Mem. Inst. Oswaldo Cruz, Rio de Janeiro, Suppl. Vol. 82, November 1987/Page 229

TDR-3

IDENTIFICATION OF THE 85 KDA SURFACE ANTIGEN GENE OF *TRYPANOSOMA CRUZI* AS A MEMBER OF A MULTIGENE FAMILY

Jerry E. Manning and David S. Peterson

Department of Molecular Biology

and Biochemistry

University of California, Irvine

Irvine, California 92672

Introduction: The parasitic protozoan Trypanosoma cruzi is the causative agent of Chagas disease, a major health problem throughout Central and South America. The parasite infects human and other vertebrate hosts by direct penetration of host cells by trypomastigotes transmitted by the insect vector. The surface antigens of the trypomastigotes have been implicated in the process of cell penetration. In particular antibody neutralization experiments show that an 85 kDa surface glycoprotein is necessary for efficient interiorization of bloodstream trypomastigotes in mammalian cells. We have recently described the molecular cloning of a genomic DNA fragment that encodes antigenic determinants present in an 85 kDa surface antigen and have confirmed previous studies that show this antigen is specific to the trypomastigote stage of the parasite. Nucleotide sequence analysis of the cloned DNA infers the presence of a nonapeptide unit that is tandemly repeated five times. Here we show that the 85 kDa gene is a member of a complex multigene family of which only three members contain the repeated sequence. Only one member appears to be expressed and it is located on a telomere. Also, the nonapeptide unit is recognized by serum antibodies of mice infected with T. cruzi.

Results and Discussion:

Genomic Organization of the 85 kDa Surface Antigen Gene:

In order to determine the size of genomic DNA restriction fragments which flank the DNA sequences present in the cloned 85 kDa sequence, designated Tcg1, a southern blot of EcoRI restricted trypomastigote genomic DNA was hybridized with Tcg1 insert DNA. Numerous EcoRI fragments hybridized (ie. 20-25), indicating that sequences within the 500 bp clone are found in multiple copies within the genome. To better define which sequences within Tcg1 are multicopy, three separate probes were constructed from the Tcg1 insert. Probe one is a 110 bp EcoRI/Sau3A fragment containing the 5' end of Tcg1. Probe 2 is a 27 base synthetic oligonucleotide representing one unit of the tandem repeat. Probe 3 is a 180 bp HhaI/EcoRI fragment containing the 3' end of Tcg1. Hybridization of these probes to a Southern blot of genomic DNA restricted with EcoRI showed that probes 1 and 3 hybridized to 10-20 different size restriction fragments while the 27 nucleotide probe detected only three restriction fragments.

Cloning of Genomic DNA Fragments Containing the 27 Nucleotide Repeat.

The above results indicate that mRNA detected by hybridization with the 500 bp insert of Tcg1 may be transcribed

by one or more members of this multigene family. To confirm that the repeat sequence is present in mRNA, the 27 nucleotide probe was shown to hybridize with a 3.8 kb RNA in both total celluiar and poly A+ RNA from trypomastigotes. No hybridization was observed with total cellular or poly A+ RNA from epimastigotes. These results are identical to those observed with the 500 bp insert in Tcg1 and clearly show that at least one of the three sites in the genome which contain the 27 bp repeat is transcriptionally active. In order to determine which of these sites are transcriptionally active we attempted to clone and characterize a representative of each.

Three separate genomic libraries were constructed in the cloning vector EMBL4. Approximately 400,000 independent recombinant lambda plaques were screened with the 27 nucleotide probe and 59 positive piaques were identified. A 4.8 kbp EcoRI fragment was observed to hybridize with the 27 nucleotide probe in 48 of the DNAs, and a 1.7 kbp EcoRI fragment hybridized in the remaining 11 DNAs. Surprisingly, no phage DNAs were observed to contain a 5.4 kbp fragment homologous to the 27 nucleotide probe.

For the 4.8 kbp and 1.7 kbp fragments the sequences homologous to the 27 bp repeat were localized to EcoRI/SmaI fragments of length 490 bp and 325 bp, respectively. The nucleotide sequence of these fragments clearly showed that Tcg1 is not present in either fragment. Tcg1 must, therefore, reside elsewhere within the genomic DNA, possibly within the 5.4 kbp

fragment.

The results above suggest that the 5.4 kbp fragment may contain the sequences found in Tcg1. If this is the case restricting genomic DNA with enzymes that cut adjacent to the repeat region should yield fragments whose sizes can be predicted from the restriction maps of Tcg1, the 4.8 and the 1.7 kbp EcoRI fragments. Restriction analysis of the three fragments revealed that each contains a HhaI site 3' of the repeat region. Digestion of the subcloned Tcg1, 4.8 and 1.7 kbp fragments with EcoRI and HhaI yielded fragments of 320, 1500, and 500 bp respectively, that hybridized with the 27 nucleotide probe. When genomic DNA was digested with EcoRI and HhaI three fragments of length 1500, 500 and 320 bp were observed to hybridize with the 27 nucleotide probe. As the 1500 and 500 bp fragments can be identified as having originated from the 4.8 and 1.7 kbp fragments respectively, the 320 bp fragment found in Tcg1 is putatively present in the 5.4 kbp fragment.

Confirmation that the 5.4 kbp fragment contains Tcg1 could be obtained by cioning, restriction mapping and sequencing the 5.4 kbp fragment. Therefore an effort was made to directly clone the 5.4 kb fragment as an EcoRI insert into the EcoRI site of lambda gt10. Total trypomastigote genomic DNA was digested with EcoRI, size fractionated by agarose gel electrophoresis and fragments of size 4.5-6.0 kbp were excised and cloned into lambda gt10. Approximately 200,000 independent clones were screened

with the 27 nucleotide probe and 69 plaques were identified and rescreened positive. In each of the 69 phage DNA hybridization with the 27 nucleotide probe was observed only to a single 1.5 kb EcoRI/Hhal fragment identical in size to the EcoRI/HhaI fragment found in the cloned 4.8 kbp fragment, and no phage were observed to contain a 5.4 kbp insert.

Nuclease Bal 31 Sensitivity

The inability of the 5.4 kbp fragment to be cloned as an EcoRI insert led us to consider the possibility that one end of the fragment was not an EcoRI termini but is found in the telomeric region of the chromosome. To test this possibility, aliquots of high molecular weight trypomastigote genomic DNA were digested for increasing times with nuclease Bal 31. The digested DNAs were then restricted with EcoRI, electrophoresed through agarose, blotted to nitrocellulose, and probed with the 27 nucleotide repeat. A progressive decrease in the size of the 5.4 kbp fragment was observed with increasing digestion with Bal 31. Neither the 4.8 kbp nor the 1.7 kbp EcoRI fragment changed significantly in size over the time course of the experiment. This result strongly implies that in the trypomastigote the 5.4 kbp fragment is telomeric.

DNase 1 Sensitivity

Sensitivity to DNase I is a characteristic of actively

transcribed genes. As a first step in determining which of the EcoRI fragments that hybridize with the 27 nucleotide probe are expressed, the relative DNase 1 sensitivity of the 3 genomic EcoRI fragments was assayed. Aliquots of trypomastigote nuclei were reacted with increasing concentrations of DNase 1, and DNA was isolated from each sample. After digestion witb EcoRI nuclease and agarose gel electrophoresis, the DNA was Southern blotted and probed with the 27 base oligonucleotide. At progressively higher concentrations of DNase I, the 5.4 kbp fragment was found to be preferentially digested. The DNA present in each hybridizing band was quantitated by densitometric measurement and the relative amounts of the three DNA bands were plotted as a function of the different DNase I concentrations. The results showed that while the amount of 4.8 kbp and 1.7 kbp EcoRI fragments changes little with respect to each other, the amount of 5.4 kbp fragment decreases markedly relative to each of these fragments. The hypersensitivity of the 5.4 kbp fragment to digestion by DNase 1 argues that it, and not the 4.8 or 1.7 kbp fragments, contains sequences which are being actively transcribed.

Nucleotide Sequence of cDNA Having Homology to the 27 bp Repeat

The results of the DNase 1 sensitivity experiment provide indirect evidence that only the 5.4 kbp fragment is transcribed. More direct evidence could be obtained by determining the nucleotide sequence of cDNA inserts selected by hybridization with the 27 bp repeat unit. Since the nucleotide sequences

adjacent to the repeat in the three genomic clones clearly differ, comparing the cDNA sequence with that of the genomic clones should identify which genomic sequence is being transcribed.

Two separate cDNA libraries were constructed from trypomastigote poly A+ RNA, the first was synthesized with oligo dT as the primer, the second used small oligonucleotides of random sequence . After the addition of EcoRI linkers, the libraries were cloned into the EcoRI site of lambda gt10. Approximately 100,000 recombinant plaques from each library were screened with the 27 nucleotide probe and 11 plaques, 8 from the random primed library and 3 from the oiigo dT prlmed library, were isolated and plaque purified. DNA was isolated from each phage and restricted with EcoRI. The cDNA inserts were sized on an acrylamide gel, subcloned into Bluescript and sequenced. All of the cDNA inserts isolated from the random primed library, as well as the largest clone from the dT primed library, matched the sequence of Tcg1 through the ninth base of the sixth repeat unit, with two notable exceptions. Each cDNA lacked two adjacent bases in the sixth repeat, a adenine and a thymidine at positions 269 and 270 respectively, found in Tcg1. One other consistent difference between the sequence of Tcg1 and that of the cDNA's, as well as that of the 490 bp EcoRI/SmaI, is the presence of a cytosine at position seven in Tcg1. This cytosine is lacking in all other fragments sequenced and is believed to be an artifact of the EcoRI linkers used in the construction of the genomic

library from which Tcg1 was isolated.

None of the nucleotide sequences of the cDNA inserts match the nucleotide sequence of the 1.7 kbp EcoRI fragment. The nucleotide sequence of Tcg1 and the 490 bp EcoRI/SmaI fragment are identical through the third repeat unit after which no homology is observed. Thus, only cDNA inserts whose sequence extends 3' downstream of the third repeat unit can be used to distinguish between Tcg1 or the 4.8 kbp fragment as the putative transcriptional template. The nucleotide sequence of ten cDNAs extends beyond the third repeat and in each case matches the sequences found in Tcg1 but not that of the 4.8 kbp EcoRI fragment. Although this result does not rule out a low level of transcription of either the 1.7 kbp or the 4.8 kbp EcoRI fragment, it does suggest that the most abundant transcript originates from a site other than of the 1.7 kbp or 4.8 kbp fragments. Based upon the results of the restriction enzyme mapping, and the Bal 31 and DNase I digestion experiments, we suggest that the transcription site for these cDNAs is the 5.4 kbp telomeric fragment.

This publication received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.