A 1.4 Å Crystal Structure for the Hypoxanthine Phosphoribosyltransferase of *Trypanosoma cruzi*[†,⊥]

Pamela J. Focia,[‡,§] Sydney P. Craig III,[‡] René Nieves-Alicea,[∥] Robert J. Fletterick,[§] and Ann E. Eakin*[,‡]

Laboratory of Molecular Parasitology & Drug Design, School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599-7360, Department of Biochemistry & Biophysics, University of California at San Francisco, San Francisco, California 94143-0448, and Department of Biochemistry, School of Medical Sciences, University of Puerto Rico, San Juan, Puerto Rico 00936-0467

Received May 6, 1998; Revised Manuscript Received September 1, 1998

ABSTRACT: The hypoxanthine phosphoribosyltransferase (HPRT) from *Trypanosoma cruzi*, etiologic agent of Chagas' disease, was cocrystallized with the inosine analogue Formycin B (FmB) and the structure determined to 1.4 Å resolution. This is the highest resolution structure yet reported for a phosphoribosyltransferase (PRT), and the asymmetric unit of the crystal contains a dimer of closely associated, nearly identical subunits. A conserved nonproline cis peptide in one active-site loop exposes the main-chain nitrogen to the enzyme active site, while the adjacent lysine side chain interacts with the other subunit of the dimer, thereby providing a possible mechanism for communication between the subunits and their active sites. The three-dimensional coordinates for the invariant Ser103−Tyr104 dipeptide are reported here for the first time. These are the only highly conserved residues in a second active-site loop, termed the long flexible loop, which is predicted to close over the active site of HPRTs to protect a labile transition state [Eads et al. (1994) *Cell 78*, 325−334]. This structure represents a major step forward in efforts to design/discover potent selective inhibitors of the HPRT of *T. cruzi*.

Metabolic differences between host and pathogen have been the subject of research for many diseases which are caused by parasites (*1*), including Chagas' disease, that affects more than 24 million people in Central and South America (*2*). Currently, there is no satisfactory drug for the treatment of Chagas' disease, resulting from an infection of the protozoan parasite *Trypanosoma cruzi*, and new therapies are urgently needed.

Mammalian cells synthesize metabolic purine nucleotides by two routes, the multistep de novo pathways and the more economical salvage pathways (*3*, *4*). In contrast, many parasites are incapable of de novo biosynthesis and are forced to rely on salvage pathways to obtain the purine nucleotides needed for cellular metabolism (*1*). In most parasites, including *T. cruzi*, hypoxanthine and guanine bases (H, G) are converted to inosine monophosphate (IMP) or guanosine monophosphate (GMP), respectively, by a single enzyme—the hypoxanthine phosphoribosyltransferase (HPRT;[1,2] EC 2.4.2.8; inosine monophosphate:pyrophosphate phosphoribosyltransferase). HPRT catalyzes the salvage of hypoxanthine and guanine bases via combination with 5′-phosphoribosyl-α-1′-pyrophosphate (PRPP), to form the corresponding nucleoside monophosphate, IMP or GMP, respectively, and inorganic pyrophosphate (PPi) (Scheme 1).

HPRTs belong to a family of phosphoribosyltransferases (PRTs) which catalyze the transfer of the ribose 5′-phosphate group from PRPP to a variety of secondary substrates in the presence of Mg^{2+}, with release of inorganic pyrophosphate. PRTs are found in the de novo purine nucleotide biosynthetic pathway, in the salvage pathways of purine, pyrimidine, and pyridine ribonucleotide metabolism and also in biosynthetic pathways for histidine and tryptophan (*7*). Two classes of PRTs have been identified structurally *(8)*. Type I PRTs include HPRTs, orotate PRTs (OPRTs), and glutamine-PRPP amidotransferases (GPATs), while quinolinic acid PRT (QAPRT) has a different protein fold and is referred to as a type II PRT.

[†] This work was supported in part by NIH Grants AI34326 (to S.P.C. III) and AI38919 (to A.E.E.).

[⊥] The coordinates and observed structure factor amplitudes for the HPRT of *T. cruzi* have been deposited in the Brookhaven Protein Data Bank under identification codes 1tc1 and R1tc1sf, respectively.

* Corresponding author. Phone: (919) 966-6422. Fax: (919) 966-6919. E-mail: eakin@unc.edu.

[‡] University of North Carolina.

[§] University of California at San Francisco.

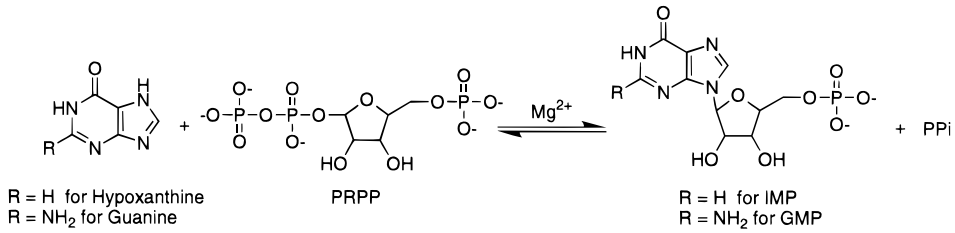
[∥] University of Puerto Rico.

[1] HPRT, hypoxanthine phosphoribosyltransferase; PRT, phosphoribosyltransferase; PRPP, 5′-phosphoribosyl-α-1-pyrophosphate; IMP, inosine 5′-monophosphate; GMP, guanosine 5′- monophosphate; OPRT, orotate phosphoribosyltransferase; XPRT, xanthine phosphoribosyltransferase; GPAT, glutamine phosphoribosylpyrophosphate amidotransferase; QAPRT, quinolinic acid phosphoribosyltransferase; ESMS, electrospray mass spectrometry; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; FmB, Formycin B, $C_{10}H_{12}N_4O_5$; MES, 2-[*N*-morpholino]ethanesulfonic acid; PEG6K, poly(ethylene glycol) MW 6000; NCS, noncrystallographic symmetry; rmsd, root-mean-square deviation.

[2] In recent years, there has been a trend for investigators to include all of the purines salvaged by the enzyme in the name of the purine phosphoribosyltransferase (i.e. HPRT, HGPRT, or HGXPRT). However, many bacterial enzymes (usually referred to as HPRTs) salvage guanine in addition to hypoxanthine. Also, several enzymes referred to as HPRTs or HGPRTs (i.e. the human enzyme) salvage xanthine, albeit at low, but detectable levels (*5*, *6*). Thus, for this paper, HPRT is used in referring to all enzymes that are genetically descended from the bacterial HPRT (encoded by the *hpt* locus).

10.1021/bi981052s CCC: $15.00
Published on Web 10/10/1998

Scheme 1: HPRT-Catalyzed Reaction



Kinetic studies of the HPRT-catalyzed reaction reveal an ordered sequential mechanism, in which PRPP binds to the enzyme first followed by the purine base (*9*−*11*). The release of products is also ordered with pyrophosphate leaving first followed by the ribonucleotide. The apparent S_N1 or S_N2 reaction (*11*) proceeds with inversion of stereochemistry at the anomeric ribose carbon C1′, and the transition state is predicted to involve a charged oxocarbonium ion (*12*). Successful completion of the chemistry requires that the transition-state species be sequestered from bulk solvent to prevent hydrolysis by water (*12*, *13*). Type I PRTs are thought to accomplish this by a long flexible loop, which in a closed conformation protects the active site from solvent during the transition state, while open conformations allow substrate binding and product release (*13*−*15*).

The crystal structure for the human HPRT provided the first structural information for an HPRT and revealed the molecular basis of several inherited mutations in the enzyme that result in human disease (*13*). A goal in solving structures of the HPRTs of parasites is to identify differences between the structures of the enzymes from the host and pathogen that could be exploited in the design of selective inhibitors. Although the HPRTs from several parasites have been proposed as attractive metabolic targets (*16*, *17*), a potent inhibitor of an HPRT has not yet been reported for the chemotherapeutic treatment of any disease caused by a parasite (*18*). A more complete understanding of the structures and chemistry of these enzymes should facilitate the achievement of this goal.

T. cruzi is a single-celled parasite belonging to the order Kinetoplastida. This order includes parasites responsible for many devastating diseases of humans, including Leishmaniasis and African Sleeping Sickness, as well as Chagas' disease. The HPRTs from these parasites are more identical in sequence to one another (47−64%) than they are with the human HPRT (30−35%). Thus, the crystal structure of the trypanosomal HPRT may facilitate the solutions of the crystal structures of HPRTs from other organisms belonging to the same order by serving as a phasing model for molecular replacement.

The HPRT of *T. cruzi* has been expressed in bacteria, purified, and kinetically characterized (*19*). The protein was crystallized and the structure solved and refined to 1.4 Å resolution using diffraction data measured at a synchrotron source from a single cryofrozen crystal. This is the first structure to be reported for an HPRT from an organism of the order Kinetoplastida and is the highest resolution structure reported for any PRT. The structure provides new information for reevaluating the functional roles of conserved and invariant residues and details interactions which may provide a mechanism for communication between subunits in dimers of the enzyme.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Crystallization. The recombinant HPRT of *T. cruzi* was overexpressed in *Escherichia coli*, purified and preliminary crystallization conditions were reported (*19*). After this initial report, limited proteolysis was shown to truncate the enzyme at the C-terminus (*20*). The catalytic activity of the truncated HPRT was indistinguishable from wild-type, but this form of the enzyme produced crystals belonging to a new monoclinic space group with superior diffraction properties (*20*). Purification procedures and crystallization conditions for generating the monoclinic crystals were identical to those used to crystallize the full-length protein in the previously reported trigonal crystal form (*19*). Electrospray mass spectrometry (ESMS) analysis revealed that up to 22 amino acids were removed from the C-terminus of the trypanosomal enzyme from which the monoclinic crystals were grown (*20*). The deleted residues comprise a significant portion of a C-terminal extension that is not present in the HPRTs of other species (alignment of Genbank data, not shown). The refined structure for the enzyme in the monoclinic crystals (presented herein) does not allow the location of protein in the electron density maps beyond residue 190. This position corresponds to just before the C-terminus of the human enzyme but is 31 amino acids away from the C-terminus of the full-length trypanosomal enzyme, as deduced from the aligned amino acid sequences. Monoclinic crystals which diffract to high resolution can be reliably reproduced with either the truncated or full-length protein.

Purified recombinant enzyme was exchanged into 20 mM Bis-Tris buffer at pH 7.0 and 6 mM $MgCl_2$, using a Centriprep10 filtration unit, and concentrated to 10−15 mg/mL, as determined using a Bradford protein assay (*21*). The protein was incubated with the ligand for 30 min prior to setting up the crystallizations. Crystals were grown by hanging drop vapor diffusion in the presence of 5−10 mM FmB, from 15 to 20% poly(ethylene glycol) MW6000 (PEG6K) in 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer over a pH range 5.5−6.0. The nucleoside analogue FmB was initially used as part of a phasing strategy because it is available in an iodinated form, although this was not successful, and may be a reflection of the relatively poor binding of FmB to the *T. cruzi* HPRT, evident in a IC_{50} of greater than 2 mM (*19*). Cocrystallization of the enzyme with FmB produced the best diffracting crystals of the initial trigonal crystal form (*19*) compared to either GMP or apo-enzyme.

Crystallographic Methods. Crystals were harvested into the precipitating solution described above, elevated in PEG concentration, and transferred in several steps into this solution with increasing amounts of glycerol, to a final

Table 1: Crystallographic Summary

	crystal 1	crystal 2
data measurement		
space group	$P2_1$	$P2_1$
a (Å)	39.4	39.5
b (Å)	93.6	94.1
c (Å)	52.2	52.1
β (deg)	95.0	94.75
resolution range (Å)	30−2.2 (2.3−2.2)[a]	10.0−1.41 (1.44−1.41)
completeness (%)[b]	100 (100)	97.7 (81.9)
R_{sym}[c]	0.072 (0.25)	0.061 (0.196)
$\langle I/\sigma(I)\rangle$	14.3 (3.5)	21.3 (4.7)
unique reflections	19 087	71 678
redundancy[d]	4.8	3.8
refinement		
resolution range (Å)		10.0−1.41
R_{cryst}[e]		0.193
R_{free}[f]		0.224
protein atoms (no.)		3085
solvent atoms (no.)		175
av B-factor (Å^2)		15.6
rmsd from target values		
bond lengths (Å)		0.010
bond angles (deg)		1.48
torsion angles (deg)		23.85

[a] Values in parentheses are the highest resolution bin for each data set. [b] Completeness: fraction of theoretically possible reflections observed at least once. [c] R_{sym}: $\sum|I_h - \langle I_h\rangle|/\sum I_h$ where $\langle I_h\rangle$ is the average intensity over symmetry equivalents. [d] Redundancy: average number of observations of each unique reflection. [e] R_{cryst}: $\sum|F_o - F_c|/\sum F_o$ calculated for all $F_o > \sigma(F_o)$. [f] R_{free}: R_{cryst} calculated for 10% of reflections omitted from the refinement.

glycerol concentration of 25%. Too rapid transfer caused diminished diffraction properties. The crystals were mounted in nylon loops and frozen directly in the nitrogen stream just prior to measurement of the data. Two data sets from monoclinic crystals were used in the structure solution. For the first data set, X-ray diffraction data were measured using a rotating anode CuKα source (50 kV, 300 mA) equipped with an R-axis IIc image plate system. A 100% complete data set to 2.2 Å was used to solve the structure and during the early rebuilding stages. The second data set was measured at the Stanford Synchrotron Radiation Laboratory beamline 7-1 (SSRL BL7-1) using a wavelength of 1.08 Å and Mar image plate detector system. Due to crystal orientation, cell dimension and time constraints, nearly complete data were measured to only 1.4 Å resolution in a single 180° sweep. After examination of the data, it was clear that the crystal diffracted well beyond 1.4 Å resolution. The statistics for the two data sets are summarized in Table 1. All data were processed using the HKL package (*22*).

The crystal structure of the HPRT from *T. cruzi* was solved by molecular replacement, using as the search probe a monomer of the HPRT of *Tritrichomonas foetus* (*23*), with which it shares ∼30% sequence identity. The solution obtained using AMoRe (*24*) and X-ray data from 15 to 3.5 Å resolution had a correlation coefficient of 27.5 and an R-factor of 49.1%. The asymmetric unit contains the proposed functional dimer and is analogous to the smaller *T. foetus* dimer; however, they differ by a small relative rotation of the monomers. The root-mean-square deviations (rmsds) for C_α atoms in common between the HPRTs from *T. foetus* and *T. cruzi* are 1.14 Å for the independent monomers and 1.64 Å for the dimers (compared to 1.48 and 1.78 Å for the human and *T. cruzi* monomers and dimers, respectively). The *T. cruzi* and *T. foetus* enzymes are the first HPRT crystal structures which do not form tetramers in the crystals (*23*).

The initial molecular replacement model was truncated to polyalanine and subjected to rigid body refinement using X-PLOR (*25*). The noncrystallographic symmetry (NCS) operator was redetermined using O (*26*). The model was refined using positional and simulated annealing protocols with X-PLOR imposing strict NCS constraints, to an R_{free} of 48.1% and R_{cryst} of 39.6% for data from 6.0 to 2.5 Å resolution. The $2F_o - F_c$ electron density maps were remarkably improved by NCS-averaging and phase refinement using the RAVE package (*27*). Over 50% of the sequence of the HPRT of *T. cruzi* could be assigned in the first pass of model rebuilding, and over 85% of the mainchain accounted for.

At this time, data were measured at the SSRL BL7-1 where the crystals diffracted far better than anticipated. The model was placed into the high-resolution data set, followed by rigid body refinement, and NCS-constrained simulated annealing, positional and isotropic B-factor refinement against data ($F > 1\sigma$) from 6.0 to 2.0 Å resolution, followed by additional fitting to averaged and unaveraged $2F_o - F_c$ electron density maps. The NCS constraints were removed, a bulk solvent correction added, and all data from the 10 Å low-resolution limit to the 1.41 Å high-resolution limit of the data set were included in the refinement protocol. The pyrazolopyrimidine group of the nucleoside analogue FmB and a MES buffer molecule were fit to the density observed in the two active sites. Water molecules were added to $F_o - F_c$ density peaks (greater than 3σ) where appropriate hydrogen bond donors and acceptors were present in the model. Both active sites of the HPRT of *T. cruzi* contain several water molecules which interact with residues from loops I and IV, although some are partially disordered.

The final model consists of residues 5−190 in the first subunit, and 5−82 plus 92−188 in the second subunit, which is missing four residues in the long flexible loop. Each active site contains one pyrazolopyrimidine group contributed by the FmB and one MES buffer molecule. Both subunits have disordered C-termini and the lack of clear density prohibits tracing the chain further, thus the proposed cleavage site for this fragment of the enzyme cannot be visualized in this structure. The two subunits are nearly identical with an overall rmsd of 0.391 Å for C_α atoms in common. The final R_{free} and R_{cryst} are 22.4 and 19.3%, respectively. Refinement statistics are included in Table 1. In the protein model of the HPRT of *T. cruzi*, 93.5% of residues fall into the most favored regions of the Ramachandran plot, and the additional allowed regions account for the remainder as analyzed by the program PROCHECK (*28*). The coordinates and structure factors for the HPRT of *T. cruzi* have been deposited with the Brookhaven Protein Data Bank (1tc1 and R1tc1sf).

RESULTS AND DISCUSSION

Subunit Structure. The crystal structure of the HPRT from *T. cruzi* was solved by molecular replacement using a monomer of the HPRT of *T. foetus* (*23*) as the search probe, and refined. The asymmetric unit of the crystal contains two subunits which form the functional dimer presumed to be active in solution (*23*). The HPRT of *T. cruzi* is an α/β protein composed of two domains, with the active site formed

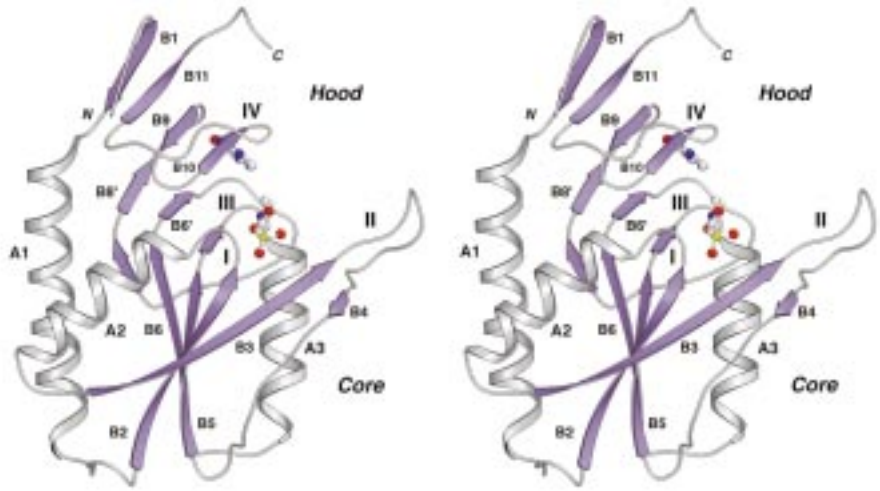


FIGURE 1: The monomer of the HPRT of *T. cruzi*. Secondary structure elements and the four active site loops are labeled. Ligands are represented as ball-and-stick models. Active site loop I is between strand B2 and helix A2, loop II between strands B3 and B4, loop III between strand B5 and helix A3 and loop IV between strands B9 and B11. Loops I−III are contributed by the core domain and loop IV by the hood domain. Secondary structure elements include these residue numbers in the *T. cruzi* sequence: B1, 7−10, A1, 15−32, B2, 44−50, A2, 54−67, B3, 72−81, B4, 93−96, B5, 106−115, A3, 118−128, B6, 134−143, B7, 147−149, B8, 155−159, B9, 164−166, B10, 172−174, B11, 182−185.

by four highly conserved loops (I−IV) and located in a cleft between the domains (Figure 1). All type I PRT crystal structures have shown a two domain architecture with structurally related core domains and divergent hood domains. The conserved core domains bind the primary PRT substrate PRPP, and the hood domains bind the secondary substrate (a purine or pyrimidine base, or glutamine) of each enzyme (*13*, *14*, *29*). The core domain of the HPRT of *T. cruzi* consists of a central β-sheet with five parallel strands (from left to right in Figure 1) (B8, B6, B5, B2, and B3) and one antiparallel edge strand (B4), with helices on both sides of the sheet. This fold and the location of phosphate-binding active-site loops provide some similarity with other mononucleotide binding proteins (*30*). Strands from the second half of the core sheet make up a short three-member parallel sheet which forms a ledge roughly perpendicular to the core sheet (strands B8′, B6′, and B5′), bisecting the two domains (Figure 1). This substructure may play an intermediary role between domains; its side chains participate in binding both the base and sugar portions of a bound nucleotide product which spans the domain interface (*13*, *23*), and this sheet includes several highly conserved residues. The hood domain stacks above the β-ledge and consists of a small antiparallel sheet (strands B1, B11, B9, and B10) (Figure 1).

The HPRT Active Site. Table 2 lists highly conserved residues and 13 residues found to be invariant in an alignment of more than 20 HPRT sequences from diverse species and compares the numbering systems for the human and trypanosomal HPRT sequences. The aligned protein sequences of 10 parasite HPRTs share from 28 to 49% identity with the human HPRT (data from GenBank entries; not shown) with much higher sequence conservation in the four active site loops (I−IV). Three of the four HPRT active-site loops (I, II, and III) are contributed by the core domain, while loop IV is contributed by the hood domain (Figure 1). Two of the core domain loops (I and III) are formed by $\beta-\alpha-\beta$ structural motifs with an active-site loop located between the first strand and the helix; B2−A2−B3 forms loop I, and B5−A3−B6 forms loop III (Figure 1). These two loops, also well conserved across the type I PRTs of known structure, are somewhat reminiscent of phosphate-binding loops found in other nucleotide binding proteins in sequence and function (*30*), but the arrangement of the two phosphate-binding loops is unique to the type I PRTs and functions to bind PRPP via its phosphate and pyrophosphate groups. The analogous loops in other type I PRTs form multiple hydrogen bonds with the phosphate groups located on either side of the substrate PRPP (*15*, *31*).

Table 2: Homologous Residues in the HPRTs of humans and *T. cruzi*[a]

human	*T. cruzi*	location
Pro24	Met1	N-terminus[b]
Val66	Val50	strand B2
Leu67	**Leu51**	**loop I**
Lys68	Lys52	loop I
Gly69	**Gly53**	**loop I**
Phe73	**Phe57**	**Helix A2**
Asp97	Glu75	strand B3
Phe98	Phe76	strand B3
Ser103	**Ser81**	**loop II**
Tyr104	**Tyr82**	**loop II**
Asp119	Asp97	strand B4
Leu130	**Leu108**	**strand B5**
Glu133	**Glu111**	**loop III**
Asp134	**Asp112**	**loop III**
Asp137	**Asp115**	**loop III**
Thr138	Thr116	loop III
Thr141	Thr119	loop III
Lys165	**Lys143**	**strand B6′**
Phe186	Phe164	loop IV
Gly189	**Gly167**	**loop IV**
Asp193	**Asp171**	**loop IV**
Glu196	Asp174	loop IV
Arg199	**Arg177**	**loop IV**

[a] The 13 invariant residues in HPRT sequences reported to GenBank are highlighted in bold, the remaining residues, with the exception of the N-terminal residue, are highly conserved. [b] Not observed in this structure.

Loop I, located between strand B2 and helix A2 in the *T. cruzi* HPRT structure (Figures 1 and 2, top), contains a

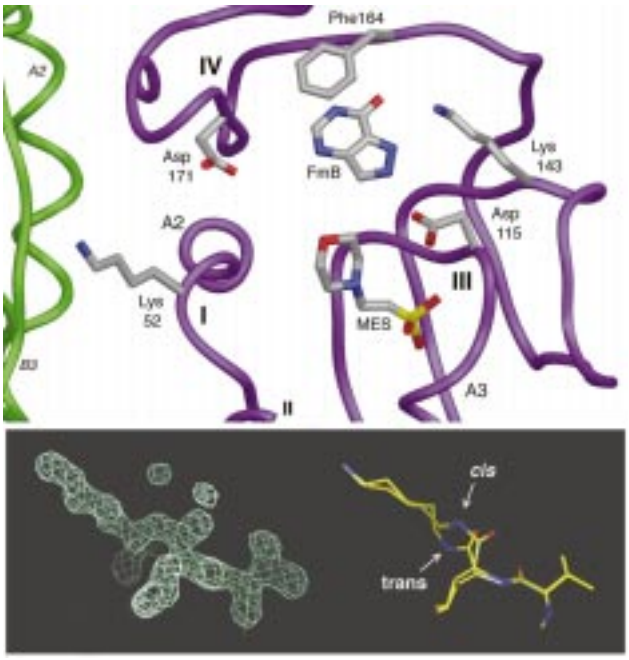


FIGURE 2: (Top) The active site of the HPRT of *T. cruzi* viewed perpendicular to the orientation shown in Figure 1. The alpha carbon trace of each subunit of the dimer is colored separately. Ligands and several conserved or invariant active site residues are displayed. (Bottom) The cis peptide in active site loop I. Omit electron density from loop I shows unambiguously the geometry of the cis peptide bond. The 1.4 Å electron density map was calculated using $F_o - F_c$ coefficients and contoured at 3 times the standard deviation of the map (3σ). The stick figure illustrates the different geometries that would be observed for either a cis or trans peptide bond in loop I. The geometry of the cis peptide bond exposes the mainchain nitrogen of Lys52 into the active site, while the side chain projects into the dimer interface.

conserved, nonproline cis peptide bond. Electron density and geometry for the cis peptide in the trypanosomal HPRT structure are shown in Figure 2, bottom. Nonproline cis peptides are extremely rare in published crystal structures (<0.05%) (*32*). When observed in other enzymes, structural roles assigned to cis peptides have included metal-binding sites and active sites (*32*−*36*). A cis peptide in the analogous active-site loop has been observed in the HPRT from *T. foetus* and the XPRT from *E. coli*, as well as the other type I PRTs, OPRT and GPAT (*15*, *23*, *37*, *38*). These observations suggest that the cis peptide in active-site loop I is structurally and functionally significant. Although all type I PRT crystal structures have not identified the cis peptide, this is likely due to the lower resolution limits of the diffraction data used to solve some of these structures. In HPRTs, the residue preceding the cis peptide is always leucine (Leu51), and the amino acid downstream is always glycine (Gly53) (glycine is also observed in loop I in OPRT). The cis peptide bond and the neighboring glycine allow a tight turn and may be involved in positioning helix A2 toward the active site (Figure 2, top) so that its dipole could interact favorably with a bound PRPP pyrophosphate group. In this structure of the *T. cruzi* enzyme, the lysine side chain adjacent to the cis peptide (Lys52) projects into the dimer interface and forms a hydrogen bond to the carbonyl group of Met74, located in strand B3 of the opposing subunit (see description below).

Loop III (B5, B5′, and the beginning of helix A3) contains the characteristic region of conserved sequence in type I PRTs referred to as the PRPP fingerprint, which encircles the active site (Figures 1 and 2, top). The sequences ^{107}VLIVEDIVDTALT^{119} in the HPRT of *T. cruzi* and ^{129}VLIVEDIIDTGKT^{141} in the human HPRT, differ primarily from the PRPP fingerprint sequences of OPRTs and GPATs (*13*, *14*, *29*) by the presence of Asp115 (human Asp137), which is invariant in HPRTs (Figure 2, top) (Table 2). Eads et al. originally hypothesized that Asp137 acts as a catalytic general base which aids in deprotonating the purine base to activate it as a nucleophile (*13*), and more recent kinetic and mutagenic studies have confirmed this predicted role for Asp137 in the human HPRT (*39*). The invariant carboxylic acid pair, Glu111−Asp112 (Asp−Asp in OPRT and GPAT), emerges from the hydrophobic floor and back wall of the active site provided by the core and ledge sheets, and has been shown to interact with either bound divalent metals, and/or the ribose hydroxyls of bound ligands in type I PRT structures (*13*−*15*, *23*, *31*, *40*). The PRPP fingerprint concludes with a loop that surrounds and binds a 5′-monophosphate group, composed of ^{116}TALT^{119} in the HPRT of *T. cruzi* and ^{138}TGKT^{141} in the human HPRT. Slight differences in sequence are found for this part of loop III in HPRTs, the first threonine is sometimes a serine, and the glycine is sometimes an alanine, while across the type I PRT families, two threonines and a glycine residue are usually present, but their order varies. This loop participates in up to seven hydrogen bonds with the 5′-phosphate group of either bound substrate or product, through main-chain nitrogens and threonine side-chain oxygens in type I PRT structures (*13*−*15*, *23*, *31*) and in the structure of the HPRT of *T. cruzi*, this loop binds in a fashion similar to a sulfate group from a MES buffer molecule used in crystallization of the enzyme (discussed below).

The remaining active-site loop, contributed by the core domain, is the long flexible loop (loop II, between strands B3 and B4, Figure 1). With the exception of the invariant Ser−Tyr dipeptide (Ser103−Tyr104 in the human HPRT; Table 2), there is little similarity in this loop among HPRT sequences. A mutation in the human enzyme which results in gout, S103R, is responsible for a 75-fold increase in the K_M for binding hypoxanthine (*41*), suggesting this residue might interact directly with purine substrates. Site-directed mutagenesis studies of the analogous Ser−Tyr dipeptide in an HPRT from *Leishmania donovani* further demonstrated the importance of these residues in catalysis by HPRTs (*42*). Although the HPRT loop II sequences differ after the Ser−Tyr dipeptide, the remaining sequence generally includes 1 or 2 charged residues, 2−4 polar residues, and often a glycine or two (alignments of data from Genbank entries; not shown). The flexible loop is poorly ordered in many HPRT crystal structures (*13*, *23*, *37*, *40*), preventing the localization of the loop's entire sequence in the electron density maps of these previous structures. In the structure of the HPRT from *T. cruzi* reported herein, well-defined electron density is present for the entire flexible loop in one subunit of the dimer and allows modeling of all residues for loop II. This is partly a consequence of stabilizing contacts with other dimers in the crystal lattice, and the position of the loop is several angstroms away from its predicted location during the transition state (Figure 1). Loop II was partially resolved in one monomer of the apo *Toxoplasma gondii* HPRT tetramer, but coordinates for the invariant Ser−Tyr dipeptide were not

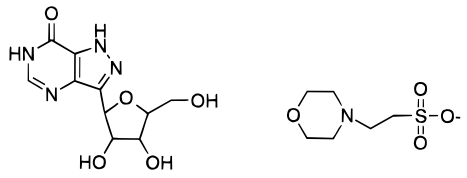
reported (*40*). Similar to the structure reported here, the remainder of the loop in the structure of the HPRT of *T. gondii* was not close enough to the active site to sequester it from solvent. The position of loop II in both structures is consistent with the characters of their active sites; the apo enzyme of *T. gondii* and the FmB/MES-bound enzyme of *T. cruzi* would be expected to have open loop II conformations. A closed conformation of loop II would be expected only if both substrates for either the forward or reverse reaction were present in the active site, as a means to protect the active-site chemistry.

A recent crystal structure of the *E. coli* GPAT with two substrate analogues bound confirms the closing of the flexible loop over the active site of this type I PRT. The closure of the active site was proposed by Krahn et al. to be driven by PRPP binding through the substrate's complementation of charged loop side chains in the closed form (*15*). Loop II in HPRTs similarly contains polar residues which might interact with PRPP atoms in the closed form. However, the invariant Ser−Tyr dipeptide in HPRTs and general sequence dissimilarity for the long flexible loop across the type I PRTs indicate that the precise geometry for shielding the active site may differ among the various enzymes.

The fourth active site loop (loop IV) is contributed by the hood domain (Figures 1 and 2, top). The type I PRT hood domains are involved in binding the various secondary substrates, and as the secondary substrates differ between the HPRTs, OPRTs and GPATs, so do the sequences and structures of their hood domains. However, loop IV is conserved in sequence among HPRTs (Table 2), and in both HPRTs and OPRTs, this domain contains a similarly located aromatic residue (Phe164 in the HPRT of *T. cruzi*) (Figure 2, top) which forms a $\pi-\pi$ stacking interaction with the aromatic ring of the bound substrate or product (*13*, *14*, *23*, *31*). Other invariant residues located in the HPRT loop IV include Gly167, Asp171, and Arg177 (human Gly189, Asp193, and Arg199) (Table 2).

The Active Site of the HPRT of T. cruzi. The HPRT from *T. cruzi* was crystallized in the presence of the inosine analogue, Formycin B (FmB; see structure 1). Figure 3, top

Structure 1: Active-Site Ligands



FmB MES

shows the 1.4 Å $2F_o - F_c$ electron density in the *T. cruzi* HPRT active site, calculated as an omit map for the ligands. The protein is well-ordered throughout, and in active site loops I, III, and IV, all side chains are clearly defined in the electron density maps. The density for the pyrazolopyrimidine group of the FmB bound in the active site's purine-binding pocket is also clearly defined, located beneath Phe164 (Figure 3, top); however, the carbon at purine position 9 and adjacent ribose ring are disordered, possibly a reflection of the poor binding of FmB to the purified *T. cruzi* HPRT (IC_{50} of greater than 2 mM) (*19*). Electron density for a second ligand was also observed bound in the two independent *T. cruzi* HPRT active sites, in the region described as the 5′-monophosphate binding loop (the end of loop III). A peak nine times the standard deviation of the electron density map corresponds to the sulfate group of a bound MES buffer molecule (see structure 1), and the density trailing from the sulfate peak corresponds to part of its hydrocarbon chain (Figure 3, top).

There is a 1.5 Å greater distance between the FmB carbonyl oxygen at purine ring position O6 and the sulfate peak of the bound MES molecule in this structure (Figure 3, middle) when compared to the observed distances between O6 and 5′-phosphate group positions of GMP bound in other HPRT structures (*13*, *23*). These longer distances indicate that the active site in the *T. cruzi* HPRT structure may be more open and solvent accessible than the product-bound HPRT active site (Figure 3, bottom). Both ligands in the *T. cruzi* HPRT active site have stable, specifically anchored regions and hydrophilic, unbound regions. The MES morpholino ring, which is also disordered, may be competing with the FmB ribose ring for the same region of space in the active site, causing steric conflict (Figure 3, top).

The sulfate group of the bound MES molecule mimics the 5′-phosphate group normally present in either the enzyme's substrate (PRPP) or product (IMP/GMP); its oxygens participate in six to seven hydrogen bonds with members of the 5′-monophosphate-binding region of loop III (Figure 3, middle), in a similar manner as observed in GMP-bound human HPRT (Figure 3, bottom). These interactions are similar to those observed between the protein and a ligand's 5′-phosphate group in several type I PRT crystal structures and also similar to protein−sulfate interactions in the sulfate-complexed *T. foetus* HPRT monomer (*11*, *13*, *14*, *23*, *31*). The small hydrophilic side chains in this loop, threonines and/or serines, appear to provide flexibility to the pattern of hydrogen-bonding interactions the loop can form with the bound sulfate or phosphate group. In HPRTs, the loop generally maintains five to seven hydrogen bonds with the charged group of the bound ligand via main-chain nitrogen and threonine side-chain interactions, with slight variations in the specific hydrogen bonds formed (Figure 3, middle and bottom panels). This region of loop III appears to have high affinity for an electronegative group, and structures missing this active-site species have shown disorder for this part of loop III (*37*, *40*). The MES buffer molecule presumably provides the negatively charged species to the active site of the HPRT of *T. cruzi* since the FmB could not.

In addition to lacking a 5′-phosphate, FmB differs from IMP/GMP in that the pyrazolopyrimidine group of FmB has nitrogens at purine ring positions 7 and 8 (rather than 7 and 9); this results in a nonhydrolyzable carbon−carbon bond between the base and ribose groups of FmB. In the structure of GMP-bound human HPRT, conserved Asp137 (*T. cruzi* Asp115) (Table 2) was observed rotated away from the binding pocket, possibly repelled by the GMP 5′-phosphate group (*13*) (Figure 3, bottom). In the trypanosomal HPRT structure, a different rotomer for the side chain of Asp115 (human HPRT Asp137) is observed, forming a 2.9 Å hydrogen bond to the unusual FmB nitrogen at purine position 8 (Figure 3, middle). The orientation for the side chain of conserved Asp115 observed in the HPRT of *T. cruzi* (Figure 3, middle) may reflect a transient conformation which

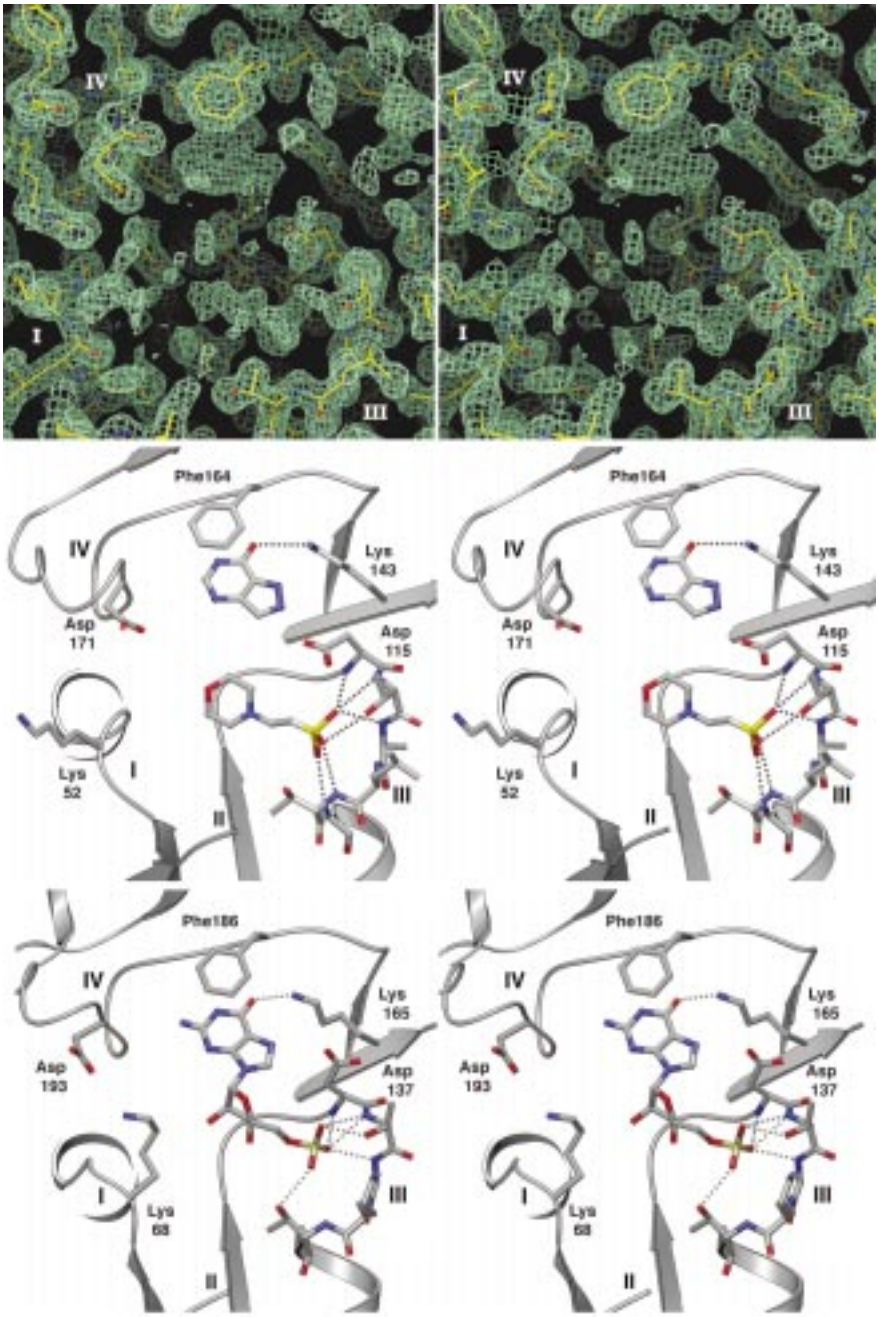


FIGURE 3: Active site detailed. (Top) The $2F_o - F_c$ electron density in the active site of the HPRT of *T. cruzi* prior to the addition of the ligands to the model (contoured at 1σ). All protein atoms from active site loops I, II, and IV are well-defined in the maps. Density for only the pyrazolopyrimidine ring of Formycin B is visible beneath the Phe164 side chain near the top of the figure. Below and to the right of center, density for the sulfate group and hydrocarbon chain of the bound MES molecule is surrounded by residues from loop III. (Middle) The active site of the HPRT of *T. cruzi* shows hydrogen bond interactions involving the bound ligands. Some mainchain interactions are omitted for clarity. The carbonyl at purine position 6 shares a 2.8 Å hydrogen bond with Lys143 (human Lys165), while the monophosphate-binding region of loop III forms 6−7 hydrogen bonds with the MES sulfate group in the two independent active sites. Asp115 (human Asp137) rotates toward the active site in this structure. Also displayed are the side chains of Lys52 (human Lys68) and Asp171 (human Asp193), as well as the side chain of Phe164 (human Phe186) under which the six-membered ring of the FmB is stacked. (Bottom) The human HPRT active site with bound GMP displays similar protein−ligand interactions to those shown in the middle panel, however Asp137 rotates away from the active site.

could facilitate its proposed role in purine deprotonation (*13*, *39*) but presumably is due to the presence of a nitrogen at purine position 8 in this structure.

Further interactions between the trypanosomal protein and the bound FmB pyrazolopyrimidine ring (Figure 3, middle) are analogous with those observed in other HPRT structures with the purine ring of the bound GMP (Figure 3, bottom) or IMP (*11*, *13*, *23*). The FmB pyrazolopyrimidine ring is held by hydrophobic forces, sandwiched between the side chains of Ile113, Val165, and Leu170 (not shown: Ile135, Val187, and Leu192 in the human HPRT) on one side of the planar aromatic ring of the ligand and a $\pi-\pi$ stacking interaction with a conserved loop IV residue Phe164 (Phe186 of the human HPRT) on the other side (*11*, *13*) (Figure 3, middle and bottom panels). The purine-binding pocket of the trypanosomal HPRT contributes two hydrogen bonds to FmB, which are similar to those observed in IMP/GMP-bound structures. The carbonyl group of Val165 (Val187 in

human HPRT) forms a 2.6 Å hydrogen bond with the FmB pyrazolopyrimidine ring nitrogen at purine position N1 (not shown), and Lys143 (Lys165 in the human HPRT) forms a 2.8 Å hydrogen bond with the FmB O6 (Figure 3, middle and bottom panels); the latter interaction, as proposed by Eads et al. (*13*), provides specificity for hypoxanthine or guanine versus adenine bases or nucleotides, which have an amino group rather than oxygen at purine position 6.

Formerly, Asp193 in the human HPRT was proposed to play a role both in base binding and in altering the conformation and/or pK_a of Lys68, which in turn was proposed to play a role in Mg^{2+}-PRPP or Mg^{2+}-PPi binding (*13*) (Asp171 and Lys52 in the trypanosomal HPRT). The mutation of Asp193 to asparagine in the human HPRT causes Lesch−Nyhan syndrome and results in increased K_M values for PRPP (500×) and hypoxanthine (100×) (*41*). When compared to the wild-type enzyme, a site-directed mutation of the corresponding residue Asp171 to asparagine in the HPRT of *T. cruzi* (D171N) results in a more than 2 orders of magnitude decrease in specific activity, while D171A reduces the specific activity by more than 3 orders of magnitude (data not shown). These mutations in the human and *T. cruzi* HPRTs are consistent with the Asp171 side chain (human Asp 193) being involved in PRPP binding; therefore, we propose that Asp 171 participates in binding PRPP and pyrophosphate via one or more magnesium ions. Additionally, the proposed role of Lys68 may be redefined by the structure of the HPRT of *T. cruzi*. Because of the unusual cis peptide in loop I (Figure 2), the main-chain nitrogen of Lys52 (Lys68 in the human enzyme) is exposed to the active site of the trypanosomal HPRT, but the Lys52 side chain points away from the active site. The lysine side chain forms a 2.7 Å hydrogen bond with the carbonyl oxygen from Met74 from strand B3 in the second subunit of the dimer (Figure 2, top, not shown in detail). Thus, in the HPRT of *T. cruzi*, an important active-site loop is linked directly to the other subunit of the dimer, thereby providing a possible mechanism for communication between subunits. This interaction has not been described formerly, and in the human HPRT product-bound structure, the side chain of the homologous residue (Lys68) was reported to be disordered (*13*). The interaction between subunits observed in the structure reported herein may therefore be absent in a product-bound state of the HPRT.

Structure of the Dimer. The HPRT from *T. cruzi* crystallized as a dimer of nearly identical subunits, shown in Figure 4, top. The human HPRT exists as a dimer or tetramer in solution depending on the pH and ionic strength (*43*) and crystallized as a tetramer, as have several other HPRTs (*13*, *40*). The trypanosomal HPRT has been reported to exist as a monomer in solution when unliganded, and as a dimer when PRPP is bound (*44*). The HPRTs of *T. cruzi* and *T. foetus*, which are both dimers in the crystal structures, lack regions corresponding to the N-terminal part of the human HPRT sequence shown to contribute to tetrameric interactions (*13*).

The dimer interface of the HPRT of *T. cruzi* involves residues from both domains of the protein (Figure 4, top). The core domain contributes helix A2, strand B3, and the nonactive site loops between helix A1 and strand B2 and between strands B4 and B5 (Figure 1). The hood domain contributes residues 167−176 (part of loop IV), Phe7 from

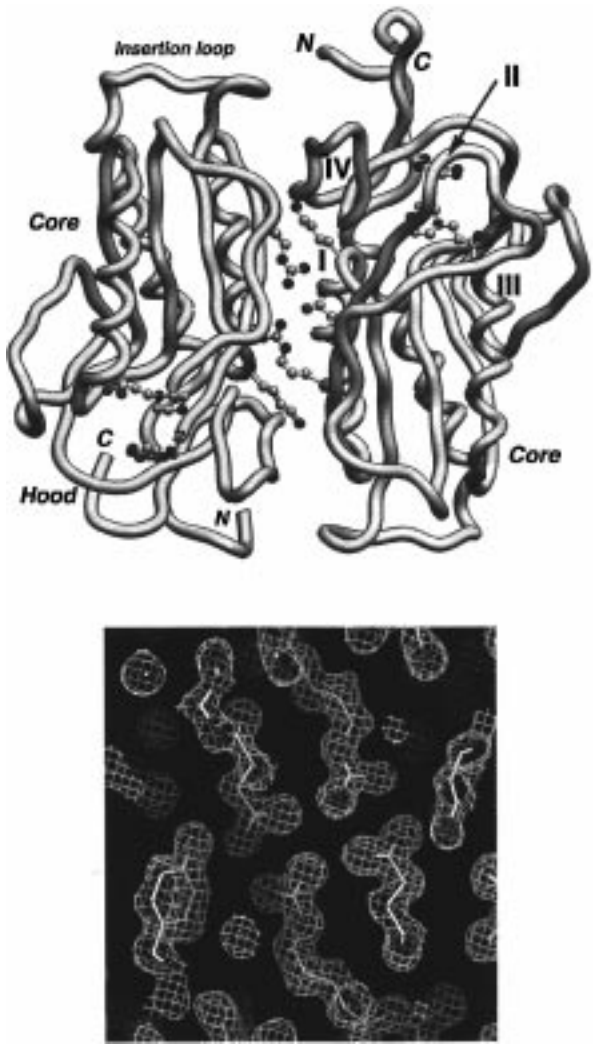


FIGURE 4: The dimer of the HPRT from *T. cruzi*. (Top) Ligands are represented as ball-and-stick models. The side chains of Lys52 (from active site loop I) and the double salt bridge residues Asp60 and Arg63 from each monomer are shown within the dimer interface. The dimer interface is mixed in character, also comprised of hydrophobic interactions contributed by the insertion loop and N- and C-terminal residues, and from regions flanking active site loops I, II, and IV. (Bottom) The $2F_o - F_c$ electron density map (contoured at 1σ) in the dimer interface corresponding to the double salt bridge.

near the N-terminus and Val184 from near the C-terminus (Figure 1). The dimer interface is mixed in character, consisting of many hydrophobic interactions, a few polar interactions, and a double salt bridge (Figure 4). The central core of the dimer interface is formed by the tight packing of the side chains of hydrophobic residues flanking active-site loop I, some of which are invariant residues in HPRTs (Leu51 and Phe57 in the *T. cruzi* enzyme and Leu67 and Phe73 in human enzyme) (Table 2). The two salt bridges are formed in the dimer interface between Asp60 and Arg63 from opposing subunits, which are located near the middle of helix A2 (Figure 4). While these charged pairs are not strictly conserved in HPRT sequences, a polar interaction in a similar region of 3D space is observed between two aspartic acid side chains and a lysine in the human HPRT dimer interface.

There is a seven residue sequence insertion (relative to the human HPRT) resulting in an extra surface loop in the

HPRT of *T. cruzi*, which is not found in other HPRT structures, although the sequence of the insert is conserved among some of the kinetidoplastida enzymes. The insertion loop of the core domain is shown at the top of the monomer on the left in Figure 4, directly across from the hood domain of the other monomer. This loop, located between helix A1 and strand B2, forms a β-turn that extends toward the dimer interface (Figure 4, top). The side chains from the insertion loop pack together with N- and C-terminal side chains from the hood domain and contribute to the hydrophobic nature of dimer interface at the periphery. These additional hydrophobic interactions contribute substantially to the buried surface area of the *T. cruzi* HPRT dimer.

In the OPRTs, the need to form dimers for catalytic activity has been clearly demonstrated and involves shared active sites (*31*, *45*). In contrast, analysis of the structures of HPRTs does not provide an obvious explanation for why these enzymes must form dimers to be catalytically active. In the structure presented here, the flexible loop II is unlikely to interact with the active site of the opposite subunit because the active sites are well separated across the dimer interface. In Figure 4, top, the two active sites of the HPRT dimer are located at the lower left and upper right quadrants. Therefore, the oligomerization of HPRTs is not likely to be driven by a need to complement the active-site ensemble.

Although there are no reports in the literature describing cooperativity between subunits in HPRTs, Xu and Grubmeyer recently discovered "that a measurable positive cooperativity exists in binding of nucleotide between two sites of the K68A mutant human enzyme" (Y. Xu and C. Grubmeyer, personal communication). This indicates for the human HPRT that the side chain of Lys68 may be responsible for masking cooperative interactions and that its removal may facilitate the detection of this positive cooperativity between subunits. These observations, together with the structural data reported here, support a hypothesis that Lys52 (human Lys68) participates in communication between subunits in dimers of the trypanosomal HPRT, but that other residues may be involved in cooperative interactions as well.

Comparisons of the active sites of the human and trypanosomal HPRTs show that many of the residues which provide key ligand-binding interactions are conserved between the two enzymes (Figure 3, middle and bottom panels). These similarities suggest that the development of parasite HPRT-specific inhibitors could be challenging and may require iterative structural comparisons of the host and parasite enzymes with lead compounds bound in the active site. Future structural analysis of the active sites of the human and trypanosomal HPRTs at different stages of catalysis also might reveal structural differences that could be exploited in future drug design efforts. Thus, the discovery of reliable crystallization conditions and the determination of this initial structure of a trypanosomal HPRT provide important steps toward inhibitor discovery that may lead to new drugs for the treatment of human diseases caused by parasites.

ACKNOWLEDGMENT

We are grateful to D. Freymann for critical reading of the manuscript and to Y. Taylor and H. Ouyang for technical assistance. This project was granted synchrotron beamline access by SSRL, program proposal number 4A16.

REFERENCES

1. Berens, R. L., Krug, E. C., and Marr, J. J. (1995) in *Biochemistry and Molecular Biology of Parasites* (Marr, J. J., and M., M., Ed.) pp 89−117, Academic Press Inc., San Diego, CA.
2. Garcia-Zapata, M. T. A., McGreevy P. B., and Marsden, P. D. (1991) in *Hunter's Tropical Medicine* (Stricklands, G. T., Ed.) pp 628−637, W. B. Saunders, Philadelphia.
3. Sculley, D. G., Dawson, P. A., Emmerson, B. T., and Gordon, R. B. (1992) *Hum. Genet. 90*, 195−207.
4. Lehninger, A. L. (1993) *Biochemistry*, 2nd ed., Worth Publishers, Inc., New York.
5. Krenitsky, T. A., Papaioannou, R., and Elion, G. B. (1969) *J. Biol. Chem. 244*, 1263−1270.
6. Kelley, W., Rosenbloom, F., Henderson, J., and Seegmiller, J. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1735−1739.
7. Musick, W. D. (1981) *CRC Crit. Rev. Biochem. 11*, 1−34.
8. Eads, J. C., Ozturk, D., Wexler, T. B., Grubmeyer, C., and Sacchettini, J. C. (1997) *Structure 5*, 47−58.
9. Giacomello, A., and Salerno, C. (1978) *J. Biol. Chem. 253*, 6038−6044.
10. Yuan, L., Craig, S. P., III, McKerrow, J. H., and Wang, C. C. (1992) *Biochemistry 31*, 806−810.
11. Xu, Y., Eads, J., Sacchettini, J. C., and Grubmeyer, C. (1997) *Biochemistry 36*, 3700−3712.
12. Goitein, R. K., Chelsky, D., and Parsons, S. M. (1978) *J. Biol. Chem. 253*, 2963−2971.
13. Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., and Sacchettini, J. C. (1994) *Cell 78*, 325−334.
14. Scapin, G., Grubmeyer, C., and Sacchettini, J. C. (1994) *Biochemistry 33*, 1287−1294.
15. Krahn, J. M., Kim, J. H., Burns, M. R., Parry, R. J., Zalkin, H., and Smith, J. L. (1997) *Biochemistry 36*, 11061−10068.
16. Ullman, B., and Carter, D. (1995) *Infect Agents Dis. 4*, 29−40.
17. Marr, J. J. (1991) *J. Lab. Clin. Med. 118*, 111−119.
18. Craig, S. P., III, & Eakin, A. E. (1997) *Parasitol. Today 13*, 238−241.
19. Eakin, A. E., Guerra, A., Focia, P. J., Torres-Martinez, J., and Craig, S. P. III (1997) *Antimicrob. Agents Chemother. 41*, 1686−1692.
20. Nieves-Alicea, R., Focia, P. J., Craig, S. P., III, & Eakin, A. E. *Biochim. Biophys. Acta* (in press).
21. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
22. Otwinowski, Z. (1993) in *Data Collection and Processing* (Sawyer, L., Isaacs, N. W., and Bailey, S., Eds.) pp 55−62, SERC Daresbury Laboratory, Warrington.
23. Somoza, J. R., Chin, M. S., Focia, P. J., Wang, C. C., and Fletterick, R. J. (1996) *Biochemistry 35*, 7032−7040.
24. Navaza, J. (1994) *Acta Crystallogr., Sect. A 50*, 157−163.
25. Brünger, A. T. (1992) *Nature 355*, 472−475.
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.
27. Kleywegt, G. J., and Jones, T. A. (1994) *Halloween....Masks and Bones*, SERC Daresbury Laboratory, Daresbury.
28. Laskowski, R. A., MacArthur, M. W., and Thorton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
29. Smith, J. L., Zaluzec, E. J., Wery, J. P., Niu, L., Switzer, R. L., Zalkin, H., and Satow, Y. (1994) *Science 264*, 1427−1433.
30. Creighton, T. E. (1993) *Proteins—Structures and Molecular Properties*, W. H. Freeman and Company, New York.
31. Scapin, G., Ozturk, D. H., Grubmeyer, C., and Sacchettini, J. C. (1995) *Biochemistry 34*, 10744−10754.
32. Stewart, D. E., Sarkar, A., and Wampler, J. E. (1990) *J. Mol. Biol. 214*, 253−260.
33. Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., and Rayment, I. (1996) *J. Biol. Chem. 271*, 21956−21968.
34. Dominguez, R., Souchon, H., Lascombe, M., and Alzari, P. M. (1996) *J. Mol. Biol. 257*, 1042−1051.
35. Hennig, M., Jansonius, J. N., Terwisscha van Scheltinga, A. C., Dijkstra, B. W., and Schlesier, B. (1995) *J. Mol. Biol. 254*, 237−246.
36. Hohenester, E., Maurer, P., Hohenadl, C., Timpl, R., Jansonius, J. N., and Engel, J. (1996) *Nat. Struct. Biol. 3*, 67−73.

37. Vos, S., de Jersey, J., and Martin, J. L. (1997) *Biochemistry*, 4125−4134.
38. Henriksen, A., Aghajari, N., Jensen, K. F., and Gajhede, M. (1996) *Biochemistry 35*, 3803−3809.
39. Xu, Y., and Grubmeyer, C. (1998) *Biochemistry 37*, 4114−4124.
40. Schumacher, M. A., Carter, D., Roos, D. S., Ullman, B., and Brennan, R. G. (1996) *Nat. Struct. Biol. 3*, 881−887.
41. Wilson, J. M., Stout, J. T., Palella, T. D., Davidson, B. L., Kelley, W. N., and Caskey, C. T. (1986) *J. Clin. Invest. 77*, 188−195.
42. Jardim, A., and Ullman, B. (1997) *J. Biol. Chem. 272*, 8967−8973.
43. Johnson, G. G., Eisenberg, L. R., and Migeon, B. R. (1979) *Science 203*, 174−176.
44. Allen, T. E., and Ullman, B. (1994) *Mol. Biochem. Parasitol. 65*, 233−245.
45. Ozturk, D. H., Dorfman, R. H., Scapin, G., Sacchettini, J. C., and Grubmeyer, C. (1995) *Biochemistry 34*, 10764−10770.

BI981052S