Crystal Structure of a Purine/Pyrimidine Phosphoribosyltransferase-Related Protein From *Thermus thermophilus* HB8

Peter H. Rehse and Tahir H. Tahirov*
RIKEN Harima Institute, Hyogo, Japan

ABSTRACT **Adenine phosphoribosyltransferase (APRTase) is a widely distributed enzyme involved in the salvage of adenine to form an adenine nucleotide. We crystallized and determined the X-ray crystallographic structure of a purine/pyrimidine phosphoribosyltransferase-related protein from the thermophilic bacterium, *Thermus thermophilus* HB8. The crystal space group was C2 with unit cell dimensions of a = 167.42 Å, b = 61.41 Å, c = 102.39 Å, β = 94.0°. Initial phases were determined to 2.6 Å using the multiple wavelength anomalous dispersion method and selenomethionine substituted protein (Se-MAD), and refined using a 1.9 Å "native" data set. The asymmetric unit contains two pairs of identical dimers, each related by noncrystallographic two-fold symmetry. The fifth monomer forms a similar dimer across a crystallographic two-fold axis. These dimers appear to be the biological unit with both monomers contributing to an unusual highly charged arginine-rich bridge region separating the two active sites. Comparison with distantly related APRTases reveal similarities and differences of the active site. Proteins 2005;61: 658–665.** © **2005 Wiley-Liss, Inc.**

Key words: adenosine phosphoribosyltransferase, X-ray

INTRODUCTION

Adenine phosphoribosyltransferase (APRTase) is responsible for the reversible production of adenosine monophosphate (AMP) and pyrophosphate from adenine and 5-phospho-α-D-ribosyl-1-pyrophosphate (PRPP).[1] The protein is a member of a broad class of proteins referred to as Type I APRTases, which are characterized by a highly conserved PRPP domain, a common α/β fold, and a structurally variable subdomain that is implicated in base recognition. Comparison of the crystallographic structures of the enzymes from a parasite *Leishmania donovani*[2] and *Saccharomyces cerevisiae*[3] indicated that the APRTases can be broadly classed into long and short forms, distinguished from each other by N- and C-terminal extensions and by sequence variation.[3] In plants, animals, yeast, and bacteria, the primary role of APRTase appears to be the salvage and recycling of adenine.[1] In humans, mutation of this protein leads to an accumulation of adenine, which ultimately leads to kidney disease.[4,5] Further, the structure of APRTase from two more parasites, *Giardia intestinalis*[1] and *Leishmania tarentolae*,[6] is also available. The human and yeast enzymes are both short forms and share a 46% identity (BLAST score 149) across the entire sequence. The long-form parasitic structures share an identity of 35–40% with the human enzyme (BLAST score 85–112). The structure of a bacterial form of this enzyme has yet to be reported.

The APRTase from the thermophilic bacteria *Thermus thermophilus* (*Tt*APRTase) does not fit easily into either form. Although its sequence length classes it as a short form, its identity with human and *L. tarentolae* is very much lower (BLAST scores 32 and 33, respectively). The region where some overlap occurs corresponds to the common α/β core and even there the *Tt*APRTase sequence varies in residues previously determined to be highly conserved. The *Tt*APRTase sequence does contain the highly conserved PRPP sequence indicative of purine/pyrimidine phosphoribosyltransferases (PRTases). The structure of a 459-amino acid-long glutamine PRTase from *Bacillus subtilus*[7] shows a similar low homology, but here too the region corresponds to the α/β core. Size considerations alone exclude the functional relationship from the *Tt*APRTase under study.

Here we report the three-dimensional (3D) structure of a putative *Tt*APRTase and discuss its relationship to APRTases from other sources based on sequence and 3D structure comparisons.

MATERIALS AND METHODS

Protein Expression and Purification

With respect to the "native" protein, the polymerase chain reaction (PCR) was used for gene amplification of *T. thermophilus* HB8 genomic DNA. The PCR product was ligated with pT7blue (Novagen) and digested with Nde I and Bgl II. The fragment was inserted into the expression vector pET-11a made linear by digestion with Nde I and Bam HI and transformed into *Escherichia coli* strain BL21 (DE3) cells. The cells were grown for 20 h at 37°C in 2.3 L medium. The cells (9.7 g) were harvested by centrifugation at 6500 rpm for 5 min and suspended in 17.5 mL of 20 m*M*

Grant sponsor: Protein 3000 Japan grant; Grant number: TT0096/HTPF00137.

*Correspondence to: Tahir H. Tahirov, RIKEN Harima Institute, 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan. E-mail: ttahirov@unmc.edu

Received 31 January 2005; Accepted 22 April 2005

Published online 8 September 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20624

Tris (2-amino-2-hydroxymethyl-propane-1,3-diol)-HCl, pH 8.0, 500 m*M* NaCl, and 5 m*M* 2-mercaptoethanol. The cells were disrupted by sonication followed by heat treatment at 70°C (*T. thermophilus* proteins are not expected to denature at this temperature) for 11.5 min. The cell debris and denatured proteins were removed by centrifugation (14,000 rpm, 30 min, 4°C) and the supernatant applied to a HiPrep 26/10 desalting column (53 mL, Amersham Biosciences) using 20 m*M* Tris-HCl, pH 8.0 (Buffer A). The elutant was applied onto a SuperQ Toyopearl 650M (30 mL, Tosoh) equilibrated in Buffer A and eluted with a 0-0.3 *M* NaCl linear gradient. The main protein peak was desalted using a HiPrep 26/10 column with Buffer A and applied to a Resource Q (6 mL, Amersham Biosciences) equilibrated in Buffer A and eluted with a 0–0.3 *M* NaCl gradient. The main protein fraction was desalted with the HiPrep 26/10 equilibrated in 10 m*M* Na phosphate, pH 7.0 (Buffer B), applied to a CHT20-I (20 mL, BIO-RAD) column and eluted with a 0.01–0.15 *M* Na phosphate gradient in Buffer B. The main protein fraction was desalted with the HiPrep 26/10 equilibrated in Buffer A and applied to a Mono Q HR 10/10 (8 mL, Amersham Biosciences) column and eluted with a 0–0.3 *M* NaCl gradient. The main protein peak was then concentrated and applied to a HiLoad 16/60 Superdex 75 (120 mL, Amersham-Biosciences) column equilibrated with 20 m*M* Tris-HCl, 50 m*M* NaCl, pH 8.0. The purified protein was homogenous on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). The protein was concentrated to 20.0 mg/mL using ultrafiltration (VIVASPIN, 5 K cut).

The selenomethionine (SeMet)-substituted protein was produced in a similar way. The gene was transformed into B834 (DE3) plysS cells and grown in 4.5 L over 4 h in medium containing SeMet and 50 mg L^{-1} amphicillin. The expression was induced by 1 m*M* isopropylthio-D-galactoside (IPTG) after which the cultivation was continued for a further 20 h, yielding 10 g of cells. Purification was essentially the same with the CHT20-I column substituted with a CHT5-I column (5 mL, BIO-RAD) and the Mono Q HR 10/10 column substituted with a Mono Q HR 5/5 (1 mL, BIO-RAD) column. The final protein concentration was 29.4 mg mL^{-1}.

Crystallization

Very small, poorly shaped protein crystals were initially obtained by the microbatch method[8] using a crystallization robot "TERA" and screening kit designed for high-throughput protein crystallization.[9] The initial conditions consisted of 100 m*M* 4-morpholineethanesulfonic acid (MES), pH 6.1, and 12.5% w/v polyethylene glycol (PEG) 4000. Diffraction-quality crystals were eventually obtained by altering the PEG polymer from 4000 to 10000 MW, optimizing for pH, and through the addition of 100 m*M* calcium acetate. The buffer remained the same. These crystals grew mainly as very thin plates but on occasion grew into large blocks (0.3 × 0.5 × 0.5 mm). The frequency of this occurrence was improved by the addition of 3% v/v MeOH into the reservoir buffer. Diffraction-quality crystals were grown using the sitting drop method mixing 1 μL of protein solution with 1 μL of a reservoir solution containing 100 mM MES, pH 5.5, 100 m*M* calcium acetate, 3% w/v PEG 10000, and 3% v/v MeOH. Crystals grew within a week. Crystals of the SeMet protein grew best at 40 m*M* calcium acetate, 1.5% w/v PEG 10000, and 3% v/v MeOH.

Data Collection

The solution surrounding the crystals was changed to a cryoprotectant solution over the course of 10 min. The buffer was identical to the reservoir buffer except for the addition of 30% v/v PEG 400. Initially 8 μL of reservoir solution was added to the drop followed by increasing amounts of cryoprotectant solution (2–5 μL) coupled with the removal of excess drop solution. The crystals were mounted onto cryoloops and flash frozen at 100 K in a N_2 stream. Multiwavelength anomalous diffraction (MAD) data were collected at the synchrotron beamline BL-26B1 at Spring-8 (Harima, Japan) using a Jupiter 210 detector. Three data sets were collected from a single crystal using wavelengths determined from a selenium absorption spectrum: the peak (0.97904 Å), the inflection point (0.97930 Å) and the low energy remote (0.98200 Å). The intensity data were indexed, integrated, and scaled with MOSFLM version 6.2.2.[10] The high-resolution native data set was collected on the same beamline using an R-axis V detector. DENZO and SCALEPACK implemented in HKL2000 program package[11] were used to process this data. Data collection statistics are summarized in Table I.

Structure Determination and Model Refinement

The initial phases for solving the X-ray structure of the protein were derived from a Se-MAD experiment to 2.6 Å. The crystal space group was *C*2 with cell dimensions of a = 168.72 Å, b = 61.65 Å, c = 102.41 Å, β = 93.89°, and five molecules in the asymmetric unit. The positions of eight, of a possible 25, seleniums were found using the Patterson function within the program SOLVE.[12] Five of the remaining sites correspond to the N-terminal methionine and were expected to be disordered. The eight sites were directly put into SHARP[13] for phasing, and the resultant phases were put into RESOLVE[14] in order to utilize the automatic chain-building function. With the use of SHARP, the remaining ordered SeMet sites were identified resulting in a much improved figure of merit (see Table I). The figures of merit produced by SOLVE, SHARP, and RESOLVE were 0.33, 0.54, and 0.57, respectively. RESOLVE placed 35% of the expected residues and built 46%, allowing rebuilding using QUANTA (Accelrys) to proceed. The initial rebuilding identified 90 residues of a single molecule that was used as a molecular replacement probe using the Se-MAD peak data set to identify a dimer. The dimer was used with the high-resolution native data set to identify four of the five molecules in the asymmetric unit. Refinement was carried out using CNS[15] with noncrystallographic averaging in the initial stages. The fifth molecule was only identified after first few cycles of refinement and rebuilding. Stereochemical analysis of the structure was performed by PROCHECK.[16] Refinement statistics

TABLE I. X-ray Data Collection, Phasing, and Refinement Statistics

	Peak	Inflection Point	Remote	Native
Resolution (Å)	29.6–2.6	29.6–2.6	29.6–2.6	50.0–1.9
Wavelength (Å)	0.97904	0.97930	0.98200	1.00000
$R(I)$ merge (%)	5.4 (24.2)	5.4 (24.6)	5.2 (27.6)	6.3 (50.1)
$\langle I/\sigma(I)\rangle$	10.2 (4.1)	9.9 (4.0)	10.4 (3.6)	19.5 (2.07)
Completeness (%)	96.9 (81.8)	96.9 (81.5)	95.6 (80.4)	98.7 (98.2)
Multiplicity	4.6 (3.5)	4.6 (3.5)	4.6 (3.5)	3.6 (3.5)
Numbers in parenthesis refer to the highest resolution shell: for Se-MAD, 2.73–2.60 Å; for Native, 1.97–1.90Å				
Phasing to 2.6 Å				
Number of selenium sites expected	25 (sequence position 1, 89, 136, 139, and 142)			
Number of selenium sites found by SOLVE	8			
Overall figure of merit after SOLVE	0.33			
Overall figure of merit after SHARP	0.54			
Overall figure of merit after RESOLVE	0.57			
Refinement using data from Native				
Cell dimensions $a = 167.42$ Å, $b = 61.41$ Å, $c = 102.39$ Å, $\beta = 93.97°$, Space group $C2$				
Resolution range (Å)	45.0–1.94			
Number of reflections	73828			
Number of reflections in Rfree	3728			
Protein atoms	6504			
Solvent atoms	389			
Other atoms	6			
R-factor (%)	23.2 (28.8)			
Rfree (%)	25.9 (33.5)			
Numbers in parentheses refer to the highest resolution shell 2.07–1.94 Å				
Model Quality				
Residues in most favored region 90.6%, additionally allowed 8.5%, generously allowed 0.7%, disallowed 0.1%				
Root-mean-square deviations bonds, 0.006 Å; angles, 1.4°				
Average B factor (Å^2): Main-chain, 43.9; Side-chain, 49.5; solvent, 48.3; others, 54.5; total, 46.7.				
Estimated coordinate error from SigmaA, 0.22 Å				

and the model quality assessment are summarized in Table I. Figures are generated with ESPript,[17] MOLSCRIPT,[18] BOBSCRIPT,[19] RASTER3D,[20] and GRASP.[21] Sequence alignments were performed using CLUSTALW.[22] Homologous proteins were found using a BLAST[23] search of sequence and structure databases.

RESULTS AND DISCUSSION

The structure was solved by the MAD method using a SeMet-substituted recombinant protein. The R-factor for the final model to 1.9 Å is 23.2%. The model consists of five (A–E) 175-residue monomers with A/B and C/D related by noncrystallographic two-fold axis, and E related to its crystallographic mate (E′) by a crystallographic two-fold axis. With respect to the latter molecule, there is some clashing at the two-fold interface, resulting in a lower quality of the model. With respect to the first four protein molecules, the structures are essentially complete, missing only a few residues in a single loop region. Molecule A misses residues 1, 100–104, and the N and Cα of Gly105; molecule B misses residues 100–104; molecule C misses 100–104, the N and Cα of Gly105, and the side-chain of Thr161; and molecule D only misses residues 102–103. The fifth molecule misses residues 1, 32–34, 73–76, 97–104, 119–120, 133–134, 140, 151–152, and the side-chains beyond Cβ for residues Ile6, Leu32, Glu52, and Leu70. In the latter case, the regions 32–34 and 73–76 are at the two-fold interface, with the two molecules very close to each other suggesting packing disorder. The model contains 6504 protein atoms, 389 water molecules, 1 calcium ion, 5 chloride ions and 4 acetate ions.

Overall Fold

The monomer structure is characterized by a complicated arrangement of α-helices and β-sheets, along with some extended loop regions. Figure 1 describes the relationship between sequence and secondary structure elements.[24] The core of the monomer structure consists of a five-stranded parallel β-sheet directly faced by two α-helices on one side (α1 and α2) and a third α-helix (α4) on the other. The core β-sheet arrangement is generated by β5-β4-β8-β9-β10 is colored red in the Figure 2(a). That core region is common among all class I PRTases. There is a small antiparallel β-sheet generated by β6 and β7 off to one side of the α4 helix with tip of the extended loop joining the two strands disordered. The corresponding region in human and yeast enzymes are referred as a "flexible" or "catalytic" loop.[3,5] Another region referred as a "hood" comprises of longer antiparallel β-sheet generated by β1 and β2, and interacting with the tip of β10, and a short β-strand (β3). Both these regions are structurally variable among the APRTases that have been solved. A very short

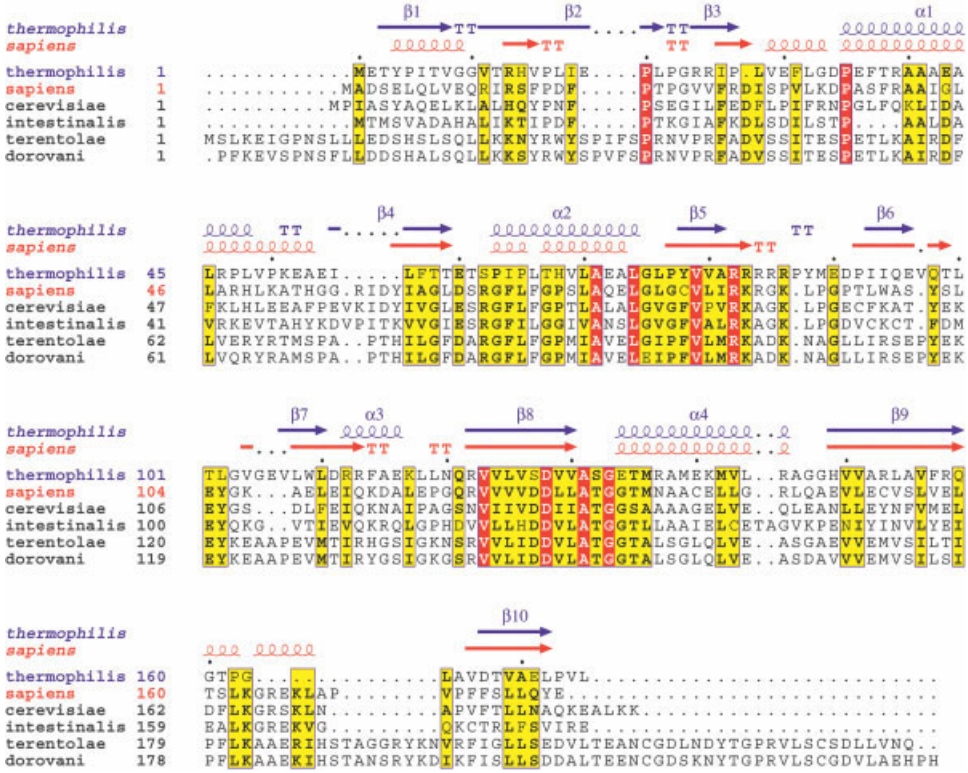

Fig. 1. Sequence alignment of *T. thermophilus* with *Homo sapiens*, *S. cerevisiae*, *G. intestinalis*, *L. tarentolae*, and *L. donovani*.

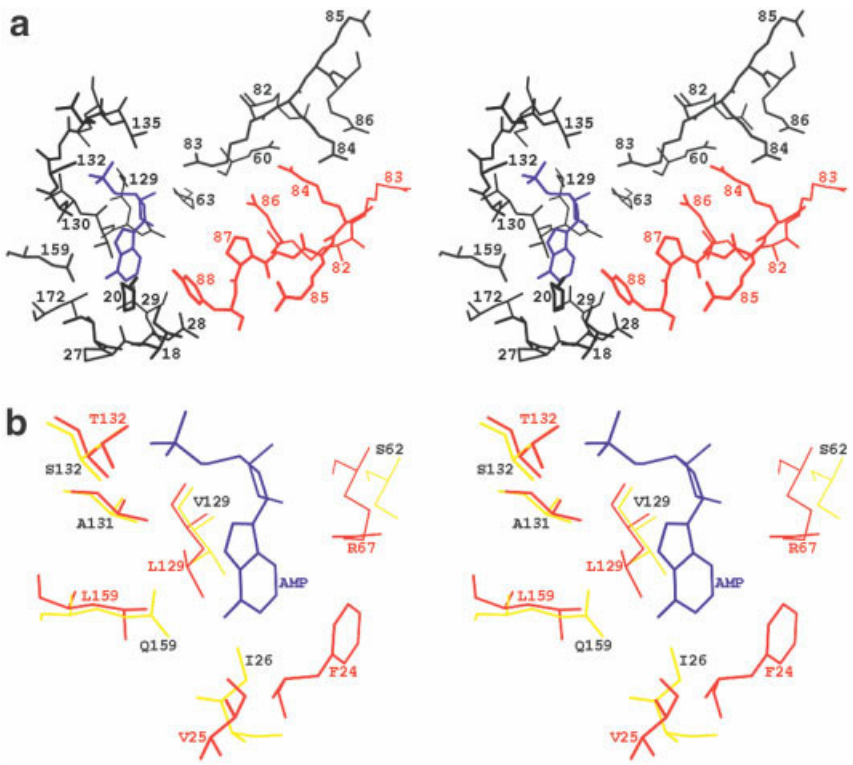

Fig. 4. (**a**) Active site including the arginine-rich region created by the two-fold axis. The modeled AMP is in blue; the A chain is in black, and the B chain is in red. (**b**) Selected active site residues from the *T. thermophilus* (yellow) and human (red) APRTases. AMP is in blue.

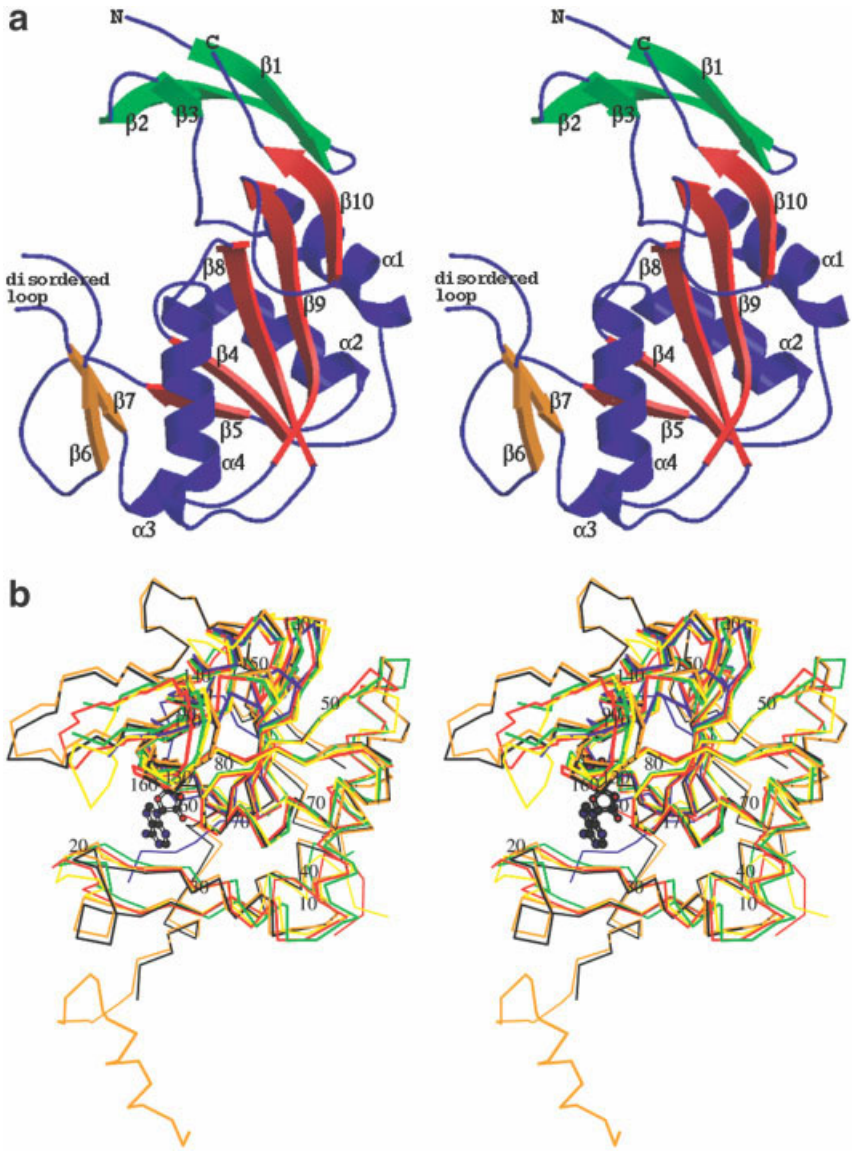

Fig. 2. (**a**) Stereo view of the monomer structure. The core β-sheet structure is in red; the small, outlying β-sheet of the catalytic domain is in orange; and a β-sheet of the hood is in green. (**b**) Overlapped Cα traces various APRTases: *T. thermophilus* (blue), *H. sapiens* (redf), *S. cerevisiae* (green), *G. intestinalis* (yellow), *L. tarentolae* (black), and *L. donovani* (orange). AMP is from the *H. sapiens* structure.

helix (α3) is inserted between the strands β7 and β8 and interacts with the C-terminal end of α4.

The overlapped monomer structures from all six sources [Fig. 2(b)] show that the core β-sheet structure is maintained, while there is significantly more deviation for the rest of the molecule. The sequence alignment shown in Figure 1 highlights the variation in secondary structure, particularly in the hood region (residues 1–34).

Biological Unit

Four of the five molecules in the asymmetric unit are arranged as dimers related by a noncrystallographic two-fold axis. The fifth molecule is related to its symmetry mate across a crystallographic two-fold axis in a near identical arrangement. The excluded surface area calculated using GRASP[21] between individual monomers (A/B) is 2586 Å^2, which is well within that expected for a stable interaction with biological significance.[25] This is further supported when one examines the surface potential model of the *Tt*APRTase [Fig. 3(b)]. Thus, this dimer appears to be the biological unit. At the center of this dimer is a four-helix bundle consisting of the α1 and α2 helices from each monomer [Fig. 3(a)]. The primary degree of interaction is between α2 from each monomer with the two-fold axis very close to the middle of this helix (residues 68–69). Generally speaking, with some adjustment, this is the assumed biological dimer arrangement described for each of the previously solved APRTase structures. In the case of *Tt*APRTase, further stabilization is provided by the N-terminal portion of α1 and part

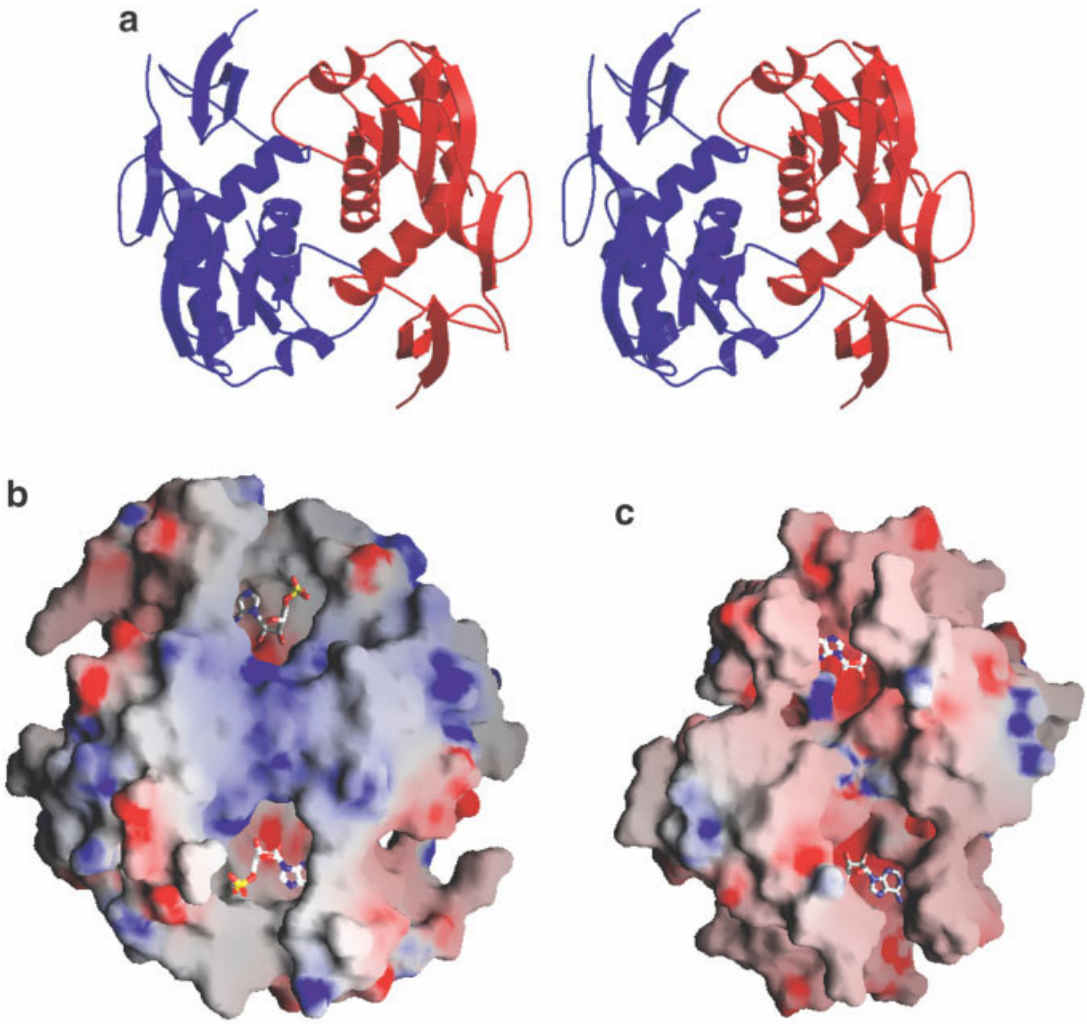

Fig. 3. (**a**) Stereo view of the biological unit. (**b**) Electrostatic potential surface of *Tt*APRTase. (**c**) Electrostatic potential surface of human APRTase.Both (a) and (b) are oriented to display the arginine-rich bridging region of the *Tt*APRTase. The AMP is modeled into the *Tt*APRTase.

of the preceding loop (residues 32–38) and the region immediately following α2 (residues 76–86). The terminal five residues of the latter region are all arginines, with the guanidine of Arg84 being very close to the two-fold axis. The result is that residues 82–86 from both monomers combine to generate a cluster of 10 arginines, with a particularly strong interaction between Arg86 and Arg84 from A subunit, and Arg84 and Arg86 from B subunit. This results in a highly positively charged patch situated across the dimer interface, an effect seen clearly when comparing the electrostatic potential surface maps of the *T. thermophilus* and human APRTases [Fig. 3(b and c, respectively)]. Interestingly the human sequence in this region is only marginally less positively charged, containing two lysines and two arginines rather than five consecutive arginines. The other sequences for which structures are available are even less positively charged. In the human APRTase structure, this region is part of a β-sheet and a turn, whereas in the *Tt*APRTase, it is less structured. The two arginine pentamers related by the two-fold axis can also be seen in the stereo diagram of the active site seen in Figure 4(a). Among APRTases, including bacterial sources, the clustering of positively charged side-chains does not appear to be the rule. Although the functional significance of the patch is unclear, it could conceivable be involved in the attraction of PRPP.

Active Site Cleft

The placement of the AMP was modeled using the human APRTase structure. It was clear that the residues binding the phosphate group were very similar both in the side-chain and the main-chain conformation among the various APRT structures. The structural alignment of this portion was optimized using just residues 52–160 of the *Tt*APRTase structure. In the *Tt*APRTase structure, a single acetate molecule binds in a very similar way to the phosphate moiety. Of minor note is the replacement of a threonine conserved in all previously solved APRTases by Ser132 in *Tt*APRTase. However, the orientation of the

side-chain is such that the hydrogen bond to the phosphate moiety is maintained [Fig. 4(a)].

Several features do present themselves, which may have implications for base recognition. The residue Ala131 is highly conserved among APRTases, whereas this residue is replaced by aspartic acid for those that bind hypoxanthine, xanthine, or guanine bases.[5] In the *Tt*APRTase structure, this residue is also an alanine supporting its role as an APRTase. In the *Tt*APRTase structure, Val129 can form van der Waals contacts to the base. This is consistent with other APRTases where this residue is either a valine, isoleucine, or leucine. This conservation of small hydrophobic residues breaks down somewhat when one examines Gln159 [Fig. 4(b)]. In the previously solved APRtases, this residue is either a leucine or isoleucine and was implicated in base recognition.[5] However, when AMP is modeled in to the *Tt*APRTase, the glutamine side-chain forms a good hydrogen bond to the N6 of adenine and therefore does not necessarily contradict selection for adenine. A BLAST search using the *Tt*APRTase sequence yielded several bacterial APRTases that conserve a FXEG motif centered at a position equivalent to Gln159. The third residue is usually a glutamic acid. The bacterial species where this motif can be found include *Deinococcus radians*, *Leuconostoc mesenteroidis*, and *Lacococcus lactis*. Excluding the phenylalanine, the E/QG motif can also be found in *Desulfotalea psychrophila*, *Fusobacterium nucleatum*, and *Lactobacillus lacterum*. In *E. coli*, which shares a 43% homology with human APRTase, this residue is a leucine.

During the modeling of the active site, the CD1 of Ile26 was close to the base. This was easily relieved by choosing a different rotomer without a clash during the transition. This residue demonstrates the structural variabilty of the "hood" region of the APRTases. In the alignment provided by Phillips et al.,[2] there is a highly conserved phenylalanine at this position. However, in the other solved APRTase structures, the side-chain is oriented directly away from the base, with the carbonyl of the previous residue forming a hydrogen bond to the N6 of adenine. It is possible to speculate that the loss of this hydrogen bond in the *Tt*APRTase structure is compensated by the hydrogen bond formed between the glutamine side-chain and the adenine N6.

In the previously solved APRTase structures, there is a highly conserved arginine that forms a hydrogen bond to the ribose O2′. In the *Tt*APRTase structure, this residue corresponds to Ser62, which has no interaction with the substrate. Once again examining the alignment of Phillips et al.,[2] the arginine is part of a highly conserved segment corresponding to the N-terminal portion of the α2 helix (Fig. 1). The *Tt*APRTase sequence has no homology to this region, although the helix is maintained. A BLAST search using the *Tt*APRTase sequence yielded no significant homology in this region.

CONCLUSION

Although the overall homology is quite low, the APRTase from *T. thermophilus* contains several features consistent with its purported role as an adenine phosphoribosyltransferase. Structural and sequence variations do occur around the active site but often compensate each other. The arginine-rich bridge between the two active sites is an interesting feature that may have implications in the attraction of the negatively charged PRPP. [The atomic coordinates were deposited in the Protein Data Bank under the code 1VCH.]

ACKNOWLEDGMENTS

We would like to thank Seiki Kuramitsu and Shigeyuki Yokoyama for the plasmid vector used in the protein expression, and Yumi Tokunaga for technical help during purification. We would also like to thank Mitsuaki Sugahara for the TERA contribution.

REFERENCES

1. Shi W, Sarver AE, Wang CC, Tanaka KSE, Almo SC, Schramm VL. Closed site complexes of adenine phosphoribosyltransferase from *Giardia lamblia* reveal a mechanism of ribosyl migration. J Biol Chem 2002;277:39981–39988.
2. Phillips CL, Ullman B, Brennan RG, Hill CP. Crystal structures of adenine phosphoribosyltransferase from *Leishmania donovani*. EMBO J 1999;18:3533–3545.
3. Shi W, Tanaka KSE, Crother TR, Taylor MW, Almo SC, Schramm VL. Structural analysis of adenine phosphoribosyltransferase from *Saccharomyces cerevisiae*. Biochemistry 2001;40:10800–10809.
4. Simmonds HA, Sahota AS, Van Acker KJ. Adenine phosphoribosyltransferase deficiency and 2,8-dihydroxyadenine lithiasis. In: Scriver CR, Beaudet WS, Sly WS, Valle D, editors. The metabolic and molecular basis of inherited disease. New York: McGraw-Hill; 1995. p 1707–1724.
5. Silva M, Silva CH, Iulek J, Thiemann OH. Three-dimensional structure of human adenine phosphoribosyltransferase and its relation to DHA-urolithiasis. Biochemistry 2004;43:7663–7671.
6. Silva M, Silva CH, Iulek J, Oliva G, Thiemann OH. Crystal structure of adenine phosphoribosyltransferase from *Leishmania tarentolae*: potential implications for APRT catalytic mechanism. Biochim Biophys Acta 2004;14:31–39.
7. Chen S, Tomchick DR, Wolle D, Hu P, Smith JL, Switzer RL, Zalkin H. Mechanism of the synergistic end-product regulation of *Bacillus subtilis* glutamine phosphoribosylpyrophosphate amidotransferase by nucleotides. Biochemistry 1997;36:10718–10726.
8. Chayen NE, Shaw Stewart PD, Maeder DL, Blow DM. An automated system for micro-batch protein crystallization and screening. J Appl Crystallogr 1990;23:297–302.
9. Sugahara M, Miyano M. Development of high-throughput automatic protein crystallization and observation system. Tanpakushitsu Kakusan Koso 2002;47:1026–1032.
10. Leslie AGW. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26. 1992.
11. Otwinowski Z, Minor W. Processing X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
12. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr 1999;55:849–861.
13. de la Fortelle E, Bricogne G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multi-wavelength anomalous diffraction. Methods Enzymol 1997;276:472–494.
14. Terwilliger TC. Maximum-likelihood density modification with pattern recognition of structural motifs. Acta Crystallogr D Biol Crystallogr 2001;57:1755–1762.
15. Brünger AT, Adams PD, Clore GM, Delano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges N, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
16. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.

17. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999; 15:305–308.
18. Kraulis J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 1991;24: 946–950.
19. Esnouf RM. Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. Acta Crystallogr D Biol Crystallogr 1999;55:938–940.
20. Merritt EA, Bacon DJ. Raster3D: photorealistic molecular graphics methods. Methods Enzymol 1997;277:505–524.
21. Nicholls A, Sharp K, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
22. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
23. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25: 3389–3402.
24. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 1983;22:2577–2637.
25. Janin J. Specific versus non-specific contacts in protein crystals. Nat Struct Biol 1997;4:973–974.