INFECTION AND IMMUNITY, July 1989, p. 1959–1967
0019-9567/89/071959-09$02.00/0

Vol. 57, No. 7

Trypanosoma cruzi Expresses Diverse Repetitive Protein Antigens

DANIEL F. HOFT,[1,2] KWANG S. KIM,[1,3] KEIKO OTSU,[1] DAVID R. MOSER,[4] W. JOHN YOST,[1] JOEL H. BLUMIN,[5] JOHN E. DONELSON,[3,4] AND LOUIS V. KIRCHHOFF[1,5]*

Departments of Internal Medicine,[1] Microbiology,[2] and Biochemistry,[3] and Diabetes and Endocrinology Research Center,[4] University of Iowa, and Veterans Administration Medical Center,[5] Iowa City, Iowa 52242*

Received 29 November 1988/Accepted 25 March 1989

We screened a *Trypanosoma cruzi* cDNA expression library with human and rabbit anti-*T. cruzi* sera and identified cDNA clones that encode polypeptides containing tandemly arranged repeats which are 6 to 34 amino acids in length. The peptide repeats encoded by these cDNAs varied markedly in sequence, copy number, and location relative to the polyadenylation site of the mRNAs from which they were derived. The repeats were specific for *T. cruzi*, but in each case the sizes of the corresponding mRNAs and the total number of repeat copies encoded varied considerably among different isolates of the parasite. Expression of the peptide repeats was not stage specific. One of the peptide repeats occurred in a protein with an M_r of >200,000 and one was in a protein of M_r 75,000 to 105,000. The frequent occurrence and diversity of these peptide repeats suggested that they may play a role in the ability of the parasite to evade immune destruction in its invertebrate and mammalian hosts, but the primary roles of these macromolecules may be unrelated to the host-parasite relationship.

Trypanosoma cruzi, the protozoan agent of American trypanosomiasis or Chagas' disease, has a complex life cycle in which it is continually exposed to immune attack in mammalian and insect hosts (4, 19). In mammals, *T. cruzi* multiplies as intracellular amastigotes that undergo transformation to the trypomastigote form prior to rupture of parasitized cells. When released, the invasive trypomastigotes enter adjacent host cells or spread hematogenously to distant tissues, in which the cycle of transformation and multiplication is repeated. Hematophagous reduviid bugs ingest circulating trypomastigotes with blood meals, which in turn differentiate into epimastigotes (Epi) that multiply in the midgut. Subsequently, in the hindgut Epi differentiate into metacyclic trypomastigotes that can infect mammalian hosts. When insect excreta containing these infective metacyclic forms come into contact with mucosal surfaces, conjunctivas, or breaks in the skin of humans or other mammals, the parasites gain entry into host cells, thus completing the cycle.

The long-term survival of *T. cruzi* depends on the establishment of infections in individual hosts of both types that persist for long periods without causing death. The permanence of *T. cruzi* in insect and mammalian hosts requires continual evasion of immune destruction, but the nature of the parasite antigens involved in this process and the host immune responses they elicit are not well understood. To provide basic information about antigens that may be involved in this complex host-parasite relationship, we used sera from several individuals chronically infected with *T. cruzi*, as well as that of a rabbit immunized with a monoclonal antibody-purified *T. cruzi* surface antigen, to isolate cDNAs from an expression library constructed by using RNAs from culture-derived metacyclic trypomastigotes (CMT). This approach yielded several dozen cDNAs, and initial DNA sequence analyses identified eight that contain regions encoding tandemly arranged peptide repeats. These repeats, which apparently are not found in closely related protozoans, are strikingly heterogeneous when viewed from several perspectives. In addition, diversity in the genes that encode the repeated peptides occurs among different *T. cruzi* isolates. The frequent occurrence and diversity of regions encoding antigenic peptide repeats suggest that they may contribute to evasion of immune destruction, but other roles unrelated to the host-parasite relationship may be of primary importance.

* Corresponding author.

MATERIALS AND METHODS

Preparation of parasites. Epi of the Corpus Christi (42), Tulahuén (T. Pizzi, Ph.D. thesis, University of Chile, Santiago, 1957), and Y (34) strains of *T. cruzi* and Sylvio X-10/4 clone (35) were maintained in logarithmic growth phase at 26°C in supplemented liver digest neutralized medium as described previously (17). Mixtures of Epi and CMT (~1:1) of the Sylvio X-10/4 clone were grown in Grace's insect medium, and purified CMT (>90%) were obtained by passage of the mixture through a DE52 column as described before (L. V. Kirchhoff and D. F. Hoft, submitted for publication). Procyclic culture forms of the IaTat1.2 clone of *T. brucei brucei* (25) were grown in BSM medium supplemented with 3 mM *cis*-aconitate and 1 mM pyruvate at room temperature (6). Promastigotes of the Morton strain of *Leishmania mexicana mexicana* (F. A. Neva, unpublished results) were maintained in logarithmic growth phase at 26°C in medium 199 supplemented with 30% fetal calf serum and 20 μg of hemin per ml.

Construction of the cDNA expression library. RNA was purified from Sylvio X-10/4 CMT (5) and cDNAs were synthesized from total RNA, without prior isolation of poly(A)$^+$ RNA, with Moloney murine leukemia virus reverse transcriptase in the BRL Synthesis System, using the protocol suggested by the manufacturer (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) (3). Following treatment with *Eco*RI methylase and attachment of *Eco*RI linkers, the cDNAs were ligated into bacteriophage λZAP (Stratagene, San Diego, Calif.) (33). After packaging of the recombinant phage (GigaPack Gold; Stratagene), a library of 6.4×10^6 independent clones was obtained, and 5×10^6 of these were amplified in *Escherichia coli* Y1090 (36).

Antibody probes and immunoscreening. Sera from Bolivian

and Salvadoran patients with xenodiagnosis-confirmed chronic *T. cruzi* infections (16, 18) were used to identify cDNAs encoding antigenic *T. cruzi* fusion proteins. In addition, a rabbit serum, designated Rα72, containing antibodies directed against an iodinatable surface protein of *T. cruzi* having an approximate M_r of 72,000 was used for this purpose. This serum was produced by hyperimmunizing a rabbit with *T. cruzi* antigen that had been affinity purified with monoclonal antibody 2D2, an antibody that binds to an iodinatable surface protein of the parasite having an M_r of 62,000. The reason for this difference in specificity is not clear. The amplified cDNA library was immunoscreened with these sera as described previously (12), using species-specific, horseradish peroxidase-conjugated goat anti-immunoglobulin G as secondary antibody.

Characterization of cDNAs. Recombinant pBluescript plasmids were recovered from purified λZAP clones identified by the anti-*T. cruzi* sera by coinfecting *E. coli* XL1-Blue with the recombinant phage and R408 helper phage (33). Nucleotide (nt) sequences of cloned cDNAs were determined by the dideoxy method (31), using the sequenase kit (U.S. Biochemicals, Cleveland, Ohio), or by the chemical cleavage procedure (22) on end-labeled fragments generated by cleavage at internal restriction sites.

Northern and Southern blots. Total RNA and genomic DNA were isolated from parasites as described previously (5, 20), and blots of purified nucleic acids were performed by standard procedures (21, 37). DNA samples were depurinated (0.25 M HCl) and denatured (1.5 M NaCl, 0.5 M NaOH) after electrophoresis. Purified cDNAs for probing blots were labeled with ^{32}P by the random oligonucleotide primer method and *E. coli* DNA polymerase I (10). The entire cDNAs were used to probe blotted nucleic acids, except in the case of the TCR27 cDNA, from which the region between the 5′ end and an *Xba*I site was isolated and labeled (see Fig. 2). Hybridization conditions for the Southern blots were 20× SET (20× SET is 3 M NaCl, 0.4 M Tris chloride [pH 7.8], and 20 mM EDTA) and 65°C.

Antibody select procedure and Western blots (immunoblots). For isolation of antibodies directed against specific fusion proteins, cloned recombinant phage were plated with *E. coli* Y1090 on agar (11). Fusion proteins were bound to isopropyl-β-D-thiogalactopyranoside-saturated nitrocellulose disks, and after blocking the disks were incubated overnight in dilute solutions of the serum originally used to identify the reactive phage. Bound antibody was then obtained by acid elution and neutralized prior to use. Antibodies selected similarly with an ovalbumin fusion protein were used as controls. To determine the specificity of the selected antibody, parasites were solubilized by boiling in 5% sodium dodecyl sulfate followed by treatment with 50 mM dithiothreitol–200 mM iodoacetamide as described before (41). Alternatively, the organisms were lysed in 2% Nonidet P-40 at room temperature. Parasite lysates were then electrophoresed on 5 to 15% sodium dodecyl sulfate-polyacrylamide gradient gels and transferred to nitrocellulose sheets as reported (39). Nitrocellulose strips (4 mm) containing the separated parasite antigens were then blocked with 20% fetal calf serum in phosphate-buffered saline and incubated in solutions containing the selected antibodies. Bound antibodies were detected by addition of species-specific, horseradish peroxidase-conjugated goat anti-immunoglobulin G and substrate.

RESULTS

Identification and characterization of *T. cruzi* cDNAs encoding tandemly repeated peptides. Human and rabbit sera having anti-*T. cruzi* antibodies were used to immunoscreen approximately 750,000 λZAP phage. Fifty-nine strongly reactive phage were identified, and after recovery of recombinant pBluescript plasmids, DNA sequences of 200 to 300 nt at both ends of each cDNA were determined. Eight of the cDNAs were found to contain regions encoding tandemly repeated peptides and further nucleotide sequencing was performed on all cDNAs in this group. A comparison of the consensus nucleotide and inferred amino acid repeat sequences of these eight cloned cDNAs is presented in Fig. 1, in which the number of amino acids in each repeat is given in parentheses. There is considerable diversity in both size and amino acid sequence among the repeats, which range from 6 to 34 amino acids in length. Several similarities also exist, as indicated by the solid squares and circles. TCR69 and TCR70 are identical, and five of the six amino acids of TCR61 are present in the seven-amino-acid sequence of TCR69. In addition, the peptide repeats encoded by TCR3 and TCR27 are both 14 amino acids long, and a 60% identity at the nucleotide level results in the same amino acid at 6 of 14, or 43%, of the positions.

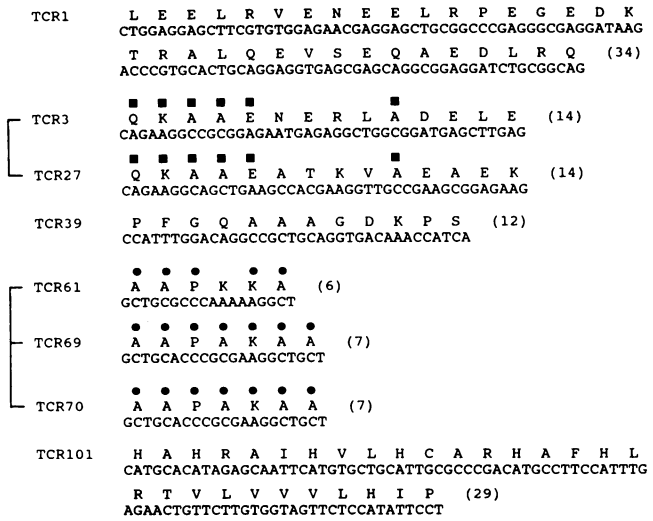

FIG. 1. Consensus nucleotide and inferred amino acid sequences of the tandemly arranged repeats identified in cDNAs derived from *T. cruzi* RNA. The number of amino acids in each repeat is given in parentheses. Solid squares and circles denote residues that are identical in the repeats grouped by the brackets.

Four of the eight cDNAs containing repetitive sequences were selected for further characterization and are presented schematically in Fig. 2. TCR3, which along with TCR1 was identified with the Rα72 serum, is a 1,195-nt insert that contains a single 777-nt open reading frame. Located at the 5′ end of this coding sequence are six and one-half copies of a 42-nt repeated sequence, followed by a 495-nt region that encodes a nonrepetitive amino acid sequence unrelated to the peptide product of the repeats. The nontranslated region downstream from the stop codon does not end in a poly(A) sequence.

The other three cDNAs depicted in Fig. 2 were identified with sera from patients with chronic *T. cruzi* infections. TCR27, a 1,660-nt insert with a single open reading frame of approximately 1,230 nt, also contains a region of 42-nt repeats, but as noted, the repeat sequence is different from that of TCR3. The DNA sequences of five or more repeats at both the 5′ and the 3′ ends of the repeated region were determined. The presence of approximately 13 additional

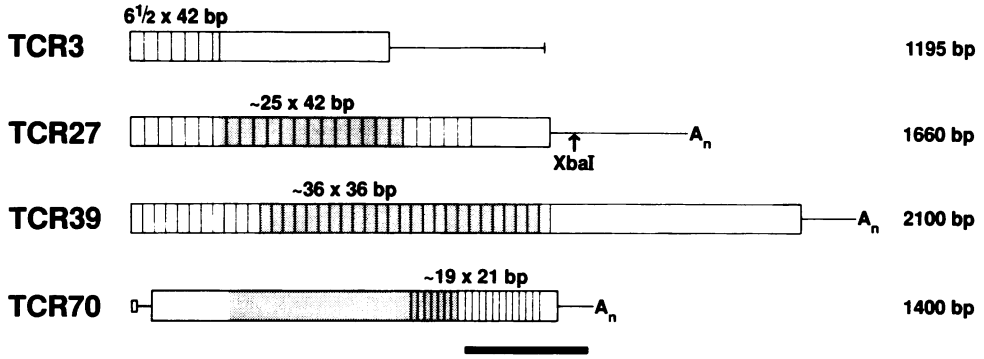

FIG. 2. Schematic representations of *T. cruzi* cDNAs identified with rabbit and human antisera. Horizontal lines indicate nontranslated nucleotide sequences. Regions encoding proteins are enclosed in large rectangular boxes. Vertical lines within these boxes delineate the positions and sizes of the units encoding repetitive peptides. The numbers above the regions of the repeats indicate the numbers and lengths of the repeats. Stippling denotes regions for which restriction maps were deduced but nucleotide sequences were not determined. Repeats in these regions were deduced by restriction endonuclease mapping. *Xba*I in the TCR27 cDNA indicates a restriction site for this enzyme, and the region to the left of this site was used as a probe in Fig. 4B. The small box at the 5′ end of the TCR70 cDNA indicates the presence of the downstream portion of the spliced leader sequence. The horizontal bar below the TCR70 cDNA denotes the location of the TCR69 sequence, which differs from the corresponding region of TCR70 only in the position of the polyadenylation site. bp, Base pairs.

repeats in the center of this region, indicated by the stippled area, was deduced from the size of the 610-base-pair DNA fragment remaining after digestion of the cDNA insert with *Pvu*II, an endonuclease that cleaves within the repeats. The open reading frame continues downstream from the repeats for 180 nt, and the nontranslated region ends in a poly(A) tail. The third cDNA is TCR39, a 2,100-nt insert that also has repeats at its 5′ end as well as a poly(A) tail. The repeats are 36 nt in length, however, and approximately 36 tandem copies are present as deduced by digestion with *Pst*I, which cleaves once in each repeat. The nucleotide sequence of the repeats and that of the downstream coding region into which the open reading frame continues are unrelated to each other.

TCR70, the fourth cDNA shown in Fig. 2, is distinct in that it is a full-length insert, as indicated by the presence, at its 5′ end, of the downstream 18 nt of the 39-nt spliced leader present on all trypanosome mRNAs (26, 40). The total length of TCR70 is 1,400 nt, and partial sequence analysis of its 5′ terminus indicates that it contains a continuous open reading frame in the same translation reading frame as the β-galactosidase gene fragment in the λZAP vector. The downstream region of the open reading frame contains an area consisting of 19 repeats which are 21 nt in length, followed by a short nonrepetitive region that encodes only five additional amino acids. The sequences of 12 of these repeats were determined, and the presence of the additional 7 repeats was deduced from the repetitive nature of faint bands on the autoradiogram of the sequencing gel. After a total of 19 repeats, the sequence no longer appeared repetitive. TCR69, a 409-nt cDNA that was completely sequenced, is identical to the segment of TCR70 indicated by the bar in Fig. 2, except that the location of its polyadenylation site differs by 5 nt.

The amino acid sequences encoded by the four cDNAs depicted in Fig. 2 are shown in Fig. 3. In general, the amino acid sequences of the repeats are highly conserved, with only occasional substitutions. However, in the last one or two repeats at the carboxyl ends of the repeated regions of each of the peptides, there is considerable degeneracy as the similarity to the consensus repeat is lost. The peptide sequences downstream from the regions of the repeats are highly heterogeneous and contain no putative metal and nucleic acid binding structures found in tail peptides of some other proteins containing repeats (2, 17).

The total number of repeats in the transcripts from which seven of the eight cloned cDNAs were derived cannot be determined from the nucleotide sequences alone because their 5′ termini are located in the repetitive regions. The exception is TCR70, which contains the entire open reading frame of the message that encodes 19 copies of the seven-amino-acid peptide repeat, as determined by partial sequence analysis. An additional feature of the sequences of the cloned cDNAs is that they apparently are all unique to *T. cruzi*. We addressed this question by probing *Rsa*I-digested DNA from *L. mexicana mexicana* and from *T. brucei brucei*, two protozoans closely related to *T. cruzi*, with the cloned cDNAs in Southern blots, and no significant hybridization was observed by using the washing conditions described above (not shown). Similarly, when the nucleotide and protein sequences of these cDNAs were compared with those in the Genbank and Swiss-Pro data bases, no close matches were found.

Analysis of *T. cruzi* genes and transcripts encoding peptide repeats. We used Northern (RNA) and Southern analyses to study in four distinct *T. cruzi* isolates the patterns of transcription and occurrence of genes related to the four cloned cDNAs depicted in Fig. 2. The results of these experiments are shown in Fig. 4, in which Northern analyses are presented in block 1 of each panel and Southern analyses are depicted in block 2. The autoradiographic patterns obtained when electrophoretically separated samples of total RNA from the Corpus Christi, Tulahuén, and Y strains, and from the Sylvio X-10/4 clone, were probed with TCR3 are shown in block 1 of panel A. In each of the four isolates, TCR3 hybridizes to two transcripts, ranging in size from 6.5 to 9 kilobase pairs (kb), but there is considerable variability in the sizes of the two transcripts from one isolate to another. These transcripts have the potential for encoding very large proteins. For example, the 7.0- and 7.8-kb transcripts in the Sylvio X-10/4 lane could encode proteins with M_rs as high as 255,000 and 285,000, respectively.

This intraspecies variability in the sizes of the TCR3 transcripts raises the question of whether they are the primary transcription products of similarly diverse genes or the result of posttranscriptional processing of a less hetero-

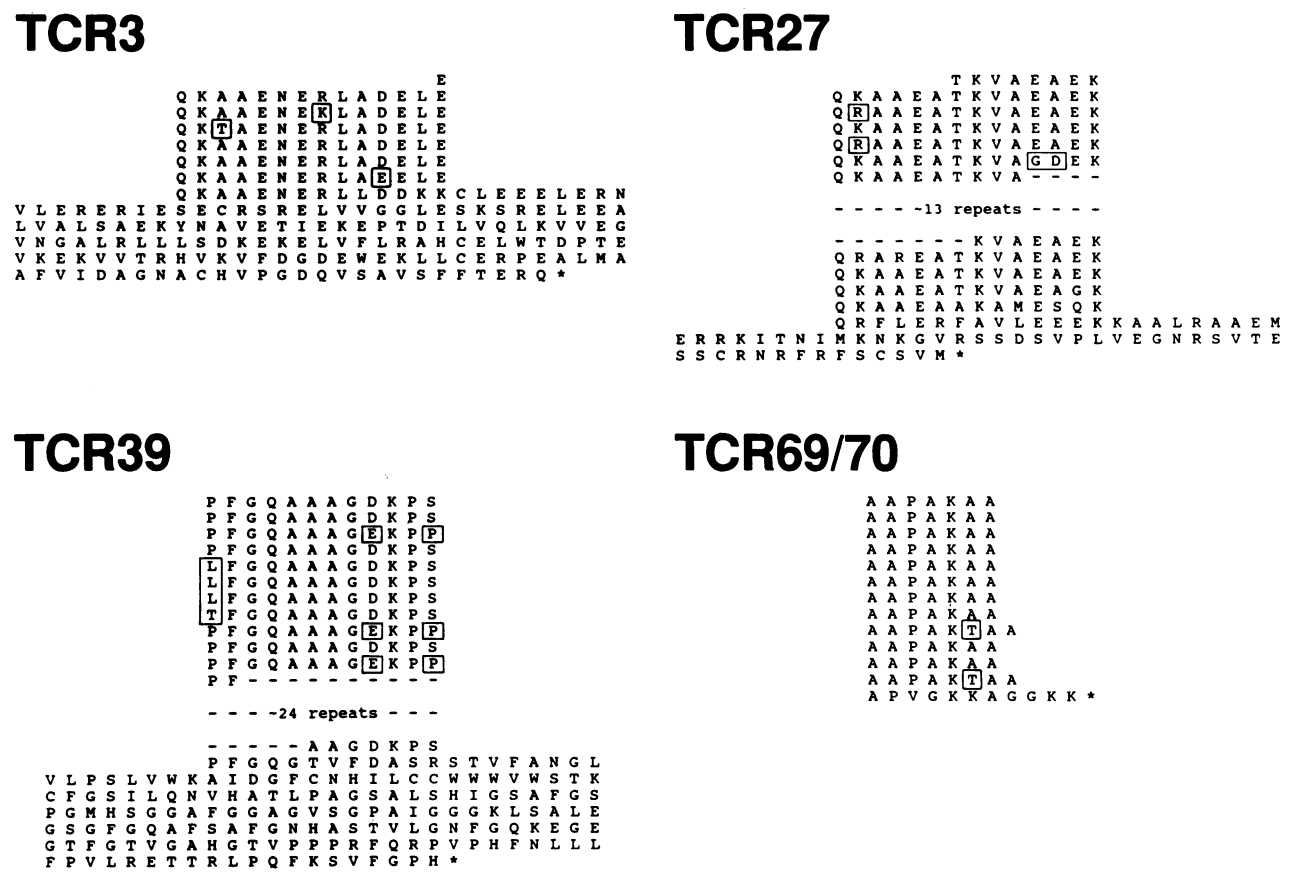

FIG. 3. Amino acid sequences containing repetitive peptides encoded by four cDNAs isolated from a *T. cruzi* expression library. The peptide repeats are aligned, and point variations from the consensus repeat sequences are indicated by the boxed residues. Nonrepetitive peptides downstream from the repeats are located in the lower portions of the sequences. C termini are indicated by asterisks.

geneous group of primary transcripts. To address this question, we digested DNA samples from the four *T. cruzi* isolates with *Rsa*I and, after electrophoresis and transfer to nitrocellulose, probed these digests with the four cDNAs depicted in Fig. 2. *Rsa*I does not cleave within the consensus repeat sequences of these cDNAs. It recognizes a 4-nt sequence (GTAC) that should occur on a random basis every 256 nt within the regions flanking the repeats. Thus, in the absence of point mutations that generate *Rsa*I sites in the repeats, this enzyme should excise intact repetitive regions of the genes from which the cDNAs were derived, flanked by the upstream and downstream regions between the repeats and the first *Rsa*I sites. In this context, the genes containing the repeats must be at least as large as the bands on the radiographs of the Southern blots probed with the cDNAs. If a repeat unit contains an *Rsa*I site generated by a point mutation, the segment of the gene consisting of the repeated units would be cleaved to yield two bands, and the parent gene would have to be larger than the smaller of the two fragments.

The results obtained when a Southern blot of *Rsa*I-digested DNA samples from the four *T. cruzi* isolates was probed with TCR3 are shown in block 2 of panel A. In a lane-by-lane comparison, there is a general correspondence between the numbers and approximate sizes of the TCR3 genes and their transcripts. This suggests that the intraspecies heterogeneity of the TCR3 transcripts results from genetic diversity rather than from posttranscriptional processing. This correlation also indicates that the size differences of the gene fragments and their transcripts observed among the four *T. cruzi* isolates results from differing numbers of the 42-nt TCR3 repeat unit. For example, the 6.3- and 7.3-kb bands present in lane S of block 2 suggest that the TCR3 cDNA was derived from a gene containing approximately either 150 or 170 repeat units. Of additional interest is the observation that the smaller of the two gene fragments in Corpus Christi strain and Sylvio X-10/4 clone (lanes C and S) are clearly smaller than the corresponding transcripts seen in the Northern blot. Most likely this results from the fact that the bands on the Southern blot reflect only the approximate sizes of the regions containing the repeats, rather than the entire genes.

The autoradiographic patterns obtained when similarly processed RNA and DNA samples were probed with TCR27 are shown in panel B. Again, each isolate has two transcripts of different sizes and two genes to which the cDNA hybridizes, but in each isolate the sizes are dissimilar from those to which the TCR3 probe hybridized. The TCR27 genes apparently encode transcripts containing far fewer repeat units than do the TCR3 genes, since the 3.0- and 4.5-kb bands in lane S of block 2 could contain no more than 75 and 105 copies of the 42-nt tandem repeats. Beyond these size differences, however, the general organization of the genes and transcripts related to the TCR3 and TCR27 cDNAs is quite similar.

The results obtained when comparable blots were probed with TCR39, the cDNA with repeats 36 nt in length, are presented in panel C of Fig. 4. The genes and transcripts to

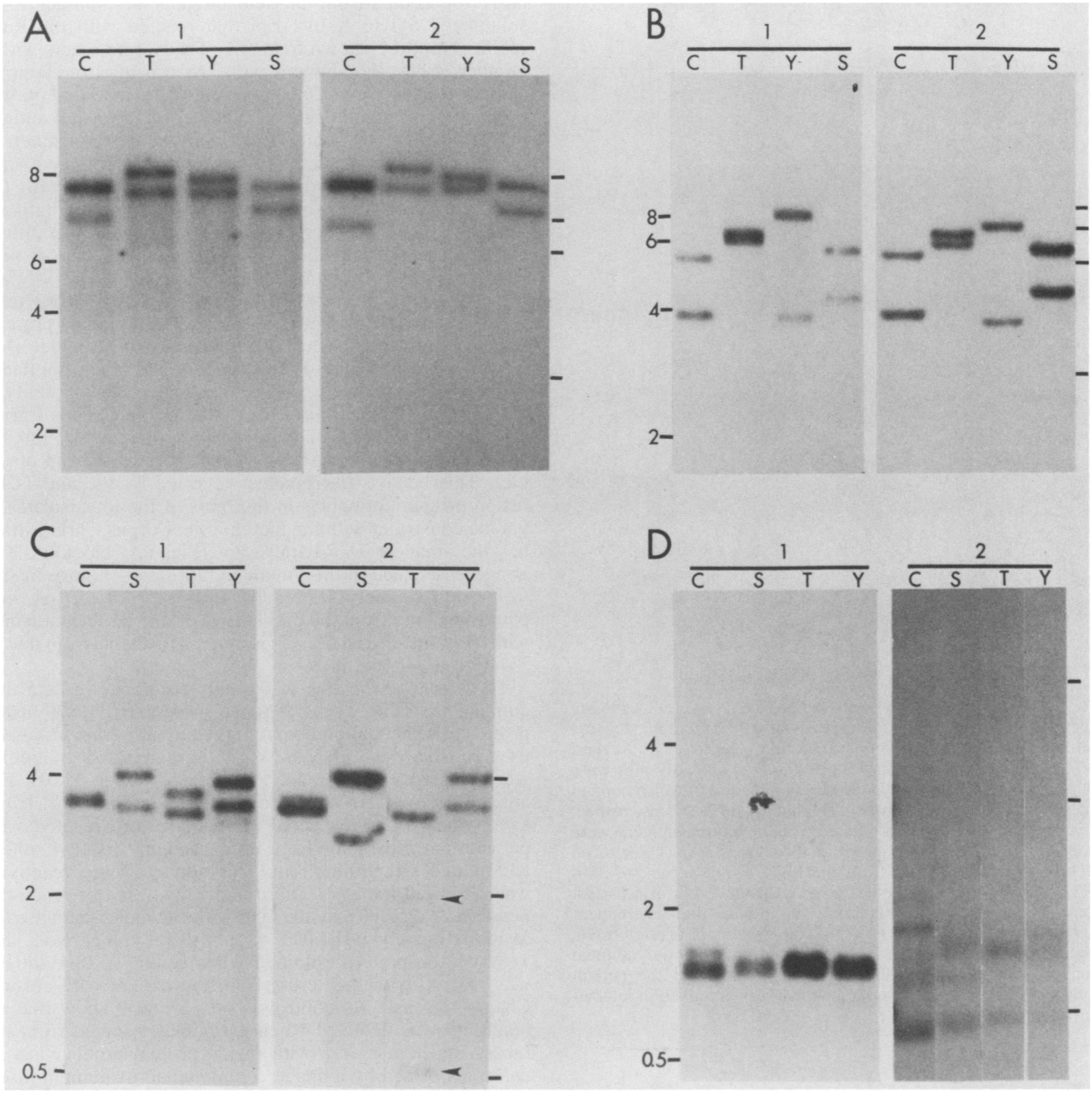

FIG. 4. Autoradiographs of Northern and Southern blots of electrophoretically separated total RNA (10 μg per lane) and *Rsa*I-digested DNA (4.5 μg per lane) from four *T. cruzi* isolates probed with radiolabeled *T. cruzi* cDNAs containing repetitive sequences. The cDNAs used as probes were as follows: (A) TCR3; (B) TCR27 (*Xba*I fragment); (C) TCR39; (D) TCR69. Northern analyses are depicted in block 1 of each panel, and Southern analyses are presented in block 2. The letter above each lane indicates the *T. cruzi* isolate used as a source of RNA or DNA: C, Corpus Christi strain; T, Tulahuén strain; Y, Y strain; and S, Sylvio X-10/4 clone. The hatchmarks to the right of block 2 indicate the same size markers in kilobases shown on the left of block 1.

which TCR39 hybridized are considerably smaller than those identified with TCR3 and TCR27, but the general patterns of correspondence of transcript and gene sizes are again similar. In the case of the Corpus Christi strain (lane C), only one transcript can be seen on the Northern blot, while two gene fragments can be distinguished on the Southern blot. The two gene fragments are quite close in size, however, indicating that the single band on the Northern blot may consist of two transcript species. In the Tulahuén strain (lane T), there are two transcripts and three gene fragments. The darker 3.0-kb fragment corresponds to the larger of the two transcripts, which is 3.5 kb in size, and the two faint fragments indicated by arrowheads may be the cleavage products of the repeat region of a second gene in which a point mutation has created an *Rsa*I site. These two fragments together constitute 2.5 kb and thus may correspond to the 3.0-kb transcript. The 3.0- and 4.0-kb gene fragments in the Sylvio X-10/4 clone (lane S) could encode transcripts containing approximately 83 and 111 repeats, and the 3.6- and 4.2-kb transcripts present in the Sylvio X-10/4 lane could

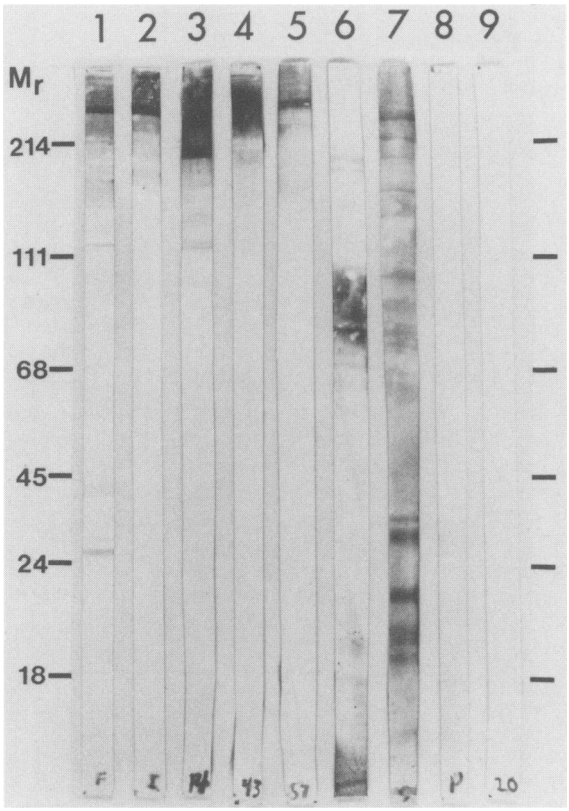

FIG. 5. Photograph of Western blots of *T. cruzi* epimastigote antigens, solubilized in 2% Nonidet P-40 or 5% sodium dodecyl sulfate prior to electrophoretic separation, reacted with several anti-*T. cruzi* antibody reagents. Blotted Sylvio X-10/4 lysates were used as target antigens in all lanes except 3, 4, and 5, which contain antigens from the Corpus Christi, Tulahuén, and Y strains, respectively. The antibody reagents reacted with the blotted antigens were as follows: 1, Rα72 rabbit serum; 2 through 5, anti-TCR3 fusion protein antibodies selected from the Rα72 rabbit serum; 6, anti-TCR39 fusion protein antibodies selected from serum of a patient chronically infected with *T. cruzi*; 7, serum from the same patient chronically infected with *T. cruzi*; 8, anti-ovalbumin fusion protein antibodies selected from the Rα72 rabbit serum; 9, anti-ovalbumin fusion protein antibodies selected from serum from the patient chronically infected with *T. cruzi*. M_r markers are given in kilodaltons.

encode proteins of approximate M_rs of 130,000 and 255,000, respectively. Panel D contains the results obtained when similar blots were probed with TCR70, the cDNA having repeat units 21 nt long. These results are striking in that the transcripts are only approximately 1.4 kb in size, and except for the Corpus Christi strain (lane C), there may be only one in each isolate. Likewise, the pattern of gene fragments is quite different from those obtained with the other three cDNA probes. There appear to be several restriction fragments in each isolate that hybridize to the cDNA, all of which are <2 kb.

Identification of parasite proteins containing repeated peptides. We attempted to react fusion protein-selected antibodies with parasite antigens in Western blots to determine the sizes of the native proteins containing the peptide repeats. Using this approach, we were able to establish the sizes of the repeat-containing proteins encoded by two of the four cDNAs chosen for extensive characterization (Fig. 5). As a basis for comparison, blotted antigens of Nonidet P-40-solubilized Sylvio X-10/4 Epi that reacted with the Rα72 rabbit serum are shown in lane 1. This serum reacts with a number of parasite antigens with M_rs ranging from approximately 30,000 to >200,000, despite the fact that it was produced by immunizing a rabbit with parasite antigen purified with a monoclonal antibody directed against an M_r-62,000 protein. Lane 2 shows the pattern obtained when antibodies selected from the Rα72 rabbit serum with the TCR3 fusion protein were reacted with the blotted antigens from the same polyacrylamide gel. Considerable diffuse reactivity is present above M_r 214,000, with one distinct band at approximately M_r 250,000. Proteins of this size are consistent with the sizes of the Sylvio X-10/4 TCR3 transcripts seen in block 1 of Fig. 4A. The TCR3 fusion protein-selected antibodies were also reacted with Nonidet P-40 lysates of Corpus Christi, Tulahuén, and Y strain Epi (lanes 3, 4, and 5, respectively). There is considerable variability in the patterns obtained, as exemplified in the Corpus Christi blot (lane 3) by the dark band at approximately M_r 180,000 and a broad area of reactivity from there to the top of the blot. This band is the smallest revealed by the anti-TCR3 fusion protein antibodies in the blots of the four isolates, a finding consistent with the fact that the Corpus Christi strain has the smallest TCR3 transcript (Fig. 4A, block 1). The antigens revealed in the Tulahuén blot (lane 4) range in size from approximately M_r 214,000 to the top of the gel, in a pattern distinct from the previous two. The pattern obtained with Y-strain antigen (lane 5) is indistinguishable from that of Sylvio X-10/4 (lane 2).

To determine the size of the native parasite protein that contains the TCR39 peptide repeats, we used TCR39 fusion protein to select antibodies from the human serum originally used to isolate the cDNA. When these selected antibodies were reacted with Nonidet P-40-treated Sylvio X-10/4 Epi antigens in Western blots, no binding was observed. However, when TCR39-selected antibodies were reacted with parasite antigens solubilized in 5% sodium dodecyl sulfate and treated with dithiothreitol and iodoacetamide, reactivity was observed (lane 6). There is a broad area of reactivity between M_r 75,000 and 105,000, with the most distinct band at approximately M_r 82,000. As a basis for comparison, lane 7 shows the pattern obtained when Sylvio X-10/4 antigen was reacted with the serum from the patient with chronic Chagas' disease. As controls, lanes 8 and 9 show that no reactivity was detected when antibodies selected from the rabbit and human sera with fusion protein produced in *E. coli* transformed by a phage containing an ovalbumin insert were used as probes.

DISCUSSION

By identifying cDNAs derived from *T. cruzi* metacyclic trypomastigote RNAs with anti-*T. cruzi* antibodies and determining their nucleotide sequences, we found that this protozoan parasite contains genes that encode a heterogeneous group of tandemly repeated antigenic peptides. The peptide repeats encoded by TCR69 and TCR70 are the same, and those of TCR3 and TCR27 are identical in 6 of 14 positions. In addition, five of six amino acids of the TCR61 repeat are present in sequence in the TCR70 repeat, although the latter is seven amino acids in length. Beyond these instances of relatedness, however, the peptide repeats encoded by the cDNAs are diverse in sequence, size, and copy number. Moreover, genes encoding these repetitive peptides are not found in closely related protozoans such as *Leish-*

mania and African trypanosomes. Expression of the peptide repeats apparently is not stage specific. The genes encoding the repeated peptides are transcribed in the Epi (insect) stage of the parasite, since this stage was the source of the RNA used for the Northern analyses. Transcription apparently also occurs in the metacyclic (infective) form, as CMT RNA was used to construct the expression library from which the repeat-containing cDNAs were cloned. In addition, the repeated peptides must be expressed in at least one of the two stages found in the mammalian host, since humans infected with *T. cruzi* have antibodies directed against them. Genes encoding each of the peptide repeats are present in all four isolates of *T. cruzi* examined, but marked intraspecies heterogeneity exists in the sizes of mRNAs encoding them and in the total number of copies encoded.

To our knowledge there have been three other reports dealing with tandemly repeated antigenic peptides in *T. cruzi*. The first of these describes the molecular cloning of a gene fragment related to an M_r-85,000 mammalian stage antigen that encodes five tandem copies of a nonapeptide repeat (27). More recently, the presence in the HSP70 gene of *T. cruzi* of a repetitive nucleotide sequence encoding 12 tandem copies of a tetrapeptide repeat was reported (29). None of the peptide repeats we describe in this report is similar to either of these two sequences. In the third report, Ibañez and co-workers describe the cloning of nine *T. cruzi* gene fragments that encode antigenic proteins, seven of which contain repetitive regions (13). Two (TCR27 and TCR39) of the eight cDNAs we identified encode peptide repeats identical to repeats described in the latter study. It is of interest to note that in both cases even the consensus nucleotide repeat sequences we found are identical to those identified by Ibañez et al. This is surprising considering the different origins of the *T. cruzi* isolates used to construct the two expression libraries from which the sequences were derived and in view of currently held concepts of heterogeneity in *T. cruzi*. The isolate we used as a source of RNA for the construction of our expression library (Sylvio X-10/4 clone) was originally obtained from a *T. cruzi*-infected individual in the Amazon region of Brazil, whereas that used as a source of DNA by the latter group (CA1 strain) came from an Argentinian patient with Chagas' disease. Examination of a sizable number of isolates of *T. cruzi* from a variety of biological and biochemical perspectives has shown that it is a highly heterogeneous species (8, 9, 15, 24, 28, 38). In this context, then, it is remarkable that the consensus nucleotide sequences of these two repeats are identical in the Sylvio X-10/4 and CA1 isolates. This finding suggests that *T. cruzi* does not tolerate even a minimal degree of diversity in the sequence of these two peptides, although, as indicated by the intraspecies variability in TCR27 and TCR39 transcript sizes (Fig. 4B and C, block 1), the number of copies of the peptide repeats can vary considerably from one *T. cruzi* isolate to another.

The findings we report here extend our knowledge of repeated peptide antigens in *T. cruzi* and lend additional support to the concept that the expression of a heterogeneous group of peptide repeats is a generalized phenomenon among protozoan parasites that persistently evade immune destruction in their hosts. The presence of genes encoding peptide repeats also has been reported in the African trypanosome *T. brucei*. Roditi and co-workers (30) described a polypeptide containing 22 tandem copies of a dipeptide repeat that is expressed only in the insect stage of *T. brucei* and therefore cannot be implicated in the interaction between the mammalian host and the parasite. More recently, Schneider et al. (32) reported the molecular cloning of a gene segment that encodes a microtubule-associated protein of *T. brucei* that contains >50 repeats of a 38-amino-acid peptide. The presence of genes encoding peptide repeats also has been reported in *L. major* (41), a human pathogen belonging to the group of organisms having a life cycle very similar to that of *T. cruzi*. In the latter report, the cloning of cDNAs that encode repeats containing 10 and 14 amino acids is described. These repeats are unrelated in sequence to those identified in our study, and no interspecies variability in the sizes of transcripts encoding these repeats was observed when two *Leishmania* species were compared, which is surprising considering the similarity of the organisms and our finding of striking intraspecies heterogeneity of repeat-encoding transcript size.

Genes encoding tandemly repeated peptides have been most extensively characterized in malarial parasites of the genus *Plasmodium* (14). In these organisms, the majority of antigens studied contain peptide repeats. The repeats, which are diverse in size, amino acid sequence, and copy number, have been found in all parasite stages, and stage-specific expression has been observed frequently. In addition, genes have been described that encode two or more blocks of tandemly repeated peptides having distinct amino acid sequences separated by unrelated and nonrepetitive peptides. It has been hypothesized that cross-reactivities among and within proteins containing repeats may cause an overstimulation of B cells, thus interfering with the normal maturation of an effective humoral response and permitting long-term evasion of immune clearance by the parasite (1, 14). Our findings suggest that the diversity of antigens containing peptide repeats in *T. cruzi* may parallel that observed in malarial parasites, and thus the ability of this trypanosome to establish lifelong infections in its hosts may also result from the expression of repeated peptides. However, there is no experimental evidence that directly supports this concept in *Plasmodium* or *T. cruzi*, and the repeated peptides may be involved in fundamental functions other than the diversion of the host immune response. Recent reports indicate that antibodies raised against synthetic peptides containing only the repeats of the circumsporozoite protein of *Plasmodium falciparum* block invasion of hepatoma cells (43) and hepatocytes (23) by sporozoites in vitro, and when used to immunize mice, they provide significant but not total protection against sporozoite challenge (7). These findings argue against a purely diversionary role for this repetitive sequence. In addition, our finding that the repeated peptides described in this report are expressed in both insect and mammalian stages of the parasite raises the question of whether they could have similar diversionary roles in hosts having such different immune systems (19). The recent report by Schneider et al. (32) describing a *T. brucei* membrane skeleton protein with a large region of repeats implicates a role other than immune evasion. Thus, parasite proteins containing repeats may have discrete physiological roles and participate in the host-parasite relationship in ways that are not clearly understood. Further characterization of the structural and functional aspects of these proteins will be necessary to resolve this issue.

ACKNOWLEDGMENTS

We thank D. M. Engman and L. V. Reddy for advice on experimental methods and helpful discussions.

This research was supported by Public Health Service grants AI24711, AI18954, and DK25295 from the National Institutes of Health, the Veterans Administration Research Service, a Syntex

Scholars Award to L.V.K., and the Molecular Parasitology Award from the Burroughs-Wellcome Foundation to J.E.D. D.F.H. is supported by Physician Scientist Program Award DK01295 from the National Institutes of Health, and W.J.Y. is a recipient of a University of Iowa College of Medicine Student Research Fellowship.

LITERATURE CITED

1. **Anders, R. F.** 1986. Multiple cross-reactivities amongst antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria. Parasite Immunol. **8**:529–539.
2. **Berg, J. M.** 1986. More metal-binding fingers. Nature (London) **319**:264–265.
3. **Bethesda Research Laboratories.** 1986. cDNA synthesis system instruction manual (catalog no. 8267SA). Bethesda Research Laboratories, Inc., Gaithersburg, Md.
4. **Brener, Z.** 1973. Biology of *Trypanosoma cruzi*. Annu. Rev. Microbiol. **27**:347–382.
5. **Chirgwin, J. M., A. E. Prybuyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
6. **Cunningham, I.** 1977. A new culture medium for maintenance of tsetse tissue and growth of trypanosomatids. J. Protozool. **24**:325–329.
7. **Egan, J. E., J. L. Weber, W. R. Ballou, M. R. Hollingdale, W. R. Majarian, D. M. Gordon, W. L. Maloy, S. L. Hoffman, R. A. Wirtz, I. Schneider, G. R. Woollett, J. F. Young, and W. T. Hockmeyer.** 1987. Efficacy of murine malaria sporozoite vaccines: implications for human vaccine development. Science **236**:453–456.
8. **Engel, J. C., J. A. Dvorak, E. L. Segura, and M. J. Crane.** 1982. *Trypanosoma cruzi*: biological characterization of 19 clones derived from two chronic Chagasic patients. I. Growth kinetics in liquid medium. J. Protozool. **29**:555–560.
9. **Engman, D. M., V. J. Reddy, J. E. Donelson, and L. V. Kirchhoff.** 1987. *Trypanosoma cruzi* exhibits inter- and intra-strain heterogeneity in molecular karyotype and chromosomal gene location. Mol. Biochem. Parasitol. **22**:115–123.
10. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
11. **Hall, R., J. E. Hyde, M. Goman, D. L. Simmons, I. A. Hope, M. MacKay, J. Scaife, B. Merkli, R. Richle, and J. Stocker.** 1984. Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria. Nature (London) **311**:379–382.
12. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Constructing and screening cDNA libraries in λgt10 and λgt11, p. 49–78. *In* D. M. Glover (ed.), DNA cloning techniques: a practical approach. IRL Press, Oxford.
13. **Ibañez, C. F., J. L. Affranchinio, R. A. Macina, M. B. Reyes, S. Leguizamon, M. E. Camargo, L. Aslund, U. Pettersson, and A. C. C. Frasch.** 1988. Multiple *Trypanosoma cruzi* antigens containing tandemly repeated amino acid sequence motifs. Mol. Biochem. Parasitol. **30**:27–34.
14. **Kemp, D. J., R. L. Coppel, H. D. Stahl, A. E. Bianco, L. M. Corcoran, P. McIntyre, C. J. Langford, J. M. Favaloro, P. E. Crewther, G. V. Brown, G. F. Mitchell, J. G. Culvenor, and R. F. Anders.** 1986. Genes for antigens of *Plasmodium falciparum*. Parasitology **91**(Suppl.):83–108.
15. **Kirchhoff, L. V., J. C. Engel, J. A. Dvorak, and A. Sher.** 1984. Strains and clones of *Trypanosoma cruzi* differ in their expression of a surface antigen identified by a monoclonal antibody. Mol. Biochem. Parasitol. **11**:81–89.
16. **Kirchhoff, L. V., A. A. Gam, and F. C. Gilliam.** 1987. American trypanosomiasis (Chagas' disease) in Central American immigrants. Am. J. Med. **82**:915–920.
17. **Kirchhoff, L. V., K. S. Kim, D. M. Engman, and J. E. Donelson.** 1988. Ubiquitin genes in Trypanosomatidae. J. Biol. Chem. **263**:12698–12704.
18. **Kirchhoff, L. V., and F. A. Neva.** 1985. Chagas' disease in Latin American immigrants. J. Am. Med. Assoc. **254**:3058–3060.
19. **Lackie, A. M.** 1988. Immune mechanisms in insects. Parasitol. Today **4**:98–105.
20. **Laurent, M., S. Van Assel, and M. Steinert.** 1971. Kinetoplast DNA. A unique macromolecular structure of considerable size and mechanical resistance. Biochem. Biophys. Res. Commun. **43**:278–284.
21. **Lehrach, H., D. Diamond, J. Wozney, and H. Boedtker.** 1977. RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry **16**:4743–4751.
22. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
23. **Mazier, D., S. Mellouk, R. L. Beaudoin, B. Texier, P. Druilhe, W. T. Hockmeyer, J. Trosper, C. Paul, Y. Charoenvit, J. Young, F. Miltgen, L. Chedid, J. P. Chigot, B. Galley, O. Brandicourt, and M. Gentilini.** 1986. Effect of antibodies to recombinant and synthetic peptides on *P. falciparum* sporozoites *in vitro*. Science **231**:156–159.
24. **Morel, C., E. Chiari, E. Plessmann, D. M. Mattei, A. J. Romanha, and L. Simpson.** 1980. Strains and clones of *Trypanosoma cruzi* can be characterized by pattern of restriction endonuclease products of kinetoplast DNA minicircles. Proc. Natl. Acad. Sci. USA **77**:6810–6814.
25. **Murphy, W. J., S. T. Brentano, A. C. Rice-Ficht, D. M. Dorfman, and J. E. Donelson.** 1984. DNA rearrangements of the variable surface antigen genes of the trypanosomes. J. Protozool. **31**:65–73.
26. **Perry, K. L., K. P. Watkins, and N. Agabian.** 1987. Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. Proc. Natl. Acad. Sci. USA **84**:8190–8194.
27. **Peterson, D. S., R. A. Wrightsman, and J. E. Manning.** 1986. Cloning of a major surface-antigen gene of *Trypanosoma cruzi* and identification of a nonapeptide repeat. Nature (London) **322**:566–568.
28. **Postan, M., J. A. Dvorak, and J. P. McDaniel.** 1983. Studies of *Trypanosoma cruzi* clones in inbred mice. Am. J. Trop. Med. Hyg. **32**:497–506.
29. **Requena, R. M., M. C. López, A. Jimenez-Ruiz, J. C. de la Torre, and C. Alonso.** 1988. A head-to-tail tandem organization of hsp70 genes in *Trypanosoma cruzi*. Nucleic Acids Res. **16**:1393–1406.
30. **Roditi, I., M. Carrington, and M. Turner.** 1987. Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei*. Nature (London) **325**: 272–274.
31. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
32. **Schneider, A., A. Hemphill, T. Wyler, and T. Seebeck.** 1988. Large microtubule-associated protein of *T. brucei* has tandemly repeated, near-identical sequences. Science **241**:459–462.
33. **Short, J. M., J. M. Fernandez, J. A. Songe, and W. D. Huse.** 1988. λ ZAP: a bacteriophage λ expression vector with *in vivo* excision properties. Nucleic Acids Res. **16**:7583–7600.
34. **Silva, L. H. P., and V. Nussenzweig.** 1953. Sobre uma cepa de *Trypanosoma cruzi* altamente virulenta para o camundongo branco. (On a strain of *Trypanosoma cruzi* highly virulent for white mice.) Folia Clin. Biol. **20**:197–207.
35. **Silveira, F. T., M. G. Dias, P. P. Pardal, A. de Oliveira Lobao, and G. de Britto Melo.** 1979. Nono caso autóctone de doença de Chagas registrado no estado do Pará, Brasil (Nota Prévia). Hileia Med. Belem **1**:61–62.
36. **Stratagene.** 1987. Predigested lambda Zap/EcoRI cloning kit instruction manual. Stratagene, San Diego, Calif.
37. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
38. **Tibayrenc, M., P. Ward, A. Moya, and F. J. Ayala.** 1986. Natural populations of *Trypanosoma cruzi*, the agent of Chagas' disease, have a complex multiclonal structure. Proc. Natl. Acad. Sci. USA **83**:115–119.

39. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
40. **Walder, J. A., P. S. Eder, D. M. Engman, S. T. Brentano, R. Y. Walder, D. S. Knutzon, D. M. Dorfman, and J. E. Donelson.** 1986. The 35-nucleotide spliced leader sequence is shared by all trypanosome messenger RNA's. Science **233**:569–571.
41. **Wallis, A. E., and W. R. McMaster.** 1987. Identification of *Leishmania* genes encoding proteins containing tandemly repeating peptides. J. Exp. Med. **166**:1814–1824.
42. **Woody, N. C., and H. B. Woody.** 1955. American trypanosomiasis (Chagas' disease): first indigenous case in the United States. J. Am. Med. Assoc. **159**:676–677.
43. **Young, J. F., W. T. Hockmeyer, M. Gross, W. R. Ballou, R. A. Wirtz, J. H. Trosper, R. L. Beaudoin, M. R. Hollingdale, L. H. Miller, C. L. Diggs, and M. Rosenberg.** 1985. Expression of *Plasmidium falciparum* circumsporozoite proteins in *Escherichia coli* for potential use in a human malaria vaccine. Science **228**:958–962.