Nucleotide sequence of a cDNA encoding another *Trypanosoma cruzi* acidic ribosomal P2 type protein (TcP2b)

Martin P.Vazquez, Alejandro G.Schijman, Alfredo Panebra and Mariano J.Levin*
Instituto de Investigaciones en Ingeniería Genética y Biología Molecular (INGEBI), Obligado 2490, 1428 Buenos Aires, Argentina

Submitted April 10 1992

EMBL accession no. X65065

We have previously described the complete amino acid sequence of the *Trypanosoma cruzi* ribosomal P-JL5 protein, TcPJL5, (1); since then it has been proved that it is an acidic ribosomal P2 type protein (2).

Screening of a λgt11 bloodstream trypomastigote cDNA library (3) with a serum from a Chagas heart disease patient allowed the identification of a recombinant, M-1, that encoded the C-terminal sequence of another acidic ribosomal P2 type protein from *T.cruzi*. To determine the complete sequence of its mRNA, we developed an RNA-PCR amplification protocol that allowed the amplification of the 5′ end of this mRNA, using an oligonucleotide derived from the *T.cruzi* spliced leader sequence, SL, (4), and another one, anti-sense, corresponding to the beginning of the cloned M-1 cDNA.

The sequences of the amplified 5′ cDNA fragment, and the M-1 cDNA span the complete mRNA sequence of the ribosomal P2 type protein TcP2b. It is 597 nucleotides long, with an ORF of 336 nucleotides encoding a protein of 112 amino acids.

The AUG initiation codon is preceded by a 63 nucleotides long non coding sequence including the first 35 bases of the *T.cruzi* SL sequence. However, a second AUG is present only two triplets further downstream. This AUG generates a better match with other ribosomal P2 proteins (Figure 1), therefore it seems necessary to examine the protein in order to determine whether translation begins at the first AUG or whether the second AUG (or both) is used. Interestingly, the use of the first AUG results in a serine at position 2, a hallmark of the acidic ribosomal P1 N-terminal sequences (see accompanying paper).

The TcP2b amino acid sequence has a molecular weight of 10965.36; as it is the case for other acidic ribosomal P proteins, it is characterized by a large number of alanines, 34 alanine residues out of the 112 total amino acids, and by the presence of 18 acidic amino acids, 8 aspartic and 10 glutamic residues, resulting in an estimated pI of 4.85.

The TcP2b amino acid sequence shows 47% homology to TcPJL5, 46% homology to the *Saccharomyces cerevisiae* YP2α (5), and 41% homology to the *S.cerevisiae* YP2β (5) (Figure 1).

Comparison of the N-terminal globular regions, frequently used to determine evolutionary relationships in this protein family (6), indicates that the two *S.cerevisiae* P2 proteins share 56% homology, while the homology between TcPJL5 and TcP2b is only 40% (Figure 1).

Interestingly, the 14 amino acid C-terminal sequence of TcP2b is 100% homologous to the corresponding sequence in TcPJL5, and contains no serine residues, a hallmark of the low molecular weight *T.cruzi* acidic ribosomal P proteins.

ACKNOWLEDGEMENTS

This work was supported by the French-Argentinian Cooperation Program (INSERM-CONICET); the United Nations Development Program-World Bank-World Health Organization Special Program for Research and Training in Tropical Diseases; and the UNIDO/ICGEB-ARG91-01 Collaborative Research Programme. This study was conducted while Alfredo Panebra was in receipt of a WHO Research Training Grant.

REFERENCES

1. Schijman,A., *et al.* (1990) *Nucleic Acids Res.* **11**, 3399.
2. Hansen,T.S., *et al.* (1991) *Gene* **105**, 143–150.
3. Levy-Yeyati,P., *et al.* (1991) *Immunol. Lett.* **31**, 27–34.
4. Milhausen,M., *et al.* (1984) *Cell* **38**, 721–729.
5. Remacha,M., *et al.* (1988) *J. Biol. Chem.* **263**, 9094–9101.
6. Hunter Newton,C., *et al.* (1990) *J. Bacteriol.* **172**, 579–588.

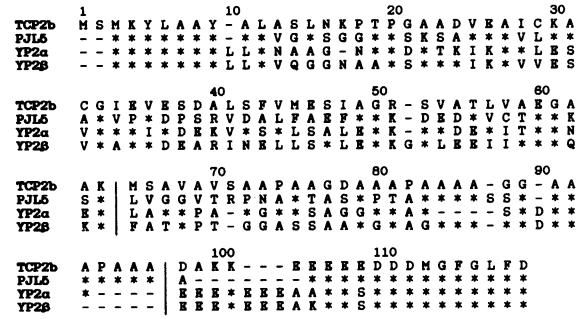

Figure 1. Amino acid sequence alignment of TcP2b with TcPJL5 (1), with YP2α (5), and YP2β (5). The N-terminal globular domain extends from positions 1 to 64; the hinge region from positions 65 to 97; and the C-terminal domain from positions 98 to 119. The stars, *, indicate identical residues; the bars indicate the limits of the three domains. The common scale is in amino acid residue positions along the linear alignment.

* To whom correspondence should be addressed