Title
Evidence for two distinct major protein components, PAR 1 and PAR 2, in the paraflagellar rod of Trypanosoma cruzi. Complete nucleotide sequence of PAR.

Journal
Journal of Biological Chemistry, 267(30)

ISSN
0021-9258

Authors
Beard, CA
Saborio, JL
Tewari, D
et al.

Publication Date
1992-10-01

DOI
10.1016/s0021-9258(19)36662-1

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 267, No. 30, Issue of October 25, pp. 21656-21662, 1992
Printed in U.S.A.

Evidence for Two Distinct Major Protein Components, PAR 1 and PAR 2, in the Paraflagellar Rod of *Trypanosoma cruzi*

COMPLETE NUCLEOTIDE SEQUENCE OF PAR 2*

(Received for publication, May 14, 1992)

Chris A. Beard, Jose L. Saborio, Devansu Tewari, Kerstin G. Krieglstein, Agnes H. Henschen, and Jerry E. Manning‡

From the Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92717

The previously identified major protein components of the paraflagellar rod in *Trypanosoma cruzi*, PAR 1 and PAR 2, were analyzed to determine if they are distinct proteins or different conformations of a single polypeptide as has been suggested for other trypanosomatids. Amino acid sequence analysis showed PAR 1 and PAR 2 to be two distinct polypeptides. Antibodies specific against either PAR 1 or PAR 2 were shown to each react with a distinct band in Western blots of paraflagellar isolates of *T. cruzi* and other trypanosomatids if rigorous protease inhibition was used. The PAR 2 message was isolated and characterized by Northern blot and nucleic acid sequence analysis. Preliminary analysis of the PAR 2 gene indicates that PAR 2 is a member of a multigene family with all members residing on a single chromosome.

Trypanosoma cruzi, a parasitic hemoflagellate, is the causative agent of American trypanosomiasis or Chagas' disease (1). This disease is a major public health problem in Central and South America and, to date, no effective chemotherapeutic agent or immunoprophylaxis has been identified. One promising line of investigation centers on the identification and characterization of cellular processes or structures that are unique to the parasite. Therapeutic agents that target these structures would hopefully function with minimal interaction with host cells.

The paraflagellar rod, a major component of the parasite flagellum, is such a unique structure. It is a complex lattice of filaments with ultrastructural characteristics unrelated to any of the major filamentous systems of the host cells, including microfilaments, microtubules, or intermediate filaments (2).

Schlaeppi *et al.* (3), working with *Trypanosoma brucei*, reported that the major protein component of the paraflagellar rod (PFR)[1] is a single polypeptide of 600 amino acids that gives two bands in PAGE due to different conformations. They also characterized the gene coding for the PFR protein and determined that there were two identical copies in the genome.

Previously, we reported the isolation of two major immunologically distinct protein components of the paraflagellar rod in *T. cruzi*, the proteins PAR 1 and PAR 2 (4). These proteins show no immunological cross-reactivity with actin, tubulin, intermediate filament proteins, or other proteins present in mammalian cells. Here, we confirm that PAR 1 and PAR 2 are distinct proteins by amino acid sequence analysis. We also report the isolation and characterization of the PAR 2 message and that the PAR 2 gene is a member of a multigene family of at least 30 members.

MATERIALS AND METHODS

Parasites—*T. cruzi* Esmeraldo clone 3 strain was obtained from James Dvorak, National Institutes of Health, Bethesda, MD. Growth and maintenance of epimastigotes and tissue culture derived trypomastigotes are as described elsewhere (5). *Leishmania brasiliensis* promastigotes (6) and *T. brucei* procyclics (7) were obtained as described.

Purification of T. cruzi Paraflagellar Proteins—Crude flagellar pellets from *T. cruzi* epimastigotes were prepared as described previously (4), except that all solutions employed contained either leupeptin (100 μg/ml), antipain (50 μg/ml), E-64 (10^{-5} M), or a mixture with these three inhibitors. In brief, about 4×10^{10} epimastigotes were harvested by centrifugation, washed twice with 0.02 M sodium phosphate, pH 7.4, 0.9% sodium chloride (phosphate-buffered saline), and lysed in 20 ml of 1% Nonidet P-40 in 0.1 M Tricine, pH 8.5. Insoluble material was collected by centrifugation at 2,000 × *g* and was extracted once more with the same buffer. The pellet was then suspended in 10 ml of 1.0 M NaCl, 0.1% Triton X-100, in 0.1 M Tricine, pH 8.5. The DNA released at this step was sheared by stirring the suspension at maximum speed for 30 s with a tissue grinder (Tissue-Tearon, Biospec Products, Bartlesville, OK). The suspension was centrifuged at 12,000 × *g*, and the resulting pellet corresponds to the crude flagellar fraction.

Crude flagellar fractions were successively extracted with 2.0 and 6.0 M urea in 10 mM Tricine, pH 8.5. As reported previously about 80% of the tubulin and 20% of the paraflagellar proteins were solubilized in 2.0 M urea, and the remaining paraflagellar proteins were solubilized in 6.0 M urea (4). The latter material was applied to a Mono-Q column equilibrated with the same buffer, and bound protein was eluted with a 0–500 mM NaCl gradient.

Polyacrylamide Gel Electrophoresis (PAGE) and Western Blot Analysis—For analysis of whole cell lysates, parasites were harvested from culture media by centrifugation, washed twice with phosphate-buffered saline, and solubilized by direct addition to boiling 2% SDS

* This work was supported by United States Public Health Service Grant AI 18873 from the National Institutes of Health and by the United Nations Development Program/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M97548.

‡ To whom correspondence should be addressed. Tel.: 714-856-7102; Fax: 714-856-8551.

[1] The abbreviations used are: PFR, paraflagellar rod; PAGE, polyacrylamide gel electrophoresis; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; HPLC, high performance liquid chromatography; PFGE, pulsed field gel electrophoresis; kb, kilobase.

with boiling continued for 5 min. These and all other samples were adjusted to the composition of the electrophoresis sample buffer (62.5 mM Tris, pH 6.8, 10% glycerol, 5% β-mercaptoethanol, 2% sodium dodecyl sulfate, 0.001% bromphenol blue) prior to analysis. One-dimensional PAGE in 0.75-mm slab gels was done according to Laemmli (8). Prestained and ^{14}C-labeled molecular weight markers (Amersham Corp.) were included in the gels. Gels were either stained with Coomassie Brilliant Blue R (9) or processed for Western blot analysis (5) as previously described. Western blots were probed with a polyclonal antibody to PAR 1 (pcAbPAR 1) (4) or with a monoclonal antibody to PAR 2 (mAbPAR 2) (4).

Amino Acid Sequencing of Paraflagellar Proteins—Selected fractions from the Mono-Q column were concentrated in Centricon tubes to a concentration of approximately 1.0 mg of protein/ml. 10 μl of β-mercaptoethanol were added to 190 μl of the protein solution and the mixture was incubated under Freon for 5 h at 50 °C. The protein solution was then made 7.5% with 4-vinylpyridine and incubation continued at room temperature for 1 h in the dark, followed by overnight dialysis against water and lyophilization.

The *S*-alkylated protein was dissolved in 100 μl of 75% formic acid and 100 μl of a freshly prepared CNBr solution (20% w/v in 75% formic acid) was added. The mixture was incubated in the dark for 2 h at room temperature, diluted with 1.0 ml of water, and concentrated by evaporation with a stream of nitrogen.

Cyanogen bromide fragments of the alkylated proteins were subjected to HPLC (LKB, Bromma, Sweden) on a reversed-phase column (Nucleosil C_{18} 300–10, column size 4 × 250 mm. Macherey-Nagel, Düren, Federal Republic of Germany) (10). The solvent system consisted of 0.1% (by volume) trifluoroacetic acid in water (A) and in acetonitrile (B). The proportion of B was increased from 0 to 50% within 110 min. The flow rate was 1.0 ml/min. The peptide fragments were monitored at 206 nm, and the peaks were collected manually.

For amino acid sequencing of cyanogen bromide fragments, the Edman degradation method was carried out in a pulsed-liquid-phase sequencer (Applied Biosystems model 477A, Foster City, CA). The phenylthiohydantoin derivatives were identified by the on-line HPLC system in which also the derivative of *S*-pyridylethyl cysteine is separated (11).

Nucleic Acid Isolation, Radiolabeling, Southern and Northern Transfer, and Restriction Enzymes—Parasites were harvested and DNA, RNA, and poly(A)$^+$ mRNA were isolated as described previously (12). Plasmid DNA was isolated by alkaline lysis miniprep as described (13). λ-Phage DNA was prepared as described (14). DNA restriction fragments were radiolabeled with [α-^{32}P]dNTP using the BRL Nick Translation Kit as recommended by the manufacturer (GIBCO BRL). Agarose gel electrophoresis of DNA, Southern transfer, prehybridization, hybridization, and filter washing were performed as described (15) except the gels were 1% agarose and the wash temperature was 68 °C. RNA was electrophoresed in a formaldehyde gel, blotted to nitrocellulose, baked, prehybridized, hybridized, and washed as described (16). All restriction enzymes were purchased from Boehringer Mannheim and used as recommended.

cDNA Library Construction and Screening—cDNA libraries were constructed in phage λgt10 or λgt11 using epimastigote poly(A)$^+$ mRNA as described (17) or by using a cDNA synthesis system according to the manufacturer's instructions (Pharmacia LKB Biotechnology Inc.). The λ-cDNA libraries were plated and transferred to nitrocellulose as described (13). Filter hybridizations were carried out as described (15) except that the wash temperature was 68 °C. The λgt10 cDNA library was screened using radiolabeled TccPar2a. This fragment was isolated from a cDNA expression library in λgt11 probed with a monoclonal antibody against the PAR 2 polypeptide (4, 17). The inserts present in phages showing positive hybridization were excised, subcloned into Bluescript KS+ (Stratagene Inc., La Jolla, CA), and characterized by restriction enzyme mapping and direct nucleotide sequence analysis.

DNA Sequencing—DNA sequence information was obtained by use of the dideoxy chain-termination method (18). Fragments to be sequenced were either subcloned into Bluescript KS+ or sequenced from the original λgt10 or λgt11 phage. Oligonucleotide sequencing primers were synthesized in the Gene Assembler Plus (Pharmacia) according to the manufacturer's instructions.

Pulsed Field Gel Electrophoresis (PFGE)—Epimastigotes were prepared and lysed in agarose blocks at a concentration of 2×10^9 cells/ml as described (19). PFGE was carried out in a CHEF-DR II system supplied by Bio-Rad. A single agarose block containing epimastigotes as described above was loaded in each well of a 1% agarose gel submerged in 0.5 X TBE (90 mM Tris base, 90 mM boric acid, 2.5 mM EDTA, pH 8.0). Chromosome separation was performed over a period of 24–48 h at 14 °C using 200 volts with a switch time of 60 and 90 s. *Saccharomyces cerevisiae* chromosomes were used as molecular weight markers. Gels were stained with ethidium bromide for 15 min and destained for 20 min before transfer of the DNA to nitrocellulose as described above.

RESULTS

Purification of Paraflagellar Proteins—A protein preparation enriched in paraflagellar proteins was obtained in the insoluble residue after successive extractions of *T. cruzi* epimastigotes with buffer solutions containing 1% Nonidet P-40, 1.0 M NaCl, and 2.0 M urea, as described under "Materials and Methods." Extraction of that insoluble residue with 6 M urea produced a soluble fraction in which approximately 50% of the total amount of protein corresponded to tubulin, and the other 50% to equimolar amounts of PAR 1 and PAR 2 (Fig. 1, *inset, lane c*).

Previously, we reported that PAR 1 and PAR 2 have slightly different isoelectric points and molecular weights and that no immunological cross-reactivity between these two polypeptides could be demonstrated (4). These observations suggested that PAR 1 and PAR 2 correspond to two different paraflagellar rod components. In an attempt to obtain additional support for this contention through separation and further characterization of these two polypeptides, the protein fraction soluble in 6 M urea was fractionated by ion-exchange chromatography on a Mono-Q column. Fig. 1 shows a typical absorbance profile and the electrophoretic pattern of some of the fractions obtained from that column. These results indicate that the paraflagellar proteins are readily separated from tubulin and that fractions containing mixtures of approximately 80% PAR 2 and 20% PAR 1 (Fig. 1, *inset, fraction 21*) and 60% PAR 1 and 40% PAR 2 (Fig. 1, *inset, fractions 25 and 26*) are obtained by this chromatographic procedure.

Amino Acid Sequencing of Paraflagellar Proteins—The fractions from the Mono-Q column enriched in either PAR 2

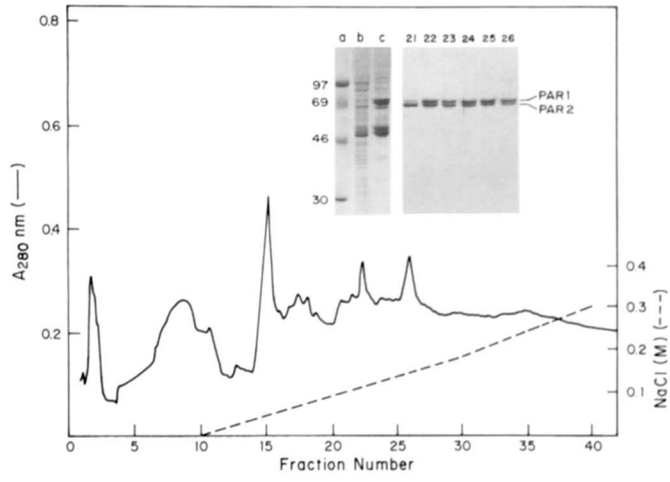

FIG. 1. **Chromatographic fractionation of paraflagellar proteins from *T. cruzi* epimastigotes.** A protein preparation solubilized with 6 M urea from crude flagellar pellets was applied to a Mono-Q column equilibrated with 6 M urea in 10 mM Tricine, pH 8.5. Protein bound to the column was eluted with a 0–500 mM NaCl gradient in the same buffer. Aliquots of some of the fractions from the column were analyzed by PAGE, and the different polypeptides were revealed by Coomassie Blue staining. The different lanes in the *inset* correspond to: *a*, molecular weight markers; *b*, epimastigote whole cell lysate; *c* 6 M urea extract from crude flagellar pellets; *21–26*, fractions from the Mono-Q column. The molecular weight markers and their corresponding molecular weights in kilodaltons are: phosphorylase *b*, 97.4; bovine serum albumin, 69; ovalbumin, 46; and carbonic anhydrase, 30.

(fraction 21) or PAR 1 (fractions 25 and 26) were subjected to mercaptolysis, alkylation, and cleavage with CNBr as described under "Material and Methods." The resulting peptides were fractionated by HPLC on a reversed-phase column.

As shown in Fig. 2, *panel A*, the HPLC profile from the preparation enriched for PAR 2 (fraction 21 from the Mono-Q column) contains 10–12 major peaks. In contrast, a more complex HPLC profile, with 18–20 major peaks, is obtained from the protein preparation containing 60% PAR 1 and 40% Par 2 (*panel B*). The positions of some of the peaks in *panel B* coincide with the positions of peaks in *panel A*, but several distinct peaks present in *panel B* are absent in *panel A*. These observations are consistent with the view that 1) PAR 1 and PAR 2 are two distinct polypeptides, 2) the major peaks observed in *panel A* represent peptide fragments derived from PAR 2, and 3) the peptides present in *panel B* but absent in *panel A* are derived from PAR 1.

The amino acid sequence of several CNBr peptides presumably derived from PAR 2 (Fig. 2, *panel A*) was determined by direct amino acid sequence analysis. Unambiguous sequences were obtained from the peaks labeled with *numbers 4, 7, 9, 11, 13, 14,* and *16* in Fig. 2, *panel A*. These sequences, which correspond to eight different peptides and a total of 217 amino acids, are shown in Fig. 3 under *PAR 2*. As will be shown in a different section of this paper, all of these sequences could be accounted for in the complete amino acid sequence of PAR 2 deduced from the nucleotide sequence of the PAR 2 gene, thus confirming the supposition that the major peptides in Fig. 2, *panel A*, are derived from PAR 2.

Amino acid sequence analysis of several peaks from the profile shown in Fig. 2, *panel B*, indicates that some of the peaks in *panel B*, selected on the basis of coincidental positions with peaks corresponding to PAR 2 polypeptides in *panel A*, indeed contain amino acid sequences corresponding

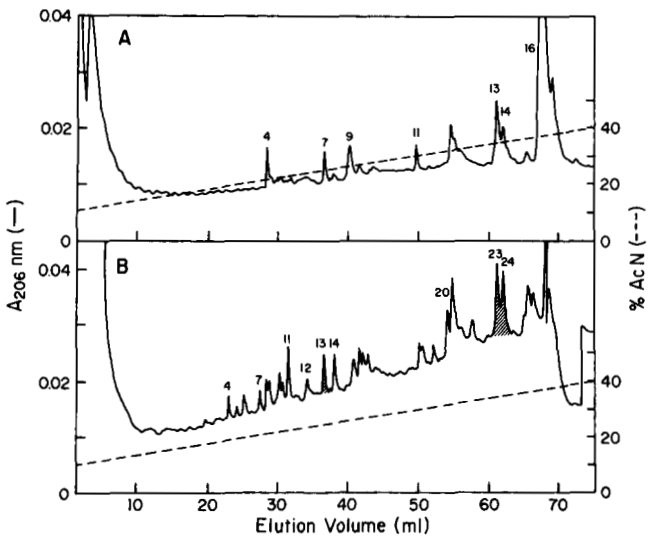

FIG. 2. **Chromatographic separation of CNBr peptides from paraflagellar proteins.** The protein in *fraction 21* and the protein in a pool of *fractions 25* and *26* from the Mono-Q column (Fig. 1) was recovered and subjected to mercaptolysis, alkylation, and CNBr cleavage as described under "Materials and Methods." The CNBr peptides were separated by reversed-phase on a Nucleosil C_{18} 300–10 (4 × 250-mm) column. The solvent system consisted of 0.1% trifluoroacetic acid in water (*A*) or in acetonitrile (*B*). The proportion of solvent B was increased from 0 to 50% within 110 min. The peaks were collected manually. *Panel A* corresponds to the CNBr peptides from PAR 2, and *panel B* to the peptides from the mixture of PAR 1 and PAR 2. Material from the *numbered peaks* was used for amino acid sequence analysis. The *shaded areas* in *panel B* are peptides identified by amino acid sequence as corresponding to PAR 2 peptides shown in *panel A*.

PAR 2

Peak #	Amino acid sequence
4	(M)VEYRAHLAKQEEVKIAAEREELKRSKTLQSQQYRGKTVQQIT (557-600)
7	(M)RVCGLQLSVRELYKPEDKP (100-119)
9	(M)AQVPVAVLKNLEE (174-187)
	(M)FGPTEDALNQAGIEFVHPAEEVEDGNLT (524-552)
11	(M)NVTVVQTALLGNEEQIKAQLAAIEKAKEIRNVAIADG (189-226)
13	(M)KRFATQKEKSEKFIQENLDRQDEAWRRIQ (316-345)
14	(M)EVVALKKTLNELKQHHNKTRTVSFTGTI (122-150)
16	(M)AIAEEQYYIKAQLLEHLVELVAD (228-251)

PAR 1 & PAR 2

Peak #	Amino acid sequence
4	(M)HYVENEERKVLEKRNVL
7	(M)VRLDTLERQARLLLRNNR
11	(M)RDAVEEL
	(M)LQYKRREKQTTSDLKNIP
12	(M)EELTADLRSY(Y/C)DEES
13	(M)PQQKHRC
	PAR 2, PEPTIDE 7 (100-119)
14	(M)EEIDRNISTTEIQLPFARSTKN
20	(M)EALRDAADSISRFAE
23	PAR 2, PEPTIDE 13 (316-345)
24	PAR 2, PEPTIDE 14 (122-150)

FIG. 3. **Amino acid sequences of CNBr peptides from paraflagellar proteins.** The amino acid sequences of all the *peaks labeled with numbers* in Fig. 2 were obtained by the Edman degradation method in a pulsed-liquid-phase sequencer. The sequences shown in this figure under *PAR 2* correspond to peaks depicted in *panel A*, Fig. 2, and those shown under *PAR 1 & PAR 2*, to *peaks* in *panel B* of the same figure. The methionines in *parenthesis* were not determined directly but are assumed to be the CNBr cleavage sites. The *numbers* in *parenthesis* after the sequences correspond to the location of those sequences in the complete sequence of 600 amino acids deduced from the nucleotide sequence of the PAR 2 gene.

to PAR 2. Thus, the amino acid sequence of peptide 7 in *panel A* was found in *peak 13* in *panel B*, while the sequences of peptides 13 and 14 in *panel A* were found in *peaks 23* and *24*, respectively, in *panel B*. The positions of these PAR 2 peptides are identified as *shaded areas* in Fig. 2, *panel B*. Unambiguous amino acid sequences were obtained from the peaks labeled with *numbers 4, 7, 11–14, 20* and *24* in Fig. 2, *panel B*. These sequences, which correspond to eight different peptides and a total of 128 amino acids, are shown in Fig. 3 under *PAR 1 & PAR 2*. None of the latter sequences correspond to PAR 2 sequences, whether determined directly or deduced from the nucleotide sequence of the PAR 2 gene. Since PAR 1 is the only observable protein other that PAR 2 in fractions 25 and 26, it is very likely that these sequences represent portions of PAR 1.

Western Blot Analysis of T. brucei, T. cruzi, and L. brasiliensis Extracts—The above results, which directly show that PAR 1 and PAR 2 are two distinct polypeptides, agree with our previous observations that indicate lack of immunological cross-reactivity between these two major components of the paraflagellar rod of *T. cruzi* (4). Others have shown that monoclonal antibodies to the paraflagellar rod of *T. brucei* (20) or *L. brasiliensis* (21) react with the two major paraflagellar polypeptides detected in each of those parasites. In *T. brucei* these results have been interpreted either as indicative of common epitopes in two different polypeptides (20) or as due to the existence of a single polypeptide with two conformational variants that exhibit slightly different electrophoretic mobilities (3). As previously reported, antibodies to *T. cruzi* PAR 1 or PAR 2 each react with a single polypeptide in extracts of that parasite (4). These latter results, however, critically depend on efficient prevention of proteolytic activity during extract preparation. If proteolysis is not completely inhibited, antibodies to PAR 1 or PAR 2 can each react with several proteolytic fragments of the corresponding polypeptides (4). In an attempt to determine whether antibodies to *T. cruzi* PAR 1 and PAR 2 each react with a single polypeptide in lysates of *T. brucei* or *L. brasiliensis*, extracts of these organisms were prepared by direct solubilization of pelleted parasites in boiling 2% SDS solution, in our hands the most

efficient procedure for immediate and efficient inactivation of proteolytic activity. These extracts, together with a *T. cruzi* extract included as a control, were fractionated by one-dimensional PAGE and processed for Western blot analysis. The results shown in Fig. 4 indicate that pcAbPAR 1 reacts with a single polypeptide of molecular mass about 70 kDa, while mAbPAR 2 reacts with a single polypeptide of molecular mass about 68 kDa in the extracts of *T. brucei* (*lanes b* and *b'*), *T. cruzi* (*lanes c* and *c'*), and *L. brasiliensis* (*lanes d* and *d'*).

Isolation of the PAR 2 Gene—To isolate a DNA fragment that encodes a portion of the PAR 2 gene, a recombinant cDNA library of epimastigote poly(A)$^+$ RNA sequences was constructed in the expression vecor λgt11 (22). Approximately 180,000 recombinant phage were screened with monoclonal antibody mAbPAR2 (4) and 11 positive plaques were identified, of which one rescreened positive. Restriction enzyme mapping analysis of the cDNA insert contained in this phage revealed a 1.1-kb insert. The cDNA insert in this phage, TccPar2a, was excised by digestion with *Eco*RI and subcloned into the plasmid vector Bluescript KS+.

In order to find a mature transcript from the PAR 2 gene (≅2.0 kb, see below) a size-selected (≥1.0 kb) Esmeraldo epimastigote cDNA library in λgt10 was screened with [^{32}P] TccPar2a. Approximately 30 positive plaques were identified from a total of 150,000 screened. Half of these rescreened positive and two that contained inserts of ≅2.0 kb were chosen for further study. The cDNA inserts of these phage, TccPar2b and TccPar2c, were excised by digestion with *Eco*RI and subcloned into the plasmid vector Bluescript KS+ for further study and sequence analysis.

Nucleotide Sequence Analysis—To confirm that the chosen cDNAs code for the PAR 2 polypeptide the sequences of TccPar2a, 2b, and 2c were determined by the dideoxy chain-termination method. Both complementary strands of the putative PAR 2 cDNAs were sequenced in the Bluescript KS+ plasmid vector using incremental oligonucleotide primers. The nucleotide sequence, with the amino acid translation, is shown in Fig. 5. The portions of the deduced amino acid sequence that have been verified by direct amino acid sequence analysis of PAR 2 are also indicated. The presence of nucleotide sequence coding for all PAR 2 peptides determined by direct amino acid sequence analysis indicates that these cDNAs do indeed code for PAR 2.

Attempts to determine the putative NH_2-terminal sequence of PAR 2 were unsuccessful, possibly because the protein is naturally blocked at the NH_2 terminus or was blocked during purification (*e.g.* carbamylation by urea). Although direct amino acid sequence analysis of the NH_2 terminus of PAR 2 could not be obtained, we believe that the coding region begins

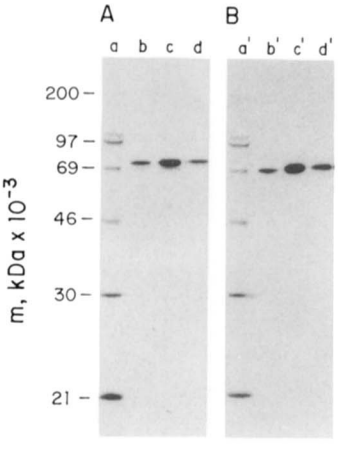

FIG. 4. **Western blot analysis of *T. brucei*, *T. cruzi*, and *L. brasiliensis* lysates.** *T. brucei* procyclics, *T. cruzi* epimastigotes, and *L. brasiliensis* promastigotes were harvested from culture media by centrifugation, washed twice with phosphate-buffered saline, and the pelleted parasites directly solubilized in boiling 2% SDS solution. Aliquots of these extracts, containing 10 μg of protein, were fractionated by one-dimensional PAGE. After electrophoresis and blotting, nitrocellulose sheets were probed with pcAbPAR 1 (*A*) or mAbPAR 2 (*B*). The different lanes correspond to: *a*, molecular weight markers; *b* and *b'*, *T. brucei*; *c* and *c'*, *T. cruzi*, and *d* and *d'*, *L. brasiliensis* lysates. The molecular weight markers and their corresponding molecular weights in kilodaltons are: myosin, 200; phosphorylase *b*, 97.4; bovine serum albumin, 69; ovalbumin, 46; carbonic anhydrase, 30; and trypsin inhibitor, 21.5.

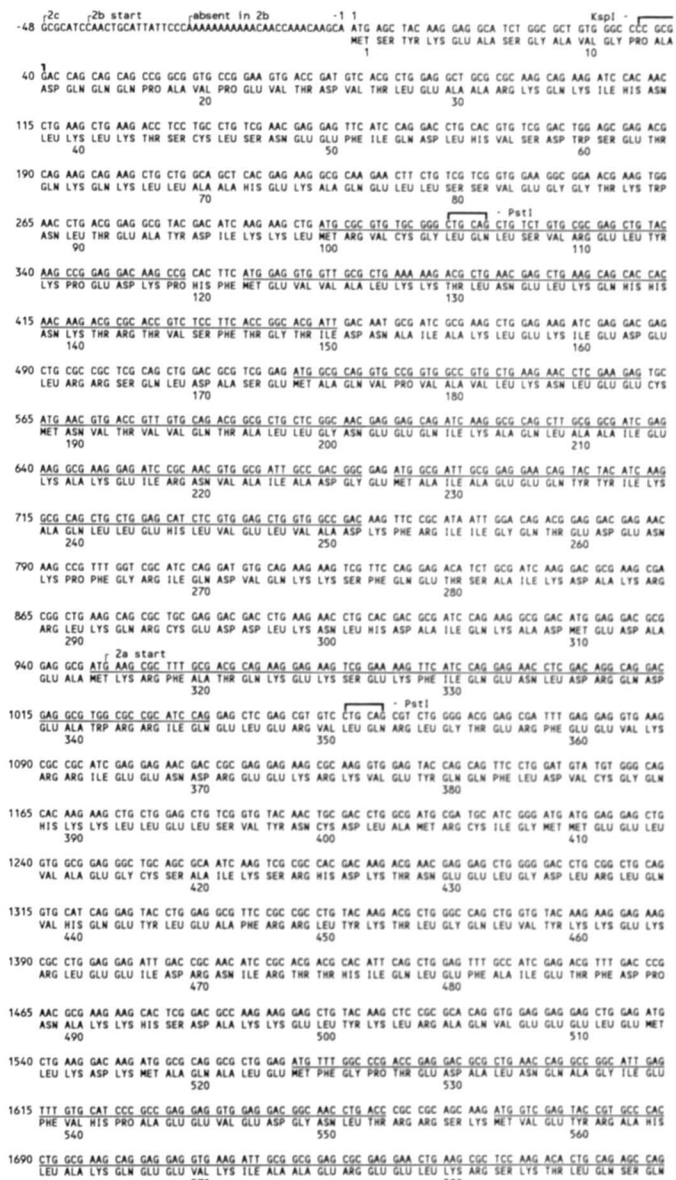

FIG. 5. **Nucleotide and deduced amino acid sequence of the cDNA coding for the PAR 2 protein.** *Numbers* above and at the ends of the sequence refer to nucleotides, *numbers* below the sequence refer to amino acids. Nucleotide 1 starts the putative initiator ATG. The location of the initially isolated cDNA clone, TccPar2a, is indicated, as is the single base difference between TccPar2b and TccPar2c. Amino acids that have been verified by direct amino acid sequence analysis are indicated by an *overline*. Significant restriction enzyme sites are also indicated.

at the indicated ATG for several reasons. 1) Translation starting at the indicated start site would produce a protein of 69,439 Da, consistent with the observed M_r of 68,000. Translation starting at the next in-frame ATG would produce a protein of only 58,450 Da. 2) The sequence environment around the next most plausible ATG (AGCTG ATG C), where direct amino acid sequence data are available, does not conform to the consensus identified for ribosome-binding sites (23) (*i.e.* a purine, usually A at the highly conserved −3 position and a purine, usually G at the +4 position). In contrast, the environment around the first ATG contains a purine in both of these positions (AAGCA ATG A). 3) The nucleotide sequence was examined using Fickett's Testcode for coding region determination (24). This analysis is based on the frequency of nucleotides in each of the three potential codon positions in all three frames. The Testcode results indicate that the entire region between the first in-frame ATG and the termination codon is probably a protein coding region (data not shown).

TccPar2a, which was initially identified by its polypeptide product in an expression vector, was determined to be the 3′ half of the gene, including a poly(A) tail. TccPar2b and TccPar2c do not contain poly(A) tails and neither included 5′ mini-exon sequence (25). The only polymorphism found was in a poly(A) region 5′ of the coding region. Clone TccPar2b contained 10 As while TccPar2c contained 11.

Stage-specific Expression of the PAR 2 Gene—To determine the developmental expression pattern of the PAR 2 gene, a Northern blot containing trypomastigote and epimastigote poly(A)$^+$ mRNA was hybridized with the ^{32}P-labeled 738-base pair *Pst*I fragment from TccPar2c (Fig. 6). A single mRNA band of approximately 2.0 kb was observed in the epimastigote (*E*) lane. A 2.0-kb band was also visible in the trypomastigote (*T*) lane, but at a greatly reduced intensity. A mRNA of this size has a theoretical coding capacity for a protein of approximately 73 kDa, in keeping with our previous studies (4) which show that the PAR 2 protein has a M_r of 68,000.

Genomic Organization of the PAR 2 Gene—The copy number of the PAR 2 gene sequence in the genome of *T. cruzi* was determined by the method previously described (17). Briefly, the ^{32}P-labeled *Pst*I fragment from TccPar2c was hybridized to a Southern blot containing trypomastigote nuclear DNA digested with *Ksp*I (Fig. 7). Included on the Southern blot was TccPar2c subcloned into Bluescript and restricted with *Ksp*I in amounts equivalent to 2, 5, 10, 20, 30, 40, and 50 copies per haploid genome. Strong hybridization of the probe was observed to a genomic fragment of 1.7 kb. When the intensity of the hybridization signal in the genomic DNA is compared to that of the various equivalents in the cloned DNA, the 1.7-kb fragment is seen to occur approximately 30 times per haploid genome.

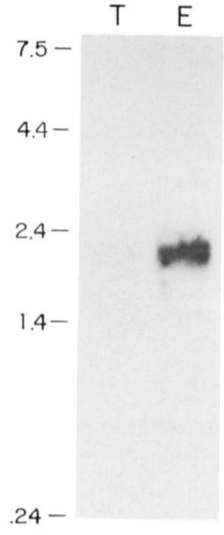

FIG. 6. **Identification of mRNA from trypomastigotes and epimastigotes complementary to the *Pst*I fragment of TccPar2c.** A Northern blot containing 2 μg of trypomastigote (*T*) or epimastigote (*E*) poly(A)$^+$ mRNA per lane was hybridized with the [^{32}P]*Pst*I fragment from TccPar2c. *Numbers* in kb on the *margin* refer to the migration of RNA molecular weight markers.

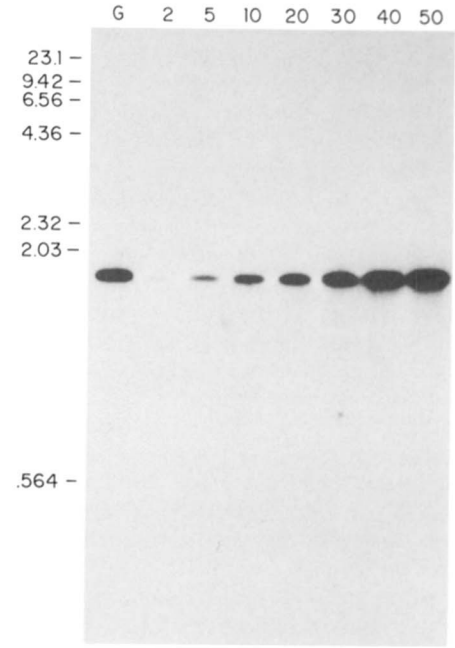

FIG. 7. **Determination of the PAR 2 gene copy number in trypomastigote DNA.** Nuclear DNA (2.5 μg) was digested with *Ksp*I and electrophoresed on a 1% agarose gel (*G*). Included in the gel was *Ksp*I-digested DNA from subclone TccPar2c containing the equivalent of 2, 5, 10, 20, 30, 40, and 50 copies per haploid genome (2–50, respectively). A Southern blot of the gel was hybridized with the [^{32}P]*Pst*I fragment from TccPar2c. *Numbers* in kb on the *margin* refer to the migration of *Hin*dIII fragments of λ phage DNA.

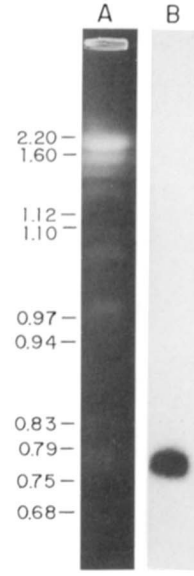

FIG. 8. **The chromosomal distribution of the PAR 2 genes.** *A*, *T. cruzi* chromosomes separated by PFGE for 48 h. *B*, Southern blot of PFGE gel described in *A* hybridized with [^{32}P]TccPar2a. *Numbers* in megabases on the *margin* refer to the migration of *S. cerevisiae* chromosome molecular weight markers.

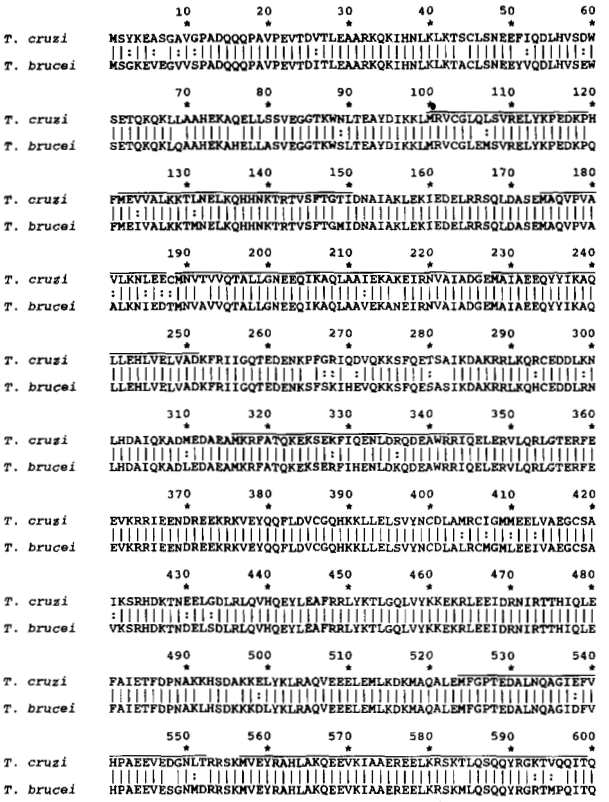

FIG. 9. **Comparison of the amino acid sequence for paraflagellar rod proteins from *T. cruzi* and *T. brucei*.** The amino acid sequence of the PAR 2 paraflagellar rod protein from *T. cruzi* and the PFR protein from *T. brucei* is compared. | indicates amino acid identity between the two proteins; (:) indicates a conservative amino acid change. Amino acids in the *T. cruzi* PAR 2 protein that have been verified by direct amino acid sequence analysis are indicated by an *overline*.

TABLE I

Comparison of DNA and amino acid homology for selected genes of T. brucei and T. cruzi

Gene	Source: organism,[a] amino acids, Refs.	% Homology		
		DNA	Amino acid	Amino acid[b]
Calmodulin	T.b. 1-149 (28) T.c. 1-149 (29)	84.3	98.7	99.3
gGAPDH	T.b. 1-359 (30) T.c. 1-359 (31)	80.8	90.2	95.3
Paraflagellar Rod Protein	T.b. 1-600 (3) T.c. 1-600 Fig 5	83.3	90.2	96.3
β-Tubulin	T.b. 203-299 (32) T.c. 1-97 (33)	81.7	84.5	91.8
Ubiquitin	T.b. 1-77 (34) T.c. 1-77 (35)	83.6	96.1	98.7

[a] T.b. = *T. brucei*, T.c. = *T. cruzi*.
[b] Amino acid homology including conservative changes.

To determine if the multiple copies are dispersed in the genome, chromosome size DNA molecules from Esmeraldo epimastigotes were separated by PFGE and blotted to nitrocellulose. Hybridization of the chromosome blots with [^{32}P] TccPar2a revealed a single band corresponding to a chromosome of 0.78 megabase size (Fig. 8), suggesting that all of the PAR 2 genes are located on a single chromosome.

Comparison of Paraflagellar Rod Proteins from Two Trypanosomes—The coding region nucleic acid and amino acid sequences of the PAR 2 protein from *T. cruzi* were compared to those of the PFR protein from *T. brucei* (3). At the nucleic acid level there was 83.3% identity between the two genes (data not shown). The results of the amino acid comparison are shown in Fig. 9. Identity was observed for 541 of 600 amino acids (90.2%) with an additional 37 conservative amino acid changes (total 96.3%).

DISCUSSION

The paraflagellar rod is a structure closely associated with the axoneme in the flagella of Trypanosomatids and Euglenoids. Although this structure has been thoroughly studied at the ultrastructural level (26), information about its molecular composition, organization, and function(s) is scant.

Two polypeptides, with molecular masses about 70 kDa, have been tentatively identified as components of the paraflagellar body of a variety of Trypanosomatids and Euglenoids (20). Whether these two putative paraflagellar polypeptides correspond to two different gene products, or to conformational variants of a single one, as suggested by Schlaeppi *et al.* (3) for *T. brucei*, has remained an open question. In a previous work (4), we directly identified the polypeptide PAR 2 as one of the major components of the *T. cruzi* paraflagellar rod and suggested that a second paraflagellar polypeptide, named PAR 1, was distinct from PAR 2. Three lines of evidence presented in this paper indicate that indeed PAR 1 and PAR 2 are two distinct paraflagellar rod components. 1) These two polypeptides can be partially separated by ion-exchange chromatography. As shown in Fig. 1, protein preparations containing about 80% PAR 2 and 20% PAR 1 (*inset, lane 21*) and 60% PAR 1 and 40% PAR 2 (*inset, lanes 25* and *26*) were reproducibly obtained. 2) The HPLC profiles of CNBr cleavage products from protein preparations enriched in PAR 2 (Fig. 2, *panel A*) or containing mixtures of PAR 1 and PAR 2 (Fig. 2, *panel B*) are clearly different. These results are only compatible with PAR 1 and PAR 2 being two different polypeptides, since the HPLC profiles of conformational variants of the same gene product would be identical. 3) The results of amino acid sequencing analysis provide definitive evidence for PAR 1 and PAR 2 being two distinct polypeptides. First, all the amino acid sequences of peptides derived from protein preparations enriched in PAR 2 (Fig. 3) can be accounted for in the complete protein sequence deduced from the nucleotide sequence of the corresponding gene (Figs. 5 and 9). The sequences determined directly correspond to the following amino acids deduced from the gene sequence: 100–119, 122–150, 174–187, 189–226, 228–251, 316–345, 524–552, and 557–600 (Fig. 3). Second, the amino acid sequences of peptides from mixtures of PAR 1 and PAR 2 (Fig. 3) include PAR 2 sequences (PAR 2 amino acids 100–119, 122–150 and 316–345 were identified), as well as sequences of at least eight major peptides unrelated to PAR 2. These latter peptides likely correspond to CNBr cleavage products from PAR 1, the only other major polypeptide detected in preparations of purified paraflagellar proteins.

These results show that PAR 1 and PAR 2, the major paraflagellar polypeptides of *T. cruzi*, are chemically distinct, in agreement with our previous results (4) which indicated these two proteins to be immunological distinct. In apparent contradiction with these data, other workers reported that monoclonal antibodies to the paraflagellar rod of *T. brucei* (20) or *Leishmania* (21) reacted with the two major paraflagellar polypeptides of each of those parasites, results that in one case were interpreted as indicative of common epitopes in those polypeptides (20). Two types of data, however, would

argue against the latter interpretation. 1) We have previously shown that mAbPAR 2, a monoclonal antibody to PAR 2, reacts in Western blots with a single polypeptide when *T. cruzi* extracts are obtained under stringent proteolysis-free conditions. In contrast, mAbPAR 2 reacts with several peptides if proteolysis is not totally inhibited (4), results that could be erroneously interpreted as indicative of the existence of different paraflagellar peptides with common epitopes. 2) In Western blots of extracts of either *T. brucei* or *L. brasiliensis* prepared under proteolysis-free conditions, each of our antibodies (4) specific for PAR 1 (pcAbPAR 1) or PAR 2 (mAbPAR 2) reacted with a single polypeptide (Fig. 4). Although the combined results of these immunological and amino acid sequencing studies provide solid evidence for two major paraflagellar proteins in *T. cruzi*, the question of possible homologies between PAR 1 and PAR 2 will only be resolved when the complete amino acid sequence of PAR 1 becomes available. Also, we cannot currently exclude the possible existence of minor polypeptides with epitopes in common with the major paraflagellar proteins. In fact, our amino acid sequencing studies of PAR 2 point to the existence of sequence microheterogeneity (to be reported elsewhere). This observation opens the possibility that more than one of the genes in the PAR 2 tandem array may be expressed.

Previous immunological studies also demonstrated cross-reactivity between paraflagellar components of different members of the Trypanosomatid family, and even with members of the more distant Euglenoid family (20). These observations predicted some degree of conservation in the primary structure of paraflagellar polypeptides, although the degree of conservation and the components involved were not identified. In this paper we present the complete amino acid sequence of *T. cruzi* PAR 2, one of the two major protein components of the paraflagellar rod of that parasite. Comparison of this sequence with the *T. brucei* paraflagellar protein studied by Schlaeppi *et al.* (3) indicated identity for 541 of 600 amino acids (90.2%) with an additional 37 conservative amino acid changes (total 96.3%). This degree of homology is consistent with that seen when other conserved genes of *T. brucei* and *T. cruzi* are compared (Table I). We anticipate that evolutionary conservation of primary structure may not be confined to PAR 2 but may extend to other paraflagellar components since, as mentioned before, pcAbPAR 1, a polyclonal antibody to *T. cruzi* PAR 1, reacts with a single polypeptide on Western blots of both *T. brucei* and *L. brasiliensis* (Fig. 4). This high degree of conservation, so far only proven for PAR 2, indicates that stringent structural requirements, relevant for either macromolecular organization or a highly specialized function(s) have been imposed upon the paraflagellar proteins. In this context, the immunological cross-reactivity of the paraflagellar components of Trypanosomatids and Euglenoids is intriguing, as the paraflagellar rod of the first group of parasites is organized as a compact filamentous structure, while in the latter group the paraflagellar rod is a hollow cylindrical body (27).

Acknowledgments—We thank Barbara Granger for assistance in maintenance and growth of parasite cultures and Dr. David Campbell from UCLA for his generous gift of *T. brucei* cultures.

REFERENCES

1. Brener, Z. (1973) *Annu. Rev. Microbiol.* **27,** 347–382
2. de Souza, W., and Souto-Padron, T. (1980) *J. Parasitol.* **66,** 229–235
3. Schlaeppi, K., Deflorin, J., and Seebeck, T. (1989) *J. Cell Biol.* **109,** 1695–1709
4. Saborio, J. L., Hernandez, J. M., Narayanswami, S., Wrightsman, R., Palmer, E., and Manning, J. (1989) *J. Biol. Chem.* **264,** 4071–4075
5. Beard, C. A., Wrightsman, R. A., and Manning, J. E. (1985) *Mol. Biochem. Parasitol.* **16,** 199–212
6. Nelson, D. J., Bugge, C. J., Elion, G. B., Berens, R. L., and Marr, J. J. (1979) *J. Biol. Chem.* **254,** 3959–3964
7. Brun, R., and Schonenberger (1979) *Acta Trop.* **36,** 289–292
8. Laemmli, U. K. (1970) *Nature* **227,** 680–685
9. Lanar, D. E., and Manning, J. E. (1984) *Mol. Biochem. Parasitol.* **11,** 119–131
10. Lottspeich, F., and Henschen, A. (1985) in *High Performance Liquid Chromatography in Biochemistry* (Henschen, A., Hupe, K.-P., Lottspeich, F., and Voelter, W., eds) pp. 139–216, VCH, Weinheim
11. Henschen, A. (1986) in *Advanced Methods in Protein Microsequence Analysis* (Wittmann-Liebold, B., Salnikow, J., and Erdmann, V. A., eds) pp. 244–255, Springer-Verlag, Berlin
12. Dragon, E. A., Sias, S. R., Kato, E. A., and Gabe, J. D. (1987) *Mol. Cell. Biol.* **7,** 1271–1275
13. Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. A., Smith, J. A., and Struhl, K. (eds) (1991) *Current Protocols in Molecular Biology,* Greene Publishing and Wiley-Interscience, New York
14. Helms, C., Graham, M. Y., Dutchik, J. E., and Olson, M. V. (1985) *DNA* **4,** 39–49
15. Fouts, D. L., and Manning, J. E. (1981) *Nucleic Acids Res.* **9,** 7053–7064
16. Rozek, C. E., and Davidson, N. (1983) *Cell* **32,** 23–34
17. Beard, C. A., Wrightsman, R. A., and Manning, J. E. (1988) *Mol. Biochem. Parasitol.* **28,** 227–234
18. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
19. Engman, D. M., Reddy, L. V., Donelson, J. E., and Kirchoff, L. V. (1987) *Mol. Biochem. Parasitol.* **22,** 115–123
20. Gallo, J.-M., and Schrevel, J. (1985) *Eur. J. Cell Biol.* **36,** 163–168
21. Ismach, R., Cianci, C. M. L., Caulfield, J. P., Langer, P. J., Hein, A., and McMahon-Pratt, D. (1989) *J. Protozool.* **36,** 617–624
22. Young, R. A., and Davis, R. W. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 1194–1198
23. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872
24. Fickett, J. W. (1982) *Nucleic Acids Res.* **10,** 5303–5318
25. Borst, P. (1986) *Annu. Rev. Biochem.* **55,** 701–732
26. de Souza, W. (1984) *Int. Rev. Cytol.* **86,** 197–283
27. Hyams, J. S. (1982) *J. Cell Sci.* **55,** 199–210
28. Tschudi, C., Young, A. S., Ruben, L., Patton, C. L., and Richards, F. F. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 3998–4002
29. Chung, S. H., and Swindle, J. (1990) *Nucleic Acids Res.* **18,** 4561–4569
30. Michels, P. A., Poliszczak, A., Osinga, K. A., Misset, O., Van Beeumen, J., Wierenga, R. K., Borst, P., and Opperdoes, F. R. (1986) *EMBO J.* **5,** 1049–1056
31. Kendall, G., Wilderspin, A. F., Ashall, F., Miles, M. A., and Kelly, J. M. (1990) *EMBO J.* **9,** 2751–2758
32. Kimmel, B. E., Samson, S., Wu, J., Hirschberg, R., and Yarbrough, L. R. (1985) *Gene (Amst.)* **35,** 237–248
33. Maingon, R., Gerke, R., Rodriguez, M., Urbina, J., Hoenicka, J., Negri, S., Aguirre, T., Nehlin, J., Knapp, T., and Crampton, J. (1988) *Eur. J. Biochem.* **171,** 285–291
34. Swindle, J., Ajioka, J., Eisen, H., Sanwal, B., Jacquemot, C., Browder, Z., and Buck, G. (1988) *EMBO J.* **7,** 1121–1127
35. Wong, S., and Campbell, D. A. (1989) *Mol. Biochem. Parasitol.* **37,** 147–150