleming, M. A., Futer, O., Raybuck, S. A., Chambers, S. P., Caron, P. R., Murcko, M. A., and Wilson, K. P. (1996) *Cell 85*, 921−930.
24. Whitby, F. G., Huete-Perez, J., Luecke, H., and Wang, C. C. (1995) *Proteins: Struct., Funct., Genet. 23*, 598−603.
25. Otwinowski, Z. (1993) Oscillation data reduction program, in *Data Collection and Processing* (Sawyer, L., Issacs, N., and Bailey, S., Eds.) pp 56−62, SERC Daresbury Laboratory, Warrington, England.
26. Collaborative Computational Project, No. 4. (1994) *Acta Crystallogr. D 50*, 760−763.
27. McRee, D. E. (1992) *J. Mol. Graphics 10*, 44−46.
28. Sheldrick, G. M. (1991) *Proceedings of The CCP4 Study Weekend*, pp 23−28.
29. Furey, W., and Swaminathan, S. (1990) *American Crystallographic Association Meeting Abstracts 18*, 73.
30. Wang, B. C. (1985) *Methods Enzymol. 155*, 90−112.
31. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A 47*, 110−119.
32. Brunger, A. T. (1989) *Acta Crystallogr. A 45*, 50−61.
33. Voelker, P. J., and McRorie, D. K. (1994) Beckman Instruments, Inc., Fullerton, CA.
34. Robertson, J. G. (1995) *Biochemistry 34*, 7533−7541.
35. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
36. Sandalova, T., and Lindqvist, Y. (1993) *FEBS Lett. 327*, 361−365.
37. Lindqvist, Y. (1989) *J. Mol. Biol. 209*, 151−166.
38. Lindqvist, Y., Branden, C. I., Matthews, F. S., and Lederer, F. (1991) *J. Biol. Chem. 266*, 3198−3207.
39. Bork, P., Gellerich, J., Groth, H., Hooft, R., and Martin, F. (1995) *Protein Sci. 4*, 268−274.
40. Deme, D., Durieu-Trautmann, O., and Chatagner, F. (1971) *Eur. J. Biochem. 20*, 269−275.
41. Riddles, P. W., Blaekeley, R. L., and Zerner, B. (1979) *Anal. Biochem. 94*, 75−81.
42. Antonino, L. C., Straub, K., and Wu, J. C. (1994) *Biochemistry 33*, 1760−1765.
43. Nimmesgern, E., Fox, T., Fleming, M. A., and Thomson, J. A. (1996) *J. Biol. Chem. 271*, 19421−19427.
44. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
45. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D 50*, 869−873.

BI9708850