Crystal Structure of the Hypoxanthine−Guanine−Xanthine Phosphoribosyltransferase from the Protozoan Parasite *Tritrichomonas foetus*[†,‡]

John R. Somoza,[§] Marian S. Chin,[||] Pamela J. Focia,[||] Ching C. Wang,*[,||] and Robert J. Fletterick[§]

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143-0448, and Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143-0446

Received December 28, 1995; Revised Manuscript Received March 28, 1996[⊗]

ABSTRACT: The crystal structure of the hypoxanthine−guanine−xanthine phosphoribosyltransferase (HGXPRTase) from *Tritrichomonas foetus* has been determined and refined against X-ray data to 1.9 Å resolution. *T. foetus* HGXPRTase crystallizes as an asymmetric dimer, with GMP bound to only one of the two molecules that form the asymmetric unit. Each molecule of HGXPRTase is formed by two lobes joined by a short "hinge" region, and the GMP binds in a cavity between the two lobes. A comparison of the two molecules in the asymmetric unit shows that the hinge region is flexible and that ligand binding affects the relative positions of the two lobes. The binding of GMP brings the two lobes closer together, rotating one lobe by about 5° relative to the other. *T. foetus* appears to depend on HGXPRTase for its supply of GMP, making this enzyme a target for antiparasite drug design. A comparison of the structures of *T. foetus* HGXPRTase and human HGPRTase reveals that, while these enzymes retain a similar polypeptide fold, there are substantial differences between the active sites of these two homologs. These differences suggest that it will be possible to find compounds that selectively inhibit the parasite enzyme.

Tritrichomonas foetus is an anaerobic protozoan parasite that causes bovine trichomoniasis, a sexually transmitted disease in cattle. In bulls, tritrichomonal infections are essentially asymptomatic, but infections in cows can lead to infertility, early embryonic death, and abortion. Bovine trichomoniasis is prevalent in beef herds throughout much of the world (Fitzgerald, 1986) and represents a substantial loss in cattle production (Speer & White, 1991). At least in the United States, there is no effective, approved systemic treatment for this disease.

Wang *et al.* (1984) have shown that *T. foetus*, like all other protozoan parasites studied to date, lacks the ability to carry out the *de novo* synthesis of purine nucleotides and instead relies primarily on hypoxanthine−guanine−xanthine phosphoribosyltransferase (HGXPRTase)[1] to salvage purine bases from the host. HGXPRTase catalyzes reactions in which the ribose phosphate of α-D-5-phosphoribosyl-1-pyrophosphate (PRPP) is transferred to hypoxanthine, guanine, or xanthine to form the corresponding ribonucleotide and pyrophosphate. Because of the central role played by HGXPRTase in providing the parasite with ribonucleotides, it is a promising target in the design of anti-tritrichomonal chemotherapeutics.

HGXPRTase forms part of a family of phosphoribosyltransferases (PRTases) that use nitrogenous bases and PRPP to form nucleotides as well as the amino acids histidine and tryptophan. Although there is a wide range of substrate specificity, substrate affinity, and kinetic mechanisms associated with the PRTases, there are enough similarities in physical and catalytic properties to suggest that the PRTases share similar three-dimensional conformations (Musick, 1981). This hypothesis has been at least partially confirmed by the recent elucidation of the crystal structures of orotate PRTase (Scapin *et al.*, 1994), glutamine PRPP amidotransferase (Smith *et al.*, 1994), and human HGPRTase (Eads *et al.*, 1994). The two PRTases, as well as the PRTase domain of the amidotransferase, share a similar α/β fold.

Clearly, any HGXPRTase inhibitor that is considered as a potential anti-tritrichomonal drug should not interfere with the mammalian HGPRTase. In humans, defects in HGPRTase are responsible for gouty arthritis and for the Lesch−Nyhan syndrome, a disease that leads to hyperuricemia, uric acid nephrolithiasis, and a number of central nervous system disorders. The sequences of all mammalian HGPRTases studied to date are essentially identical.

A comparison of the physical properties of the tritrichomonal and human purine salvage PRTases reveals a number of differences that could be exploited to design inhibitors specific to the tritrichomonal enzyme. About 30% of the *T. foetus* HGXPRTase sequence is identical to that of the human HGPRTase (Figure 1). Aside from the differences in primary structure, these two enzymes have different substrate specificity. The most significant difference is that the *T. foetus* HGXPRTase accepts xanthine as a substrate, while the human enzyme does not. This difference in substrate specificity may be especially relevant to finding or designing compounds that inhibit the *T. foetus* enzyme but not its mammalian counterpart.

[†] This work was supported by National Institutes of Health Grant AI 19391.

[‡] The coordinates and observed structure factor amplitudes for *T. foetus* HGXPRTase have been deposited in the Brookhaven Protein Data Bank under identification codes 1HGX and R1HGXSF, respectively.

* To whom correspondence should be addressed.

[§] Department of Biochemistry and Biophysics.

[||] Department of Pharmaceutical Chemistry.

[⊗] Abstract published in *Advance ACS Abstracts,* May 15, 1996.

[1] Abbreviations: HGXPRTase, hypoxanthine−guanine−xanthine phosphoribosyltransferase; HGPRTase, hypoxanthine−guanine phosphoribosyltransferase; PRTase, phosphoribosyltransferase; PRPP, α-D-5-phosphoribosyl-1-pyrophosphate; SDS, sodium dodecyl sulfate; GMP, guanosine 5′-monophosphate; XMP, xanthosine 5′-monophosphate; IMP, inosine 5′-monophosphate.

S0006-2960(95)03072-8 CCC: $12.00

```
            1         10         20         30         40         50
H.sapiens   MATRSPGVVI SDDEPGYDLD LFCIPNHYAE DLERVFIPHG LIMDRTERLA RDVMKEMGGH
                                               |||||       |   |   ||
T.foetus                                 MTETPMMD DLERVLYNQD DIQKRIRELA AELTEFYEDK
                                         1         9         19         29

            60         70         80         90         100        110
H.sapiens   HIVALCVLKG GYKFFADLLD YIKALNRNSD RSIPMTVDFI RLKSYCNDQS TGDIKVIGGD
              |  ||| |    |  |||                    | |    ||    | ||    |  |
T.foetus    NPVMICVLTG AVFFYTDLLK HLD.........FQLEPDYI ICSSYSGTKS TGNL.TISKD
            39         49         59           62         70         80

            120        130        140        150        160        169
H.sapiens   DLSTLTGKNV LIVEDIIDTG KTMQTLLSLV RQYNPKMVKV ASLLVKRT.P RSVGYKPDFV
                  |  | | ||||||||  ||  ||       |   ||   |  |           |
T.foetus    LKTNIEGRHV LVVEDIIDTG LTMYQLLNNL QMRKPASLKV CTLCDKDIGK KAYDVPIDYC
            89         99         109        119        129        139

            179        189        199        209
H.sapiens   GFEIPDKFVV GYALDYNEYF RDLNHVCVIS ETGKAKYKA
            ||         ||  |      | |        |
T.foetus    GFVVENRYII GYGFDFHNKY RNLPVIGILK ESVYT
            149        159        169        179
```

FIGURE 1: Alignment of the sequences of the human HGPRTase and the HGXPRTase from *T. foetus*.

EXPERIMENTAL PROCEDURES

Purification. The *T. foetus* HGXPRTase has been cloned and expressed in *Escherichia coli* (Chin & Wang, 1994). The protein used in this study was purified by ammonium sulfate precipitation, followed by anion-exchange chromatography and isoelectric focusing chromatography, using a protocol similar to the one described by Chin (1995). The HGXPRTase purified by this procedure was active and greater than 95% pure, as judged by silver-stained SDS−polyacrylamide gel electrophoresis and by Coomassie-stained isoelectric focusing gel electrophoresis. The mass of the protein, which was measured by electrospray ionization mass spectrometry, is 21 091.04 ± 1.48 Da for the protein with the N-terminal methionine and 20 960.50 ± 1.11 Da for the protein (about 20−30% of total) without the methionine. The measured mass is in close agreement with the mass calculated from the amino acid sequence: 21 091.43 and 20 960.23 Da for the protein with and without the methionine, respectively.

Crystallization. The "sparse matrix" strategy (Jancarik & Kim, 1991) was used as the starting point for the crystallization of HGXPRTase. Crystals suitable for high-resolution structural studies were obtained by sitting drop vapor diffusion at room temperature (approximately 22 °C). The reservoir contained 0.2 M ammonium sulfate, 0.1 M bis-tris buffer (pH between 6.3 and 6.8), and 28−32% (w/v) PEG 8000. The drops contained an equal mixture of the reservoir solution and a protein solution containing 10 mg/mL HGXPRTase, 50 mM bis-tris (pH 6.8), and 6 mM $MgCl_2$. Under these conditions, two crystal forms appeared. First, a large number of non-birefringent, cubic-shaped crystals were obtained but diffracted to only 10 Å resolution (using X-rays from a rotating anode generator and collecting data on an RAXIS II C imaging plate). A second crystal form was also grown under these conditions. These crystals, which grew as either thin plates or rods, diffracted to at least 1.9 Å resolution. This second crystal form rarely appears spontaneously under the conditions described above, but these crystals can be readily grown by microseeding the drops. This was done by crushing up a crystal and transferring microcrystals to a preequilibrated drop with a cat's whisker. The crystals grow over the course of several days to several weeks.

Data Collection. Two data sets were used in the structure determination and refinement. Both sets of diffraction data were measured on an R-AXIS II C imaging plate system using X-rays from a Rigaku RU-200 rotating anode generator operating at 15 kW (50 kV, 300 mA). The data were processed with the DENZO and SCALEPACK software packages (Otwinowski, 1988). A data set to 2.4 Å resolution was obtained from a crystal that had been obtained with no ligand added to the crystallization conditions. The resulting data set is 88.9% complete to 2.4 Å resolution and is 54.1% complete in the resolution shell between 2.4 and 2.5 Å. The $R_{scale}(I)$ for these data is 5.9%; for all observations $I > 0.0\sigma(I)$. The autoindexing routine showed that these crystals belong to the primitive monoclinic Bravais lattice, and an inspection for systematic absences indicated that they belong to the space group $P2_1$. This space group was later confirmed by the structure determination. The cell parameters are $a = 48.8$ Å, $b = 74.7$ Å, $c = 55.3$ Å, and $\beta = 110.8°$, and the volume of the unit cell suggested that two molecules formed the asymmetric unit. A second data set was gathered from a crystal that had been obtained after adding 10 μM GMP to the crystallization conditions. This data set is 91.2% complete to 1.9 Å resolution and is 55.6% complete in the shell between 1.96 and 1.90 Å resolution. The $R_{scale}(I)$ for this data set is 6.3% [based on all observations $I > -3\sigma(I)$]. The crystals obtained with and without added GMP are isomorphous.

Structure Determination and Refinement. The positions of the two molecules in the asymmetric unit were identified by carrying out rotation and translation searches with the program AMORE (Navaza, 1994), using the structure of the human HGPRTase (Eads *et al.*, 1994) as a search probe. The molecular replacement search led to a solution with a correlation coefficient of 0.296, and a crystallographic *R*-factor of 49.3%, for data between 15.0 and 3.5 Å.

The electron density maps obtained directly from the results of the molecular replacement search were not clear enough to allow refitting of the model, and two steps were taken that led to interpretable maps. First, the model itself was gradually improved by replacing the side chains corresponding to the human HGPRTase sequence with the corresponding residues from the *T. foetus* sequence. This process was accompanied by several rounds of refinement with the program X-PLOR v. 3.1 (Brünger, 1992), using the parameters described by Engh and Huber (1991). The second step was the improvement in the electron density by making use of the noncrystallographic symmetry (NCS) of the asymmetric unit, as well as our knowledge of the solvent

region. A mask encompassing one of the molecules in the asymmetric unit was made using the program MAMA (Kleywegt & Jones, 1993), and 30 cycles of electron density averaging and phase recombination were carried out using the program RAVE (Jones, 1992) and parts of the CCP4 software package (Collaborative Computational Project. Number 4, 1994). Even though the two molecules were later shown to be not identical, most of the electron density in the maps obtained from the NCS averaging procedure was unambiguously interpretable.

The refinement of the model consisted of cycles of manual fitting of the maps and simulated annealing refinement using X-PLOR v. 3.1 (Brünger, 1992). The manual fitting was carried out with the aid of the program O (Jones *et al.*, 1991). Throughout the refinement, the quality of the model was monitored by calculating a free *R*-factor, based on approximately 10% of the reflections. During the initial stages of the refinement, the two molecules in the asymmetric unit were forced to maintain perfect noncrystallographic symmetry, and only one of the two molecules was fit to density. Also, at each cycle of refinement, the maps were improved by NCS averaging using RAVE. In the latter stages of refinement, the two molecules were refined independently, the individual isotropic atomic temperature factors were refined, and GMP, sulfate, and water molecules were incorporated into the model. The model was refined against the 2.4 Å data set until the model reached an *R*-factor of 16.2% and a free *R*-factor of 26.6%. Further refinement was then carried out against the second data set, to 1.9 Å resolution. At the current stage of refinement, the model includes 2 molecules of HGXPRTase, 134 water molecules, 1 sulfate ion, and 1 molecule of GMP. There are some residues in each of the two molecules in the asymmetric unit for which there is poor or uninterpretable electron density. The following residues have not been modeled: in monomer 1, residues 1−6, 74−82, and 180−183; in monomer 2, residues 1−6 and 72−84. In addition, residues 7, 8, and 179 of monomer 1 and residues 7, 71, 85, and 86 of monomer 2 have been tentatively modeled in electron density. However, the temperature factors for these residues are high, and the placement of these residues should be viewed with skepticism. Finally, there are some side-chain atoms that have weak or nonexistent density. These are primarily long, hydrophilic side chains at the surface of the protein.

The crystallographic *R*-factor for the current model is 17.5% (for all data between 6.0 and 1.9 Å resolution), with a free *R*-factor of 23.4%. The rms deviations from ideal bond lengths and angles are 0.008 Å and 1.42°, respectively. The ϕ and ψ angles for 92.2% of the residues are within the most favorable region of the Ramachandran plot, and no residues are within the disallowed region. The statistics for the current model are summarized in Table 1.

Analysis of the Structure. The stereochemistry of the HGXPRTase model was examined with X-PLOR v. 3.1 (Brünger, 1992) and with PROCHECK (Laskowski *et al.*, 1993). The assignment of the secondary structure of HGXPRTase was carried out using the algorithm described by Kabsch and Sander, as implemented by the program DSSP (Kabsch & Sander, 1983). This program was also used for the solvent accessibility calculations that were used to characterize the dimer interface. These calculations were made with a solvent probe of radius 1.40 Å. The analysis of the hydrogen-bonding network between the GMP and enzyme was aided by the program HBPLUS (McDonald & Thornton, 1994), using the geometic criteria suggested by Baker and Hubbard (1984). The analysis of the conformational changes that stem from the GMP binding was carried out with INSIGHT II (Biosym, Inc.) and with GEM (E. Fauman, in preparation).

Table 1: Data and Refinement Statistics

data	
resolution range (Å)	∞−1.9
no. of observations	97259
no. of unique reflections	27314
completeness (%)	
∞−1.9 Å	91.2
1.96−1.9 Å	55.6
R-merge (%)	
∞−1.9 Å	6.3
1.96−1.9 Å	38.0
$\langle I/\sigma(I)\rangle$	8.1
model	
residues missing from model	
monomer 1	1−6, 74−82, 180−183
monomer 2	1−6, 72−84
no. of nonprotein atoms included in model	
waters	134
sulfate	1
GMP	1
resolution used for refinement (Å)	6.0−1.9
σ cutoff [$F/\sigma(F)$]	0.0
R-factor (%)	17.5
free *R* (%)	23.4
rms deviations from ideal geometry	
bonds (Å)	0.008
angles (deg)	1.42

RESULTS AND DISCUSSION

The HGXPRTase Structure. The structure of *T. foetus* HGXPRTase includes nine β-strands and three α-helices (Figure 2). There is also a small 3_{10}-helix at the C-terminus of the enzyme, although it is seen in only one monomer in the asymmetric unit; the other monomer has a disordered C-terminus. Throughout the remainder of this discussion, the elements of regular secondary structure will be numbered according to their appearance in the sequence (Figure 2). The first and the last two of the strands interact to form a small antiparallel β-sheet, and the remaining strands form a large, twisted, mostly parallel β-sheet that transects the enzyme, separating helices 1 and 2 from helix 3. Although the general fold of the enzyme is similar to that of human HGPRTase (Figure 3), the *T. foetus* protein has no counterpart to the first strand or to the first helix that was observed in the human structure, and the second helix and third β-strand of the *T. foetus* enzyme are shorter than the corresponding elements in the human HGPRTase structure. The net effect of these differences is to make the *T. foetus* HGXPRTase more compact than the human enzyme.

The *T. foetus* HGXPRTase structure, like that of human HGPRTase, consists of two lobes. The larger of these lobes, which will be referred to as lobe 1, contains all three of the helices as well as the larger of the β-sheets. The smaller lobe (lobe 2) contains the small antiparallel β-sheet. The two lobes are connected by a "hinge" region formed by the residues at the N-terminus of helix 1 and at the C-terminus of strand 7. The GMP binds in a shallow cavity formed by residues from both lobes (Figure 4).

The phosphate moiety of GMP binds to residues in the loop between strand 5 and helix 3 in lobe 1. The backbone

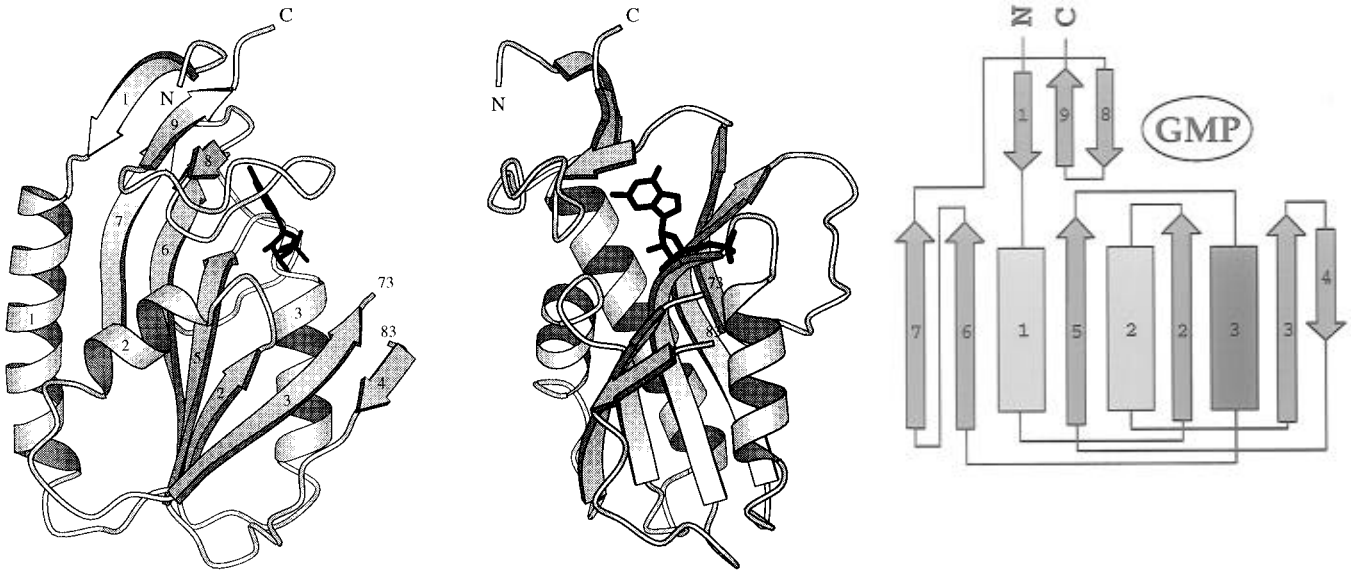

FIGURE 2: Schematic views of the structure of *T. foetus* HGXPRTase. The two views in (a, left) and (b, center) are related by a 90° rotation around a vertical axis. The arrows represent β-strands and the coils represent α-helices. In (a), the elements of regular secondary structure are numbered, with the helices and strands numbered separately. These figures were generated using the program MOLSCRIPT (Kraulis, 1991). The topology of HGXPRTase is depicted in (c, right), with the β-strands represented as arrows and the helices as rectangles. The two lighter helices lie above the β-sheet, and the darker helix lies below the sheet.

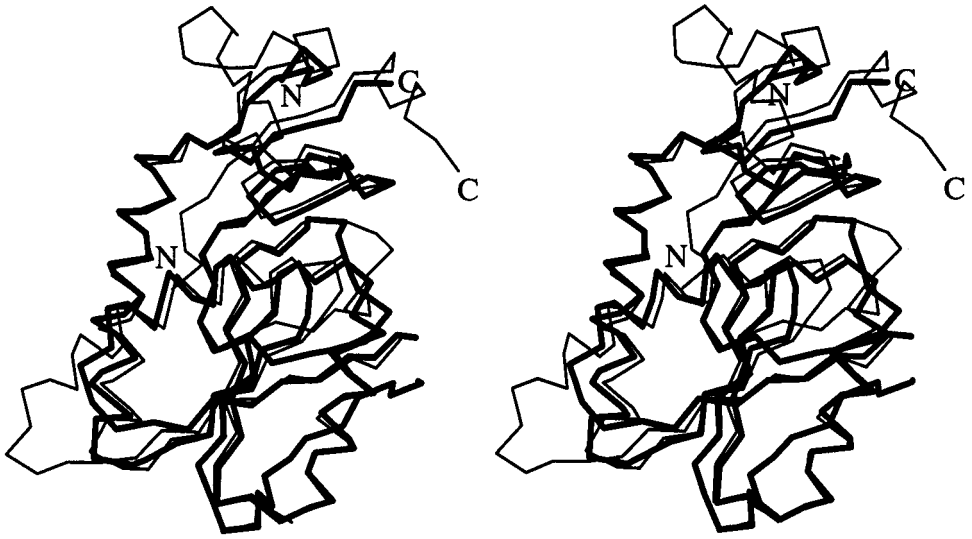

FIGURE 3: Stereo superposition of the α-carbon traces of *T. foetus* HGXPRTase (thick lines) and human HGPRTase (thin lines). This figure was prepared by superimposing the α-carbons from residues 17−32, 50−59, 96−107, 109−120, and 124−134 of the HGXPRTase with the α-carbons from residues 38−53, 71−80, 127−138, 140−151, and 155−165 of the human HGPRTase. The rms deviation between these sets of atoms is 0.78 Å.

amides of residues 106, 107, 108, and 110 are all within hydrogen-bonding distance of the phosphate oxygens, as are the side-chain hydroxyls of threonines 107 and 110 (Figure 5). There are also a number of solvent molecules in the vicinity that may be contributing to the phosphate binding. Finally, it is likely that additional stabilization of the phosphate is provided by the dipole moment of helix 3.

The O2′ and O3′ oxygens of the ribose moiety of the GMP are in direct or indirect (through water molecules) hydrogen-bonding contact with glutamate 102 and aspartate 103. Asp 103 is positioned to form a hydrogen bond with the O2′ hydroxyl, while glutamate 102 appears to form a hydrogen bond to a water molecule, which in turn forms a hydrogen bond with the ribose O3′ atom.

There are a number of hydrogen bonds formed between residues in lobe 2 and atoms of the purine base. The amino group of Lys 134, which is conserved throughout the purine salvage PRTases, interacts with the O6 of the purine ring, and presumably functions to distinguish between the purine bases with carbonyl oxygens at the C6 position (guanine, hypoxanthine, and xanthine) and those with an amino group at this position (adenine). The backbone amide of Ile 157 is also in position to hydrogen bond to O6 and may be contributing to the discrimination between purine bases. The purine N1 nitrogen is within hydrogen-bonding distance of the backbone carbonyl of Ile 157, and the amino substituent at the C2 position of GMP can potentially form hydrogen bonds with the main-chain oxygen of aspartate 163, the side-chain hydroxyl of tyrosine 156, a water molecule, and, possibly, the backbone carbonyl of Ile 157. Aside from the hydrogen bonds, the GMP binding is stabilized by hydrophobic interactions with the protein, most notably with Ile

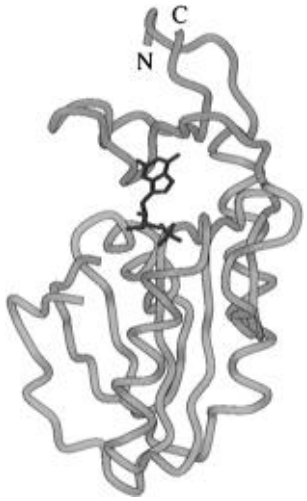

FIGURE 4: α-Carbon trace of the *T. foetus* HGXPRTase showing the relation between lobe 1 (yellow), lobe 2 (green), and the GMP.

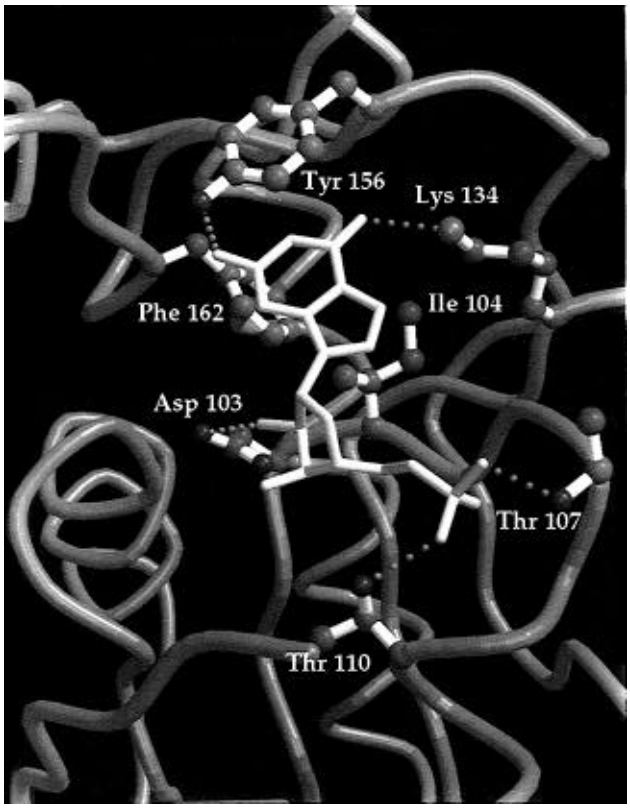

FIGURE 5: α-Carbon trace of the active site of the *T. foetus* HGXPRTase (blue) with bound GMP (yellow). The side chains that are shown are those that directly interact with the GMP. Potential hydrogen bonds are shown as green dotted lines.

104 and Phe 162. These residues primarily contact atoms in the purine base.

The active site cleft extends significantly beyond the region where the GMP binds. Judging from the structure of OPRTase in complex with PRPP and orotate (Scapin *et al.*, 1995), the pyrophosphate moiety of PRPP extends into the unoccupied region of the active site, toward the loop that connects strand 2 with helix 2. Interestingly, in the *T. foetus* HGXPRTase structure this loop contains a *cis* peptide bond between residues Leu 46 and Thr 47; c*is* amide bonds are very rare in published structures (Stewart *et al.*, 1990).

Comparison with the Active Site of Human HGPRTase. Given the similarities in overall structure between the *T. foetus* HGXPRTase and the human HGPRTase, the active sites of these enzymes are surprisingly different. Figure 6 shows a comparison of the active sites of the tritrichomonal and human enzymes after superimposing the two molecules of GMP. The two phosphate binding loops make similar hydrogen bonds to the phosphate, although the positions of the residues involved are not entirely conserved. A similar situation is seen with the glutamate (*T. foetus*, 102; human, 133) and aspartate (*T. foetus*, 103; human, 134) that contact the ribose hydroxyls. While the aspartates are both able to hydrogen bond to the O2′ hydroxyl, the two glutamates superimpose poorly (their α-carbons are approximately 3 Å apart), and only the glutamate from the human enzyme is able to directly contact the ribose.

A comparison of the way the two enzymes interact with the purine base shows that they both have a lysine (*T. foetus*, 134; human, 165) able to hydrogen bond to the GMP O6 carbonyl, as well as an aromatic side chain that stacks on top of the base (*T. foetus*, Tyr 156; human, Phe 186). However, aside from these similarities, the two enzymes make use of different interactions to bind the ribonucleotide (Figure 6). Presumably, some of the differences in the vicinity of the C2 position of the purine base explain the differences in xanthine affinity between HGPRTase and HGXPRTase (xanthine differs from guanine and hypoxanthine only in the identity of the substituent attached to the C2 carbon of the purine base: xanthine has a carbonyl at this position, guanine has an amino group, and hypoxanthine has no substituent). For example, the side-chain Oη of tyrosine can act as both a hydrogen bond donor and acceptor, so Tyr 156 of the *T. foetus* enzyme could potentially hydrogen bond to either guanine or xanthine.

The active sites of the tritrichomonal and human enzymes also show major differences in regions not involved in binding to the ribonucleotide, but which presumably bind to PRPP. The most dramatic difference involves residues 46 through 49 in the loop that joins strand 2 and helix 2 in the *T. foetus* HGXPRTase and the corresponding loop in human HGPRTase (residues 67−70). There is ample evidence indicating that this loop is important for catalysis. The loop is a prominent feature of the active site clefts of both enzymes, and the residues forming this loop are highly conserved across the purine salvage PRTases that have been studied to date. The residue at position 46 is a leucine in 9 out of 10 sequences examined;[2] position 47 is almost always lysine, although not in the *T. foetus* HGXPRTase or in the *G. lamblia* GPRTase (J. Sommer, personal communication), where this residue is threonine, and *E. coli*, where it is an arginine (Pratt & Subramani, 1983). Position 48 is always glycine, and position 49 is usually either glycine or serine, but in the *T. foetus* and *G. lamblia* enzymes it is alanine. Eads *et al.* (1994) have proposed that the conserved lysine in this loop (Lys 68 in the human HGPRTase, which corresponds to Thr 47 of the *T. foetus* HGXPRTase) binds Mg^{2+} pyrophosphate. This hypothesis is supported by the

[2] Sequences from purine salvage PRTases from the following organisms were examined: *Tritrichomonas foetus* (Chin & Wang, 1994), *Homo sapiens* (Jolly *et al.*, 1983), *Plasmodium falciparum* (King & Melton, 1987), *Toxoplasma gondii* (Vasanthakumar *et al.*, 1994), *Schistosoma mansoni* (Craig *et al.*, 1988), *Trypanosoma brucei* (Allen & Ullman, 1993), *Trypanosoma cruzi* (Allen & Ullman, 1994), *Leishmania donovani* (Allen *et al.*, 1995), *Escherichia coli* (Pratt & Subramani, 1983), and *Giardia lamblia* (J. Sommer, unpublished work).

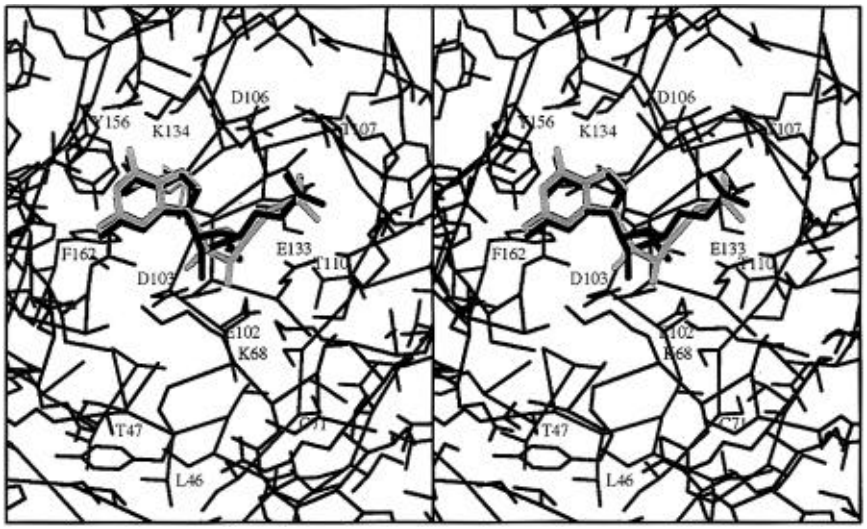

FIGURE 6: Stereo comparison of the active sites of the *T. foetus* HGXPRTase (blue) and human HGPRTase (red), based on a superposition of the two molecules of GMP.

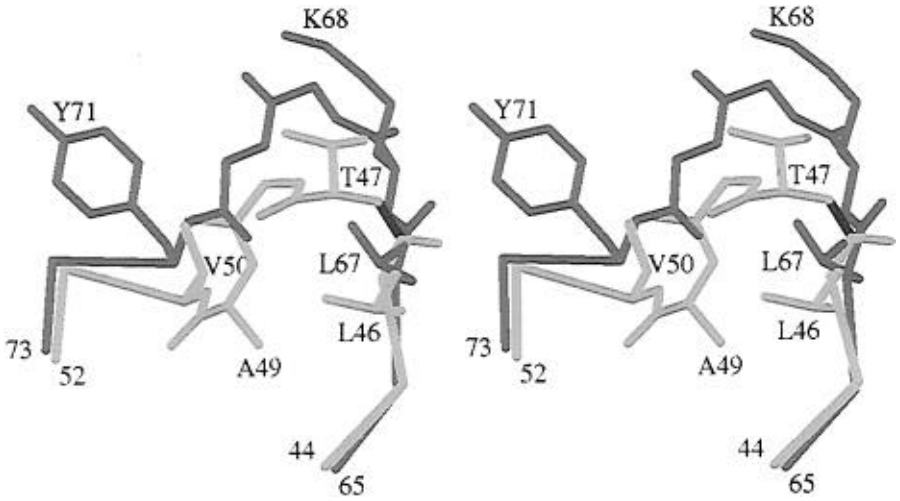

FIGURE 7: Stereo comparison of the loop between strand 2 and helix 2 of HGXPRTase (blue) with the corresponding residues of human HGPRTase (red). The residues shown explicitly are 46−49 of HGXPRTase and 67−70 of HGPRTase. The *cis* peptide bond between Leu 46 and Thr 47 in the HGXPRTase structure is shown in black.

crystal structure of the OPRTase/PRPP/orotate complex (Scapin *et al.*, 1995); in OPRTase there is a lysine at an equivalent position (Lys 73), and it is positioned to hydrogen bond with the β-phosphate oxygens of the pyrophosphate. To further underscore the functional relevance of this loop, several of the inherited mutations in the human HGPRTase gene that lead to the Lesch−Nyhan or Kelley−Seegmiller disorders involve two of the residues in this loop (Gly 69 and Gly 70) (Wilson *et al.*, 1986; Bouwens-Rombouts *et al.*, 1993).

In spite of the importance of this loop and the conservation of the leucine and glycine between the *T. foetus* and human enzymes, the loop adopts completely different conformations in the two enzymes (Figures 6 and 7), with the loop protruding further out into the active site in the HGPRTase structure than it does in the HGXPRTase. This conformational difference, along with the fact that the tritrichomonal enzyme has no equivalent to the catalytically important lysine that is generally present in this loop, suggests that the interaction between this enzyme and PRPP (and presumably PP_i) is very different than in the human enzyme and probably different from what will be seen in the majority of the other purine salvage PRTases as well. It is likely that these loop differences explain why the K_m for PP_i is an order of magnitude higher for the *T. foetus* enzyme than for the human enzyme (Chin, 1995; Giacomello & Salerno, 1978). If this hypothesis is correct, we should also expect to see a high K_m for pyrophosphate in the *G. lamblia* GPRTase, which closely matches the sequence of the *T. foetus* enzyme in this region. This enzyme has recently been cloned and expressed in *E. coli* (J. Sommer, personal communication), so this prediction will soon be tested.

Quaternary Structure. HGXPRTase crystallizes with a dimer in the asymmetric unit (Figure 8). Two characteristics of this dimer strongly suggest that it is physiologically significant and not an artifact of the crystallization. First, the monomers are closely packed, with approximately 1080 Å^2 of the solvent-accessible surface of each monomer (or about one-tenth of the total surface of each monomer) buried in the dimer interface. More importantly, the spatial arrangement of the two monomers is very similar to the relation between the monomers that form the dimer seen in the structure of human HGPRTase (Figure 8a), which is known to function as either a dimer or tetramer, depending on the

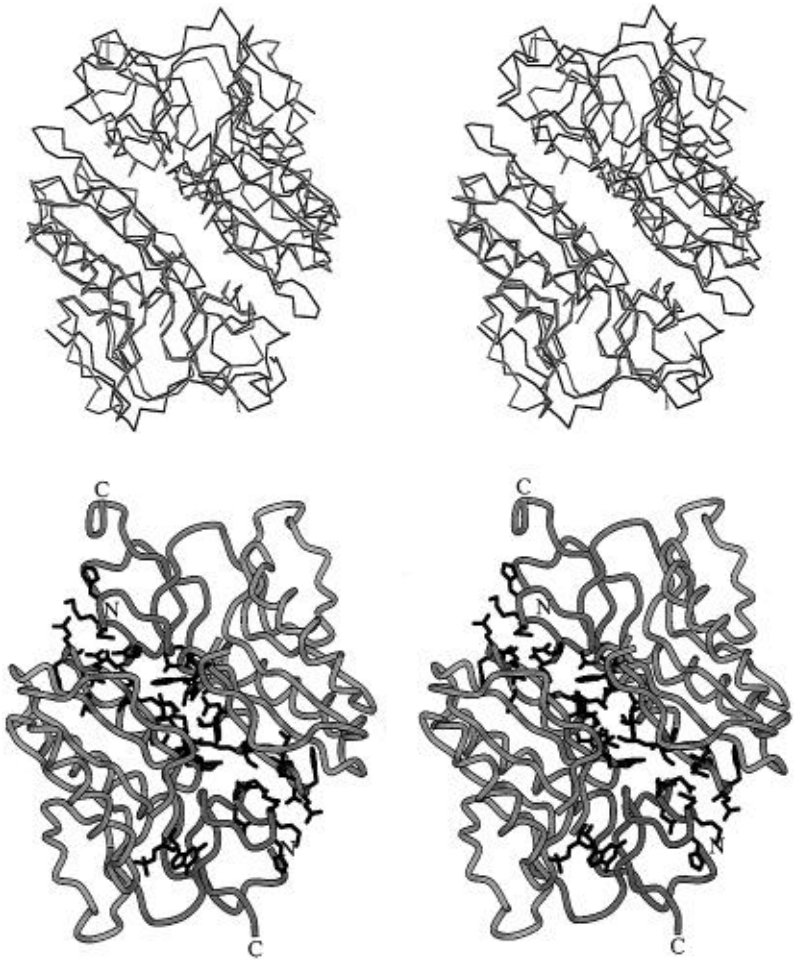

FIGURE 8: (a, top) Stereo figure comparing the *T. foetus* HGXPRTase dimer (blue) with the human HGPRTase dimer (red). (b, bottom) Stereo figure showing an α-carbon trace of the HGXPRTase dimer. Lobe 1 is shown in yellow and lobe 2 in green. The only residues shown explicitly are those that contribute at least 5 Å^2 to the dimer interface.

pH and ionic strength of the solution (Johnson *et al.*, 1979). The existence of the dimer was somewhat surprising in light of the work of Beck and Wang (1993) and Chin (1995), who studied the oligomerization state of *T. foetus* HGXPRTase using gel filtration chromatography and native gel electrophoresis. Under the conditions used in those studies, which did not include ligand, HGXPRTase was present primarily as a monomer, although a small percentage of the protein did form dimers. It appears that either we are selectively crystallizing the dimers, or dimerization is being favored by the high protein concentrations used in the crystallization trials or by the addition of GMP. The role of dimer formation in the activity of the enzyme will be examined in future studies.

The dimer interface involves residues from both lobes of each monomer and is formed primarily by residues in helix 2, in the loop between strand 2 and helix 2, in strands 3 and 4, in the loop between helix 2 and strand 3, and in the section of random coil between amino acids 165 and 171 (Figure 8b). Although the location of the dimer interface is conserved between the human and tritrichomonal enzymes, neither the identities nor the positions of the residues making up this interface are well conserved. Of the 50 residues that contribute at least 5 Å^2 to the dimer interface in the *T. foetus* enzyme, only 9 are identical to the corresponding residues in the human HGPRTase sequence.

A second surprise was that the HGXPRTase dimer is asymmetric, with GMP bound to only one of the two monomers. The structure of this dimer gives us the first view of a purine salvage PRTase without bound ligand and allows us to examine the structural effects of ribonucleotide binding. The most striking structural consequence of GMP binding appears to be a change in the relative positions of the two lobes that form HGXPRTase. The product binding has led to a closing of the HGXPRTase structure, with the two lobes moving toward each other to enclose the ribonucleotide. The conformation change can best be described as a hinge movement in which the ribonucleotide binding leads to a rotation of approximately 5° of one lobe relative to the other around the hinge axis. This rotation moves lobe 2 residues involved in product binding approximately 1 Å closer to lobe 1 and into position for interacting with the ribonucleotide.

Individually, the two lobes undergo only small changes upon the binding of GMP. The rms deviation in the positions of α-carbon atoms between lobe 1 with and without bound nucleotide is only 0.35 Å (the residues used for the superposition were 17−71 and 85−150 of monomer 1 and the corresponding residues of monomer 2) (Figure 9a). A similar analysis of the changes in lobe 2 leads to an rms deviation of 0.36 Å between corresponding pairs of α-carbon atoms (using residues 7−16 and 151−179) (Figure 9b). The

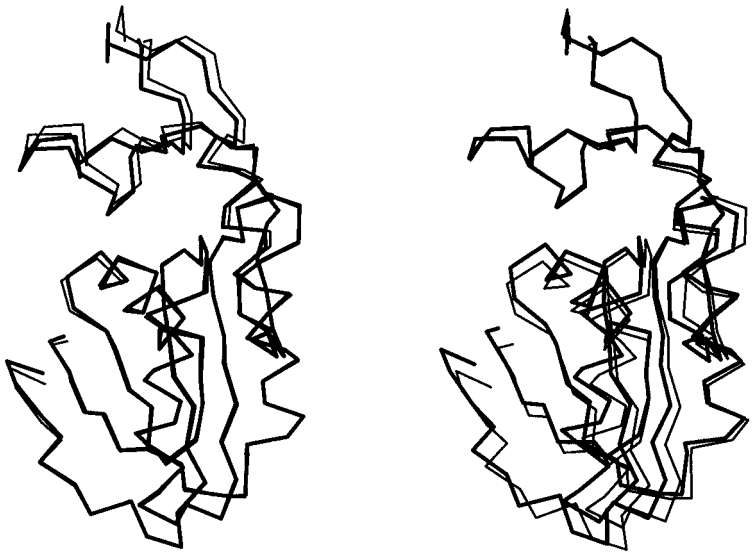

FIGURE 9: Comparison of the α-carbon traces of the two HGXPRTase molecules in the asymmetric unit. The molecule with bound GMP is the one drawn in thicker lines. In (a, left), the superposition is based on α-carbon atoms from lobe 1 (residues 17−71 and 85−150). In (b, right), the α-carbons from lobe 2 are superimposed (residues 7−16 and 151−179). Only residues that were traced in both monomers are shown.

GMP binding does lead to a few structural changes in the binding pocket and in the interface between the lobes. For example, in the open (unbound) form of the enzyme, the side chain of Tyr 156 has rotated away from the position that it occupies when stacking with the purine base. There is also a small but visible change in the structure of the loop between strand 2 and helix 2. Finally, it should be noted that in the monomer containing the bound nucleotide we were able to trace more of the electron density corresponding to the loop between strands 3 and 4 than in the monomer without the nucleotide. There are two reasonable explanations for this extra electron density. On one hand, in monomer 1 this region is in contact with a symmetry-related molecule, and it is possible that this packing interaction is providing additional stability to this area of the structure. A second possibility is that the GMP binding is stabilizing this loop.

A sequence comparison of the PRTases suggests that these enzymes share the bilobal structure seen in the human HGPRTase and *T. foetus* HGXPRTase crystal structures. Therefore, it is tempting to suggest that the lobe movement observed in HGXPRTase is common to other purine salvage PRTases. One caveat to this proposal is that some of the PRTases, like human HGPRTase, have a longer N-terminus than the *T. foetus* enzyme, and at least in the case of the human enzyme, this N-terminal tail stretches between the two lobes, creating a thicker hinge (Figure 3). Thus, enzymes with the longer N-terminus may have more rigid hinges and might be less likely to undergo lobe movement.

It is not clear whether the asymmetry in the dimer reflects some aspect of the enzyme's activity or whether it is simply due to packing requirements of the crystal lattice. At the moment there is no independent evidence confirming the existence of an asymmetric dimer in solution or that dimerization plays a role in enzyme activity. However, there are features of the crystal structure that suggest that dimer formation is involved in the regulation of the enzymatic activity of the HGXPRTase. Most notably, the fact that the dimer interface is formed by residues from both lobes (Figure 8b) suggests that dimer formation restricts the hinge movement between the lobes that is presumably necessary for catalysis.

CONCLUSIONS

On the basis of the structures of *T. foetus* HGXPRTase and human HGPRTase, and on sequence comparisons with other purine salvage PRTases, it is likely that even the most distantly related enzymes of the purine PRTase family will share the bilobal structure that has been described above and by Eads *et al.* (1994). In spite of the similarities in the fold of HGXPRTase and HGPRTase, the active sites of these two enzymes are substantially different. The observed differences between the active sites provide a first step toward understanding the variation in purine base specificity seen throughout the purine salvage PRTase group of enzymes and also suggest that there will be significant variations in how these enzymes interact with PRPP. A major long-term goal of our work on the *T. foetus* HGXPRTase is to use the three-dimensional structure of this enzyme to identify compounds that will inhibit this enzyme but not its mammalian counterpart. The differences in the active sites of the two enzymes suggest that this is a realistic goal.

The crystal structure of the asymmetric dimer gives us our first look at a purine PRTase without bound ligand and allows us to witness the structural consequences of GMP binding. The hinge motion that the HGXPRTase undergoes between the bound and unbound states is likely to be a property of at least a subset of the (H)(G)(X)PRTases. Also, the way in which the two monomers interact to form the dimer suggests that dimer formation may be involved in the regulation of the enzyme's activity. This possibility needs to be tested, but if correct, these observations may have general implications for the mechanism of regulation of other proteins that undergo hinge movement.

ACKNOWLEDGMENT

We thank Dr. J. Eads and Prof. J. Sacchettini for providing us with the atomic coordinates of human HGPRTase. We

also thank Dr. David King for the mass spectrometry data, Dr. Henry Bellamy for help in collecting data at beamline 1−5 at SSRL, and Dr. Jaru Jancarik for advice on the crystallization of this enzyme.

REFERENCES

Allen, T. E., & Ullman, B. (1993) *Nucleic Acids Res. 21*, 5431−5438.
Allen, T. E., & Ullman, B. (1994) *Mol. Biochem. Parasitol. 65*, 233−245.
Allen, T. E., Hwang, H.-Y., Jardim, A., Olafson, R., & Ullman, B. (1995) *Mol. Biochem. Parasitol. 73*, 133−143.
Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−197.
Beck, J. T., & Wang, C. C. (1993) *Mol. Biochem. Parasitol. 60*, 187−194.
Bouwens-Rombouts, A. G. M., van den Boogaard, M. J. H., Puig, J. G., Mateos, F. A., Hennekam, R. C. M., & Tilanus, M. G. J. (1993) *Hum. Genet. 91*, 451−454.
Brünger, A. T. (1992) *X-PLOR, version 3.1. A system for x-ray crystallography and NMR*, Yale University Press, New Haven, CT.
Chin, M. S. (1995) The hypoxanthine−guanine−xanthine phosphoribosyltransferase in *Tritrichomonas foetus*, Ph.D. Thesis, The University of California, San Francisco.
Chin, M. S., & Wang, C. C. (1994) *Mol. Biochem. Parasitol. 63*, 221−229.
Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
Craig, S. P., III, McKerrow, J. H., Newport, G. R., & Wang, C. C. (1988) *Nucleic Acids Res. 16*, 7087−7101.
Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., & Sacchettini, J. C. (1994) *Cell 78*, 325−334.
Engh, R. A., & Huber, R. (1991) *Acta Crystallogr. A47*, 392−400.
Fitzgerald, P. R. (1986) *Veterinary Clinic of North America*: *Food Animal Practice* Vol. 2, pp 277−283, W. B. Saunders, Philadelphia, PA.
Giacomello, A., & Salerno, C. (1978) *J. Biol. Chem. 253*, 6038−6044.
Jancarik, J., & Kim, S.-H. (1991) *J. Appl. Crystallogr. 24*, 409−411.
Johnson, G. G., Eisenberg, L. R., & Midgeon, B. R. (1979) *Science 203*, 174−176.
Jolly, D. J., Okayama, H., Berg, P., Esty, A. C., Filpula, D., Bohlen, P., Johnson, G. G., Shively, J. E., Hunkapillar, T., & Friedmann, T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 477−481.
Jones, T. A. (1992) in *Molecular Replacement* (Dodson, E. J., *et al.*, Eds.) pp 92−105, SERC Daresbury Laboratory, Warrington, U.K.
Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
King, A., & Melton, D. W. (1987) *Nucleic Acids Res. 15*, 10469−10481.
Kleywegt, G. J., & Jones, T. A. (1993) *ESF/CCP4 Newsl. 28*, 56−59.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
McDonald, I. K., & Thornton, J. M. (1994) *J. Mol. Biol. 238*, 777−793.
Musick, W. D. (1981) *CRC Crit. Rev. Biochem. 11*, 1−34.
Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
Otwinowski, Z. (1988) *DENZO. A program for automatic evaluation of film densities*, Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT.
Pratt, D., & Subramani, S. (1983) *Nucleic Acids Res. 11*, 8817−8823.
Scapin, G., Grubmeyer, C., & Sacchettini, J. C. (1994) *Biochemistry 33*, 1287−1294.
Scapin, G., Ozturk, D. H., Grubmeyer, C., & Sacchettini, J. C. (1995) *Biochemistry 34*, 10744−10754.
Smith, J. L., Zaluzec, E. J., Wery, J. P., Niu, L., Switzer, R. L., Zalkin, H., & Satow, Y. (1994) *Science 264*, 1427−1433.
Speer, C. A., & White, M. W. (1991) *Large Anim. Vet. 46*, 18−20.
Stewart, D. E., Sarkar, A., & Wampler, J. E. (1990) *J. Mol. Biol. 214*, 253−260.
Strauss, M., Behlke, J., & Goerl, M. (1978) *Eur. J. Biochem. 90*, 89−97.
Vasanthakumar, G., van Ginkel, S., & Parish, G. (1994) *Gene 147*, 153−154.
Wang, C. C., Wang, A. L., & Rice, A. (1984) *Exp. Parasitol. 57*, 68−75.
Wilson, J. M., Stout, J. T., Palella, T. D., Davidson, B. L., Kelley, W. N., & Caskey, C. T. (1986) *J. Clin. Invest. 77*, 188−195.

BI953072P