ELSEVIER

International Journal for Parasitology 29 (1999) 601–605

INTERNATIONAL Journal for PARASITOLOGY

Research note

Cloning and characterisation of a cDNA encoding the *Trypanosoma brucei* ribosomal protein L24☆

Laurie K. Read *, Kevin T. Militello, George E. Nerantzakis

Department of Microbiology and Center for Microbial Pathogenesis, SUNY Buffalo School of Medicine, Buffalo, NY 14214, USA

Received 6 October 1998; received in revised form 4 December 1998; accepted 7 December 1998

Abstract

A cDNA encoding ribosomal protein L24 was amplified by PCR from the protozoan parasite, *Trypanosoma brucei*. The 621 nucleotide cDNA had an open reading frame of 375 nucleotides, predicting a highly basic protein of 125 aa. Database searches revealed 33–40% identity between the *T. brucei* RPL24 protein and several eukaryotic RPL24 homologues. Southern blot analysis indicated that the gene was present as a single copy, and a transcript of approximately 620 nucleotides was detected in procyclic forms of the parasite. Interestingly, *T. brucei* PRL24 is the smallest eukaryotic RPL24 protein described to date. It is also the most divergent of the known kinetoplastid ribosomal proteins.

Keywords: Kinetoplastid; Ribosome; Trypanosome

Ribosomes are composed of both proteins and RNAs. Peptidyl transferase activity apparently resides in an RNA component of the ribosome, since in vitro synthesised *Escherichia coli* 23S rRNA has been shown to catalyse peptide bond formation in the absence of any proteins [1]. Indeed, this enzymatic activity can be reproduced using only domain V of the 23S rRNA, and it is stimulated markedly by addition of domain VI [2]. It has been suggested that some protein components of ribosomes may function to stabilise interactions between rRNA domains [3]. One protein that is localised in the peptidyl transferase centre of eukaryotic cytoplasmic ribosomes, and thus could serve such a stabilisation function, is ribosomal protein L24 (RPL24) [4]. Genes encoding identical RPL24 proteins have been cloned from human and rat ([5, 6]; P38663 and X78443, respectively). Homologues of this protein have also been identified in three plant species (*Hordeum vulgare*, P50888; *Arabidopsis thaliana*, P38665; and *Cicer arietinum*, AJ225027 [7]), three yeast species (*Saccharomyces cerevisiae*, P04449 and P24000 [8], *Schizosaccharomyces pombe*, Z81317 and Q10353; and *Kluveromyces lactis*, P38665 [9]), and three

☆ Note: Nucleotide sequence data reported in this paper are available in the GenBank, EMBL, and DDBJ databases under the accession number AF011787.

* Corresponding author. Present address: Department of Microbiology, SUNY Buffalo School of Medicine, 138 Farber Hall, Buffalo, NY 14214, USA. Tel: 716-829-3307; Fax: 716-829-2158; e-mail: lread@acsu.buffalo.edu

0020-7519/99/$20.00
PII: S0020-7519(99)00011-9

archaebacteria (*Methanococcus jannaschii*, P54064 [10]; *Pyrococcus horkoshii*, AP000006 [11], and *Archaeoglobus fulgidus*, AE001051 [12]). We report here the sequence of the RPL24 homologue from the protozoan parasite, *Trypanosoma brucei*. The predicted protein has significant homology to all reported RPL24 proteins. We have also determined the RPL24 genomic organisation and transcript size to gain insight into expression of the RPL24 gene in this parasite.

A cDNA with homology to RPL24 was identified during a PCR amplification. Briefly, oligo-d(T)-primed cDNA was generated from *T. brucei* procyclic form RNA and amplified using a primer corresponding to the *T. brucei* spliced leader sequence [13], X-SL24 (TATCTAGAACAGTTTCTGTACTATATTG; *Xba*I restriction site is underlined), and a second, degenerate oligonucleotide corresponding to a region of homology between eukaryotic poly(A) polymerases, (PAP-6; GCGGATCC(G/A)TG(G/T/A/C)AC(G/T/A/C)CC(G/T/A/C)A(G/A)(G/T/A/C)C(G/T)(G/A)TA; *Bam*HI restriction site is underlined). Polymerase chain reaction was performed under the following conditions: 5 cycles of 94°C 1 min, 35°C 30 s, 60°C 1 min, and 35 cycles of 94°C 1 min, 50°C 30 s, 72°C 1 min. A 389 nucleotide product was isolated, ligated into the *Xba*I/*Bam*HI site of pBluescript II SK−, and two clones sequenced in both directions by automated *Taq* cycle sequencing. The sequences of these cDNAs were identical, and the proteins predicted from their sequences revealed significant homology to a portion of several RPL24 genes when searched against protein databases. In order to amplify the remainder of the RPL24 cDNA, a primer based on the cDNA sequence (RPL24-1; GCGAATTCGCTCTCCTACATTCAAGAAGTGCG; *Eco*RI restriction site is underlined) was used in conjunction with oligonucleotide XSC-d$(T)_{17}$ (GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT; *Xho*I, *Sal*I, and *Cla*I restriction sites are underlined) to amplify *T. brucei* procyclic form cDNA under the following conditions: 4 cycles of 94°C 1 min, 45°C 30 s, 72°C 1 min, and 26 cycles of 94°C 1 min, 55°C 30 s, 72°C 1 min. Polymerase chain reaction products were digested and ligated into the *Eco*RI/*Xho*I site of pBluescript II SK−, and two clones were isolated and sequenced in both directions. The sequences were 339 and 343 nucleotides in length, including 3′ poly(A) tracts of 38 and 42 nucleotides, respectively. The complete 617–621 nucleotide RPL24 cDNA predicts a 375 nucleotide open reading frame (ORF) encoding a 125 aa protein with a M_r of 14 595 and a P.I. (isoelectric point) of 12.09 (Fig. 1). There are 10 nucleotides between the spliced leader and the AUG start codon and a 177 nucleotide 3′ untranslated region. Comparison of the cDNA sequence with the PAP-6 primer revealed complementarity between 11 out of 12 of the 3′-most bases of the primer and the cDNA sequence, accounting for the initial amplification of this cDNA.

A BLAST search of the protein sequence predicted from the ORF indicated that the protein was highly homologous to RPL24 genes from a wide range of organisms. Multiple sequence alignment of *T. brucei* RPL24 with other RPL24 sequences was performed with the PILEUP and PRETTY programs from the Genetics Computer Group (GCG) package [14] (Fig. 2). Pairwise sequence alignments between RPL24 protein sequences were performed with the GAP program from the GCG package. *Trypanosoma brucei* RPL24 showed the highest homology with the RPL24 homologue (called RPL30 in yeast [8, 9]) from *K. lactis* (39.7% identity; 54.6% similarity). It also showed similar levels of homology with RPL24 homologues from *S. cerevisiae* (37.2%; 53.7%), *H. sapiens* (36.3%; 51.24%), *A. thaliana* (35.5%; 45.5%), and *H. vulgare* (33.1%; 45.5%). The *T. brucei* RPL24 homologue is the smallest of the eukaryotic RPL24 proteins described to date, being about 80% the size of the other predicted proteins. Interestingly, the *T. brucei* protein lacks sequence at its C terminus, which is the region least conserved among all RPL24 proteins [6]. The archaebacterial *M. jannaschii* RPL24 homologue, which is 70 aa in size, is significantly smaller than all of the eukaryotic proteins, and is 26% identical and 35% similar to the *T. brucei* protein. Surprisingly, given that *T. brucei* is an evolutionarily ancient eukaryote [15], the *M. jannaschii* RPL24 homologue showed the

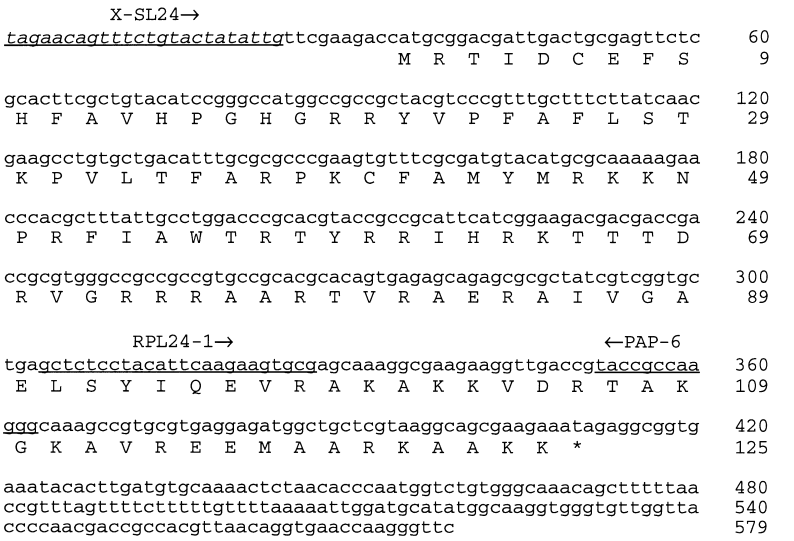

Fig. 1. RPL24 cDNA and predicted aa sequences. Nucleotide and aa residue numbers are indicated at the right. Positions of oligonucleotides used in PCR amplification are underlined, and their names and orientations shown above the sequence. Italicised nucleotides represent spliced leader sequence.

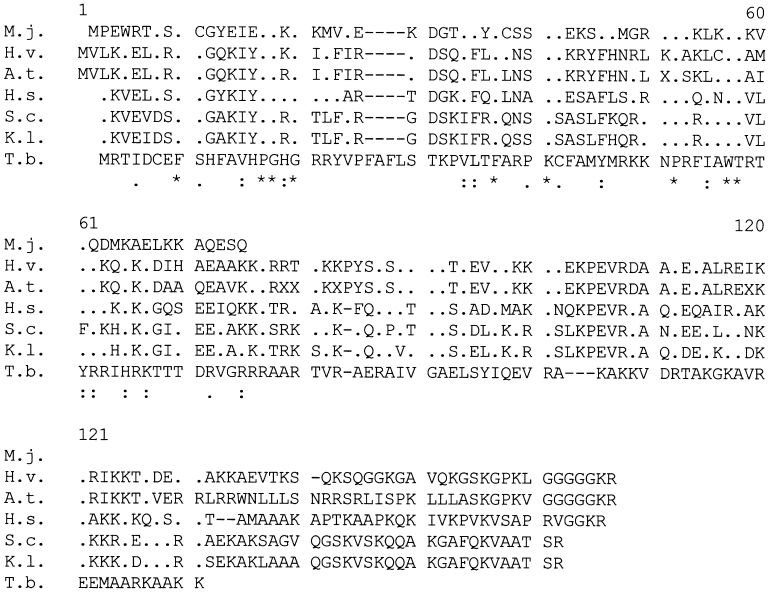

Fig. 2. Comparison of the *T. brucei* RPL24 aa sequence with those of other RPL24 homologues. Species names are abbreviated at the left and represent *Methanococcus jannaschii*, *Hordeum vulgare*, *Arabidopsis thaliana*, *Homo sapiens*, *Saccharomyces cerevisiae*, *Kluveromyces lactis*, and *Trypanosoma brucei*. Note that the *H. sapiens* sequence is identical to that of *Rattus norvegicus*, and that there is an additional, highly similar RPL24 homologue in *S. cerevisiae* (P04449 [22]). Dots within the aa sequence indicate residues identical to those in *T. brucei*. Dashes indicate gaps that have been inserted for optimal alignment. Symbols below the alignment refer to homologies between all seven RPL24 proteins as follows: *, identical aa; :, conservative substitutions; ●, semi-conservative substitutions. Sequence alignment was performed using the PILEUP and PRETTY programs from the Genetics Computer Group package. Amino acid residue numbers are indicated at the beginning and end of each line.

second lowest homology with the *T. brucei* protein in pairwise comparisons between itself and all reported eukaryotic RPL24 proteins. The *T. brucei* protein is similar to other reported RPL24 proteins in that it is highly basic even for a ribosomal protein, with lysine and arginine constituting approximately 27% of the aa residues.

Southern blot analysis was performed using *T. brucei* genomic DNA which had been digested with *Eco*RI, *Bam*HI, or *Xho*I, each of which cuts outside of the cDNA sequence (Fig. 3A). The blot was probed with a random primed probe corresponding approximately to the 3′ half of the RPL24 cDNA. In all three digests, only one band was observed, suggesting that the RPL24 gene is present in a single copy. To rule out the presence of tandemly repeated genes [16], we amplified *T. brucei* genomic DNA with two sets of RPL24-specific primers that were completely degenerate at the 3′ position. Even under annealing conditons of only moderate stringency and extension times up to 5 min, no products with a size and restriction digestion pattern that could be interpreted as tandemly repeated copies of the RPL24 gene were observed (data not shown). Thus, we conclude that the *T. brucei* RPL24 gene is present as a single copy. The size of the RPL24 cDNA including the complete spliced leader, but excluding the poly(A) tail, is 590 nucleotides. Northern blot analysis of *T. brucei* procyclic form RNA reveals a transcript of the expected size (Fig. 3B). The smearing of the signal may indicate heterogeneity in poly(A) tail length, although this was not directly evaluated.

Several ribosomal protein sequences have been reported from kinetoplastid species [17–21]. These proteins are generally more closely related to eukaryotic than to prokaryotic homologues as was observed for the *T. brucei* RPL24 protein. Levels of homology between kinetoplastid ribosomal proteins and their eukaryotic homologues are typically very high, ranging from approximately 46–69% identity. Thus, although the level of conservation between the RPL24 protein of *T. brucei* and those of several evolutionarily distant organisms is quite high, the *T. brucei* RPL24 protein is the least conserved of the kinetoplastid ribosomal proteins described to date.

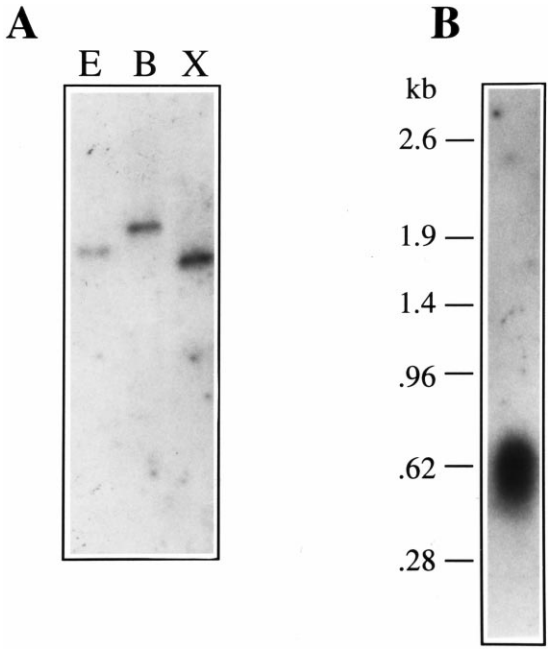

Fig. 3. Southern and northern blot analysis of the RPL24 gene. (A) Genomic DNA (10 μg) from *T. brucei* was digested with *Eco*RI (E), *Bam*HI (B), or *Xho*I (X), electrophoresed on a 1% agarose gel, and transferred to Nytran. The blot was hybridised with a random primed probe corresponding to the 3′ 286 bp of the RPL24 gene overnight at 42°C in 50% formamide, 6×SSC, 0.5% SDS, and 100 μg ml^{-1} salmon sperm DNA. Washes were performed with 2×SSC, 0.1% SDS for 15 min at room temperature, 0.1×SSC, 0.5% SDS for 30 min at 37°C, and 0.1×SSC, 0.5% SDS for 30 min at 68°C. (B) Ten micrograms of *T. brucei* procyclic form RNA was electrophoresed on a 1.5% agarose gel containing 0.66 M formaldehyde and 20 mM MOPS (pH 7.0) and transferred to Nytran. The blot was hybridised with the same probe as described above for the Southern blot overnight at 42°C in 50% formamide, 5×SSPE, 5×Denhardt's solution, 1% SDS, and 100 μg ml^{-1} salmon sperm DNA. Washes were performed with 2×SSPE, 0.1% SDS twice for 5 min at room temperature, 0.2×SSPE, 0.1% SDS twice for 5 min at room temperature, 0.2×SSPE, 0.1% SDS twice for 15 min at 42°C, and 0.1×SSPE, 0.1% SDS twice for 15 min at 68°C.

Acknowledgements

We are grateful to Drs Ken Stuart and Peter Myler for their input during the early stages of this work, and to Dr Ira Wool for advice on ribosomal protein nomenclature. We also thank Mark Hayman for critical reading of the manu-

script. L.K.R. is a Burroughs Wellcome Fund New Investigator in Molecular Parasitology.

References

[1] Nitta I, Ueda T, Watanabe K. Possible involvement of *Escherichia coli* 23S ribosomal RNA in peptide bond formation. RNA 1998;4:257–67.
[2] Nitta I, Kamada Y, Noda H, Ueda T, Watanabe K. Reconstitution of peptide bond formation with *Escherichia coli* 23S ribosomal RNA domains. Science 1998;281:666–9.
[3] Schimmel P, Alexander R. All you need is RNA. Science 1998;281:658–9.
[4] Reyes R, Vazquez D, Ballesta JPG. Peptidyl transferase center of rat-liver ribosome cores. Eur J Biochem 1977;73:25–31.
[5] Chan Y-L, Olvera J, Wool IG. The primary structure of rat ribosomal protein L24. Biochem Biophys Res Commun 1994;202:1176–80.
[6] Johnson KR. Characterization of cDNA clones encoding the human homologue of *Saccharomyces cerevisiae* ribosomal protein L30. Gene 1993;123:283–5.
[7] Estebsan R, Labrador E, Dopico B. Ribosomal protein L24 hololog ue is expressed in *Cicer arietinum* L. epicotyls. Plant Physiol 1998;117:717–8.
[8] Baronas-Lowell DM, Warner JR. Ribosomal protein L30 is dispensible in the yeast *Saccharomyces cerevisiae*. Mol Cell Biol 1990;10:5235–43.
[9] Eng FJ, Warner JR. Structural basis for the regulation of splicing of a yeast messenger RNA. Cell 1991;65:797–804.
[10] Bult CJ, White O, Olsen GJ et al. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 1996;273:1058–73.
[11] Kawarabayasi Y, Sawada M, Horikawa H et al. Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res 1998;5:55–76.
[12] Klenk HP, Clayton RA, Tomb J et al. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature 1997;390:364–70.
[13] Parsons M, Nelson RG, Watkins KP, Agabian N. Trypanosome mRNAs share a common 5′ spliced leader sequence. Cell 1984;38:309–16.
[14] Devereux J, Haeverli P, Smithies O. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 1984;12:387–95.
[15] Sogin ML, Gunderson JH, Elwood HJ, Alonso RA, Peattie DA. Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science 1989;243:75–7.
[16] Van Hamme L, Pays E. Control of gene expression in trypanosomes. Microbiol Rev 1995;59:223–40.
[17] Tebabi P, Halleux S, Pays E. Nucleotide sequence of a full-length cDNA coding for the ribosomal L44 protein of *Trypanosoma brucei*. Nucleic Acids Res 1990;18:2809.
[18] Bard E, Janeczko R. Sequence and genomic locus of a *Leishmania* protein homologous to rat ribosomal protein S8. Nucleic Acids Res 1992;20:369.
[19] Metzenberg S, Joblet C, Verspieren P, Agabian N. Ribosomal protein L25 from *Trypanosoma brucei*: phylogeny and molecular co-evolution of an rRNA-binding protein and its rRNA binding site. Nucleic Acids Res 1993;21:4936–40.
[20] Myler PJ, Tripp CA, Thomas L, Venkataraman GM, Merlin G, Stuart KD. The LD1 amplified element from *Leishmania infantum* encodes a homolog of ribosomal protein L37. Mol Biochem Parasitol 1993;62:147–52.
[21] Love DC, Wilson ME, Mosser DM. Cloning and characterization of the ribosomal *l11* gene from *Leishmania* spp. Mol Biochem Parasitol 1995;71:261–4.
[22] Mitra G, Warner JR. A yeast ribosomal protein gene whose intron is in the 5′ leader. J Biol Chem 1984;259:9218–24.