Nucleotide sequence of a cDNA encoding a *Trypanosoma cruzi* acidic ribosomal P1 type protein

Martin P.Vazquez, Alejandro G.Schijman 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. X65025

Several highly antigenic λ gt11 cDNA recombinant phage were isolated from a *Trypanosoma cruzi* bloodstream trypomastigote cDNA library by immunoscreening (1). One of these, L-6, encoded the C-terminal portion of a *T.cruzi* acidic ribosomal P protein. Using an oligonucleotide derived from a specific internal region of the L-6 cDNA, and another one corresponding to the first 20 nucleotides of the *T.cruzi* splice leader sequence, SL (2), we amplified, by RNA-PCR, the 5' end of this mRNA.

The amplified cDNA, together with the L-6 cDNA, span the complete mRNA sequence of an acidic ribosomal P1 type protein, TcP1. It is 541 nucleotides long, and includes an ORF of 330 nucleotides encoding a 109 amino acid polypeptide, with an estimated pI of 5.13. It shares 51% and 40% homology with the corresponding human P1 (3) and *Saccharomyces cerevisiae* P1 α proteins (4), respectively (Figure 1), and it shows typical N-terminal features of the ribosomal P1 proteins such as the serine residue in alignment position 3, and the amino acid sequence LAC, alignment positions 8–10 (Figure 1, and reference 5). Moreover, the TcP1 sequence presents an internal amino acid motif GG(V/P)APAAG, alignment positions 67–74, only found in P1 proteins of higher eukaryotic organisms (5).

It is noteworthy that TcP1 also shares a great degree of homology with the two other small molecular weight *T.cruzi* ribosomal P2 type proteins, TcPJL5, 45%, and TcP2b, 47%, (6). All three *T.cruzi* proteins conserve: a) the characteristic P1 amino acid motif KAXGXXV beginning in alignment position 32 (Figure 1), b) a large number of alanines residues clustered in the alanine-rich hinge region, and c) the serine free C-terminal sequence EEEEDDD-MGFGLFD (Fig. 1).

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.

REFERENCES

1. Levy-Yeyati, P., Bonnefoy, S., Mirkin, G., Debrabant, A., Lafon, S., Panebra, A., Gonzalez-Cappa, E., Dedet, J.P., Hontebeyrie-Joskowicz, M. and Levin, M. (1991) *Immunol. Lett.* **31**, 27–34.
2. Milhausen, M., Nelson, R.G., Sather, S., Selkirk, M. and Agabian, N. (1984) *Cell* **38**, 721–729.
3. Rich, B.E. and Steitz, J.A. (1987) *Mol. Cell. Biol.* **7**, 4065–4074.
4. Hunter Newton, C., Shimmin, L.C., Yee, J. and Dennis, P.P. (1990) *J. Bacteriol.* **172**, 579–588.
5. Hansen, T.S., Andreasen, P.H., Dreisig, H., Hojrup, P., Engberg, J. and Kristiansen, K. (1991) *Gene* **105**, 143–150.
6. Schijman, A., Dusetti, N., Vazquez, M., Lafon, S., Levy-Yeyati, P. and Levin, M. (1990) *Nucleic Acids Res.* **11**, 3399.

	1	10	20	
TCP1	M - S S K Q Q L A C T Y A A L I L A D S G K T D M D S - L L			
HP1	* A * V - S E * * * I * S * * Y * H * D E V * V E T D K I N			
YP1 α	* - * T E S A * - - S * * * * * * * * E T E I S S E K * *			
	31	40	50	
TCP1	K V T K A A G V D V S K G M A S A F A S I L K N V D I N D V			
HP1	A L I * * * * N * G P F W P G L * * K A * A * * N * G S L			
YP1 α	T L * N * * N * P D E N I W * D I * * K A * D G Q N L K * I			
	61	70	80	
TCP1	L S K V S F G G V A P A A G G A T A A P A A A A A A A P			
HP1	I C N * G A * * P * * * * A * P * G G P * P S T * * * *			
YP1 α	* V N T S A - - - - G * A A P A G V A G G V * G G E * G			
	91	100		
TCP1	A A A A A - - - K K E E - E E E D D D M G F G L F D			
HP1	* E E K K V E A * * * * S * * S * * * * * * * * * *			
YP1 α	E * E * E K E E E E A K - * * S * * * * * * * * * *			

Figure 1. Amino acids sequence alignment of TcP1 with Hs P1 (3) and Sc P1 α (4). The N-terminal globular domain extends from positions 1–64; the hinge region from positions 65–95, and the C-terminal domain from positions 96–116. 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