Crystal Structure of *Escherichia coli* Xanthine Phosphoribosyltransferase†

Siska Vos,‡ John de Jersey,‡ and Jennifer L. Martin*,§

Centre for Protein Structure, Function and Engineering, Department of Biochemistry, and Centre for Drug Design and Development, The University of Queensland, Brisbane, Queensland 4072, Australia

Received October 21, 1996; Revised Manuscript Received January 2, 1997⊗

ABSTRACT: Xanthine phosphoribosyltransferase (XPRT; EC 2.4.2.22) from *Escherichia coli* is a tetrameric enzyme having 152 residues per subunit. XPRT catalyzes the transfer of the phosphoribosyl group from 5-phospho-α-D-ribosyl 1-pyrophosphate (*P*Rib-*PP*) to the 6-oxopurine bases guanine, xanthine, and hypoxanthine to form GMP, XMP, and IMP, respectively. Crystals grown in the absence of substrate or product were used to determine the structure of XPRT at a resolution of 1.8 Å, by multiple isomorphous replacement. The core structure of XPRT includes a five-stranded parallel β-sheet surrounded by three α-helices, which is similar to that observed in other known phosphoribosyltransferase (PRTase) structures. The XPRT structure also has several interesting features. A glutamine residue in the purine binding site may be responsible for the altered 6-oxopurine base specificity seen in this enzyme compared to other 6-oxopurine PRTases. Also, we observe both a magnesium ion and a sulfate ion bound at the *P*Rib-*PP* binding site of XPRT. The sulfate ion interacts with Arg-37 which has a *cis*-peptide conformation, and the magnesium ion interacts with Asp-89, a highly conserved acidic residue in the *P*Rib-*PP* binding site motif. The XPRT structure also incorporates a feature which has not been observed in other PRTase structures. The C-terminal 12 residues of XPRT adopt an unusual extended conformation and make interactions with a neighboring subunit. The very last residue, Arg-152, could form part of the active site of a symmetry-related subunit in the XPRT tetramer.

Xanthine phosphoribosyltransferase (XPRT,[1] EC 2.4.2.22) from *Escherichia coli* is a purine salvage enzyme which catalyzes the reaction of guanine, xanthine, and to a lesser extent hypoxanthine, with 5-phospho-α-D-ribosyl 1-pyrophosphate (*P*Rib-*PP*), to form GMP, XMP, or IMP, respectively (Deo *et al*., 1985). It is one of two enzymes responsible for the salvage of 6-oxopurines in *E. coli*, the second being hypoxanthine phosphoribosyltransferase (HPRT) (Jochimsen *et al*., 1975).

XPRT is a member of a group of enzymes, known as the phosphoribosyltransferases (PRTases), that catalyze the formation of purine, pyrimidine, and pyridine nucleotides and, in bacteria and lower eukaryotes, the formation of histidine and tryptophan (Musick, 1981). Although there is very little protein sequence similarity between these enzymes, a highly conserved stretch of residues has been identified in many of the PRTases (Hershey & Taylor, 1986) and in the *P*Rib-*PP* synthetases (Hove-Jensen *et al*., 1986). This sequence, termed the *P*Rib-*PP* binding site motif, includes two adjacent acidic residues surrounded by hydrophobic residues and followed by smaller residues, of which one is often a glycine.

The three-dimensional structures of several PRTases have recently been determined (Table 1). They share a common core region consisting of five parallel β-strands surrounded by three or four α-helices. The *P*Rib-*PP* binding site motif is located at the C-terminal end of the third β-strand in the core β-sheet. In those PRTase structures with substrate or product bound, residues from the *P*Rib-*PP* motif interact with the 5′-phosphate group of *P*Rib-*PP* and encircle the ribose ring of either *P*Rib-*PP* or nucleotide. Specifically, the crystal structure of *Salmonella typhimurium* orotate PRTase (OPRT) complexed with *P*Rib-*PP* (Scapin *et al*., 1995) shows that the first of the two adjacent aspartates interacts via a carboxylate oxygen with the O3 group on the ribose ring of *P*Rib-*PP* with a hydrogen bond distance of 3.1 Å. In the structures of human hypoxanthine PRTase (HPRT) complexed with GMP, *Tritrichomonas foetus* HPRT complexed with GMP, and *Bacillus subtilis* glutamine *P*Rib-*PP* amidotransferase (GAT) complexed with AMP, the carboxylate oxygen of the equivalent aspartate residue interacts via a hydrogen bond (2.5−2.9 Å) with either the O2 or O3 hydroxyl groups on the ribose ring of the bound nucleotide (Eads *et al*., 1994; Somoza *et al*., 1996; Smith *et al*., 1994). However, in the crystal structure of *S. typhimurium* OPRT complexed with OMP, no interaction between this aspartate and the ribose ring is observed (Scapin *et al*., 1994). Instead, the ribose ring is reoriented in the active site, so that the ribose hydroxyls of OMP interact with the side chain Nζ of Lys-26. The carboxylate oxygens of the two adjacent

† This work was supported by a grant from the Australian Research Council and a Queen Elizabeth II fellowship to J.L.M. Coordinates and structure factors for the XPRT structure have been deposited with the Brookhaven Protein Data Bank (accession code 1NUL). Structure factors will be on hold for 4 years from the date of publication.

* Author to whom correspondence should be addressed [telephone, +61 7 3365 4942; Fax, +61 7 3365 1990; e-mail, J.Martin@mailbox.uq.oz.au].

‡ Centre for Protein Structure, Function and Engineering, Department of Biochemistry.

§ Centre for Drug Design and Development.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1997.

[1] Abbreviations: EMP, ethylmercuric phosphate; GAT, glutamine *P*Rib-*PP* amidotransferase; GlutHg, glutaraldehyde cross-linked and mercury-soaked crystal; Glut, glutaraldehyde cross-linked crystal; GMP, guanosine 5′-monophosphate; HPRT, hypoxanthine phosphoribosyltransferase; IMP, inosine 5′-monophosphate; OPRT, orotate phosphoribosyltransferase; PEG 4000, poly(ethylene glycol) of average molecular weight 4000; *P*Rib-*PP*, 5-phospho-α-D-ribosyl 1-pyrophosphate; PRTase, phosphoribosyltransferase; Se-Met, selenomethionine; XMP, xanthosine 5′-monophosphate; XPRT, xanthine phosphoribosyltransferase.

S0006-2960(96)02640-2 CCC: $14.00

Table 1: Summary of PRTase Structures Determined to Date

source	PRTase	no. of residues per subunit	ligand	resolution (Å)	reference	PDB code
S. typhimurium	OPRT	213	OMP	2.6	Scapin *et al.*, 1994	1STO
S. typhimurium	OPRT	213	*P*Rib-PP + orotate	2.3	Scapin *et al.*, 1995	1OPR
B. subtilis	GAT	465	AMP	3.0	Smith *et al*., 1994	1GPH
human	HPRT	217	GMP	2.5	Eads *et al*., 1994	1HMP
T. foetus	HPRT	183	±GMP	1.9	Somoza *et al*., 1996	1HGX
E. coli	OPRT	213	sulfate	2.4	Henrikson *et al*., 1996	1ORO
T. gondii	HPRT	231	Mg^{2+}	2.4	Schumacher *et al*., 1996	1DBR
			XMP	2.9		
E. coli	XPRT	152	sulfate/Mg^{2+}	1.8	this paper	1NUL

aspartates in both *S. typhimurium* and *E. coli* OPRT (Henriksen *et al*., 1996) are within hydrogen-bonding distance (2.7−3.1 Å), suggesting that one of the two acidic groups is protonated. In the crystal structure of *Toxoplasma gondii* HPRT, both the acidic groups interact with a bound magnesium (Schumacher *et al*., 1996).

The PRTases require a divalent metal ion such as magnesium for catalysis. The precise role of the metal ion in catalysis is unknown. It is thought that the Mg/*P*Rib-*PP* complex is the substrate species which undergoes catalysis in the PRTases (Bhatia & Grubmeyer, 1993). However, at least some of the PRTases are able to bind free magnesium ion (Kosaka *et al*., 1977; Victor *et al*., 1979; Ali & Sloan, 1983), implying that free *P*Rib-*PP* may also be a substrate for these enzymes (Kosaka *et al*., 1977).

The reaction catalyzed by the PRTases is thought to proceed via a transition state with partial positive charge on the ribose ring oxygen of *P*Rib-*PP* (Tao *et al*., 1996). While several residues on *S. typhimurium* OPRT have been implicated in binding and catalysis (Ozturk *et al*., 1995a; Tao *et al*., 1996), the function of the conserved acidic residues in the *P*Rib-*PP* binding motif remains unclear. It has been postulated that the conserved acidic residues could stabilize the positively charged transition state intermediate of the reaction (Tao *et al*., 1996). Site-directed mutagenesis experiments on *E. coli P*Rib-*PP* synthetase have suggested that the second of these conserved acidic residues may have a role in metal ion binding (Willemoës *et al*., 1996). This role is supported by the recently reported structure of *T. gondii* HPRT (Schumacher *et al*., 1996), which has a magnesium interacting with this residue.

We are investigating the structure and properties of *E. coli* XPRT in a comparative study with *E. coli* HPRT, human HPRT, and *Plasmodium falciparum* HPRT. We have cloned the gene encoding XPRT into a high-yielding expression vector and have purified and characterized the recombinant protein. A mutant of XPRT, C59A, has also been prepared which is more stable than the wild-type protein while retaining similar kinetic properties. XPRT C59A has been crystallized into three different crystal forms, orthorhombic ($P2_12_12$), monoclinic (C2), and tetragonal ($P4_12_12$ or $P4_32_12$) (Vos *et al*., 1996). Orthorhombic form crystals have also been obtained for wild-type XPRT and these diffract to 1.8 Å resolution. We now report the crystal structure of the orthorhombic crystal form of XPRT and its comparison with other published PRTase structures.

EXPERIMENTAL PROCEDURES

Purification of XPRT. In this study the pT7-7, pGP1-2 dual plasmid expression system (Tabor & Richardson, 1985) was used in *E. coli* SØ606 cells (*ara*, Δ*pro-gpt-lac*, *thi*, *hpt*, F^-) for the overexpression of XPRT. The gene encoding XPRT, *gpt*, was obtained in the vector pSV2*gpt* from the American Type Culture Collection and, after the introduction of appropriate restriction sites into the flanking regions of the gene, was cloned into the pT7-7 vector. Recombinant protein was expressed as previously described for human HPRT (Free *et al*., 1990). XPRT was purified to homogeneity (as assessed by SDS−PAGE) by two column chromatography steps: DEAE-Trisacryl at pH 7.0 with elution by a 0−0.3 M KCl gradient and Sephacryl CL-4B gel filtration. The native molecular mass of XPRT was determined by chromatography on a 2.4 mL Superose-12 column using a Pharmacia SMART system and comparison of the retention time of XPRT with those of molecular mass standards (bovine γ-globulin, 158 kDa; ovalbumin, 44 kDa; equine myoglobin, 17 kDa; and cobalamin, 1350 Da) chromatographed under identical conditions.

Kinetic Characterization of XPRT. The specificity of XPRT for the purine bases, guanine, xanthine, and hypoxanthine, was determined spectrophotometrically by following their conversion to GMP at 257.5 nm ($\Delta\epsilon = 5817$ M^{-1} cm^{-1}), XMP at 255 nm ($\Delta\epsilon = 4685$ M^{-1} cm^{-1}), and IMP at 245 nm ($\Delta\epsilon = 2439$ M^{-1} cm^{-1}) at 25 °C, respectively (Keough *et al*., 1987). Assays were performed in Tris-HCl at pH 8.5 and 100 mM $MgCl_2$. The concentration of *P*Rib-*PP* used in this assay was determined by the method of Kornberg *et al.* (1955). One unit of enzyme activity is defined as 1 μmol of product (IMP, GMP or XMP) formed per minute.

Mass Spectrometry and Protein Sequencing. Mass spectrometry was performed using a PE SCIEX API III triple quadrupole mass spectrometer equipped with an ion spray atmospheric pressure ionization source, used in the positive mode. Full scan data were acquired by scanning quadrupole 1 from *m/z* 700 to 2100 with a scan step size of 0.2 Da and a dwell time of 2 ms. Samples of XPRT C59A at approximately 10 μM in water were injected with a rate of infusion of 3−5 μL/min.

Intact and degraded samples of XPRT C59A were sequenced at the Brisbane Centre for Protein and Nucleic Acid Research by Edman degradation using an API 470A protein sequencer.

Crystallization and Data Measurement. Orthorhombic crystals of XPRT were grown using the hanging drop vapor diffusion method from 18% to 23% PEG 4000, 0.1 M Tris-HCl, pH 8.5, and 0.1 M Li_2SO_4 (Vos *et al*., 1996). The space group of the crystal is $P2_12_12$, and the unit cell parameters are $a = 59.2$ Å, $b = 92.7$ Å, and $c = 53.2$ Å. The crystals diffract to beyond 1.8 Å and are stable in the

Table 2: Kinetic Constants for Wild-Type XPRT

substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\times 10^6$ M^{-1} s^{-1})
guanine[a]	4.3 ± 0.3	112.0 ± 1.5	26.0 ± 2.2
xanthine[a]	30.5 ± 2.6	150.1 ± 2.9	4.9 ± 0.5
hypoxanthine[a]	90.8 ± 11.3	54.8 ± 4.3	0.60 ± 0.12
*P*Rib-*PP*[b]	139 ± 16	112 ± 2	8.1 ± 1.0

[a] Measured in the presence of 1.7 mM *P*Rib-*PP*. [b] Measured in the presence of 72.9 μM guanine.

X-ray beam. There are two subunits in the asymmetric unit, giving a solvent content of 43%, which is within the normal range for proteins (Matthews, 1968).

The structure of *E. coli* XPRT was solved by multiple isomorphous replacement (MIR) methods using three heavy atom derivatives. The first heavy atom derivative was obtained by soaking crystals of wild-type XPRT in 1.1 mM ethylmercuric phosphate (EMP) for 28 days (EMP was added directly to the crystallization drop after the crystal had grown). A second mercury derivative was obtained after cross-linking an XPRT crystal with glutaraldehyde. A crystal of XPRT was transferred to a stabilizing solution (30% PEG 4000, 0.1 M Tris-HCl, pH 8.5, and 0.1 M Li_2SO_4) containing 0.5% glutaraldehyde and equilibrated for 18 h. A native data set for glutaraldehyde cross-linked XPRT (Glut) was measured. A glutaraldehyde cross-linked crystal was then soaked in stabilizing solution containing 0.25 mM mercuric acetate for 1 h prior to data collection (GlutHg). A third heavy atom derivative was obtained using a selenomethionine derivative of the C59A variant of XPRT (Se-Met), as described previously (Vos *et al*., 1996).

Crystallographic data were measured using a Rigaku R-AXIS IIC imaging plate area detector, with a Rigaku RU-200 copper target rotating anode X-ray source operating at 5.4 kW. Data were autoindexed and processed on a Silicon Graphics Indigo R4000 using the PROCESS package (Higashi, unpublished) provided by Rigaku Corp. for the processing of R-AXIS imaging plate data. Unit cell, crystal system, and space lattice information was confirmed with DENZO (Z. Otwinowski and W. Minor), and the space group was identified by systematic absences in the raw data. Data collection statistics for all native and derivative data sets are summarized in Table 3.

Structure Determination. The three heavy atom derivatives are isomorphous with their respective native form crystals. The R_{merge} values (based on $|F^2|$) are 20.0% for EMP-soaked versus wild-type data, 20.1% for GlutHg versus glutaraldehyde cross-linked data, and 12.4% for Se-Met versus C59A data. The position of one of the two mercury atoms in the EMP derivative was determined by difference Patterson synthesis. The positions of the remaining heavy atoms in both mercury derivatives as well as six of the eight selenium atoms in the Se-Met derivative were determined by difference Fourier techniques and refined using the program package PHASES (Furey, 1990). Phases were estimated for native reflections to a resolution of 2.5 Å, with a mean figure of merit of 0.523 for 9141 phased reflections (see Table 3). Solvent flattening (Wang, 1985), as implemented in PHASES, was performed to yield improved phases with a mean figure of merit of 0.760. The resulting MIR electron density map revealed the presence of a right-handed helix.

Model building was performed using the program O (Jones & Kjeldgaard, 1992) on a Silicon Graphics Indy R4400. When the structure of XPRT had been partially modeled, the noncrystallographic symmetry (NCS) relationship was determined by relating the two monomers in the asymmetric unit. NCS averaging of the electron density map increased the mean figure of merit to 0.825 (Table 4). When differences between the subunits of XPRT became discernible, the NCS restraint was released and MIR phases were extended to include reflections up to 1.8 Å. The final XPRT model is the result of several rounds of model building and refinement, utilizing maps calculated first from combined MIR and model-based phases and then by use of $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps. Omit maps (Bhat & Cohen, 1984), calculated by omitting 10 residues at a time, were used to assess the final model.

Structure Refinement. Initial models of XPRT were refined using the simulated annealing "slowcool" procedure in X-PLOR 3.1 (Brünger, 1992). Subsequent overall *B*-factor refinement improved the R_{free} from 46.1% to 39.7% (R_{crys} 35.3%). Several more cycles of modeling and refinement reduced the R_{free} to 37.4% and R_{crys} to 31.4%. Individual *B*-factor refinement further improved the R_{free} to 34.9% and

Table 3: Data Collection Statistics for Native and Derivative Data Sets of XPRT

	wt	Hg	Glut	GlutHg	C59A	Se
unit cell parameters						
a (Å)	59.2	59.7	59.3	59.5	59.2	59.2
b (Å)	92.7	94.1	92.7	93.3	92.9	93.1
c (Å)	53.2	53.7	53.3	53.4	53.2	53.2
α (deg)	90.0	90.0	90.0	90.0	90.0	90.0
β (deg)	90.0	90.0	90.0	90.0	90.0	90.0
γ (deg)	90.0	90.0	90.0	90.0	90.0	90.0
maximum resolution (Å)	1.8	2.1	2.08	2.1	2.0	1.8
no. of observations	77 524	41 310	37 383	28 812	90 011	70 272
no. of unique reflections	25 622	16 223	15 375	12 833	19 047	23 842
no. of rejected observations	20	370	807	315	409	280
mosaicity (deg) (refined value)	0.35	0.55	0.59	0.59	0.4	0.37
completeness of data (%)	91.8	87.8	84.1	69.3	92	87.9
(in highest resolution shell)[a]	(81.9)	(76.6)	(66.6)	(52.4)	(83)	(78.5)
$I/\sigma I$	14.8	12.0	5.6	6.4	13.0	12.7
(in highest resolution shell)[a]	(3.0)	(3.4)	(2.3)	(2.5)	(3.5)	(3.4)
R_{merge} (%)	6.0	8.9	11.4	10.9	6.4	7.3
(in highest resolution shell)[a]	(26.3)	(23.5)	(30.2)	(28.6)	(25.4)	(27.5)

[a] Highest resolution shell for wt is 1.8−2.0 Å, for Hg is 2.1−2.25 Å, for Glut is 2.08−2.25 Å, for GlutHg is 2.1−2.25 Å, for C59A is 2.0−2.05 Å, and for SeMet is 1.8−1.9 Å.

Table 4: Heavy Atom Derivatives Used in the Structure Determination of XPRT

derivative	site	*x*	*y*	*z*	occ[a]	*B*[b]	R_c[c]	R_k[d]	no. of reflections	phasing power[e]	*D* (Å)[f]	FOM[g]
EMP	Hg1	0.131	0.162	0.208	1.000	20.00	0.615	0.090	7685	1.53	2.50	0.325
	Hg2	0.262	0.250	0.303	0.255	28.03						
GlutHg	gHg1	0.111	0.166	0.234	1.000	20.00	0.613	0.086	4404	1.60	2.50	0.388
	gHg2	0.267	0.261	0.278	0.645	2.81						
Se-Met	Se1	0.075	0.011	0.434	1.000	20.00	0.613	0.150	8285	1.57	2.50	0.379
	Se2	0.668	0.205	0.539	1.051	44.83						
	Se3	0.653	0.036	0.184	1.069	26.29						
	Se4	0.516	0.252	0.099	1.330	30.08						
	Se5	0.548	0.103	0.391	1.074	30.97						
	Se6	0.069	0.211	0.030	0.837	21.50						
overall figure of merit												0.523
figure of merit after solvent flattening												0.760
figure of merit after map averaging												0.825

[a] Relative occupancy. [b] Isotropic temperature factor. [c] R_c, Cullis *R*-factor, $\sum|||F_{PH}|_{obs}| \pm |F_P|_{obs}| - |F_H|_{calc}|/\sum||F_{PH}|_{obs} \pm |F_P|_{calc}|$, where F_P, F_{PH}, and F_H are the native structure factor, the protein + heavy atom structure factor, and the heavy atom structure factor, respectively. [d] R_k, Kraut *R*-factor, $\sum||F_{PH}|_{obs} - |F_P|_{calc}|/\sum|F_{PH}|_{obs}$. [e] The phasing power, defined as the mean value of the heavy atom structure factor divided by the residual lack-of-closure error. [f] Maximum resolution used for phasing. [g] Figure of merit.

R_{crys} to 29.1%. Further rounds of modeling and refinement, addition of water molecules, and modeling of sulfate and magnesium ions lowered the R_{free} to a final value of 23.4% and R_{crys} to a value of 19.2%. The final model of XPRT was validated using PROCHECK (Laskowski *et al*., 1993) and WHATIF (Vriend & Sander, 1993).

The N-terminal methionine of the XPRT enzyme is enzymatically cleaved, so that there are 302 residues in the dimer which comprises the crystallographic asymmetric unit (subunits A and B are numbered A2−A152 and B2−B152, respectively). The crystal structure of XPRT reported here includes 280 of the possible 302 residues, plus 215 water molecules, 2 sulfate ions, and 2 magnesium ions (each with 5 octahedrally coordinated waters). The first residue (Ser-A2 and Ser-B2) and residues in a mobile loop (Tyr-A63 to Glu-A70, Ser-B61 to Lys-B72) in the XPRT structure are disordered in both subunits. Similarly, the two glycine residues A94-Gly-Gly-A95 and B94-Gly-Gly-B95 have poor density in subunit A and no density in subunit B. In addition, the side chains of residues Lys-A4, Lys-A106, Lys-B4, and Arg-B53 are terminated at the Cδ atom and the side chains of residues Lys-A30 and Lys-A72 are modeled only to the Cγ atom. No side chain densities were observed for residues Glu-A3, Lys-A18, Ser-A62, Glu-A78, Trp-A134, Glu-B3, and Thr-B93 so these are modeled as alanine. The ($2F_o - F_c$) and ($F_o - F_c$) electron density maps indicated dual conformations for Ser-A21, Met-A141, Met-B141, Val-B143, and Val-B144, and two alternate positions have been modeled for these residues. The statistics for the final XPRT model are summarized in Table 5.

Table 5: Refinement Statistics

resolution range (Å)	8.0−1.8 Å
no. of reflections ($I > 1\sigma(I)$)	25 198
R_{crys}[a] ($I > 1\sigma(I)$)	19.2%
R_{free}[b] ($I > 1\sigma(I)$)	23.4%
RMS deviations from ideal	
bond length	0.007 Å
bond angle	1.37°
dihedral angle	25.0°
improper angle	1.19°
average *B*-factor[c]	25.5 Å^2
average *B*-factor[d]	23.4 Å^2
average *B*-factor[e]	19.5 Å^2
mean coordinate error[f]	0.19−0.23 Å
mean coordinate error[g]	0.23−0.30 Å
Ramachandran plot statistics	
residues in most favored regions[h]	94.4%
residues in additionally allowed regions	5.6%
residues in disallowed regions	0.0%

[a] $R_{crys} = \sum|F_o - F_c|/\sum F_o$. [b] R_{free} as defined by Brünger (1992), using 10% of the data. [c] Including waters. [d] Excluding waters. [e] From Wilson Plot (Wilson, 1949). [f] Coordinate error from R_{crys} vs resolution Luzzati plot (Luzzati, 1952). [g] Coordinate error calculated from R_{free} vs resolution Luzzati plot. [h] Excluding glycine residues.

RESULTS AND DISCUSSION

Subunit Structure. The XPRT subunit incorporates a core structure consisting of five parallel β-strands surrounded by three α-helices (Figure 1). The structural core resembles the well-known dinucleotide-binding fold (Branden & Tooze, 1991) but with one less β-strand. This same core structure has been observed in all the PRTase enzymes listed in Table 1. *E. coli* XPRT is the smallest of the proteins known to incorporate this fold (Table 1 and Figure 2), with 122 of its 151 residues contributing to the PRTase core.

All the PRTase crystal structures listed in Table 1 have a mobile loop at the cross-over connection between the two halves of the core β-sheet. In XPRT, as described above, 10−12 residues of this loop are disordered. In addition, the conformation of the residues immediately following the mobile loop varies considerably in these PRTase structures (Figure 2). In XPRT, a short β-strand (β4) formed from residues Val-A73 to Lys-A75 makes antiparallel interactions with strand β2 of the core β-sheet (Figure 1). The short β-strand in XPRT is followed by several residues with random coil conformation. In *T. foetus* HPRT, *T. gondii* HPRT, and *E. coli* OPRT the equivalent stretch of residues also forms an antiparallel β-strand, but in human HPRT and *B. subtilis* GAT, it is helical, and in *S. typhimurium* OPRT, the residues adopt a random coil conformation (Figure 2).

The β-strands in the second or C-terminal portion of the core β-sheet (β5−β7) of XPRT have a distinct bend creating a short secondary β-sheet which is oriented perpendicular to the central β-sheet (Figure 1). A similar structural feature is observed in human HPRT (Eads *et al*., 1994) where the secondary β-sheet is extended to four strands through interaction with another β-strand from the second lobe of that structure. In XPRT, the second β-sheet incorporates only the three β-strands of the core region.

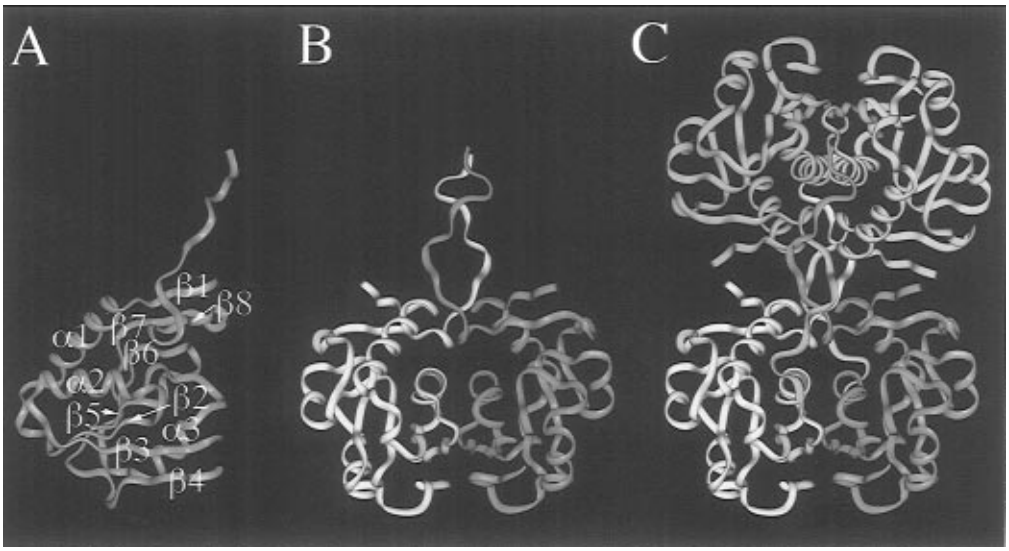

FIGURE 1: Subunit structure and oligomeric arrangement of *E. coli* XPRT. (A) The XPRT subunit structure. Strands β2, β3, β5, β6, and β7 form the core β-sheet which is surrounded by helices α1, α2, and α3. In this orientation, the short antiparallel β-strand (β4) is located below the core β-sheet. Strands β1 and β8 probably form the base specificity determining region of XPRT. The C-terminal 12 residues extend away from the core of the subunit. (B) Subunits A (light green) and B (white) of the XPRT dimer in the crystallographic asymmetric unit. This view is rotated 90° around the vertical axis compared with that in (A). (C) The XPRT tetramer (subunit C is in light blue and subunit D is in yellow) showing the interaction between the C-terminal extended arm of each subunit and a neighboring subunit. This view is in the same orientation as (B). This figure was generated using the program InsightII (Biosym Corp.).

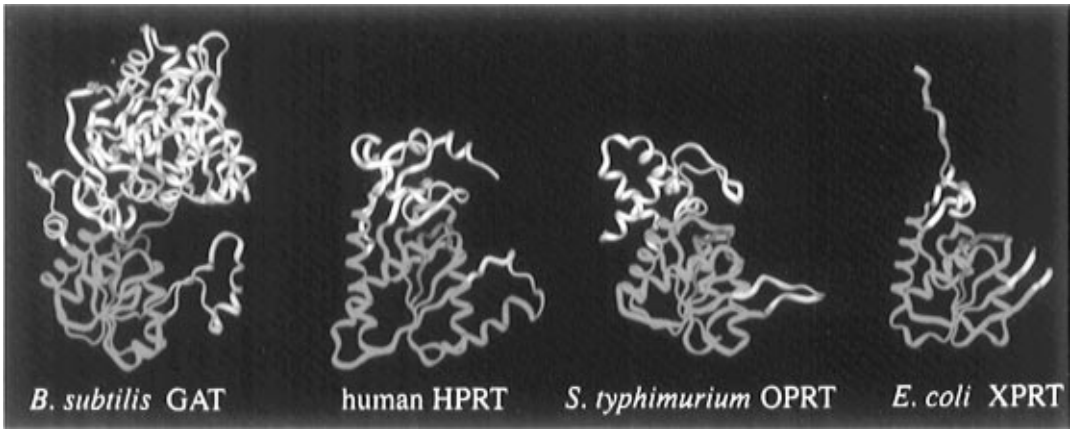

FIGURE 2: Comparison of the subunit structures of several PRTase proteins: *B. subtilis* GAT, *S. typhimurium* OPRT, human HPRT, and *E. coli* XPRT. The core structure is shown in green, the base specificity determining region (and additional domain of GAT) in white, the flexible loop region in yellow, and the *P*Rib-*PP* binding motif in pink. The C-terminal tail of XPRT, not seen in other PRTase structures, is shown in light blue. This figure was generated using the program InsightII (Biosym Corp.).

Most of the XPRT sequence is utilized in forming the core PRTase structure. Those parts that do not contribute to the core region are the N-terminal 7 residues and the C-terminal 22 residues. Together, these 29 residues form two short parallel β-strands (β1 and β8) and an extended arm (Figure 1). The two β-strands form a region that is quite separate from the core and is reminiscent of the second lobes of human, *T. foetus*, and *T. gondii* HPRTs. This lobe is involved in substrate recognition and, as might be expected, is quite variable in structure and sequence among these PRTases (Figures 2 and 3). The extended arm of XPRT, comprising the last 12 residues in the sequence (Met-141 to Arg-152), has no counterpart in the known structures of other PRTases.

XPRT Dimer and Tetramer. The native molecular mass of XPRT, determined by comparison of retention times of marker proteins with that of XPRT on a 2.4 mL Superose-12 gel filtration column, was found to be 65 kDa, suggesting that this enzyme behaves as a tetramer in solution. The orthorhombic crystal form of XPRT has two subunits (subunits A and B, Figure 1) in the asymmetric unit, therefore corresponding to half of the XPRT tetramer observed in solution. The structure of the XPRT tetramer is generated from the dimer by crystallographic symmetry. The two subunits of the asymmetric unit are essentially identical, with minor differences in the lengths of strands and helices and small variations in the position of the N and C termini. However, subunit B has somewhat higher loop mobility than subunit A. For instance, residues flanking the large mobile loop (residues Ser-B61, Ser-B62, and Lys-B72) and the two glycine residues, Gly-B94 and Gly-B95, are not modeled into subunit B but are included (though poorly ordered) in subunit A. The RMS difference for Cα atoms of the 136 residues common to both subunits is 0.88 Å. If a total of 8 residues from the N and C termini are excluded (residues 3−4 and 147−152) the RMS difference for the Cα atoms of the remaining 129 residues drops to 0.38 Å.

```
                                                                              β1      α1                      β2         α2
                                                                             <->  <--------------->          <---->    <---------->
E. coli XPRT               ---------------------------------MSEKYIVTWDMLQIHARKLASRLMPSE--QWKGIIAVSRGGLVPGALLARELG-
human HPRT                 ----ATRSPGVVISDDEPGYDLDLFCIPNHYAEDLERVFIPHGLIMDRTERLARDVMKEMGGHHIVALCVLKGGYKFFADLLDYIKA
T. foetus HPRT             -------------------------MTETPMMDDLERVLYNQDDIQKRIRELAAELTEFYEDKNPVMICVLTGAVFFYTDLLKHLD-
S. typhimurium OPRT        KPYQRQFIEFALSKQVLKFGEFTLKSGRKSPYFFNAGLFNTGRDLALLGRFYAEALVDSGIE-FDLLFGPAYKGIPIATTTAVALAE
E. coli OPRT               KPYQRQFIEFALNKQVLKFGEFTLKSGRKSPYFFNAGLFNTGRDLALLGRFYAEALVDSGIE-FDLLFGPAYKGIPIATTTAVALAE
B. subtilis GAT    (N-terminal domain)-NINRSICSMEYIYFSRPDSNIDGINVHSARKNLGKMLAQESAV----EADVVTGVPDSSISAAIGYAEATGE
                                                                                            *

                                β3               β4                         β5          α3             β6
                              <---->            <->                      <------->    <----->         <---->
E. coli XPRT               ---------IRHVDTVCISSYDHDNQRELKVLKRAEGD------------GEGFIVIDDLVDTGGTAVAIREMY-----PKAHFVTIF
human HPRT                 LNRNSDRSIPMTVDFIRLKSYCNDQSTGDIKVIGGDDLSTL--------TGKNVLIVEDIIDTGKTMQTLLSLVRQYNPKMVKVASLL
T. foetus HPRT             --------FQLEPDYIICSSYSGTKSTGNLTISKDLKT---NI------EGRHVLVVEDIIDTGLTMYQLLNNLQMRKPASLKVCTLC
S. typhimurium OPRT        HHDKD----LPYCFSRKEAKDHGEGGSLVGSA-----------------LQGRVMLVDDVITAGTA-RESMEIIQAHGATLAGVLISL
E. coli OPRT               HHDLD----LPYCFNRKEAKDHGEGGSLVGSA-----------------LQGRVMLVDDVITAGTAIRESMEIIQANGATLAGVLISL
B. subtilis GAT            ---------IPYELGLIKNRYVGRTVIQPSQALREQGVRMKLSAVRGVVEGKRVVMVDDSIVRGTTSRRIVTMLREAGATEVHVKISS
                                                                                    *    *

                                                                  β7       β8
                                                                 <--->     <->
E. coli XPRT               AK-----------------------PAGRPL-----VDDYVVDIPQDTWIEQPWDMGVVFVPPISGR
human HPRT                 VKRTPRSVGYK-------------------------PDFVGFEIPDKFVVGYALDYNEYFRDLNHVCVISETGKAKYKA
T. foetus HPRT             DKDIGKKAYDVP------------------------IDYCGFVVENRYIIGYGFDFHNKYRNLPVIGILKESVYT
S. typhimurium OPRT        DRQERGRGEIS--------------AIQEVERDYGCKVISIITLKDLIAYLEEKPEMAEHLAVKAYREEFGV
E. coli OPRT               DRQERGRGEIS--------------AIQEVERDYNCKVISIITLKDLIAYLEEKPEMAEHLAVKAYREFFGV
B. subtilis GAT            PPIAHPCFYGIDTSTHEELIASSHSVDEIRQEIGADTLSFLSVEGLLKGIGRKYDDSNCGQCLACFTGKYPTEIYQDTVLPHVKEAVLTK
```

FIGURE 3: Primary sequence alignment of the PRTases based on their tertiary structures. The 72 residues of the core PRTase structure common to all the proteins are in bold. The *P*Rib-*PP* sequence motif is underlined, and helices and strands of XPRT are labeled. Asterisks denote the conserved aspartate, glycine, and alanine residues.

The dimer of XPRT buries approximately 1400 Å^2 of the total 8000 Å^2 of surface area per monomer, which is similar to that observed for other dimeric molecules (Janin *et al.*, 1988). The two dimers that form the tetramer are related by a crystallographic 2-fold axis. Subunit D is identical to subunit A, and subunit C is identical to subunit B (Figure 1). Upon tetramer formation, the last 12 residues of each subunit, which form the extended arm, reach across and interact with an adjacent subunit in the tetramer (see below). Tetramer interactions bury an additional 1400 Å^2 of surface area of each monomer. Thus, upon tetramer formation, approximately 35% (i.e., 2800 Å of 8000 Å) of the total surface area of each monomer is buried. This is within the usual range for oligomeric proteins (Janin *et al*., 1988).

Comparison of PRTase Core Structures. A primary sequence alignment of the PRTases based on their tertiary structures (not including *T. gondii* HPRT) identifies 72 residues that are common to all these proteins in the conserved core region (Figure 3). The RMS differences in Cα positions of the 72 residues in human HPRT, *T. foetus* HPRT, *S. typhimurium* OPRT, *E. coli* OPRT, and *B. subtilis* GAT compared to those in XPRT are 1.6, 1.7, 2.7, 2.8, and 2.0 Å, respectively. The high values for the comparison of XPRT and both OPRTases are due to the rotation of a helix (corresponding to α1 in XPRT) in both these structures. A lower RMS difference of 2.4 Å for 56 residues is obtained for both these OPRTases when residues from this helix are excluded from the calculation.

The sequence alignment shown in Figure 3 highlights the low level of sequence identity among these proteins. The only region of high sequence identity is the *P*Rib-*PP* binding site motif previously identified by Hershey and Taylor (1986). In XPRT, this motif consists of the residues 85-IVIDDLVDTGGT-96. As in the previously reported PRTase structures, these residues form the C-terminal end of the third *β*-strand in the core *β*-sheet (*β*5 in XPRT).

The XPRT Active Site. The active site of XPRT was identified by superimposition of the core regions of the structures of *S. typhimurium* OPRT, human HPRT, and *B. subtilis* GAT with that of XPRT and then identification of the positions of substrates and products bound at the active sites of the OPRT, HPRT, and GAT structures. The active site of XPRT was thus shown to be located between the C-terminal part of the core *β*-sheet and the two short *β*-strands above the core region. The conserved *P*Rib-*PP* binding site sequence forms part of the active site, and its relative position is completely conserved among these PRTase structures (Figure 2).

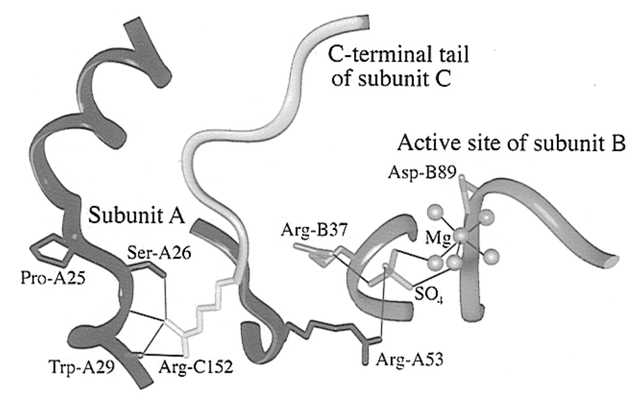

FIGURE 4: Active site of XPRT. The interaction of sulfate and magnesium ions with active site residues in subunit B is shown. Arg-B37 and Asp-B89 are from subunit B, shown in medium gray, and Arg-A53 is from subunit A, shown in black. The C-terminal arginine from subunit C, Arg-C152, shown in white, interacts with Ser-A26 and Trp-A29 of subunit A but is in the vicinity of the active site of subunit B. This figure was generated using InsightII (Biosym Corp.).

Although substrate and product were not added to the crystallization medium, XPRT crystals were grown in the presence of sulfate and magnesium. In both subunits of the resulting XPRT crystal structure, we observed a sulfate ion and a hydrated magnesium ion bound at the active site (Figure 4). Comparison of the active sites of the *S. typhimurium* OPRT/*P*Rib-*PP* complex (Scapin *et al*., 1995) and XPRT reveals that the position occupied by the bound

sulfate ion corresponds closely to that of the β-phosphate group of *P*Rib-*PP*. A sulfate ion has also been observed bound at a similar position in the active site of *E. coli* OPRT (Henriksen *et al*., 1996), while a magnesium ion has been observed bound at a similar position in the active site of *T. gondii* HPRT (Schumacher *et al*., 1996). However, this appears to be the first example of a PRTase with *both* sulfate and magnesium bound at the active site.

The sulfate ion in XPRT is located at the N-terminal end of helix α2 and may be stabilized by the helix dipole. The sulfate also interacts with Arg-37 (3.2−3.3 Å to Nϵ, 3.3−3.4 Å to NH2, and 2.9−3.1 Å to N), Gly-38 (2.8−3.5 Å to N), and several water molecules in the active site. In addition, Arg-A53 (3.2 Å to Nϵ, 3.4 Å to NH2) of subunit A interacts with the sulfate ion bound in the active site of subunit B. However, in the active site of subunit A, the side chain of the equivalent residue (Arg-B53 from subunit B) is disordered.

The backbone amide of Arg-37, one of the positively charged residues which interacts with the sulfate ion, has a *cis*-peptide conformation. Although rarely observed in residues other than proline, this *cis* conformation is also found at the same relative position of the core PRTase structure in both *T. foetus* HPRT (Leu-46 and Thr-47) and *E. coli* OPRT (Ala-71 and Tyr-72). The *cis*-peptide conformation enables the formation of a hairpin turn at this point in the backbone of these proteins. In both *E. coli* OPRT and XPRT, the bound sulfate ion forms a hydrogen bond (2.9−3.1 Å) with the backbone nitrogen of the *cis*-peptide residue and with neighboring basic residues.

The sulfate ion at the XPRT active site also interacts, through two water molecules, with the bound magnesium ion. The magnesium ion and its surrounding ligands adopt a typical octahedral arrangement with bond lengths varying between 2.0−2.35 Å. In XPRT, the ligands are five water molecules and the Oδ1 atom of Asp-89 (Figure 4). Asp-89 is one of the conserved aspartate residues of the *P*Rib-*PP* binding site motif which could have a role in stabilizing the positively charged transition state intermediate (Tao *et al.*, 1996). The equivalent residue in *P*Rib-*PP* synthetase has also been postulated to have a role in metal ion binding (Willemoës *et al*., 1996). In the structure of *T. gondii* HPRT, a magnesium ion is observed in the active site where it interacts with both conserved acidic residues (Schumacher *et al*., 1996), further supporting a metal binding role for these residues. The interaction observed between the magnesium ion and the second conserved acidic residue in the *P*Rib-*PP* binding site motif in the crystal structure of XPRT supports the possibility of a metal ion binding role for this residue.

The crystal structure of *S. typhimurium* OPRT complexed with orotate and *P*Rib-*PP*, however, shows the magnesium ion to be positioned differently, interacting with the first of the two conserved acidic residues via a water molecule as well as with the ribose and pyrophosphate groups of *P*Rib-*PP* (Scapin *et al*., 1995). It has also been proposed that direct binding of the magnesium ion to the conserved acidic residues in the *P*Rib-*PP* motif may explain the observed inhibitory effect of high concentrations of divalent cations on PRTase catalysis (Schumacher *et al*., 1996).

Other Conserved Residues. In the alignment of the PRTase sequences based on the known structures (Figure 3), the glycine residue corresponding to Gly-94 in XPRT is completely conserved. This residue is the first glycine in the 94-Gly-Gly-95 loop of XPRT and is found toward the end of the *P*Rib-*PP* binding site motif. It is curious that although this residue is conserved, the backbone structure is not conserved, at least not in XPRT. The backbone conformation of this part of the PRTase core structure is similar for the other PRTase structures listed in Table 1, but the conformation observed for 94-Gly-Gly-95 in subunit A of XPRT is quite different. Thus in the other PRTase structures the conserved glycine residue marks the first turn of a helix (α3 in XPRT) while in XPRT this helix is unwound by two residues forming instead a type II β-turn, and the helix does not begin until Thr-A96. There is a marked shift in position (4−6 Å) for the Cα atoms of Gly-A94 and Gly-A95 of XPRT compared with equivalent residues in the other PRTase structures. The reason for this structural difference involving the 94-Gly-Gly-95 loop is not yet understood. The flexibility of the loop (these two residues are poorly ordered in subunit A and disordered in subunit B), however, suggests a conformational change could occur in XPRT upon ligand binding at the active site. Indeed, if the subunit A conformation of 94-Gly-Gly-95 in XPRT were fixed, it would preclude the binding of substrate or product since there is a steric clash between Gly-A95 and the 5′-phosphate group of substrate or product (using the bound conformation observed in the other PRTase structures). The 94-Gly-Gly-95 loop in XPRT corresponds to the 5′-phosphate binding region of *S. typhimurium* OPRT (Scapin *et al*., 1994), human HPRT (Eads *et al*., 1994), and *B. subtilis* GAT (Smith *et al*., 1994).

Primary sequence alignments of all the known PRTases in the region of the *P*Rib-*PP* binding motif have indicated that several HPRTs, such as those from *Crithidia fasciculata*, *Leishmania donovani*, *Trypanosoma cruzi*, and *Trypanosoma brucei* (Ullman & Carter, 1995), do not have an equivalent glycine to either Gly-94 or Gly-95 of XPRT. Furthermore, in the structures of both the *T. foetus* and *T. gondii* HPRTs there does not appear to be a high degree of flexibility in the equivalent glycine residues when the substrate or products are not bound. These observations suggest that if the 94-Gly-Gly-95 loop of XPRT is flexible, it may be a unique feature among the PRTase family.

A third residue, Ala-20 in XPRT, is also completely conserved among the six PRTases in the sequence alignment based on their structures (Figure 3). However, a structural analysis of this region gives no clue as to a critical functional role for this residue. It is not located near the active site, and the side chain at this position is not limited to a methyl group by steric constraints. It does not appear to have a structural role since it is located in the middle of an α-helix and its backbone ϕ,ψ angles correspond to the α-helical region of the Ramachandran plot. Furthermore, sequence alignments suggest that *Schistosoma mansoni* HPRT may have a serine residue in this position.

The PRTase Flexible Loop. The flexible loop of XPRT (61-SSYDHDNQRELK-72) is found at the cross-over point between the two halves of the core β-sheet in the same relative position as the loops of previously reported PRTase structures (Figure 2). It has been postulated that the flexible loop of the PRTases could play a role in shielding the active site from solvent molecules (Ozturk *et al*., 1995b), thereby preventing reaction of water with the transition state species. Indeed, the structure of *T. gondii* HPRT in its unliganded

form reveals that the loop is poised directly over the active site (Schumacher *et al*., 1996).

Several residues in the flexible loop have been proposed to be involved in binding and catalysis. In the structure of GAT, the positively charged loop residues Lys-305 and Arg-307 interact with the bound nucleotide AMP (Smith *et al*., 1994), and in one of the *E. coli* OPRT subunits, the positively charged loop residue Lys-103 forms a hydrogen bond to the sulfate ion bound at the active site of an adjacent subunit (Henriksen *et al*., 1996). In addition, several positively charged residues in the flexible loop of *S. typhimurium* OPRT have been implicated in catalysis (Lys-103; Ozturk *et al*., 1995a) and *P*Rib-*PP* binding (Arg-99 and Lys-100; Scapin *et al*., 1995).

In *T. gondii* HPRT, closure of the loop brings two conserved residues, 117-Ser-Tyr-118, into the vicinity of the active site. These residues are conserved in the HPRTases (and XPRT), and their presence in the active site of *T. gondii* HPRT suggests a binding and/or catalytic role for this dipeptide (Schumacher *et al*., 1996).

The loop of XPRT includes two positively charged residues, Arg-69 and Lys-72 (both of which are poorly ordered in this XPRT structure), in addition to the conserved dipeptide 62-Ser-Tyr-63. Arg-69 is disordered in both subunits and Lys-72 is disordered in subunit B and poorly ordered in subunit A (where the side chain in the structural model is truncated to Cγ). If it is involved in binding or catalysis, a conformational change would be necessary to place Lys-72 in the vicinity of the active site. Ser-62 is disordered in subunit B and poorly ordered in subunit A while Tyr-63 is disordered in both subunits.

A binding or catalytic role for the XPRT loop is supported by our observations that removal of loop residues causes almost complete loss of activity. Some purified and concentrated samples of XPRT were found to rapidly lose activity due to degradation by a contaminating protease (Figure 5A). Mass spectrometric analysis of both native and degraded protein indicated that the subunit mass had changed from a single species of 16 805 Da (Figure 5B) to two species of 7143 and 8302 Da (Figure 5C), respectively. Subsequent N-terminal protein sequencing revealed that the protease activity had caused the release of 11 residues corresponding to the sequence 66-DNQRELKVLKR-76. The resulting XPRT protein had no catalytic activity but was able to bind to GMP−Sepharose, indicating a functional GMP binding site. Furthermore, protease degradation was slowed by addition of *P*Rib-*PP*: XPRT stored in the presence of 9 mM *P*Rib-*PP* at 4 °C for 5 days retained 65% of its activity, while enzyme stored without *P*Rib-*PP* retained only 17% of the original activity. These results suggest that some or all of the 11 residues cleaved by the protease form part of the *P*Rib-*PP* binding site.

Inspection of the XPRT structure shows that the protease cleavage points, His-65·Asp-66 and Arg-76·Ala-77, both occur in loops in the protein structure. His-65 is part of the flexible loop described above and is disordered in this XPRT structure. Arg-76 is in a loop that follows strand β4. Thus, cleavage at these positions results in the removal of almost all the cross-over connection residues including most of the mobile loop and all of strand β4. The resultant loss in catalytic activity may therefore be due to an inability to shield the active site from solvent or the removal of any or all of the positively charged residues, Arg-69, Lys-72, Lys-75, and

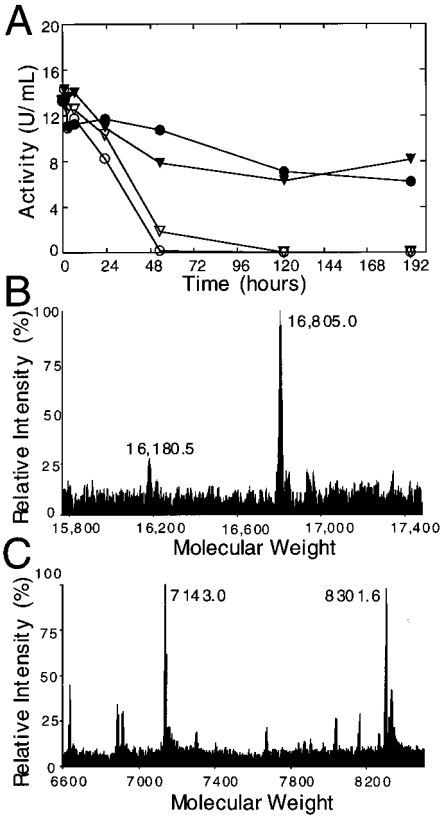

FIGURE 5: Protease degradation of XPRT C59A. (A) Storage of XPRT C59A in the presence and absence of protease inhibitors. Protein samples containing XPRT C59A at 45 μg/ml, 0.05 M Tris-HCl, pH 7.6, 0.01 M $MgCl_2$, and either 1 mM EDTA (▽), 1 mM *p*-CMBS (●), 1 mM PMSF (▼), or no inhibitor (○) were stored at 37 °C, and aliquots were assayed at various times. (B) Mass spectrometric analysis of intact XPRT C59A. The peak at 16 181 may correspond to protein from which the first six amino acid residues (MSEKYI) have been cleaved (predicted mass is 16 186 Da). (C) Mass spectrometric analysis of protease-degraded XPRT C59A. The peak at 7143.0 corresponds to the Ser-2 to His-65 fragment of XPRT (predicted mass is 7139 Da), while the peak at 8301.6 corresponds to the Ala-77 to Arg-152 fragment (predicted mass is 8298 Da).

Arg-76. The residues Ser-62 and Tyr-63 were not removed by protease action and so were still available for catalysis.

Base Specificity. *E. coli* XPRT catalyzes the transfer of a phosphoribosyl group to 6-oxopurine bases, discriminating against the 6-aminopurine base, adenine. The kinetic constants we have obtained for XPRT are summarized in Table 2. Comparison of k_{cat}/K_m values for the 6-oxopurine bases reveals that XPRT is able to utilize bases in the order guanine > xanthine >>> hypoxanthine. Both *T. foetus* and *T. gondii* HPRTs have high specificity for hypoxanthine and guanine but lower specificity for xanthine, while human HPRT has specificity for hypoxanthine and guanine only. The specificity for the 6-oxopurines in the three HPRTs appears to be defined by three different interactions: (1) an interaction between a positively charged protein side chain (Lys-134 in *T. foetus* HPRT, Lys-178 in *T. gondii* HPRT, and Lys-165 in human HPRT) with the 6-oxo group on the purine base (thereby selecting against the 6-aminopurines), (2) a stacking interaction between an aromatic side chain (Phe-162 in *T. foetus* HPRT, Trp-199 in *T. gondii* HPRT, and Phe-186 in human HPRT) and the purine base, and (3) an interaction that defines the enzyme specificity for guanine, hypoxanthine, or xanthine (these bases have respectively an

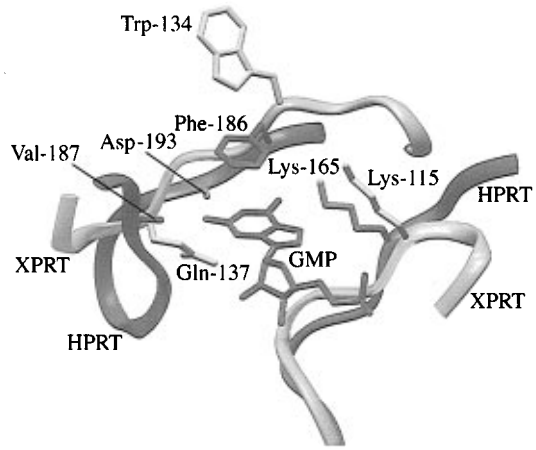

FIGURE 6: Superimposition of the human HPRT/GMP complex (dark shading) with the structure of XPRT (light shading) in the base-specifying region of human HPRT. In human HPRT, specificity for hypoxanthine and guanine is achieved through the backbone carbonyl oxygens of Val-187 and Asp-193. These interact with the exocyclic amino group in position 2 of the guanine ring of GMP. Lys-165 of HPRT interacts with the 6-oxo group of the guanine ring, thus conferring 6-oxopurine base specificity. The aromatic side chain of Phe-186 in human HPRT interacts with the guanine ring of GMP (Eads *et al.*, 1994). This figure was generated using the program InsightII (Biosym Corp.).

amino, a hydrogen, or a carbonyl oxygen atom at the 2-position of the 6-oxopurine base system). In the structure of the human HPRT/GMP complex, this third specificity interaction appears to be defined by the main chain carbonyl groups of Val-187 and Asp-193 which interact favorably with the 2-amino group of the guanine base to promote its interaction with the enzyme (Figure 6). These same main chain oxygens would also allow space for the binding of hypoxanthine but would not favor the binding of the 2-oxo group of xanthine (Eads *et al.*, 1994). In *T. foetus* HPRT, the side chain OH group of Tyr-156 interacts with the 2-amino group of guanine. Since this group can act as both a hydrogen bond donor and acceptor, this residue could hydrogen bond to either the guanine amino group or the xanthine carbonyl oxygen (Somoza *et al*., 1996). In *T. gondii* HPRT, the mechanism for base discrimination is not yet clear, but it appears that the backbone conformation differs sufficiently from that of human HPRT to allow xanthine to bind (Schumacher *et al*., 1996).

Although XPRT was crystallized in the absence of purine base or nucleotide, comparison of the XPRT crystal structure with that of the human HPRT/GMP complex allows the identification of residues that could define the binding specificity of XPRT for 6-oxopurines (Figure 6). Thus, the positively charged residue in XPRT that selects for 6-oxopurines (and against 6-aminopurines) is probably Lys-115, which occupies a similar position to Lys-165 of human HPRT. The residue in XPRT equivalent to the aromatic residue that interacts with the purine base in HPRT (Phe-186) could be Trp-134. The side chain of Trp-A134 in subunit A of XPRT is solvent exposed and disordered, while the conformation of the side chain in subunit B (Trp-B134) is different from that of Phe-186 of HPRT. If this residue is involved in binding the substrate, the correct side chain conformation must be stabilized upon purine binding.

Comparison of the human HPRT and *E. coli* XPRT structures reveals that Gln-137 could be the residue in XPRT which discriminates between the three 6-oxopurine bases (Figure 6). Thus, binding of guanine could be stabilized by XPRT through interaction of the 2-amino group of the substrate with Gln-137 Oϵ2. Similarly, binding of xanthine could be stabilized through interaction of its 2-oxo group with Gln-137 Nϵ2. The lower affinity of XPRT for hypoxanthine (Table 2), which has no exocyclic group at the 2-position, might be explained by the loss of interaction with Gln-137 compared with the other 6-oxopurines. In the conformation observed in the unliganded crystal structure of XPRT, Gln-137 interacts via a hydrogen bond with the Oδ2 group of Asp-89. For optimum interaction with groups at the 2-position of bound 6-oxopurines, it would be necessary for Gln-137 to adopt an alternate conformation. We intend to investigate the role of Gln-137 in substrate binding by site-directed mutagenesis and structural studies with XPRT/ligand complexes.

The C-Terminal Extended Arm. The 12 residues at the C terminus of the XPRT subunit adopt a random coil conformation that resembles an arm extending away from the subunit (Figure 1), a feature that is not observed in other PRTase structures. The "arm" extends across the tetrameric XPRT complex to interact with a crystallographically related subunit. For example, the C-terminal residues of subunit C interact with backbone and side chain atoms of residues in subunit A (the C-terminal residues of each of the other three subunits interact in the same way with symmetry-related subunits, but here we describe the interactions for the C-terminal arm from subunit C only). For the most part, the stabilizing hydrogen bond interactions of the 12 C-terminal residues involve backbone atoms, with the exception of the last residue, Arg-C152 (Figure 4). The side chain of Arg-C152 forms several hydrogen bonds with residues in subunit A—specifically with the backbone O atom of Trp-A29 (2.9−3.0 Å to NH2, 2.9−3.2 Å to NH1) and the backbone O (2.7−2.8 Å to NH1) and side chain Oγ atoms (2.8−3.0 Å to NH1) of Ser-A26. The position of Ser-A26 (and several following residues) and its interaction with Arg-C152 appear to be dependent upon the *cis*-peptide conformation of Pro-A25 which disrupts helix α1.

Since both a *cis*-peptide conformation of Pro-A25 and the consequent disruption of helical hydrogen bond interactions have energy costs, it is possible that the interaction between Ser-A26 and Arg-C152 could be important in the context of substrate binding and/or catalysis. We suggest that this interaction may be necessary to hold Arg-C152 (of subunit C) in close proximity to the active site of subunit B. In the absence of bound *P*Rib-*PP*, the positive charge in the active site from residues Arg-B37 and Arg-A53 would not favor the presence of Arg-C152. Upon substrate binding, the side chain of Arg-C152 could swing into the active site to stabilize the negative charge of *P*Rib-*PP*. This scenario could provide at least a partial explanation for the requirement of XPRT tetramer formation for catalytic activity. Arg-152 is preceded in the XPRT sequence by Gly-151 so a conformational change required to bring Arg-C152 into the active site of subunit B is possible. Our initial modeling suggests that the changes required involve a ϕ,ψ backbone rotation at Gly-C151 and a change in the χ_2 dihedral of the Arg-C152 side chain. In addition, hydrogen bonds between the Arg-C152 side chain and Ser-A26 and Trp-A29 of subunit A would be broken, though these hydrogen-bonding groups would become accessible for interaction with solvent. The role of

Arg-152 in catalysis and substrate binding is currently being investigated by site-directed mutagenesis. In this context, it is also interesting to note that Arg-A53 of subunit A binds to the sulfate ion bound in the active site of subunit B. It will be of considerable interest to see whether substrate/product binding and catalyis at the XPRT active site involve residues from more than one subunit as has been observed in *E. coli* and *S. typhimurium* OPRT and *B. subtilis* GAT.

Summary. XPRT is the smallest of the PRTase structures determined to date, the PRTase core accounting for most of the XPRT sequence. The structure reveals several interesting and unique features, the most intriguing of which is the C-terminal tail that extends away from the PRTase core and is positioned close to the active site of an adjacent subunit. The orientation and placement of the tail suggests a binding and/or catalytic role for the C-terminal residue, Arg-152.

The purine base specificity determining region of XPRT is small compared with those of other PRTases. However, by comparison with the HPRT structures, two of the three important base-determining residues may be conserved in XPRT. Thus, a lysine residue (Lys-115) could interact with the 6-oxo functionality, and an aromatic residue (Trp-134) could stabilize the purine ring system. However, the specificity of XPRT for guanine and xanthine does not appear to be due to the main chain oxygen interactions seen in some of the HPRTs but rather may be uniquely determined through interaction with a glutamine (Gln-137).

Structures of other PRTase enzymes have shown that sulfate or magnesium can bind to the active site. In this XPRT structure, we show that magnesium AND sulfate can bind together in the active site, in the absence of other products or substrates. The magnesium interacts with a highly conserved aspartate residue, and the sulfate binds to a site equivalent to that which interacts with the β-phosphate group of *P*Rib-*PP* (from the complex with *S. typhimurium* OPRT).

These features of the XPRT structure will be investigated further using both site-directed mutagenesis and structural studies of substrate/inhibitor complexes.

NOTE ADDED IN PROOF

A recent paper by Eads *et al.* (1997) reveals the structure of QAPRTase, a PRTase enzyme that lacks the *P*Rib-*PP* binding site motif, to have a completely different structural fold to the previously reported PRTase enzyme structures.

ACKNOWLEDGMENT

We thank Professor Greg Petsko for his kind donation of EMP and also for many helpful discussions. We also thank Mr. Alun Jones for performing the mass spectrometry experiments and Dr. Coral Wynter for performing the protein sequencing. Thanks to Dr. Dianne Keough for her valuable comments and to Dr. Janina Eads and Dr. Janet Smith for providing coordinates for human HPRT and *B. subtilis* GAT prior to release in the Brookhaven protein data bank.

REFERENCES

Ali, L. Z., & Sloan, D. L. (1983) *Biochemistry 22*, 3419−3424.
Bhat, T. N., & Cohen, G. H. (1984) *J. Appl. Crystallogr. 17*, 244−248.
Bhatia, M. B., & Grubmeyer, C. (1993) *Arch. Biochem. Biophys. 303*, 321−325.
Branden, C., & Tooze, J. (1991) in *Introduction to protein structure*, Garland Publishing Inc., New York.
Brünger, A. T. (1992) *X-PLOR Version 3.1 Manual: A system for crystallography and NMR*, Yale University, New Haven, CT.
Brünger, A. T. (1993) *Acta Crystallogr. D49*, 24−36.
Deo, S. S., Tseng, W. C., Saini, R., Coles, R. S., & Athwal, R. S. (1985) *Biochim. Biophys. Acta 839*, 233−239.
Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., & Sacchettini, J. C. (1994) *Cell 78*, 325−334.
Eads, J. C., *et al.* (1997) *Structure 5*, 47−58.
Free, M. L., Gordon, R. B., Keough, D. T., Beacham, I. R., Emmerson, B. T., & de Jersey, J. (1990) *Biochim. Biophys. Acta 1087*, 205−211.
Furey, W., & Swaminathan, S. (1990) *Am. Crystallogr. Assoc. Meet. Abstr., Ser. 2 18*, 73.
Henriksen, A., Aghajari, N., Jensen, K. F., & Gajhede, M. (1996) *Biochemistry 35*, 3803−3809.
Hershey, H. V., & Taylor, M. W. (1986) *Gene 43*, 287−293.
Hove-Jensen, B., Harlow, K. W., King, C. J., & Switzer, R. L. (1986) *J. Biol. Chem. 261*, 6765−6771.
Janin, J., Miller, S., & Chothia, C. (1988) *J. Mol. Biol. 204*, 155−164.
Jochimsen, B., Nygaard, P., & Vestergaard, T. (1975) *Mol. Gen. Genet. 143*, 85−91.
Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Keough, D. T., Emmerson, B. T., & de Jersey, J. (1987) *Clin. Chim. Acta 163*, 301−308.
Kornberg, A., Lieberman, I., & Simms, E. S. (1955) *J. Biol. Chem. 215*, 389−402.
Kosaka, A., Spivey, H. O., & Gholson, R. K. (1977) *Arch. Biochem. Biophys. 179*, 334−341.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Luzzati, V. (1952) *Acta Crystallogr. 5*, 802−810.
Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
Musick, W. D. L. (1981) *CRC Crit. Rev. Biochem. 11*, 1−34.
Ozturk, D. H., Dorfman, R. H., Scapin, G., Sacchettini, J. C., & Grubmeyer, C. (1995a) *Biochemistry 34*, 10755−10763.
Ozturk, D. H., Dorfman, R. H., Scapin, G., Sacchettini, J. C., & Grubmeyer, C. (1995b) *Biochemistry 34*, 10764−10770.
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.
Schumacher, M. A., Carter, D., Roos, D. S., Ullman, B., & Brennan, R. G. (1996) *Nat. Struct. Biol.* 3, 881−887.
Smith, J. L., Zaluzec, E. J., Wery, J.-P., Niu, L., Switzer, R. L., Zalkin, H., & Satow, Y. (1994) *Science 264*, 1427−1433.
Somoza, J. R., Chin, M. S., Focia, P. J., Wang, C. C., & Fletterick, R. J. (1996) *Biochemistry 35*, 7032−7040.
Tabor, S., & Richardon, C. C. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1074−1078.
Tao, W., Grubmeyer, C., & Blanchard, J. S. (1996) *Biochemistry 35*, 14−21.
Ullman, B., & Carter, D. (1995) *Infect. Agents Dis. 4*, 29−40.
Victor, J., Leo-Mensah, A., & Sloan, D. L. (1979) *Biochemistry 18*, 3597−3604.
Vos, S., de Jersey, J., & Martin, J. L. (1996) *J. Struct. Biol. 116*, 330−334.
Vriend, G., & Sander, C. (1993) *J. Appl. Crystallogr. 26*, 47−60.
Wang, B.-C. (1985) *Methods Enzymol. 115*, 90−112.
Willemoës, M., Nilsson, D., & Hove-Jensen, B. (1996) *Biochemistry 35*, 8181−8186.
Wilson, A. J. C. (1949) *Acta Crystallogr. 2*, 318−321.

BI962640D