Short sequence-paper

Cloning, characterization and preliminary crystallographic analysis of *Leishmania* hypoxanthine–guanine phosphoribosyltransferase

Paulo S. Monzani\(^a\), Juan D. Alfonzo\(^b\), Larry Simpson\(^b\), Glaucius Oliva\(^a\), Otavio H. Thiemann\(^a,\*\)

\(^a\) Laboratory of Protein Crystallography and Structural Biology, Physics Institute of São Carlos, University of São Paulo—USP, Av. Trabalhador Sãocarlense 400, 13566-590, São Carlos—SP, Brazil

\(^b\) Howard Hughes Medical Institute, MacDonald Building, 90095-1662, UCLA, Los Angeles, CA, USA

Received 21 November 2001; accepted 4 February 2002

Abstract

Hypoxanthine–guanine phosphoribosyltransferase (HGPRT) (EC 2.4.2.8) is an important enzyme involved in the recycling of purine nucleotides in all cells. Parasitic protozoa of the order Kinetoplastida are unable to synthesize purines de novo and use the salvage pathway for the synthesis of nucleotides; therefore, this pathway is an attractive target for antiparasitic drug design. The \(hgprt\) gene was cloned from a *Leishmania tarentolae* genomic library and the sequence determined. The *L. tarentolae* \(hgprt\) gene contains a 633-nucleotide open reading frame that encodes a 23.4-kDa protein. A pairwise alignment of the different HGPRT’s sequences revealed a 26%–53% sequence identity with the *Leishmania* sequences and 87% identity to the HGPRT of *Leishmania donovani*. A recombinant protein was expressed in *Escherichia coli*, purified to homogeneity and found to retain enzymatic activity. The steady-state kinetic parameters were determined for the recombinant enzyme and the enzyme is active as a homodimer in solution. Single crystals were obtained for the *L. tarentolae* HGPRT representing the first *Leishmania* HGPRT crystallized and initial crystallographic data were collected. The crystals obtained belong to the orthorhombic space group (\(P2_12_12_1\)) with unit cell parameters \(a = 58.104\ \text{Å}, b = 85.443\ \text{Å}\) and \(c = 87.598\ \text{Å}\) and diffract to a resolution of 2.3 Å. The availability of the HGPRT enzyme from *Leishmania* and its crystallization suitable for X-ray diffraction data collection should provide the basis for a functional and structural analysis of this enzyme, which has been proposed as a potential target for rational drug design, in a *Leishmania* model system. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Crystallization; *Leishmania tarentolae*; Hypoxanthine–guanine phosphoribosyltransferase; Trypanosomatida

Most organisms synthesize purine nucleotides by both the de novo and the salvage pathways. In contrast, protozoan parasites are strict purine nucleotide auxotrophs because of the absence of a purine de novo biosynthetic pathway [1]. Enzymes of the phosphoribosyltransferase (PRTase) family catalyze the biosynthesis of purine nucleotides. Three PRTase enzymes are known to be involved in the recycling of purine nucleotides by the salvage pathway in the kinetoplastid protozoa from the genus *Leishmania*, hypoxanthine–guanine PRTase (HGPRT) (EC 2.4.2.8), adenine PRTase (APRT) (EC 2.4.2.7) and xanthine PRTase (XPRT) (EC 2.4.2.22) [2]. HGPRT is responsible for catalyzing the conversion of guanine and hypoxanthine and \(\alpha\)-\(\beta\)-5-phosphoribosyl 1-pyrophosphate (PRPP) into guanine-5-monophosphate (GMP), inosine-5-monophosphate and pyrophosphate (PPi) [3]. The only \(hgprt\) gene that has been cloned and characterized from the *Leishmania* genus is that from *Leishmania donovani* [4] and no structural data are known for the enzyme of this parasite genus. The availability of the *Trypanosoma cruzi* HGPRT structure [5] is of great comparative importance and will be used for the molecular substitution approach.

*Leishmania tarentolae* has been exploited as a model *Leishmania* for a variety of molecular, biochemical and evolutionary studies. The evolutionary data support the inclusion of *L. tarentolae* as a monophyletic clade branching between the *Viannia* and *Leishmania* subgenera [6–8]. As in the case of other trypanosomatids, *L. tarentolae* is a purine nucleotide auxotroph [9]. The ease of cell culture and genetic analysis of *L. tarentolae* should facilitate its use for site-directed mutagenesis of the \(hgprt\) gene as well as for functional complementation and testing of inhibitory substrates for the rational drug design for future leishmaniasis chemotherapy. We describe in this paper the cloning,
expression, characterization and the first crystallization of a Leishmania HGPRT protein. In view of the close phylogenetic relationship, the results will be of general significance as a model for the pathogenic Leishmania species.

A 207-bp fragment of L. tarentolae hgprt gene was PCR amplified from genomic DNA with degenerate oligodeoxynucleotides for two conserved regions, one comprising the purine site YLLCVLKGS and the other ILIVEDIV corresponding to the purine/pyrimidine phosphoribosyl transfersase signature at positions 121–128 from the L. donovani HGPRT sequence. The amplified DNA contained an open reading frame of 69 amino acids with sequence identity to other HGPRT polypeptides. The full-length hgprt gene was isolated from a L. tarentolae UC strain Lambda ZAP Express BamHI–SalI genomic library by hybridization screening with the 207-bp PCR-amplified L. tarentolae-specific probe. A 1.5-kbp genomic fragment was sequenced and found to contain a 633-nucleotide open reading frame that encodes a protein of 210 amino acids of 23.4 kDa expected molecular mass and a calculated pI of 7.64 (Fig. 1). The L. tarentolae HGPRT predicted amino acid sequence shares 87% identity with the HGPRT of L. donovani [4], divergence being mostly in the N- and C-terminus regions. Inspection of the sequence alignments in Fig. 1 reveals the conserved purine and PRPP binding domains [4,10] flanked by regions without significant sequence identity among the HGPRTs. The prediction of protein folding motifs from the individual polypeptides and the aligned sequences indicates a conservation of α-helical (H) and β-sheet (B) structural motifs that has allowed the molecular modeling of the HGPRT from L. tarentolae (data not shown). A pairwise alignment of the different sequences revealed a 26–53% sequence identity with the Leishmania sequences (data not shown).

The hgprt open reading frame was amplified by a PCR reaction with oligodeoxynucleotides specific for the L. tarentolae hgprt gene 5' and 3' ends. Oligodeoxynucleotide primers for PCR amplification (GIBCO-BRL) introduce both NdeI and Xhol restriction sites at its 5' and 3' end, respectively ( 5'-AGCTCCATATGAGCAACTCAGTC-3'; 5'-AGCTAAGCTTGATGTCGAACGAGGGCG-3'). The PCR reaction containing 2 pmol of each primer and approximately 0.5 ng of the 1.5-kbp genomic fragment was carried out in a GeneAmp 2400 thermocycler (Perkin-Elmer Cetus) with 2.5 U of AmpliTaq DNA polymerase (Promega) according to manufacturer's conditions. The sample was subjected to 2 min denaturation at 94 °C followed by 30 cycles of denaturation at 94 °C for 0.5 min, annealing at 40 °C for 0.5 min, and extension at 72 °C for 1 min. A DNA band of approximately 650 bp in length was gel purified by the NaI-glass powder method [12]. The purified DNA was digested with the restriction enzymes NdeI and Xhol, cloned into the pET29a(+) expression vector (Novagen). The PCR reaction containing 2 pmol of each primer and approximately 0.5 ng of the 1.5-kbp genomic fragment was carried out in a GeneAmp 2400 thermocycler (Perkin-Elmer Cetus) with 2.5 U of AmpliTaq DNA polymerase (Promega) according to manufacturer's conditions. The sample was subjected to 2 min denaturation at 94 °C followed by 30 cycles of denaturation at 94 °C for 0.5 min, annealing at 40 °C for 0.5 min, and extension at 72 °C for 1 min. A DNA band of approximately 650 bp in length was gel purified by the NaI-glass powder method [12]. The purified DNA was digested with the restriction enzymes NdeI and Xhol, cloned into the pET29a(+) vector into the Ndel/XhoI sites and transformed into BL21(DE3) competent cells. The cloned hgprt

![Fig. 1. Multiple alignment of representative HGPRT sequences. The amino acid sequences of several HGPRT proteins are shown aligned with the L. tarentolae and L. donovani sequences. The amino acid positions are indicated at the top of the alignment. The consensus of the alignment is indicated. “*” indicates amino acid identities, “+” and “-” indicate conserved substitutions. The boxes indicate the predicted purine binding domain (Site I) and predicted PRPP binding domain (Site II). The sequences are as follows: Ldon (Leishmania donovani L25412), Ltar (Leishmania tarentolae AF139722), Smans (Schistosoma mansoni X13531), Homo sapiens (Homo sapiens NM_000194), Rattus norvegicus (Mus musculus M86443), Rbuc (Trypanosoma brucei L10721), Tcruz (Trypanosoma cruzi L07486), Tric (Trichomonas foetus L06822), Lact (Lactococcus lactis X67015), Tgond (Toxoplasma gondii U09219) and Clong (Cricetulus longicaudatus X17656). The sequences were aligned by the CLUSTALX [11] with further manual refinements.](image-url)
sequence identity was confirmed by the BigDye-termination method and the T7 and T7-terminator sequencing primers in an ABI377 DNA sequencer (Perkin-Elmer). Cells from a single colony were grown overnight at 37 °C, 250 rpm in 5 ml LB medium containing 25 μg/ml kanamycin. A larger cell culture was grown in 2XYT with 25 μg/ml kanamycin.

Fig. 2. (A) Expression of \textit{L. tarentolae} \textit{hgp}rt in \textit{E. coli} and purification of the recombinant protein. The \textit{L. tarentolae} \textit{hgp}rt gene was subcloned into the pET29a (+) expression vector and overexpressed in \textit{E. coli} BL21(DE3). The proteins were separated in a 15% SDS-PAGE gel. Lanes: 1—crude lysate, 2—clarified extracts (S20000 supernatant), 3—S20000 protein pellet, 4—ammonium sulfate precipitate fraction, 5—ammonium sulfate supernatant fraction, 6—pooled fractions from POROS 20HQ passage. (B) Crystallization of \textit{L. tarentolae} HGPRT. (1) HGPRT crystals obtained in 20% PEG 6000, 100 mM citric acid and pH 5.0; (2) 15% PEG 6000, 100 mM citric acid and pH of 5.1, at 18 °C; (3) 17% PEG 4000, 5% glycerol, 19% isopropanol, pH 5.6; and (4) in crystals obtained in 19% isopropanol, 19% PEG 4000, 5% glycerol, pH 5.6, at 18 °C. Diffraction data were collected from crystals of condition 4. Interestingly, no diffraction pattern was detectable from the crystals from conditions 1 and 2.
Table 1
Data collection statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2_12_1</td>
</tr>
<tr>
<td>Unit cell parameters in Å</td>
<td>a = 58.104</td>
</tr>
<tr>
<td></td>
<td>b = 85.443</td>
</tr>
<tr>
<td></td>
<td>c = 87.598</td>
</tr>
<tr>
<td>Resolution</td>
<td>20 – 2.3 Å</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>18,895</td>
</tr>
<tr>
<td>Rmerge* (last shell) (%)</td>
<td>8.1 (25.0)</td>
</tr>
<tr>
<td>Completeness (last shell)</td>
<td>94.5 (96.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.67</td>
</tr>
<tr>
<td>&lt;I/σ&gt;</td>
<td>9.9</td>
</tr>
</tbody>
</table>

$R_{\text{merge}} = \Sigma |I - \langle I\rangle| / \Sigma I$.

at 37 °C, 250 rpm, until OD$_{600}$ = 0.6. The 250-ml culture was then induced for 4 h at 37 °C, 250 rpm, with 0.5 mM IPTG.

The purification of recombinant HGPRT protein from *L. donovani* has been previously reported [4] using an affinity purification method with a GMP-Agarose column. Interestingly, the *L. tarentolae* HGPRT protein binds with very low affinity to either the GMP or the GTP-Agarose columns. As a result, we developed a new purification protocol. All purification procedures were carried out at 4 °C unless specified. The *Escherichia coli* BL21(DE3)-induced cells were harvested by centrifugation at 4000 × g for 15 min. The cells were washed in 100 mM Tris–HCl, 10 mM MgCl$_2$, pH 7.5 (buffer TMD). Cell lysis was obtained by five cycles of freeze–thaw in TMD buffer plus 1% Triton X-100 followed by six 1-min cycles of sonication in an ice bath. The crude extract was clarified by centrifugation (20,000 × g, 30 min) and brought to 60% (w/v) ammonium sulfate for 20 min with slow shaking. The suspension was separated by centrifugation for 20 min at 20,000 × g. The protein pellet was dissolved in 8 ml TMD and dialyzed against four changes of 250 ml TMD. The HGPRT from the ammonium sulfate fraction was loaded in a 1.6-ml POROS-20HQ column, equilibrated in TMD buffer adjusted to pH 9.0 with a 1-ml/min flux at room temperature. The HGPRT eluted in the void volume of the column with only trace amounts of contaminants (Fig. 2A). The protein fractions eluted along the NaCl gradient (0–1 M) did not reveal the presence of HGPRT protein by SDS-PAGE or enzymatic activity assay (data not shown). For crystallization purposes, the fractions eluted from the POROS-20HQ column were pooled, concentrated in a Centriprep 10 (10 kDa cutoff) to a volume of approximately 6 ml and loaded in a second pass trough the POROS-20HQ chromatography, as described above. The void volume of the second chromatography contained highly purified HGPRT from *L. tarentolae* (Fig. 2A) yielding approximately 70 mg HGPRT per liter of culture and over 95% purity based on SDS-PAGE analysis (Fig. 2A). The recombinant HGPRT protein migrates as a 23-kDa protein in 15% SDS-PAGE (Fig. 2A) and retained enzymatic activity and substrate specificity (Table 2). The *L. tarentolae* recombinant HGPRT is a homodimer of approximately 50 kDa in the presence of GMP, hypoxanthine or the free enzyme form (Fig. 3C). Samples of the purified *Leishmania* HGPRT were analyzed by isoelectric focusing in a FAST System (Pharmacia Biotech) and a pI of 8.20 was observed (data not shown). This pI is consistent with the result from the POROS-20HQ chromatography where the recombinant HGPRT eluted in the column void volume. The fractions collected in the void volume of the second chromatography were pooled and concentrated to 7 mg/ml in a Centriprep 10 (Millipore).

The crystallization conditions for the *L. tarentolae* HGPRT protein were screened by the hanging-drop vapor-diffusion method. The sparse-matrix kits Crystal Screen I, Crystal Screen II, Crystal Screen PEG6K, Ammonium Sulfate, Cryo-Crystal Screen and Crystal Screen MPD from Hampton Research were used to test initial conditions at 4 and 18 °C. A 3-µl HGPRT solution containing 7 mg/ml enzyme, in TMD buffer pH 7.5, was mixed with an equal volume (3 µl) of the well solutions (500 µl/well) to form the drop. Screening different concentrations of precipitating agent and pH further optimized the initial crystallization conditions. The best crystallization conditions so far obtained for the *L. tarentolae* HGPRT were at 18 °C in 19% iso-propanol, 19% PEG 4000, 5% glycerol, pH 5.6 and in 17% PEG 4000, 5% glycerol (Fig. 2B). Minor variations in the crystallization conditions, showed no significant improvements in the crystal quality as shown in Fig. 2B. Crystals of long orthorhomobic habit diffracted at 2.3 Å. In 20% PEG 6000, 100 mM citric acid and pH of 5.0 and in 15% PEG 6000, 100 mM citric acid and pH of 5.1 at 18 °C (Fig. 2B) several hexagonal shape HGPRT crystals were obtained. Hexagonal crystals were also obtained at 4 °C with subtle differences in the overall crystal shape. These crystals did not diffract properly and were discarded.

Crystals were mounted in nylon loops (Hampton Research) after quick soaking in a cryo-protectant solution

---

**Fig. 3.** Determination of the *K*_m for guanine, hypoxanthine and PRPP of HGPRT and molecular mass. (A) Double reciprocal plots of the HGPRT reaction for guanine (○) and hypoxanthine (○). The concentration of PRPP was kept constant at 1 mM. Concentrations were held constant at 3, 5, 10, 15, 20, 30 and 40 µM for guanine and 3, 5, 10, 15, 20, 50 and 100 µM for hypoxanthine. (B) Double reciprocal plots of the HGPRT reaction for PRPP at constant concentrations of 40 µM guanine (○) and 100 µM hypoxanthine (○) while the concentration of PRPP was 10, 30, 50, 100, 200, 500 and 1000 µM. (C) Sizing of the HGPRT. Recombinant *L. tarentolae* HGPRT was chromatographed on a Superose 12 column as described. Different fractions are plotted as a function of OD 280. Proteins of known molecular mass (1 – 66 kDa bovine serum albumin, 2 – 45 kDa ovalbumin, 3 – 29 kDa carbonic anhydrase and 4 – 13 kDa cytochrome C) were used as standards in the calibration of the column. HGPRT free enzyme and with GMP or PRPP bound (H) elute between the BSA (1) and ovalbumin (2) standards, corresponding to approximately a 46-kDa homodimer.
consisting of the same reservoir solution and 15% ethylene glycol. Diffraction data of the long orthorhombic crystals were collected at the Protein Crystallography beamline (P212121) at the National Synchrotron Light Laboratory (LNLS, Campinas, Brazil) using monochromatic X-rays set on 1.537 Å. A total of 75 frames were collected at 100 K, with an exposure time of 5 min and a crystal-to-detector distance of 155 mm. The collected diffraction data were processed using DENV0 and SCALEPACK [13]. The crystals belong to the primitive orthorhombic space group P212121 with unit cell parameters of a = 58.104 Å, b = 85.443 Å, c = 87.958 Å. Statistical analysis of the merged reflections indicated that the collected data set is better described as having a resolution of 2.3 Å (Table 1). Such resolution range is sufficient for initial structure determination and modeling.

The kinetic parameters were measured at 25 °C following the consumption of guanine at 255 nm and hypoxanthine at 243 nm for 1 min in a 1-ml reaction volume containing 100 mM Tris-HCl, 5 mM MgSO4, pH 7.4, 1 mM PRPP and 0.04 mM guanine or 0.1 mM hypoxanthine [14]. An extinction coefficient of 4.2 for guanine and of 2.2 for hypoxanthine was used. To calculate the first-order rate constants (Km and Vmax) for guanine, hypoxanthine or PRPP, the other components were kept constant at 40 μM for guanine, 100 μM for hypoxanthine and 1 mM for PRPP according to the experiment (Table 2). The substrate concentrations for the Km and Vmax measurements were of 3–40 μM guanine, 3–100 μM hypoxanthine and 10–1000 μM PRPP. The kinetic parameters of L. tarentolae HGPRT were analyzed and Km and Vmax values of 2.8 μM/10.9 μM/min for guanine and 4.4 μM/5.7 μM/min for hypoxanthine were obtained. The PRPP Km and Vmax were measured with guanine or hypoxanthine as a substrate and the following results were obtained: Km and Vmax of 127.1 μM/10.9 μM/min with guanine and 137.6 μM/5.2 μM/min with hypoxanthine (Table 2).

The ability to overexpress and purify the recombinant L. tarentolae protein in two chromatographic steps, independent of its binding to GMP or GTP columns, should allow the purification of HGPRT mutants with different affinities for the substrate that would not bind with the same affinity to those columns. This is an interesting feature for the structural–functional study of the enzyme.

We have characterized the kinetic parameters of the recombinant L. tarentolae HGPRT. The values of Km obtained for the substrate guanine and hypoxanthine, 2.8 and 4.4 μM, respectively, are lower than the values described for the L. donovani homologue of 3.8 and 7.6 μM, respectively. Such differences in Km values can be due to small structure differences between the two enzymes. It is interesting to note that the L. tarentolae HGPRT is a homodimer either in the free enzyme form or combined with the product GMP or the substrate PRPP as determined by size exclusion chromatography.

Finally, the successful crystallization of the L. tarentolae HGPRT enzyme, suitable for structure determination, should allow us to compare the Leishmania enzyme with the other HGPRTases crystallized to date. Such knowledge may be valuable for future structure-based drug design strategies using this enzyme as a model system for the Leishmania genus. Interestingly, we have observed differences in Km and Vmax between the L. tarentolae and L. donovani enzymes in spite of the high sequence conservation observed between those enzymes. The structural and functional comparison between the Leishmania HGPRT and the available structures of T. cruzi [5,15], human [16–18] and other HGPRTases should allow us to investigate this important enzyme from Leishmania and possibly explain how subtle structural differences contribute to the enzyme catalytic activity. We are currently refining the structure.

**Acknowledgements**

The work was supported in part by a Young Scientist Research Grant 98-14979-7 to O.H. Thiemann from the Fundação de Amparo a Pesquisa no Estado de São Paulo (FAPESP). G. Oliva is an International Scholar from the Howard Hughes Medical Institute. Paulo S. Monzani is a FAPESP doctorate fellow. We thank the members of the Protein Crystallography and Structural Biology Group (IFSC-USP) and the Simpson’s laboratory (HHMI-UCLA) for helpful discussions in the course of this work. We express our gratitude also to: Dr. Beatriz D. Lima (UnB-Brasilia) for the L. tarentolae UC strain Lambda ZAP Express genomic library used in this study, José Brandão (LNLS, Campinas) and Humberto d’Muniz Pereira (IFSC-USP, São Carlos) for technical support in the data collection.

**Table 2**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Km L. tarentolae (μM)</th>
<th>Km L. donovani* (μM)</th>
<th>Vmax L. tarentolae (μM/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guanine–PRPP</td>
<td>3.8</td>
<td>2.8</td>
<td>10.9</td>
</tr>
<tr>
<td>Hypoxanthine–PRPP</td>
<td>7.6</td>
<td>4.4</td>
<td>5.7</td>
</tr>
<tr>
<td>PRPP–Guanine</td>
<td>ND</td>
<td>127.1</td>
<td>10.9</td>
</tr>
<tr>
<td>PRPP–Hypoxanthine</td>
<td>ND</td>
<td>137.6</td>
<td>5.2</td>
</tr>
</tbody>
</table>

ND = Not determined.

* Published data from L. donovani [14].

**References**


[2] B. Ullman, D. Carter, Molecular and biochemical studies on the hy-


