Volume 257, number 2, 365–368 FEB 07794 November 1989

Variable number of repeat units in genes encoding *Trypanosoma cruzi* antigens

Roberto A. Macina, José L. Affranchino, Guido D. Pollevick, Elena E. Jazin and Alberto C.C. Frasch

Instituto de Investigaciones Bioquímicas Fundación Campomar, Antonio Machado 151, 1405 Buenos Aires, Argentina

Received 11 September 1989

The genes encoding a protein antigenic during the acute phase (SAPA) and another one antigenic during the chronic phase (antigen 30) of human Chagas' disease were analyzed in 15 *Trypanosoma cruzi* isolates and clones collected in distant geographical regions. These two genes had tandem repeats which were present in all parasites tested. However, large differences in the length of restriction fragments were observed among isolates. This was readily explained by variations in the number of repeat units present in homologous genes. This result was confirmed after analysis of 3 members of the SAPA gene family. In the case of antigen 30, we propose that differences in the number of repeat units result in size differences in the corresponding RNAs and proteins, which explains the large size heterogeneity in otherwise homologous *T. cruzi* antigens.

Gene; Tandem repeat; RNA; Antigen; (*Trypanosoma cruzi*)

1. INTRODUCTION

Several proteins containing repeated amino acid sequences have been described in eucaryotic cells. However, the most extraordinary cases are present in parasites. Plasmodium spp. were the first protozoa where a large variety of antigens containing tandem repeats of amino acid sequences was found (see [1] for a review). Some antigens contained repeats conserved among parasites while other antigens showed a large variation in amino acid sequences and length in different Plasmodium spp. and among parasites of the same species [1,2]. The presence of such an extraordinary number of repeat-containing antigens, some of which have intra- and intermolecular cross-reacting epitopes, could result in dramatic immunological consequences [1]. *Trypanosoma cruzi*, the agent of Chagas' disease, is the other protozoan parasite where a number of proteins containing amino acid repeats have been described [3,4]. 8 out of 10 DNA clones encoding different parasite antigens contained non-homologous tandem repeats ranging in size from 15 to 204 base pairs [4–6]. One of these antigens, named SAPA for shed-acute-phase-antigen, was a shed protein and a good marker of the acute phase of Chagas' disease because antibodies against it were detected in 93% of sera from acute human infections [6]. Other antigens, like no.30, were markers of the chronic stage of human infections; they were detected by a large proportion of sera collected from patients during the chronic phase of the disease [4]. In this paper we show that sequences homologous to genes encoding antigens SAPA and no.30 are present in all 15 parasite isolates and clones collected in distant geographical regions. Also, we show that the large heterogeneity in the size of genes, RNAs and proteins can be explained by variations in the number of repeat units.

Correspondence address: A.C.C. Frasch, Instituto de Investigaciones Bioquímicas Fundación Campomar, Antonio Machado 151, 1405 Buenos Aires, Argentina

2. MATERIALS AND METHODS

Parasites used were epimastigotes grown in liquid media [5]. Non-cloned isolates and clones were collected in the indicated countries and were numbered as follows (see figs): (1) CA1-72, Argentina, clone; (2) Tulahuen 0, Chile, isolate; (3) Tulahuen 2, Chile, isolate; (4) 3/1, Argentina, isolate; (5) RA, Argentina, isolate; (6) P/11, Bolivia, isolate; (7) P/185, Bolivia, isolate; (8) P/209, Bolivia, isolate; (9) P/255, Bolivia, isolate; (10) MS, Brazil, clone; (11) AC/125, Brazil, clone; (12) GER, Brazil, clone; (13) Y, Brazil, isolate; (14) CA1-69, Argentina, clone; (15) Miranda/80, Argentina, clone; (16) Sonia, Brazil, isolate; (17) AWP, Argentina, isolate; (18) 40/3, Argentina, isolate; (19) Corpus Christi, USA, isolate.

Nuclear DNA and RNA were purified from epimastigotes as described [5]. The two DNA probes containing fragments encoding two different *T. cruzi* antigens (SAPA and 30) were sequenced [4,6]. Nuclear DNA from parasites was digested with the indicated restriction enzyme, run on 0.8% agarose gels and blotted onto nitrocellulose filters. Hybridization with nick-translated DNA probes was performed under stringent conditions as described [5]. SAPA clone was obtained from *T. cruzi* Miranda/76 and completely sequenced [6]. Two other fragments homologous to SAPA (clones 4.1 and 4.2) were rescued from the same parasite using SAPA DNA as labelled probe. The boundaries of the region containing repeated motifs were located by end labelling of the cloned inserts and partial digestion with the restriction enzyme *Rsa*I that cuts the repeat units.

Total protein extracts from epimastigotes were prepared [5] using protease inhibitors: PMSF (phenylmethylsulfonyl fluoride) 0.5 mM,

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

TLCK (tosyl-lysine chloromethyl ketone) 2 mM, leupeptin 5 μM, pepstatin 5 μM and E-64 5 μM. The extracts were loaded on 7.5% polyacrylamide gels and electrophoresed [7]. Protein blotting onto nitrocellulose filters was performed essentially as described by Burnette [8]. Antibodies against antigen 30 were prepared in rabbits as described [4]. This serum was shown to specifically detect antigen 30 and was used in 1/2000 dilution for reaction with Western blots [5]. Molecular mass markers were: myosin (205 kDa), β-galactosidase (116 kDa), phosphorylase B (97.4 kDa) and bovine serum albumin (66 kDa).

3. RESULTS AND DISCUSSION

3.1. *Sequences homologous to genes SAPA and 30 were present in all parasite strains but they differed in the number of repeat units*

All 15 isolates collected in Bolivia, Brazil and Argentina had sequences homologous to SAPA (fig.1). However, the hybridization pattern obtained was complex in all parasites due to the presence of a family of SAPA genes (see below). Several bands that greatly differed among isolates were observed with *Hpa*II and *Mbo*I, both are restriction enzymes that do not cut the repeat unit of SAPA (see sequence in [6]). Since similar or identical patterns were observed with both enzymes in each isolate, it is likely that each of these bands are mainly made up of repeated units. Further proof comes from experiments in which DNAs were digested with *Rsa*I, a restriction enzyme that cuts the repeat units of SAPA gene [6] (fig.1). Under these conditions, few bands of small sizes were detected in all isolates and the large heterogeneity among parasites was no longer observed. These small bands correspond to the non-repeated 5′ and 3′ regions that flanked the repeat units in the SAPA probe (see fig.2). These results showed a large variation in the size of the fragments containing the repeat units. The *Rsa*I site was conserved in the repeat units of all isolates tested. Therefore, variations in the size of restriction endonuclease fragments are most probably due to variations in the number of repeat units. In order to confirm the presence of different numbers of repeated motifs among genes and also the existence of a SAPA gene family, homologous genomic DNA fragments corresponding to three SAPA genes were isolated from one parasite clone. SAPA clone was previously sequenced, it contained 14 copies of 36 base pairs tandem repeats [6]. The other two genomic clones were analyzed by restriction mapping and partial digestion with *Rsa*I and they roughly contained 33 and 47 copies of the repeats (fig.2). Thus, the difference in size of the region containing repeat units can be as large as 1200 base pairs among these cloned sequences. Restriction sites around the repeated region were in most cases conserved (fig.2).

Given the variations in the number of repeat units among *T. cruzi* isolates in otherwise homologous genes, a main prediction from these results is that antigens encoded by these genes must differ in their molecular weights. In the case of SAPA, one parasite isolate had at least 4 RNAs and 5 proteins homologous to SAPA that differed in molecular weight [5,6]. Unfortunately, SAPA was detected only in the blood stage of this parasite [6] and preparation of this parasite stage from several *T. cruzi* isolates is difficult. Consequently, we decided to analyze a second gene, number 30, whose protein is expressed in a parasite stage (epimastigote) which can be readily grown in culture media.

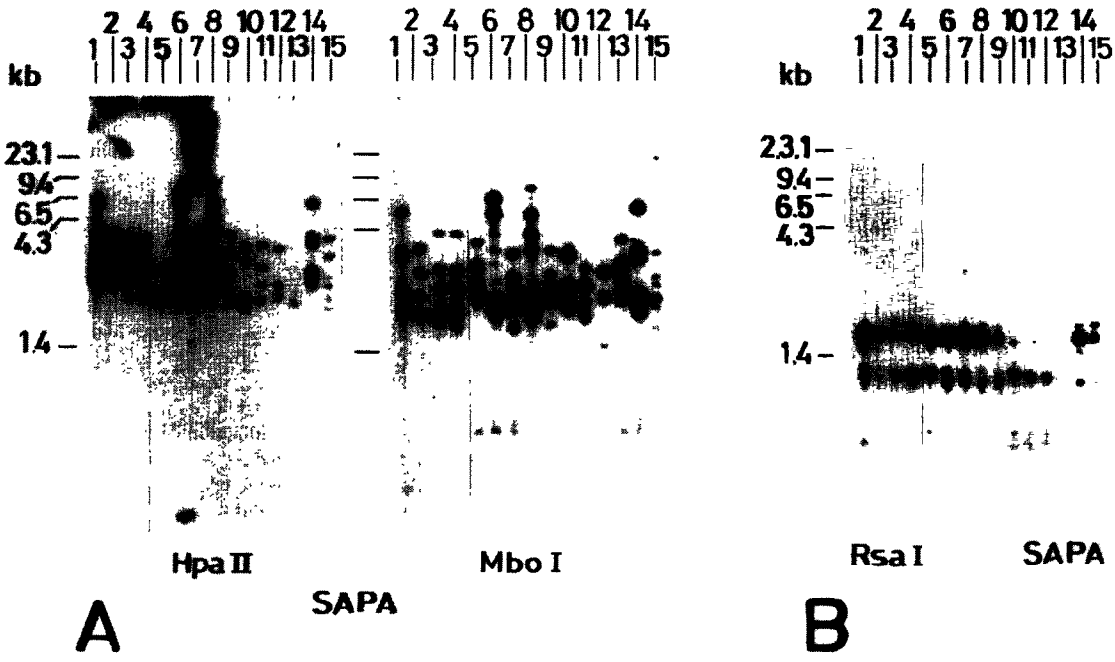

Fig.1. Hybridization pattern of DNAs from parasite isolates and clones using SAPA probe. See section 2 for parasite isolates and clones origin.

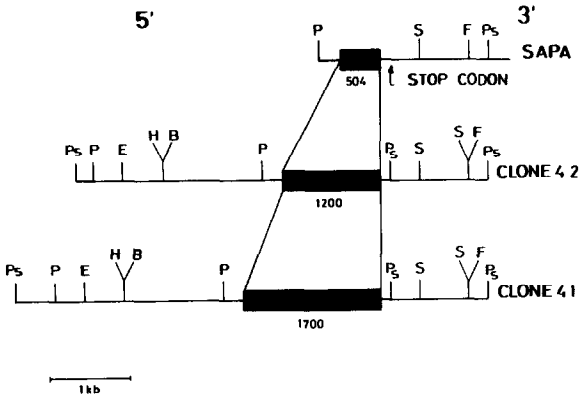

Fig.2. Restriction endonuclease maps of the 3′ regions of three SAPA genes. Boxes indicate regions containing repeated units that are cut by the enzyme *Rsa*I. Numbers below boxes indicate length in base pairs of regions containing repeated units in each clone. P, *Pvu*II; E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; Ps, *Pst*I; S, *Ssp*I; F, *Fsp*I.

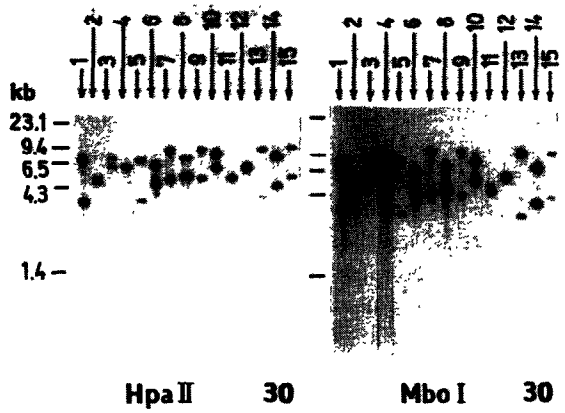

Fig.3. Hybridization pattern of DNAs from parasite isolates and clones using probe 30.

Clone 30 probe was entirely made up of repeat units [4] and it detected homologous sequences in all 15 isolates (fig.3). The size of the fragments detected differed greatly among parasites. Two fine-cutter restriction enzymes that do not hinder the repeat units of clone 30 (*Hpa*II and *Mbo*I), gave patterns which were almost identical in each parasite isolate or clone (fig.3). Thus, with few exceptions, two bands of almost identical size were observed with both enzymes in each parasite. Moreover, a third enzyme that did not cut the repeat units of clone 30 (*Taq*I) also gave rise to two bands in 3 parasites (CA1-72, Miranda and Tul2), which were similar in size to those obtained with *Hpa*II and *Mbo*I (not shown). It is very unlikely that the sites for these three enzymes are linked. The most plausible conclusion is that, as for the cloned SAPA genes (fig.2), the size differences detected with probe 30 are due to variations in the number of repeat units. Furthermore, an enzyme that cuts the repeats of gene 30 (*Hae*III, [4]), completely cuts both bands in most of the isolates tested, thus showing that *Hae*III sites were conserved in most parasites (not shown). The two bands observed in each parasite could correspond to duplicated genes or to alleles of a single gene. We have been unable to analyze whole cloned fragments containing the repeat units together with flanking regions, they tend to be partially deleted in *E. coli* RecA.

3.2. *Proteins and RNAs 30 differed in size among parasite isolates*

We predicted that protein 30 should vary in size among isolates because they should contain a different number of amino acids repeats. Antigen 30 is expressed in the epimastigote stage, and consequently, it was possible to grow several parasite isolates in culture to analyze their antigens. Total proteins from 10 isolates and clones were probed with polyclonal antibodies against antigen 30 (fig.4). A main protein band was observed. Its apparent molecular mass varied among isolates, ranging from 130 to 180 kDa. Other weak bands were present in some lanes. This variable pattern among isolates is not likely to be due to proteolytic activity because several protease inhibitors were used. Moreover, patterns could be reproduced in different experiments. If size differences in antigen 30 were due to variations in the length of the coding region and not to other protein modification, it would be expected that RNAs encoding antigen 30 should also differ in size. A parasite expressing a small antigen 30 (RA) and another one expressing a large antigen 30 (Tul2) were selected to analyze their corresponding RNAs (fig.4). Two RNA bands were observed with probe 30 in each parasite isolate. The bigger RNAs were similar in size (8.5 kb) in both isolates and may be the ones encoded by the putative larger gene 30 as seen in Southern blots (fig.3). The smaller RNA bands differed in size between both isolates (4.3 kb in RA and 7.0 kb in Tul2) and they may be the ones encoded in the putative smaller gene 30. The 4.3 kb and 7.0 kb RNA bands were large enough as to encode proteins of 130 kDa and 180 kDa in RA and Tul2 isolates, respectively. Those were the sizes of

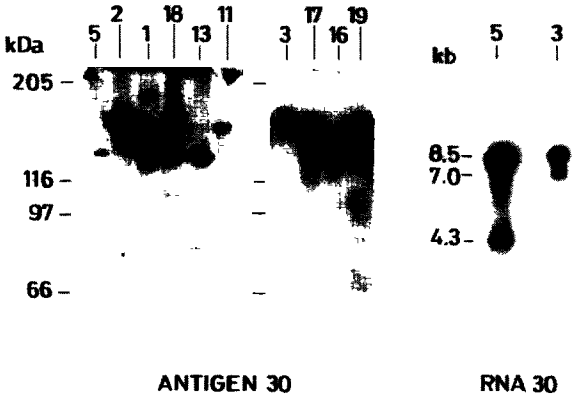

Fig.4. Variations in apparent molecular weight of antigen 30 and in RNAs sizes of *T. cruzi* isolates and clones.

the proteins detected in Western blots (fig.4). Furthermore, there was a good correspondence between the size of the smaller band in Southern blots (fig.3) and the apparent molecular weight of the protein (fig.4) for all isolates tested.

Genes for antigens SAPA and 30 as well as nos 1, 2, 13 and 36 not included in this paper [4], seem to be always present in distantly collected parasite isolates. These genes code for proteins antigenic during the acute (SAPA) or chronic (antigen 30) human infections [4,6] and can be used in the diagnosis of Chagas' disease (Reyes, M.B. et al., submitted). The second main observation made in this paper is the variation in the number of repeat units in *T. cruzi* genes. Heterogeneity in *T. cruzi* antigens was described. Furthermore, in some cases homologous antigens varied in size among isolates [9–11]. We are now proposing that size variations in some antigens is due to large differences in the number of repeat units present in the corresponding genes. Since several *T. cruzi* antigens are made up of tandem repeated motifs [4,6], it is expected that size variations in homologous antigens will be frequently observed in this parasite.

Acknowledgements: We thank Dr C.F. Ibañez for his help in preliminary experiments, Dr J. Peralta for providing us with Brazilian clones MS, AC/125 and GER, Dr F. Brenière for providing parasites from Bolivia, Dr J.J. Cazzulo and Dr U. Pettersson (Uppsala, Sweden) for helpful discussions, and Dr Berta Franke for providing epimastigotes of *T. cruzi*. This work was supported by grants from UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the Rockefeller Foundation, Swedish Agency for Research Cooperation with Developing Countries (SAREC) and Consejo Nacional de Investigaciones Cientificas y Tecnicas.

REFERENCES

[1] Kemp, D.J., Coppel, R.L. and Anders, R.F. (1987) Annu. Rev. Microbiol. 41, 181–208.
[2] Nussenzweig, V. and Nussenzweig, R.S. (1985) Cell 42, 401–403.
[3] Peterson, D.S., Wrightsman, R.A. and Manning, J.E. (1986) Nature 322, 566–568.
[4] Ibañez, C.F., Affranchino, J.L., Macina, R.A., Reyes, M.B., Leguizamon, S., Camargo, M.E., Aslund, L., Pettersson, U. and Frasch, A.C.C. (1988) Mol. Biochem. Parasitol. 30, 27–34.
[5] Ibañez, C.F., Affranchino, J.L. and Frasch, A.C.C. (1987) Mol. Biochem. Parasitol. 25, 175–184.
[6] Affranchino, J.L., Ibañez, C.F., Luquetti, A.O., Rassi, A., Reyes, M.B., Macina, R.A., Aslund, L., Pettersson, U. and Frasch, A.C.C. (1989) Mol. Biochem. Parasitol. 34, 221–228.
[7] Laemmli, U.K. (1970) Nature 227, 680–682.
[8] Burnette, W.N. (1981) Anal. Biochem. 112, 195–203.
[9] Plata, F., Pons, F.G. and Eisen, H. (1984) Eur. J. Immunol. 14, 392–397.
[10] Wrightsman, R.A., Leon, W. and Manning, J.E. (1986) Infect. Immunol. 53, 235–239.
[11] Alves, M.J.M., Abuin, G., Goçalves, M.F., Kuwajima, V.Y. and Colli, W. (1987) in: Molecular Strategies of Parasitic Invasion, vol.42 (Agabian, N., Goodman, E. and Nogueira, N. eds) UCLA Symposium on Molecular and Cellular Biology New Series, pp.323–332, Alan R. Liss, New York.