ELSEVIER

Molecular and Biochemical Parasitology 95 (1998) 141–146

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

Short communication

Cloning and characterization of *Leishmania tarentolae* adenine phosphoribosyltransferase

Otavio H. Thiemann [a], Juan D. Alfonzo [b], Larry Simpson [a,b,c,*]

[a] *Department of Molecular, Cell and Developmental Biology, Howard Hughes Medical Institute, UCLA School of Medicine, UCLA, Los Angeles, CA 90095-1662, USA*

[b] *Howard Hughes Medical Institute, UCLA School of Medicine, UCLA, Los Angeles, CA 90095-1662, USA*

[c] *Department of Medical Microbiology, Immunology and Molecular Genetics, Howard Hughes Medical Institute, UCLA School of Medicine, UCLA, Los Angeles, CA 90095-1662, USA*

Received 15 June 1998; accepted 17 June 1998

Keywords: Leishmania tarentolae; APRT; Adenine phosphoribosyltransferase; Kinetoplastid protozoa

The biosynthesis of purine, pyridine and pyrimidine nucleotides is catalyzed by enzymes of the phosphoribosyltransferase (PRTase) family. Most organisms synthesize adenine nucleotides by both the de novo and the salvage pathways. In contrast, all protozoan parasites examined to date are strict purine nucleotide auxotrophs because of the absence of a purine de novo biosynthetic pathway [1]. Kinetoplastid protozoa from the genus *Leishmania* possess three enzymes involved in the recycling of purine nucleotides by the salvage pathway, adenine PRTase (APRT) (EC 2.4.2.7), hypoxanthine-guanine PRTase (HGPRT) (EC 2.4.2.8) and xanthine PRTase (XPRT) (EC 2.4.2.22) [2]. APRT is responsible for catalyzing the conversion of adenine and α-D-5-phosphoribosyl 1-pyrophosphate (PRPP) into adenosine-5-monophosphate (AMP) and pyrophosphate (PPi) by the anomeric inversion of the ribofuranose ring [3]. The only aprt gene that has been cloned and characterized from a kinetoplastid protozoan is that from *Leishmania donovani* [4].

Leishmania tarentolae, a parasite of the gecko, has been exploited in our laboratory as a model trypanosomatid for a variety of molecular, biochemical and evolutionary studies. In this communication we describe the cloning and expression of the *aprt* gene from this organism. It has been recently shown that the lizard *Leishmania* forms a monophyletic clade branching between the *Viannia* and *Leishmania* subgenera [5,6]. These data confirm the inclusion of *L. tarentolae* in the *Leishmania* genus, which was previously

* Corresponding author. Tel.: +1 310 8254215; fax: +1 310 2068967; e-mail: simpson@hhmi.ucla.edu

0166-6851/98/$ - see front matter © 1998 Elsevier Science B.V. All rights reserved.

PII S0166-6851(98)00089-9

proposed by kinetoplast DNA and lipid analyses [7]. As in the case of other trypanosomatids, *L. tarentolae* is a purine nucleotide auxotroph [8]. The ease of cell culture and genetic analysis of *L. tarentolae* should facilitate its use for site-directed mutagenesis of the *aprt* gene as well as for functional complementation and testing of inhibitory substrates. In view of the close phylogenetic relationship, the results should be of general significance for the pathogenic *Leishmania* species.

A 230 bp fragment of *L. tarentolae aprt* gene was PCR-amplified from genomic DNA with degenerate oligodeoxynucleotides for two highly conserved regions (DARGFLFG and VVLIDDVL) of known *aprt* genes, S-2181 (5′-ACCGGAATTCGATGCTCGCGGCTTCCTCTTCGGCC-3′) and S-2182 (5′-TGCTAAGCTTCCCACCCGTGTTGCCAGGACGTCATCGATTAGCACCAC-3′) (the underlines indicate added *Eco*RI and *Hin*dIII sites). The amplified DNA contained a deduced open reading frame of 76 amino acids with sequence identity to other APRT polypeptides. The full-length *aprt* gene was isolated from a *L. tarentolae* UC strain λZAP express *Bam*HI-*Sal*3A I genomic library by hybridization screening with the PCR amplified *L. tarentolae*-specific probe. A 3 kbp genomic fragment was sequenced and found to contain a 708 nucleotide open reading frame that encodes a protein of 236 amino acids with a molecular weight of 25 kDa.

The 5′ RACE of the *aprt* transcript using the 3′ primer S-2488 (5′-CAGTAAAGAGTTGGGTCCGA-3′) which is antisense to the *aprt* mRNA, and the 5′ primer, S-2273 (5′-AACTAACGCTATATAAGTATCAGTTTCTGTACTTTATTG-3′), which is specific to the *L. tarentolae* spliced leader RNA, indicated the presence of one splice acceptor site located at position −155 (SAS in Fig. 1) from the adenylate residue of the predicted methionine initiation codon (underlined in Fig. 1). The splice acceptor site is preceded by a polypyrimidine track of 28 pyrimidines at positions −238– −210. The *aprt* gene is preceded by an in-frame stop codon at position −294.

The mapping of the 3′ UTR of the *aprt* transcript was performed by 3′ RACE. cDNA synthesis from polyA+ RNA was performed with an oligo-dT primer. The cDNA was PCR amplified with the primers, S-2272 (5′-TTGAATTCGCATTGAGCACCTGC-3′) and S-2181. An aliquot of the first PCR reaction was subjected to a second PCR amplification with the primers S-2272 and S-2184 (5′-ACCGGAATTCGTGGTGCTAATCGATGACGTCTCGCAACGGGTGGGAC-3′). The amplified product was cloned into the pCR 2.1-TOPO vector (Invitrogen) and transformed into *Escherichia coli* DH5α. Sequencing identified a single polyadenylation site at position 856, which is 148 nucleotides from the TAG stop codon. These results would indicate a mature mRNA transcript of 1050 nt.

The *L. donovani* and *L. tarentolae* sequences share 87% amino acid identity with sequence divergence being mostly in the N- and C-terminus regions. Inspection of the sequence alignments in Fig. 2 reveals the conserved purine and PRPP binding domains [9] flanked by regions without significant sequence identity. The prediction of protein folding motifs from the individual polypeptides and the aligned sequences indicate a conservation of α-helical (H) and β-sheet (B) structural motifs (Fig. 2). The *Leishmania* sequences are ≈50 amino acids longer than the other APRT sequences, and the *L. tarentolae* sequence is 9 amino acid residues shorter at the C-terminus than the *L. donovani* homologue [4]. The function of these extended C-terminal sequences is not understood. A pairwise alignment of the different sequences revealed a 29–36% sequence identity with the *Leishmania* sequences (data not shown). The measurements of distances between the sequences showed that the APRT genes from *Leishmania* (*L. tarentolae* and *L. donovani*) together with the APRT of the methanogenic bacterium, *Methanococcus jannaschii*, are the most rapidly evolving sequences within the APRT family (data not shown).

Southern analysis of genomic DNA indicated that the *L. tarentolae aprt* gene is single copy (data not shown). Contour-clamped homogeneous field electrophoresis analysis (CHEF), run in two different pulse and field strength conditions, was used to localize *aprt* to a 1.2 Mbp chromosome (data not shown).

```
-390 CTGCAGAGTCGCAGCCTGACCGCATCACACATcAAGCGTTACAGCCTCTCCGTCTTTGCT -331
-330 GCACCCAAGGTATTCGTGTGACCGCGTTGAGGGATCTAATGGATTCGGAACTTTGGTCGT -271
-270 TTGGTCTCTCATGCTGCCCAATCCCATGCGCTTTTCTCCGTGCCTCTCTGTCTCCCCTTC -211

                                                           SAS
-210 CGTTCTATTTGGACAACGATGTGTACTGTAATGGTGCGCCATCTGACAACGAATAGATCA -151
-150 GCACGATTCGCACACTTGCACATACCCAGTGAAGCTTTTGTGTCTGTCGTATTGACAACA -91
 -90 CCGACTGCAACAAGGTGTAGATAGAAGTTGGCCTTCTCGCTCGCTCGCACGCTCTTCACG -31

 -30 CTCCTGCTTTCCTTGCTGTGCCTTGCCACCATGTCCCTCAAGGAAATCGGACCCAACTCT 33
                                   M  S  L  K  E  I  G  P  N  S   11

  34 TTACTGCTCGAGGATTCGCACTCCTTGtCGCAACTGCTAAAGAAGAACTACCGCTGGTAC 93
  12 L  L  L  E  D  S  H  S  L  S  Q  L  L  K  K  N  Y  R  W  Y   31

  94 TCCCCAATCTTCTCCCCACGCAACGTCCCTCGCTTTGCCGATGTCAGCAGCATCACCGAA 153
  32 S  P  I  F  S  P  R  N  V  P  R  F  A  D  V  S  S  I  T  E   51

 154 TCTCCAGAGACGTTGAAAGCTATTCGTGACTTCCTTGTCGAGCGGTACCGCACCATGTCG 213
  52 S  P  E  T  L  K  A  I  R  D  F  L  V  E  R  Y  R  T  M  S   71

 214 CCAGCCCCAACGCATATCCTAGGCTTCGATGCTCGCGGCTTTCTCTTTGGCCCCATGATC 273
  72 P  A  P  T  H  I  L  G  F  D  A  R  G  F  L  F  G  P  M  I   91

 274 GCCGTCGAGCTCGGAATCCCGTTCGTTCTGATGCGCAAGGCGGACAAGAATGCTGGTCTC 333
  92 A  V  E  L  G  I  P  F  V  L  M  R  K  A  D  K  N  A  G  L   111

 334 CTTATCCGAAGTGAACCGTACGAAAAGGAGTACAAGGAGGCCGCGCCGGAGGTCATGACG 393
 112 L  I  R  S  E  P  Y  E  K  E  Y  K  E  A  A  P  E  V  M  T   131

 394 ATCCGCCACGGCAGCATCGGCAAGAATTCGCGTGTGGTGCTAATCGATGACGTTCTGGCA 453
 132 I  R  H  G  S  I  G  K  N  S  R  V  V  L  I  D  D  V  L  A   151

 454 ACGGGTGGAACGGCGCTGTCGGGGTTGCAGCTGGTGGAGGCAAGCGGTGCAGAGGTTGTC 513
 152 T  G  G  T  A  L  S  G  L  Q  L  V  E  A  S  G  A  E  V  V   171

 514 GAGATGGTCTCGATCTTGACCATTCCTTTCCTGAAGGCGGCGGAGAGGATTCACTCTACC 573
 172 E  M  V  S  I  L  T  I  P  F  L  K  A  A  E  R  I  H  S  T   191

 574 GCTGGCGGTCGATACAAGAACGTTAGGTTTATCGGTCTTCTCTCTGAAGACGTGCTGACA 633
 192 A  G  G  R  Y  K  N  V  R  F  I  G  L  L  S  E  D  V  L  T   211

 634 GAGGCGAACTGTGGAGACTTGAACGACTACACTGGTCCCCGTGTGCTGAGCTGCAGCGAT 693
 212 E  A  N  C  G  D  L  N  D  Y  T  G  P  R  V  L  S  C  S  D   231

 694 CTACTGGTGAACCAGTAGCGCTACCTGTCACATTACATGTCTCTCTGAGTTGAATCTTTG 753
 232 L  L  V  N  Q  Stop                                           236

 754 TCTTTGCTGTGCGGCAGTGTGCTTGCTGCATGTCTGTGCCAACTTCCTGGCCGGCGCGTC 813
 814 TATGCGTGTCTCGATGCACAGGAAGGAGGAGCTGTGATGAAAG                  856
```

Fig. 1. Nucleotide and predicted amino acid sequences of the *L. tarentolae* APRT gene. A 1246 bp sequence of the 3 kbp genomic fragment is shown. The nucleotides within the protein coding region and the predicted amino acid residues are enumerated from the methionine start codon (bold face underlined ATG). A termination codon (TAG) is indicated in bold at position +708. Nucleotides within the 5′ untranslated region are numbered negatively starting from the start codon. The G residue corresponding to the slRNA splice-acceptor site is indicated (SAS). The tentative polypyrimidine tracks identified 5′ to each splice-acceptor site are boxed, and the polyadenylation site identified in the transcript is indicated at position +856.

Northern analysis revealed two low abundance *aprt* transcripts of ≈1.1 and 7.5 kb in length (data not shown). The 1.1 kb band is consistent with the mature *aprt* mRNA which was deduced from the sequence and from 5′ and 3′ RACE experiments. We speculate that this band represents an accumulation of pre-mRNA transcript, possibly previous to trans-splicing and/or polyadenylation.

```
                                                                                                               Purine binding domain
          1.......10 ......  20 ........30 ........40 ........50 ........60 ........70 ........80 ........90 .......100 .......110 .......120
Aprt12    .......... ...TATAQQL EYLKNSIKSI .QDY.PKPGI LFRDVTSLLE DPKAYALSID LLVERYKNAG IT....KVVG TEARGFLFGA PVALGLGVGF VPVRKPGKLP RETI.SETYD
Aprt4     .......... .....MTTQL DLIKSSIKSI .PNY.PKEGI IFRDITTLLE VPAAFKATID LIVEQYRDKG IT....KVLG TESRGFIFGA PVALALGLPF ELVRKPKKLP RETI.SQSYQ
Aprt3     .......... ..MTEPTGIT ELLLSRIRDV .ADY.PEPGV VFKDITPLLA DPGAFAALTD ALAEAAGRTG AT....KVVG LEARGFILGA PVALRAGLGF IPVRKAGKLP GATL.SQAYD
Aprt15    .......MAT ED..VQDPRI AKIASSIRVI .PDF.PKPGI MFQDITTLLL DTEAFKDTIA LFVDRYKDKG IS....VVAG VEARGFIFGP PIALAIGAKF VPMRKPKKLP GKVI.SEEYS
Aprt2     .......... ....ASDGRV ERIASSIRAI .PNF.PKPGI LFQDITTLLL DPQAFRDTTD LFVERYKDKD IT....VVAG VEARGFIFGP PIALAIGAKF VPIRKPKKLP GEVI.SEEYS
Aprt11    ......MFAV ENGLQGDPRL KAISDAIRVI .PHF.PKTGI MFQDITTLLL DPVAFKHVVD IFVDRYKHMN IS....LVAG VEARGFIFGP PIALAIGAKF VPLRKPGKLP GRVI.SEEYE
Aprt13    .......... ....MSESEL QLVARRIRSF .PDF.PIPGV LFRDISPLLK DPDSFRASIR LLAGHLKSTH GG.KIDYIAG LDSRGFLFGP SLAQELGVGC VLIRKRGKLP GPTV.SASYS
Aprt8     .......... ....MAESEL QLVAQRIRSF .PDF.PIPGV LFRDISPLLK DPASFRASIR LLASHLKSTH GG.KIDYIAG LDSRGFLFGP SLAQELGLGC VLIRKRGKLP GPTV.SASYA
Aprt1     .......... .....SEPEL QLVARRIRSF .PDF.PIPGV LFRDISPLLK DPDSFRASIR LLASHLKSLH GG.KIDYIAG LDSRGFLFGP SLAQELGVGC VLIRKRGKLP GPTL.SASYA
Aprt14    .......... .....ADSEL QLVEQRIRSF .PDF.PTPGV VFRDISPVLK DPASFRAAIG LLARHLKATH GG.RIDYIAG LDSRGFLFGP SLAQELGLGC VLIRKRGKLP GPTL.WASYS
Aprt10    .........M SPSISAEDKL DYVKSKIGEY .PNF.PKEGI LFRDIFGALT DPKACVYLRD LLVDHIRESA P..EAEIIVG LDSRGFLFNL LIATELGLGC APIRKKKKLA GEVV.SVEYK
Aprt16    .......... ..MSISESYA KEIKTAFRQF .TDF.PIEGE QFEDFLPIIG NPTLFQKLVH TFKTHLEEKF AKEKIDFIAG IEARGLLFGP SLALALGVGF VPIRRVGKLP GECA.SITFT
Aprt9     .......... ..MSIAS.YA QELKLALHQY .PNF.PSEGI LFEDFLPIFR NPGLFQKLID AFKLHLEEAF PEVKIDYIVG LESRGFLFGP TLALALGVGF VPVRKAGKLP GECF.KATYE
Aprt6     MSLKEIGPNS LLLEDSHSLS QLLKKNYRWY SPIFSPRNVP RFADVSSITE SPETLKAIRD FLVERYRTMS PAP..THILG FDARGFLFGP MIAVELGIPF VLMRKADKNA GLLIRSEPYE
Aprt7     MPFKEVSPNS FLLDDSHALS QLLKKSYRWY SPVFSPRNVP RFADVSSITE SPETLKAIRD FLVQRYRAMS PAP..THILG FDARGFLFGP MIAVELEIPF VLMRKADKNA GLLIRSEPYE
Consensus                                : : *        * *.          .            :          : *   ::**:::.    :*           :*:  * .          :
Fold                          HHHH HHHHHHHHH    LLLLLLL          LL L HHHHHHHH HHHHHHHH L LLL  BBBBB B                    LLL BBBB LLLLL LL

                     PRPP binding domain
          .......130 .......140 .......150 .......160 .......170 .......180 .......190 .......200 .......210 .......220 .......230 .......240
Aprt12    LEY...GTDQ LEIHVDAIK. PGDKVLVVDD LLATGGTIEA TVKLIRRLGG EVADAAFIIN LFDLGGEQRL E..K....QG ITSYSLVPFP GH........ .......... ..........
Aprt4     LEY...GQDT LEMHVDAIS. EGDNVLIIDD LLATGGTVEA TVKLVQRLGG AVKHAAFVIN LPELGGEKRL N..N....LG VDCYTLVNFE GH........ .......... ..........
Aprt3     LEY...GSAE IEVHAEDLT. AGDRVLVVDD VLATGGTAEA SLELIRRAGA EVAGLAVLME LGFLGGRARL EPAL....AG APLEALLTVZ .......... .......... ..........
Aprt15    LEY...GTDT IEMHVGAVE. PGERAIIIDD LIATGGTLAA AIRLLERVGV KIVECACVIE LPELKGKEKL G........E TSLFVLVKSA A......... .......... ..........
Aprt2     LEY...GTDK IEMHVGAVQ. PNDRVLIVDD LIATGGTLCA AAKLIERVGA KVVECACVIE LPELKGRDKL G........D MPVFVLVQAD ESV....... .......... ..........
Aprt11    LEY...GRDC LEMSVEAVK. SEERALIIDD LVATGGTLSA SINLLERAGA EVVECACVVG LPKFKGQCKL K........G KPLYVLVEPN QFDELTL... .......... ..........
Aprt13    LEY...GKAE LEIQKDALE. PGQKVVIVDD LLATGGTMCA ACELLSQLRA EVVECVSLVE LTSLKGREKL GP........ VPFFSLLQYE .......... .......... ..........
Aprt8     LEY...GKAE LEIQKDALE. PGQKVVVVDD LLATGGTMCA ACELLGQLQA EVVECVSLVE LTSLKGREKL GS........ VPFFSLLQYE .......... .......... ..........
Aprt1     LEY...GKAE LEIQKDALE. PGQRVVIVDD LLATGGTMCA ACELLNQLRA EVVECVSLVE LTSLKGRERL GP........ IPFFSLLQYD .......... .......... ..........
Aprt14    LEY...GKAE LEIQKDALE. PGQRVVVVDD LLATGGTMNA ACELLGRLQA EVLECVSLVE LTSLKGREKL AP........ VPFFSLLQYE .......... .......... ..........
Aprt10    LEY...GSDT FELQKSAIKQ PGQKVVVVDD LLATGGSLVA ATELIRKVGG VVVESLVVME LVGLEGRKRL DG........ CKVHSLIKY. .......... .......... ..........
Aprt16    KLD...HEEI FEMQVEAIP. FDSNVVVVDD VLATGGTAYA AGDLIRQVGA HILEYDFVLV LDSLHGEEKL S......... APIFSILHS. .......... .......... ..........
Aprt9     KEY...GSDL FEIQKNAIP. AGSNVIIVDD IIATGGSAAA AGELVEQLEA NLLEYNFVME LDFLKGRSKL N......... APVFTLLNAQ KEALKK.... .......... ..........
Aprt6     KEYKEAAPEV MTIRHGSIG. KNSRVVLIDD VLATGGTALS GLQLVEASGA EVVEMVSILT IPFLKAAERI HSTAGGRYKN VRFIGLLSED VLTEANCGDL NDYTGPRVLS C.........
Aprt7     KEYKEAAPEV MTIRYGSIG. KGSRVVLIDD VLATGGTALS GLQLVEASDA VVVEMVSILS IPFLKAAEKI HSTANSRYKD IKFISLLSDD ALTEENCGDS KNYTGPRVLS CGDVLAEHPH
Consensus            ; :    :     ...:::** ::****:  :     *:         :     ::  :  :  .  ::                ::
Fold      HHHH                 LLLL BBBBB         HHH HHHHHHH      BBBBBBBBB B LLLL                BBBBBB
```

Fig. 2. Multiple alignment of representative APRT sequences. The amino acid sequences of several APRT 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 predicted secondary-structure fold, which was utilized as a guide to refine the sequence alignment, is shown: H, α-helices; L, loops; and B, β-sheets. The shaded boxes indicate the predicted purine binding domain and the boxed sequence indicates the predicted PRPP binding domain. The sequences are as follows: *Drosophila melanogaster* (APRT10, S34831), *Homo sapiens* (APRT14, P07741), *Saccharomyces cerevisiae* APT1 (APRT9, S49755) and APT2 (APRT16, L14434), *Rattus norvegicus* (APRT13, P36972), *Arabidopsis thaliana* APT1 (APRT15, P31166) and APT2 (APRT11, Q42563), *Triticum aestivus* (APRT2, U22442), *Leishmania donovani* (APRT7, L25411), *Escherichia coli* (APRT12, M14040, M25902), *Haemophilus influenzae* (APRT4, U32748, L42023), *Streptomyces coelicolor* (APRT3, X87267), *Cricetulus griseus* (APRT8, S36334), *Mastomys hildibrantii* (APRT1, U28722) and *L. tarentolae* (APRT6, AF060886). The sequences were aligned by the CLUSTAL_X program [10] with further manual refinements, taking into consideration the predicted secondary structures. The predict protein program (EBI at http://www.ebi.ac.uk/searches/searches.html) and the PHDsec program [11–13] were used to deduce the secondary structure predictions.

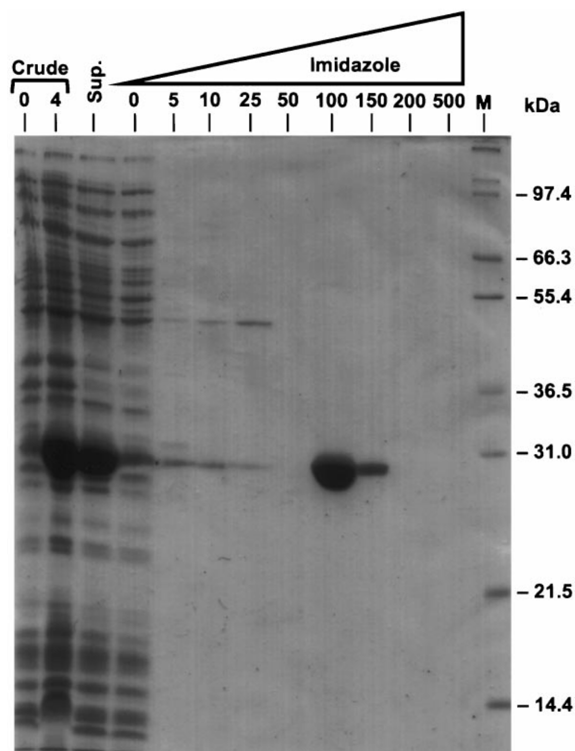

Fig. 3. Expression of *L. tarentolae aprt* in *E. coli* and purification of the recombinant protein. The *L. tarentolae aprt* gene was subcloned into the pQE-30 expression vector and overexpressed in *E. coli*. The proteins were separated in a SDS-PAGE gel. The 0 and 4 lanes represent lysates of non-induced and 4 h IPTG-induced *E. coli*, respectively. Sup., the supernatant of the 4 h-induced crude lysate. The 0–500 lanes represent fractions eluted from the Ni^{2+}-NTA column with increasing concentrations of imidazole (0–500 mM). M, molecular size markers in kDa.

The *aprt* open reading frame was amplified by a PCR reaction with oligodeoxynucleotides S-2284 (5′-CGTGGGATCCATGTCCCTCAAGGAAA-TCGGACCCAC-3′) and S-2286 (5′-GCTAAT-TAAGCTTCTGGTTCACCAGTAGATCGCCG-3′) (added *Bam*HI and *Hin*dIII restriction sites are underlined). The 728 bp PCR fragment was then cloned into the pQE-30 overexpression vector (Qiagen). The IPTG-induced transformed *E. coli* cells produced high levels of soluble, enzymatically active *L. tarentolae* APRT, which was purified to homogeneity on a Ni^{2+}-NTA column by elution with a step gradient of imidazole (Fig. 3). The expressed APRT eluted in the 100–150 mM imidazole fractions, together with enzymatic activity. The recombinant APRT protein with an N-terminal hexahistidine tag migrates as a 26 kDa protein in 10% SDS-PAGE. The recombinant APRT retained enzymatic activity and substrate specificity (data not shown). The purification of recombinant APRT protein from *L. donovani* has been previously reported [4] by a method which utilized the affinity of APRT for the AMP product and which required an enzymatically active protein. The ability to overexpress and purify the recombinant *L. tarentolae* protein in a single chromatographic step, independent of the active binding of the APRT to the substrate, should allow the purification of APRT mutants with different affinities for the substrate. This is a prerequisite for a structural-functional study of the enzyme.

Size exclusion chromatography of the purified *L. tarentolae* APRT indicates that the protein migrates as a dimer of 52 kDa in the absence of PRPP (data not shown). This is an interesting difference from the *L. donovani* homologue, which has been shown to elute as a monomer in the presence or absence of PRPP [4]. It is possible that the few amino acid differences seen in these two APRT proteins, principally at the N- and C-terminus sequences, may play a role in the formation of higher order structure.

Acknowledgements

The work was supported in part by a research grant AI-09102 to L.S. from the National Institutes of Health. We would like to thank the members of the Simpson laboratory for helpful discussions in the course of this work. We would like to express our gratitude to Dr James A. Lake for helpful discussion on the phylogeny of the APRT genes and to Dr Beatriz D. Lima for the *L. tarentolae* λZAP express genomic library.

References

[1] Berens R, Krug R, Marr JJ. Purine and pyrimidine metabolism. In: Marr JJ, Muller M, editors. Biochemistry and Molecular Biology of Parasites. London: Academic Press, 1995:89–118.

[2] Ullman B, Carter D. Molecular and biochemical studies on the hypoxanthine-guanine phosphoribosyltransferases of the pathogenic haemoflagellates. Int J Parasitol 1997;27:203–13.

[3] Musick WD. Structural features of the phosphoribosyltransferases and their relationship to the human deficiency disorders of purine and pyrimidine metabolism. CRC Crit Rev Biochem 1981;11:1–34.

[4] Allen T, Hwang HY, Wilson K, Hanson S, Jardim A, Ullman B. Cloning and expression of the adenine phosphoribosyltransferase gene from *Leishmania donovani*. Mol Biochem Parasitol 1995;74:99–103.

[5] Croan D, Ellis J. Phylogenetic relationships between *Leishmania*, *Viannia* and *Sauroleishmania* inferred from comparison of a variable domain within the RNA polymerase II largest subunit gene. Mol Biochem Parasitol 1996;79:97–102.

[6] Croan DG, Morrison DA, Ellis JT. Evolution of the genus *Leishmania* revealed by comparison of DNA and RNA polymerase gene sequences. Mol Biochem Parasitol 1997;89:149–59.

[7] Gomez-Eichelmann MC, Holz G Jr, Beach D, Simpson AM, Simpson L. Comparison of several lizard *Leishmania* species and strains in terms of kinetoplast minicircle and maxicircle DNA sequences, nuclear chromosomes, and membrane lipids. Mol Biochem Parasitol 1988;27:143–58.

[8] Trager W. Nutrition of a hemoflagellate (*Leishmania tarentolae*) having an interchangeable requirement for choline or pyridoxal. J Protozool 1957;4:269–76.

[9] Hershey HV, Taylor MW. Nucleotide sequence and deduced amino acid sequence of *Escherichia coli* adenine phosphoribosyltransferase and comparison with other analogous enzymes. Gene 1986;43:287–93.

[10] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876–82.

[11] Rost B, Sander C. Combining evolutionary information and neural networks to predict protein secondary structure. Proteins 1994;19:55–72.

[12] Rost B, Sander C. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Proc Natl Acad Sci USA 1993;90:7558–62.

[13] Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 1993;232:584–99.