The Major Component of the Paraflagellar Rod of *Trypanosoma brucei* Is a Helical Protein That Is Encoded by Two Identical, Tandemly Linked Genes

Kathrin Schlaeppi, Judith Deflorin, and Thomas Seebeck

Institut für Allgemeine Mikrobiologie, CH-3012 Bern, Switzerland

Abstract. The flagellum of the parasitic hemoflagellate *Trypanosoma brucei* contains two major structures: (*a*) the microtubule axoneme, and (*b*) a highly ordered, filamentous array, the paraflagellar rod (PFR). This is a complex, three-dimensional structure, of yet unknown function, that extends along most of the axoneme and is closely linked to it. Its major structural component is a single protein of 600 amino acids. This PFR protein can assume two different conformations, resulting in two distinct bands of apparent molecular masses of 73 and 69 kD in SDS-gel electrophoresis. Secondary structure predictions indicate a very high helix content. Despite its biochemical similarity to the intermediate filament proteins (solubility properties, amino acid composition, and high degree of helicity), the PFR protein does not belong in this class of cytoskeletal proteins. The PFR protein is coded for by two tandemly linked genes of identical nucleotide sequence. Both genes are transcribed into stable mRNAs of very similar length that carry the mini-exon sequence at their 5′ termini.

TRYPANOSOMA brucei, the causative agent of human sleeping sickness and a complex of veterinary diseases in large parts of Africa, is a uniflagellate protozoon (of the *Kinetoplastida* order and *Trypanosomatida* family). Its single flagellum, which arises from within the cell body, emerges through a "flagellar pocket" and extends along the outer cell surface toward and beyond the anterior end of the cell. The flagellum contains a microtubular axoneme of the canonical nine plus two configuration. Besides this axoneme, the trypanosomal flagellum contains a second prominent structure, the paraflagellar rod (PFR)[1] complex (Fuge, 1969; Vickerman and Preston, 1976).

PFR structures have been described for three large groups of flagellates: the kinetoplastids, the euglenoids, and the dinoflagellates (Cachon et al., 1988). They are always highly ordered lattices of fibrous proteins that are located inside the flagellum and assume a fixed orientation with respect to the microtubular axoneme (Souto-Padron et al., 1984). However, preliminary structural studies have already indicated substantial differences in the architecture of the PFR complexes from kinetoplastids and euglenoids (DeSouza and Souto-Padron, 1980; Hyams, 1982).

Despite its structural prominence, nothing is known about the function of PFR (Cachon et al., 1988). Piccinni et al. (1975) observed an ATPase activity in the PFR of *Euglena gracilis* and proposed a function of this structure in flagellar motility. For trypanosomatids, such a function appears less probable. Two species of trypanosomatids, *Crithidia deanei* and *C. oncopelti*, have been observed that lack the PFR structure entirely (Freymuller and Camargo, 1981). However, a comparison of the wave patterns of beating flagella from these organisms with those from a *Crithidia* species that does contain a PFR complex revealed no differences, suggesting that PFR is unlikely to be involved in flagellar mechanics (Goldstein et al., 1970; Johnson et al., 1979).

The overall structure of the PFR complex was found to be very similar in many trypanosomatids such as *T. congolense* (Evans et al., 1979), *Herpetomonas mariadeanei* (Freymuller and Camargo, 1981), *C. fasciculata* (Russell et al., 1983), *Phytomonas davidi*, *H. megaseliae* (Farina et al., 1986), *T. cruzi*, *H. samuel pessoai*, *Leptomonas samueli*, *C. harmosa* (DeSouza and Souto-Padron, 1980), and *T. brucei* (Schneider et al., 1987).

Biochemical and immunological analyses of the PFR structure of several trypanosomatids by different authors have suggested that its major components are two immunologically closely related proteins of apparent molecular masses of ~73 and 69 kD. Two proteins of similar molecular masses have also been detected in *E. gracilis*, another organism containing a PFR structure (Hyams, 1982). These two proteins are immunologically related to the trypanosomal PFR proteins (Gallo and Schrevel, 1985).

The present report describes the isolation and characterization of the genetic locus coding for the major PFR protein. The results demonstrate that, in contrast to the general belief, the major structural component of the PFR is a single protein (PFR protein) of 600 amino acids, which corre-

1. *Abbreviations used in this paper*: PFR, paraflagellar rod.

sponds to the lower molecular mass (69 kD) PFR component described in the literature. The second, more slowly migrating (73-kD) protein band is a reduction-induced derivative thereof. The PFR protein is coded for by two closely linked, tandemly arranged genes of identical nucleotide sequence.

Materials and Methods

Growth of Cells

Procyclic trypanosomes (*T. brucei brucei*, stock STIB 366) were grown in SDM-79 medium (Brun and Schönenberger, 1979) at 26°C. The other *Trypanosoma* strains—*T. brucei rhodesiense*, *T. brucei gambiense*, *T. congolense*, *T. simiae*, *T. cruzi*, and *T. rangeli*—as well as *Leishmania major*, *L. donovani*, and *C. fasciculata* were gifts from the Tropeninstitut, (Basel, Switzerland). *C. deanei* (ATCC 30255) and *C. oncopelti* (ATCC 12982) were purchased from the American Type Culture Collection (Rockville, MD). They were grown at 25°C in media according to the manufacturer. *E. gracilis* was a gift from M. Schärer and *Chlamydomonas reinhardtii* was a gift from A. Boschetti (Departments of Plant Physiology and Biochemistry, respectively, University of Bern, Bern, Switzerland). *Tetrahymena pyriformis* was provided by R. Peck (Department of Protozoology, University of Geneva, Geneva, Switzerland). *Dictyostelium discoideum* was from R. Parish and *Giardia lamblia* was from B. Gottstein (Departments of Plant Biology and Parasitology, respectively, University of Zurich, Zurich, Switzerland).

Isolation of Proteins

For total proteins, 10 ml of cells were harvested during exponential growth: i.e., at densities of ~5×10^6 to 1×10^7 per ml. The cells were washed once in 10 mM morpholino propane sulfonic acid, pH 6.9, 1 mM EGTA, 1 mM $MgSO_4$, 250 mM sucrose and then solubilized in 500 µl sample buffer (Laemmli, 1970). After boiling for 10 min, 20–30 µl were analyzed on polyacrylamide gels.

Cell fractionations for isolating cytoskeletal and flagellar proteins were performed as described (Schneider et al., 1987, 1988). The resulting protein fraction consists predominantly of the axonemal α- and β-tubulin and of two proteins that migrate in SDS-gel electrophoresis with apparent molecular masses of 73 and 69 kD (Gallo and Schrevel, 1985). The bands corresponding to these two proteins were cut out of a preparative 7.5% polyacrylamide gel. The proteins were eluted with PBS, pH 7, 0.1% SDS, precipitated with acetone, and resuspended in PBS. This preparation was used for immunization.

Reduction and Carboxymethylation of Cell Extracts and Flagella

Cells or isolated flagella were solubilized in 6 M guanidinum chloride, 50 mM morpholino propane sulfonic acid, pH 7.5, and incubated at 37°C for 40 min. Proteins were reduced with 5 mM DTT or with 0.8 mM tributylphosphine (Rüegg and Rudinger, 1977). Oxidation of sulfhydryls was catalyzed by 1 mM *o*-phenanthroline and 0.5 mM $CuSO_4$. Blocking of free sulhydryls was done with 20 mM iodoacetamide. After reaction, proteins were extracted with chloroform/methanol (Wessel and Flügge, 1984) and were dissolved by boiling in sample buffer containing 20 mM iodoacetamide.

Antibodies

Polyclonal antibodies were raised against the 73- and 69-kD proteins in rats. The antiserum was entirely specific as tested by immunoblotting of whole-cell extracts and was used for immunoscreening without further purification. A monoclonal antibody against the PFR proteins was a gift from J.-M. Gallo (University of Poitiers, Poitiers, France) (Gallo and Schrevel, 1985). Peroxidase-conjugated swine antibodies to rat or mouse immunoglobulins (DAKOPATTS, Copenhagen, Denmark) were used to detect bound primary antibodies on Western blots. Goat anti-rat IgGs coupled to 15-nm gold for immunoelectron microscopy were purchased from Janssen Life Science Products (Beerse, Belgium).

Immunogold Electron Microscopy

For electron microscopy, Lowicryl-embedded insect and bloodstream form trypanosomes were provided by U.-P. Modespacher (Tropeninstitut, Basel, Switzerland). Thin sections were mounted on parlodion-carbon-coated nickel grids. After blocking with 5% FCS in 20 mM Tris-HCl, pH 8.2, 150 mM NaCl (TBS), the grids were incubated for 4 h at room temperature with the fusion protein affinity-purified polyclonal anti-PFR antibodies, diluted 1:5 in TBS, 5% FCS. After rinsing in TBS, 5% FCS, the grids were incubated for 2 h with 15-nm gold-labeled goat anti-rat antibody diluted 1:2 in TBS, 5% FCS. Grids were postfixed for 5 min with 1% glutaraldehyde in TBS, rinsed, and contrasted with 2% uranyl acetate and lead citrate. Control incubations received only the second gold-labeled antibody.

Restriction Mapping and Southern Blot Analysis

High molecular mass DNA was prepared as described (Seebeck et al., 1983). DNA was digested with restriction endonucleases (Boehringer Mannheim GmBH, Mannheim, FRG) under the conditions suggested by the suppliers. After electrophoretic separation on agarose gels, Southern blots of each gel were prepared (Smith and Summers, 1980) and analyzed with appropriate nick-translated or kinase-labeled DNA or RNA probes following procedures described by Maniatis et al. (1982).

Hybridization Conditions

DNA-containing filters were prehybridized for 2 h at 37°C with 50 µl/cm² of 0.6 M NaCl, 4 mM EDTA, 0.08 M Tris-HCl, pH 7.8, 0.1% BSA, 0.01% Ficoll, 0.01% polyvinylpyrrolidone, 50% (vol/vol) formamide, 0.2% $NaDodSO_4$, 0.1% sodium-pyrophosphate, 100 µg/ml depurinated calf thymus DNA. Subsequent hybridization was carried out with 10^6–10^7 cpm of ^{32}P-labeled DNA per filter in 30 µl/cm² of the above buffer overnight at 37°C. After hybridization, the filters were washed twice successively in 0.3 M NaCl, 30 mM sodium citrate, pH 7.2 (2× SSC), 0.1% $NaDodSO_4$ and twice in 0.2× SSC, 0.1% $NaDodSO_4$. Filters containing DNA hybridized with homologous probes were washed at 60°C, whereas filters containing heterologous probes were washed at room temperature.

Constructing and Screening of a Genomic DNA Library from T. brucei in Lambda gt 11

High molecular mass DNA from *T. brucei* stock STIB 366 was prepared as described (Seebeck et al., 1983). It was sheared by sonication to a length of 0.2–6 kb, treated with T_4 DNA polymerase, methylated with Eco RI methylase, and completed with Eco RI linkers before ligation into the unique dephosphorylated Eco RI site of lambda gt 11 (Young and Davis, 1983). The recombinant viral DNA molecules were packaged and plated on *Escherichia coli* strain Y 1088, giving a total of 6×10^6 recombinants. Before screening, the library was amplified once on *E. coli* Y 1088. 6×10^5 phages of the amplified library were screened, using the polyclonal PFR antiserum at a dilution of 1:500 in 20 mM Tris-HCl, pH 7.4, 150 mM NaCl (TBS) containing 5% horse serum following published procedures (Young and Davis, 1983). Bound antibodies were detected by incubation with anti-rat IgG conjugated to horseradish peroxidase at a dilution 1:400 in TBS, pH 7.4, 5% horse serum. Peroxidase activity was visualized with a freshly prepared solution of 0.5 mg/ml diaminobenzidine, 4 mM H_2O_2 in TBS, pH 7.4.

Constructing and Screening of a Genomic DNA Library from T. brucei in Lambda EMBL 4

A genomic library from *T. brucei* stock STIB 366 in the lambda vector EMBL 4 (Frischauf et al., 1983) was constructed using published procedures (Kaiser and Murray, 1985). 5×10^5 independent recombinant phages were obtained, and the library was then amplified once on *E. coli* NM 539. A 180-bp insert of a positive gt 11 phage was isolated and subcloned into a pEP-30 plasmid vector, yielding plasmid pEP72(180). The pEP-30 vector is a modification of pGEM-3 vector (Promega Biotec, Madison, WI) containing the single strand replication origin from pEMBL 9 (Dente et al. 1983; Imboden et al., 1987). The resulting pEP72(180) was used to screen the trypanosomal genomic library by DNA hybridization procedures described by Ozaki and Traub-Cseko (1984).

Affinity Purification of the Antiserum on Fusion Proteins

Fusion proteins of positive gt 11 phages were produced as described (Huynh

et al., 1985) using lysogenic strains on *E. coli* Y 1089. The cells were harvested and suspended in 100 mM Tris-HCl, pH 7.4, 2 mM EDTA and frozen down quickly in ethanol-CO_2. A crude lysate was obtained by slow thawing of the cells. As a purification step, an ammonium sulfate precipitation with a final concentration of 33% $(NH_4)_2SO_4$ was carried out. The resulting pellet was suspended in 100 mM Tris-HCl, pH 7.4, and the dissolved proteins were precipitated with a 10% final concentration of TCA. The proteins were separated on a preparative 8–15% polyacrylamide gradient gel and transferred onto nitrocellulose. After staining with 0.04% Ponceau S (Fluka AG, Buchs, Switzerland) in 10% acetic acid, the band corresponding to the fusion proteins was excised. The strip was washed twice in 10 mM Tris-HCl, pH 7.4, 150 mM NaCl (TBS), 0.05% Tween 20 for 10 min and incubated overnight at 4°C with the polyclonal PFR antiserum, diluted 1:5 in TBS, 0.05% Tween 20, 3% BSA (Fluka AG). After washing three times for 10 min in TBS, 0.05% Tween 20, elution of the adsorbed antibodies was done essentially as described by Goldstein et al. (1986). The affinity-purified antibodies were tested on Western blots and immunofluorescence for their ability to bind back to PFR proteins. These antibodies were then used for immunogold microscopy.

DNA Sequencing

A 6.4-kb Kpn I–Kpn I fragment containing the genes for PFR1 and PFR2 was subcloned into the Kpn I site of the Bluescript plasmid (Stratagene), yielding plasmid pTbPFR-A. Ordered deletions were produced by exonuclease III digestion following the procedures described in the Stratagene manual. Single and double strand sequencing was performed with the Sequenase kit (United States Biochemical Corp., Cleveland, OH) following the dideoxy procedure of Sanger et al. (1977) and according to the instructions of the manufacturer. The sequence was analyzed using the DNAstar software (Madison, WI).

Hybrid Selection and In Vitro Translation

mRNA coding for PFR proteins was hybrid selected from total RNA using pTbPFR-A DNA spotted onto nitrocellulose following the procedure of Parnes et al. (1981). The hybrid-selected RNA was subsequently used for in vitro translation in a reticulocyte lysate (New England Nuclear, Boston, MA), using [^{35}S]methionine as labeled amino acid, according to the instructions of the manufacturer. After 1 h of incubation at 37°C, cold methionine was added from a stock solution to a final concentration of 1 mM, and the incubation was continued for another 10 min at 37°C. 15 μl of the translation products was precipitated with 1 ml 10% TCA and washed twice with acetone. The resulting pellet was dissolved in 10 μl of 2% (wt/vol) NP-40, 2% (vol/vol) ampholines, pH 3.5–10, 5% (vol/vol) 2-mercaptoethanol, and 9.5 M urea. As internal markers, nonlabeled purified 73- and 69-kD PFR proteins (1 μg in 2 μl of the solution described above) were added. This protein mixture was resolved by two-dimensional gel electrophoresis and transferred onto nitrocellulose. The blots were examined by immunostaining with PFR antibody and by autoradiography.

Primer Extension Sequencing

Primers specific for the divergent upstream regions of the mRNAs for PFR-A (5′-CTTTATCTACAAAGTTAATT-3′) and for PFR-B (5′-TCTCTTAAATTTGGGATAAC-3′) were synthesized on a synthesizer (Applied Biosystems, Inc., Foster City, CA) and were purified on NAC-10 minicolumns (Pharmacia Fine Chemicals, Piscataway, NJ).

Primer extension was performed essentially as described (Imboden et al., 1986). 100 ng of 5′ labeled primer (100,000 cpm) were annealed with 20 μg total trypanosomal RNA in 10 mM Tris-HCl, pH 7.5, 0.25 M KCl, 1 mM EDTA in a final volume of 20 μl at 37°C for 2 h. Then 40 μl of a mixture containing 75 mM Tris-HCl, pH 7.5, 15 mM DTT, 12 mM $MgCl_2$, 3 μg actinomycin D (Sigma Chemical Co., St. Louis, MO), 12 μl RNasin (Promega Biotec), 20 μl AMV reverse transcriptase (Super RT; Anglian Biotechnology Ltd., Cambridge, UK), and the nucleotide triphosphates was added. Nucleotide triphosphates were present each at 500 μM for the full reaction. For each sequencing reaction, the dideoxy derivative was used at 10 μM and the corresponding deoxy-triphosphate at 20 μM, while the other three triphosphates were present each at 40 μM. The reaction mixtures were incubated at 37°C for 1 h. Nucleic acids were then precipitated with ethanol, resuspended in 10 μl loading buffer (96% formamide, 20 mM EDTA, pH 8, 0.1% bromophenol blue, 0.1% xylene cyanol FF), heated for 3 min at 80°C, and analyzed on 6% polyacrylamide gels. An end-labeled Hpa II digest of pBR322 was used as a set size marker.

Other Procedures

One-dimensional gel electrophoresis was performed according to Laemmli (1970). Two-dimensional gel electrophoresis was carried out by the original procedure (O'Farrell, 1975) with the following modifications: (*a*) the mixture of ampholines used was a 1:1 ratio of pH 4–6/pH 5–8; (*b*) the prefocusing run was omitted—electrophoresis was done on a mini-gel apparatus at a constant 215 V for 3 h and was then continued for an additional hour at 430 V; and (*c*) a 10% polyacrylamide gel was used for the second dimension. Transfer of polypeptides to nitrocellulose and immunostaining with poly- and monoclonal antibodies were done as described in Towbin et al. (1979).

Results

The PFR Proteins

The PFR structure of the trypanosomal flagellum consists of a highly ordered, three-dimensional fiber network. It runs parallel to the microtubular axoneme and is closely linked to the latter through a complex network of protein linkers (Fig. 1). The PFR of isolated flagella is highly resistant to extraction by nonionic detergents, low salt buffers, or 1 M NaCl (Schneider et al., 1987), but it can be solubilized with 6 M urea. SDS-gel electrophoresis analysis demonstrated that, in accordance with the literature, the major structural components of the PFR were two proteins with apparent molecular masses of 73 and 69 kD. Attempts to separate the two

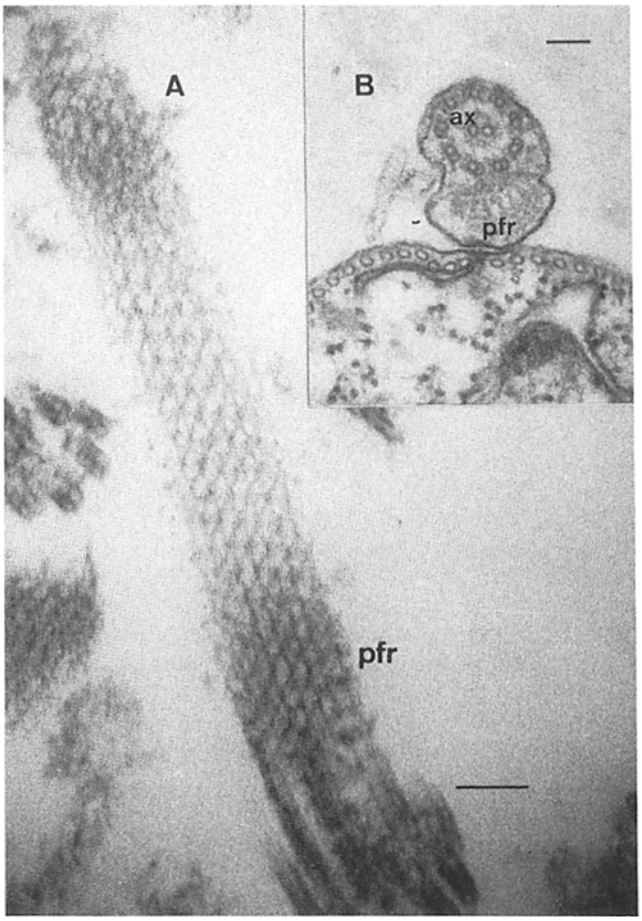

Figure 1. (*A*) Section through a flagella preparation of *T. brucei*. (*B*) Cross section through a flagellum. *ax*, axoneme; *pfr*, paraflagellar rod. Bars, 0.1 μm.

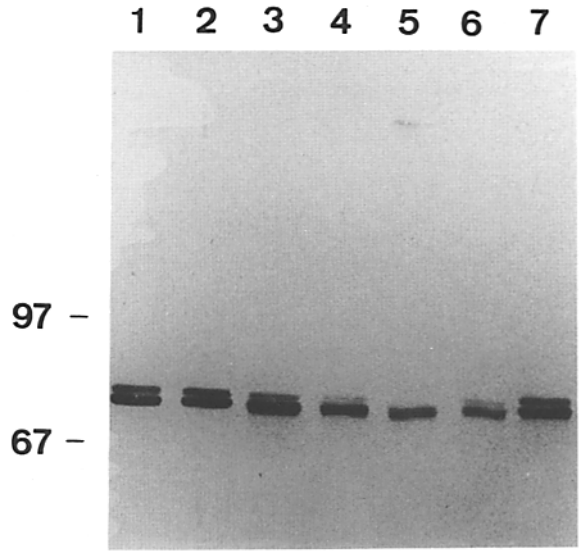

Figure 2. Migration of PFR protein is dependent on its reduction status. (Lane *1*) Whole-cell lysate in 6 M GuCl and 20 mM iodoacetamide; (lane *2*) same as lane *1*, but oxidized with phenanthroline/copper at 25°C for 30 min; (lane *3*) whole-cell lysate in 6 M GuCl; (lane *4*) same as lane *3*, but incubated at 25°C for 30 min; (lane *5*) same as lane *3*, but oxidized with phenanthroline/copper at 25°C for 30 min; (lane *6*) same as lane *5*, but subsequently reduced with 5 mM DTT at 50°C for 30 min; (lane *7*) whole-cell lysate in 6 M GuCl and 5 mM DTT. Each slot received the equivalent of 1.3 × 10^6 cells. Proteins were blotted and visualized by immunostaining with PFR antiserum.

Table I. The Amino Acid Composition of PFR1 and PFR2 as Compared with Intermediate Filaments

Amino acid	PFR protein*	PFR protein‡	Keratin‡	Desmin‡	Intermediate filament *Helix pomatia**
	%	%	%	%	%
Ala	10	8.8	6.4	8.9	9.1
Arg	6	6.3	6.8	9.8	7.1
Asn + Asp	10	9.0	10.8	8.4	9.6
Cys	–	1.2	0.5	0.2	–
Gln + Glu	18	20.0	16.9	20.9	17.1
Gly	9	3.0	9.3	4.3	7.1
His	2	3.0	1.0	1.6	3.0
Ile	4	4.8	3.7	3.9	4.9
Leu	8	9.8	11.2	9.3	9.0
Lys	8	9.8	5.4	4.6	7.0
Met	1	3.2	3.2	2.5	1.2
Phe	3	2.3	2.4	3.2	2.0
Pro	1	1.7	0.7	1.6	2.8
Ser	6	4.3	8.3	6.4	8.1
Thr	5	3.8	4.6	5.4	5.1
Trp	–	0.5	0.5	0.2	–
Tyr	1	2.2	2.9	2.5	3.1
Val	6	6.0	5.4	6.2	3.9
Mass (kD)	69/73	69.6	50.0	49.4	66/53/52

Values are presented as percentages.
* Biochemical determination (intermediate filament *Helix pomatia* from Bartnik et al. [1985]).
‡ Calculated from the deduced protein sequence (vimentin from Hanukoglu and Fuchs [1982] and desmin from Geisler et al. [1982]).

species from each other by a variety of chromatographic procedures failed (our unpublished frustrations), suggesting that they might be very similar. This notion was supported by the observation that a monoclonal antibody raised against PFR (Gallo and Schrevel, 1985) is capable of detecting both proteins.

Electrophoretic analysis of the PFR proteins indicated the presence in *T. brucei* of two distinct species of PFR proteins with apparent molecular masses of 73 and 69 kD. These observations are in full agreement with much published evidence (Cachon et al., 1988; Cunha et al., 1984; Gallo and Schrevel, 1985; Russell et al., 1983; Saborio et al., 1989) which supported the concept of two distinct, though related, proteins as the major structural component of PFR.

However, a different picture emerges when whole cells or isolated flagella are analyzed under various conditions of oxidation and reduction (Fig. 2). When cells are lysed either in the presence of iodoacetamide (lane *1*) or 5 mM DTT (lane *7*), PFR protein migrates as two distinct bands. Oxidation of carboxymethylated PFR does not change this migration pattern (lane 2). In contrast, oxidation of noncarboxymethylated PFR in the absence (lane *4*) or in the presence (lane *5*) of *o*-phenanthroline/copper as catalyst shifts its migration towards a single band. Reduction of such oxidized PFR leads to the reappearance of the more slowly migrating band (lane *6*).

These experiments indicate that the two clearly distinguishable bands of PFR protein that are present in the cell (see Fig. 2, lane *1*) represent in fact two conformations of a single polypeptide chain.

The amino acid composition of purified PFR protein was found to bear similarity to that of intermediate filaments (Table I), suggesting that the proteins might contain a large proportion of helical structure. In agreement with this assumption, limited digestions of PFR protein with trypsin or chymotrypsin produce a protease-resistant fragment of 65 kD, suggesting the presence of a tightly organized, protease-resistant core flanked by more loosely structured, protease-sensitive NH_2- and COOH-terminal domains (not shown).

PFR Protein Is Exclusively Found in Species Containing a PFR

Rat antiserum was raised against gel-purified PFR protein (see Materials and Methods). Antibodies were affinity purified before further use by adsorption to a PFR-β-galactosidase fusion protein from phage TbPFR72 (see below). When retested on immunoblots of authentic PFR protein, the affinity-purified antibody still was able to recognize both the 73- and the 69-kD bands. This observation again supported the notion of common epitopes in both bands of PFR protein. The binding specificity of the affinity-purified antibody was further established on the structural level by immune electron microscopy. The electron micrographs shown in Fig. 3 demonstrate that the PFR antibody does in fact bind specifically to the PFR structure.

When lysates from different species of lower eukaryotes were probed with the polyclonal PFR antiserum, a clear pattern emerged. As expected, the antibody recognized similar proteins in all those genera of hemoflagellates that are known to contain a PFR structure (*Trypanosoma*, *Leishmania*, *Crithidia*) (Fig. 4). However, the antibody clearly discriminates

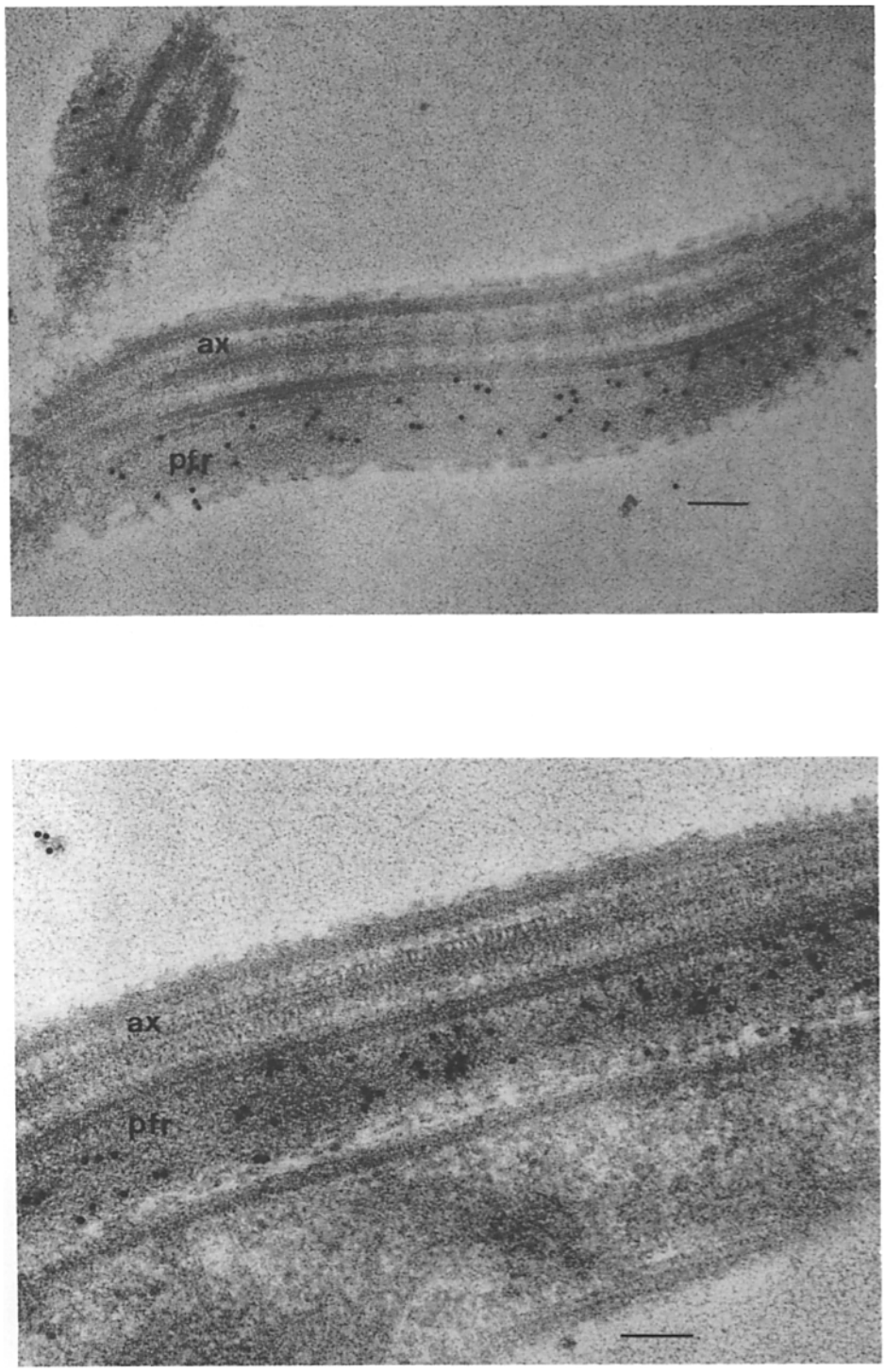

Figure 3. Sections of Lowicryl-embedded procyclic trypanosomes after immunogold staining by anti-PFR antibody affinity purified on the β-galactosidase–PFR fusion protein of TbPFR 72. Bar, 0.1 μm.

between different trypanosomal subgenera and species. All species of the subgenus *Trypanozoon*, (*T. b. brucei*, *T. b. rhodesiense*, and *T. b. gambiense*) react equally well with the polyclonal antiserum (Fig. 3, lanes *1–3*). Similarly, representatives of the subgenus *Nannomonas* (*T. congolense*, stock K44, and *T. simiae*; lanes *4* and *7*) also react strongly. In contrast to these salivarian trypanosomes, the stercorarian species *T. cruzi* (*Schizotrypanum*) (lane *5*) and *T. rangeli* (*Herpetosoma*) (lane *6*) exhibit a much weaker reaction. In all these species, the proteins detected by the antibodies migrate in the same molecular mass range as those of *T. brucei*. This is in agreement with a previous study (Gallo and Schrevel, 1985) that described two paraflagellar proteins in *T. brucei* of 75 and 72 kD. The same pattern was also obtained when the bloodstream forms of *T. brucei* were stained with the same antibody (not shown). The apparent molecular masses of the PFR proteins of the more distantly related kinetoplastids *C. fasciculata*, *L. major*, and *L. donovani* (lanes *8*, *10*, and *11*) are sightly different from those of *T. brucei*, in agreement with Russell et al. (1983) who determined appar-

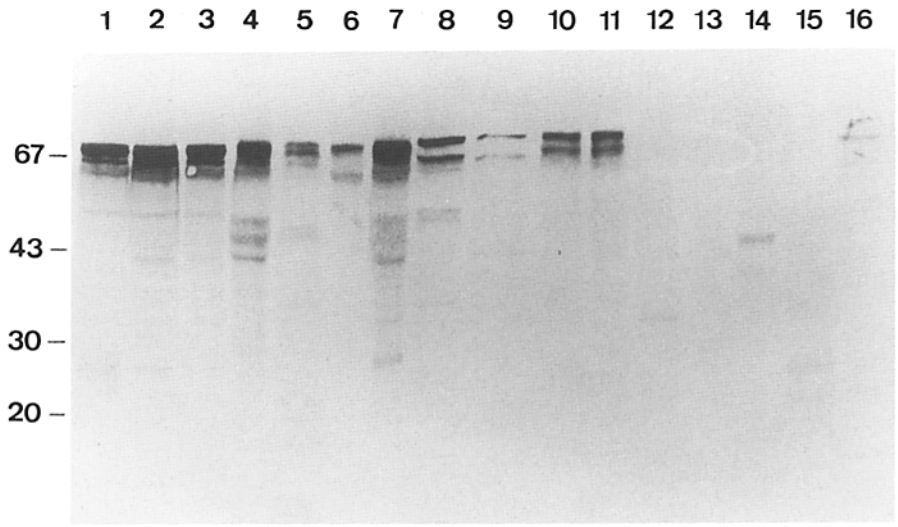

Figure 4. Presence of related PFR proteins in various protozoa. Immunostaining of cell lysates with the polyclonal antibody against PFR proteins. (Lane *1*) *T. brucei brucei*; (lane 2) *T. brucei rhodesiense*; (lane *3*) *T. brucei gambiense*; (lane *4*) *T. congolense*; (lane *5*) *T. cruzi*; (lane *6*) *T. rangeli*; (lane *7*) *T. simiae*; (lane *8*) *C. fasciculata*; (lane *9*) *C. deanei*; (lane *10*) *L. donovani*; (lane *11*) *L. major*; (lane *12*) *G. lamblia*; (lane *13*) *Tet. pyriformis*; (lane *14*) *D. discoideum*; (lane *15*) *C. reinhardtii*; (lane *16*) *E. gracilis*. All lanes contain similar amounts of total cellular protein.

ent molecular masses of 76 and 68 kD for the corresponding proteins in *C. fasciculata*.

In contrast to *C. fasciculata*, only a very weak reaction was observed with *C. deanei* (lane 9), and no reaction was detectable in *C. oncopelti* (not shown). The absence of PFR staining in these species is in agreement with, and extends, the earlier observations of Freymuller and Camargo (1981) who have shown that neither of these species contains a PFR structure that is detectable by electron microscopy. The presence of traces of PFR protein in *C. deanei* may suggest that this species still contains a low level of PFR protein, though it is apparently not organized into a discernible PFR structure.

Among the nonkinetoplastid organisms tested, a very weakly positive signal was obtained with *E. gracilis*. According to Hyams (1982), *E. gracilis* contains a rod-like structure within its flagellum which is made up predominantly of two polypeptides of 80 and 69 kD. The PFR antibody in fact does stain two polypeptides at the appropriate molecular mass range (lane *16*). Besides *Euglena*, none of the other protozoa tested (*G. lamblia*, *Tet. pyriformis*, *D. discoideum*, and *C. rheinhardtii*) reacted with the PFR antibody.

The PFR Protein Is Coded for by Two Genes That Are Tandemly Linked

The PFR antiserum was used for screening a genomic lambda phage gt 11 expression library of *T. brucei*. The 180-bp insert from a recombinant phage (λTbPFR72) isolated from this screening was subcloned into the pEP 30 vector

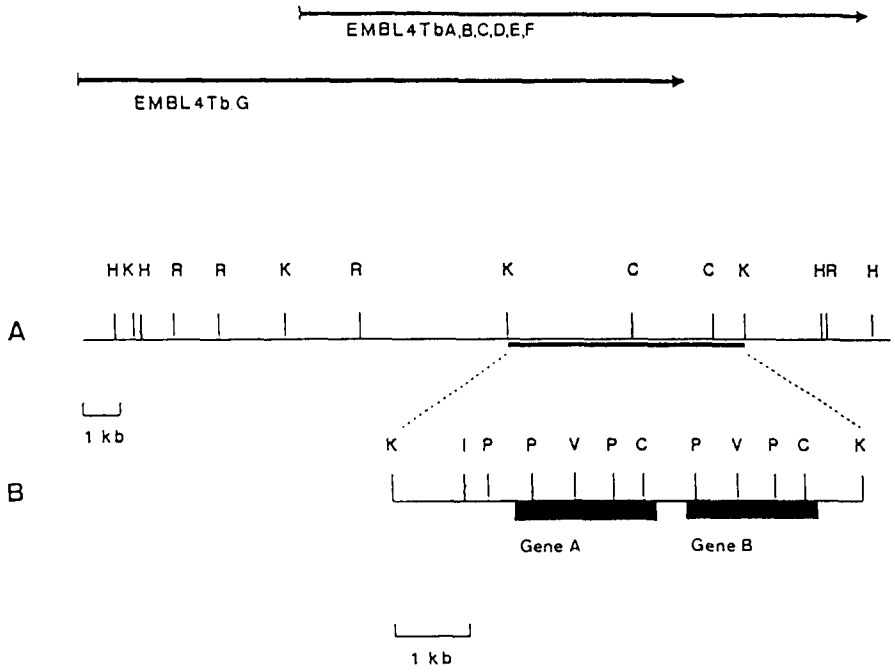

Figure 5. Physical map of the PFR locus. (*A*) A partial restriction map of the seven overlapping bacteriophage recombinants (λEMBL4Tb A–G) covering a total of 23 kb. (*B*) A 6.4-kb Kpn I fragment hybridizes with the coding sequence of pEP72(180), and the 5-kb Hind II–Kpn I fragment has been sequenced. The positions of the two coding regions are indicated by solid boxes. The direction of transcription is from left to right. *H*, Hind III; *K*, Kpn I; R, Eco RI; *C*, Cla I; *I*, Hind II; *P*, Pst I; *V*, Pvu II.

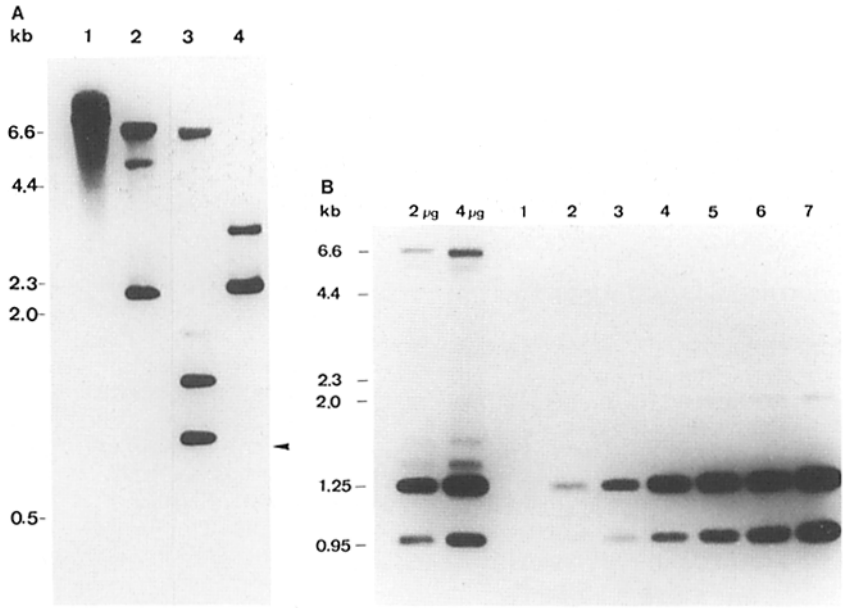

Figure 6. (*A*) Genomic hybridization. Southern blot of genomic DNA from *T. brucei* (2 μg/lane) digested with restriction enzymes and hybridized with labeled pTbPFR-A. (Lane *1*) Kpn I; (lane *2*) Cla I; (lane *3*) Cla I–Pvu II; (lane *4*) Kpn I–Cla I. The arrowhead indicates a barely visible band in lane *4*. (*B*) Copy number of the 0.95-kb Cla I–Pvu II fragment of the coding regions of genes A and B. The isolated 2.2-kb Cla I probe, digested with Pvu II, was hybridized to a Southern blot of *T. brucei* genomic DNA (2 and 4 μg) restricted with Cla I and Pvu II. Carefully calibrated amounts of the purified 2.2-kb Cla I fragment, restricted with Pvu II, are present in lanes *1–7*. (Lanes *1–5*) 0, 110, 220, 550, and 770 pg, respectively; (lanes *6* and *7*) 1.1 and 2.2 ng, respectively. Lanes *1–7* represent, respectively, 0, 1, 2, 5, 7, 10, and 20 copies of the 2.2-kb fragment relative to the *T. brucei* genome content in 2 μg DNA.

and the resulting pEP72(180) was then used to screen a genomic DNA library in λ EMBL4. Seven individual phages were purified, and restriction analysis of their isolated DNA demonstrated that these seven recombinants all contained overlapping DNA regions from the same genetic locus, covering a total of 23 kb. Within the recombinant EMBL phage DNA, a 6.4-kb Kpn I fragment was identified that carried PFR coding sequence. This Kpn I fragment was subcloned into a Bluescript vector, and the resulting plasmid pTbPFR-A was analyzed by restriction mapping and hybridization. The detailed physical map of the PFR locus is presented in Fig. 5. The locus contains two closely spaced regions that both code for PFR. A degree of similarity between the two is indicated by a number of restriction sites that are identical between the two regions.

The total number of PFR loci in the genome of *T. brucei* was determined by restriction enzyme analysis of genomic DNA. Hybridization of restriction digests of genomic DNA with the PFR-specific probe pTbPFR-A (see Fig. 5) revealed a band pattern (Fig. 6 *A*) that is fully compatible with the restriction map of the PFR locus given in Fig. 5. Thus, no PFR genes are located elsewhere in the trypanosomal genome. This conclusion was further corroborated by titration hybridization of genomic DNA digests. Genomic DNA (2 and 4 μg) was doubly digested with Cla I and Pvu II. This digestion produces one copy of a 1.25-kb fragment and two copies of an 0.95-kb fragment which can be detected by a PFR-specific hybridization probe (see Fig. 5). As a titration standard, the 2.2-kb Cla I fragment isolated from pTbPFR-A was digested with Pvu II, and a concentration series of this digest was loaded onto the gel beside the genomic DNA digests. After gel electrophoresis and transfer to nitrocellulose, the DNA was hybridized with the 2.2-kb Cla I fragment isolated from pTbPFR-A. The resulting hybridization pattern (Fig. 6 *B*) indicates that the 0.95-kb fragment, originating exclusively from the PFR coding region, is present in a frequency of about two copies per haploid genome equivalent. This estimate is based on a haploid genomic DNA content of 4×10^7 bp (i.e., 0.1 pg DNA) (Borst et al., 1982) and on the assumption of a diploid genome (Borst et al., 1982; Gibson et al., 1985). Assuming an equivalent extent of hybridization of 2 μg trypanosomal DNA and 550 pg of isolated Cla I fragment (Fig. 6 *B*, lane *4*), the results presented in Fig. 6 support the final conclusion that the PFR protein is coded for by a single locus that contains two closely linked and very similar coding regions.

DNA Sequence and the Deduced Amino Acid Sequence

From the physical map of the PFR locus given in Fig. 5, the 5-kb Hind II–Kpn I fragment was selected for sequencing. Its complete nucleotide sequence is given in Fig. 7.

The analysis of the sequence reveals the presence of two closely linked, tandemly repeated open reading frames (PFR-A and PFR-B) of identical length (1,800 bp) and of identical nucleotide sequence, each coding for the identical protein of 600 amino acids. Outside these open reading frames, the similarity between the two genes rapidly degenerates. The nucleotide sequence of the PFR gene fragment of the λTbPFR72 phage originally identified by the immunoscreening is represented by nucleotides 1,449–1,619 and 3,603–3,773 within the coding areas of PFR-A and PFR-B in Fig. 7.

The calculated molecular mass of the protein coded for by PFR-A and PFR-B (69.9 kD) and its calculated isoelectric point (5.87) correspond well with the values determined experimentally for nonreduced PFR protein (see above; Rindisbacher, L., and T. Seebeck, unpublished observations). Similarly, the calculated amino acid composition fully agrees with the values determined biochemically (Table I). Codon usage in the PFR open reading frames is similar to that of other trypanosomal genes, such as the tubulin genes (Kimmel et al., 1985) or fructose bisphosphate aldolase (Clayton, 1985). The trypanosomal codon usage differs from that of most eukaryotes by a strong preference for the argi-

```
                                                                    62
CAAAACTTGA AGACCACCAC ATTGAGGCGA CTGCGTGTTG GAGTCTACTT GAACCATTAA
                                                                   122
CTGAACACAG CGAGTTAGCG ATAACATGGC TGCAGGAATC ACTGTTTCAT ACTCCTCAAG
                                                                   182
GACTGCTCGC ATGGATCATC GCCTGCACTA ACAACAATTG CAACCAAACG GCTTTTGTGA
                                                                   242
TGGCTGCATG GGTGACACTT TTATGCCGCA CTGAAAAGGT AGGTGGGTCA TTCAATAAAC
                                                                   302
ATGTCCAACA ACATCTACGG TCCAAATGCA CAAAAAAACA AGTGGAACAA CTTCTGTTGC
                                                                   362
TTGATCACTG CAATTAAGAT CCAAAGAAAG CGCGGACTTG CATAGCGCAT CTACAAGGGA
                                                                   422
AAACGAAACG CATAAAAAAA ATTAATTATA AAAAAAATTA ATTATAAAAA AAAGAAATAA
                                                                   482
CTTCCACAAC CAAAATATAA AAAAAAACAT GAATACCACC ATATAAATCA ATCACTTTTT
                                                                   542
TTTCACCAAC ATATAGTATC AAATAATAAG AGAATTAACT TTGTAGATAA AGAAAGCAAT
                                                572                602
AAAGCATCA   ATG AGC GGA AAG GAA GTT GAA GGT GTT GTG AGT CCT GCG GAC CAG CAG CAG
            MET SER GLY LYS GLU VAL GLU GLY VAL VAL SER PRO ALA ASP GLN GLN GLN
                                                632                662
CCA GCC GTC CCG GAG GTA ACA GAT ATC ACG CTG GAG GCC GCC CGC AAG CAG AAA ATT CAC
PRO ALA VAL PRO GLU VAL THR ASP ILE THR LEU GLU ALA ALA ARG LYS GLN LYS ILE HIS
                                                692                722
AAC CTG AAG TTG AAG ACC GCC TGC CTT TCG AAT GAG GAA TAT GTC CAG GAC CTG CAC GTA
ASN LEU LYS LEU LYS THR ALA CYS LEU SER ASN GLU GLU TYR VAL GLN ASP LEU HIS VAL
                                                752                782
TCC GAG TGG AGT GAG ACG CAG AAG CAG AAG CTG CAG GCT GCA CAC GAG AAA GCG CAT GAA
SER GLU TRP SER GLU THR GLN LYS GLN LYS LEU GLN ALA ALA HIS GLU LYS ALA HIS GLU
                                                812                842
TTG CTT GCC TCA GTG GAG GGT GGG ACG AAG TGG AGC CTG ACA GAG GCG TAT GAC ATC AAG
LEU LEU ALA SER VAL GLU GLY GLY THR LYS TRP SER LEU THR GLU ALA TYR ASP ILE LYS
                                                872                902
AAG CTG ATG CGC GTC TGT GGT CTT GAG ATG TCT GTG CGT GAA CTG TAC AAG CCG GAG GAC
LYS LEU MET ARG VAL CYS GLY LEU GLU MET SER VAL ARG GLU LEU TYR LYS PRO GLU ASP
                                                932                962
AAG CCA CAG TTC ATG GAG ATT GTT GCA CTC AAG AAG ACA ATG AAC GAA CTG AAG CAA CAT
LYS PRO GLN PHE MET GLU ILE VAL ALA LEU LYS LYS THR MET ASN GLU LEU LYS GLN HIS
                                                992               1022
CAC AAC AAG ACT CGC ACG GTG TCT TTC ACC GGC ATG ATC GAC AAT GCC ATC GCC AAA CTG
HIS ASN LYS THR ARG THR VAL SER PHE THR GLY MET ILE ASP ASN ALA ILE ALA LYS LEU
                                               1052               1082
GAG AAA ATC GAA GAC GAA CTG CGC CGG TCC CAG CTC GAC GCT TCT GAG ATG GCG CAA GTT
GLU LYS ILE GLU ASP GLU LEU ARG ARG SER GLN LEU ASP ALA SER GLU MET ALA GLN VAL
                                               1112               1142
CCT GTG GCT GCA CTG AAG AAT ATT GAG GAC ACG ATG AAC GTG GCT GTT GTG CAG ACG GCT
PRO VAL ALA ALA LEU LYS ASN ILE GLU ASP THR MET ASN VAL ALA VAL VAL GLN THR ALA
                                               1172               1202
CTT CTT GGG AAC GAG GAG CAG ATC AAA GCC CAA CTT GCA GCC GTT GAG AAG GCG AAC GAA
LEU LEU GLY ASN GLU GLU GLN ILE LYS ALA GLN LEU ALA ALA VAL GLU LYS ALA ASN GLU
                                               1232               1262
ATC CGT AAT GTT GCC ATT GCC GAT GGT GAG ATG GCG ATT GCT GAG GAA CAG TAT TAC ATT
ILE ARG ASN VAL ALA ILE ALA ASP GLY GLU MET ALA ILE ALA GLU GLU GLN TYR TYR ILE
                                               1292               1322
AAG GCG CAG CTG TTG GAG CAC CTT GTG GAG CTT GTG GCC GAC AAG TTT CGC ATC ATT GGG
LYS ALA GLN LEU LEU GLU HIS LEU VAL GLU LEU VAL ALA ASP LYS PHE ARG ILE ILE GLY
                                               1352               1382
CAA ACT GAG GAT GAG AAT AAG AGC TTC AGT AAG ATC CAC GAG GTA CAG AAG AAG TCA TTT
GLN THR GLU ASP GLU ASN LYS SER PHE SER LYS ILE HIS GLU VAL GLN LYS LYS SER PHE
                                               1412               1442
CAG GAA TCT GCC TCA ATC AAG GAC GCG AAG CGC CGC CTT AAG CAA CAC TGC GAG GAC GAC
```

Fig. 7

```
GLN GLU SER ALA SER ILE LYS ASP ALA LYS ARG ARG LEU LYS GLN HIS CYS GLU ASP ASP
                                    1472                                    1502
CTA CGT AAC CTT CAC GAT GCC ATC CAG AAA GCT GAC TTG GAG GAC GCC GAA GCC ATG AAA
LEU ARG ASN LEU HIS ASP ALA ILE GLN LYS ALA ASP LEU GLU ASP ALA GLU ALA MET LYS
                                    1532                                    1562
CGG TTC GCC ACG CAG AAG GAG AAG TCG GAG CGG TTC ATC CAC GAG AAC CTC GAC AAA CAG
ARG PHE ALA THR GLN LYS GLU LYS SER GLU ARG PHE ILE HIS GLU ASN LEU ASP LYS GLN
                                    1592                                    1622
GAC GAG GCA TGG CGT CGC ATT CAG GAA CTG GAG CGC GTG TTG CAG CGC CTT GGG ACG GAG
ASP GLU ALA TRP ARG ARG ILE GLN GLU LEU GLU ARG VAL LEU GLN ARG LEU GLY THR GLU
                                    1652                                    1682
CGT TTT GAA GAG GTG AAG CGC CGT ATT GAG GAG AAC GAC CGC GAG GAG AAG CGT AAG GTG
ARG PHE GLU GLU VAL LYS ARG ARG ILE GLU GLU ASN ASP ARG GLU GLU LYS ARG LYS VAL
                                    1712                                    1742
GAG TAC CAA CAG TTC CTC GAT GTA TGT GGC CAG CAT AAA AAG CTG CTG GAA CTG TCT GTG
GLU TYR GLN GLN PHE LEU ASP VAL CYS GLY GLN HIS LYS LYS LEU LEU GLU LEU SER VAL
                                    1772                                    1802
TAC AAC TGC GAC CTT GCG CTT CGC TGC ATG GGT ATG CTG GAG GAG ATC GTA GCC GAG GGC
TYR ASN CYS ASP LEU ALA LEU ARG CYS MET GLY MET LEU GLU GLU ILE VAL ALA GLU GLY
                                    1832                                    1862
TGC AGT GCC GTC AAG TCA CGC CAT GAC AAG ACG AAC GAT GAG TTG TCT GAC CTT CGG CTG
CYS SER ALA VAL LYS SER ARG HIS ASP LYS THR ASN ASP GLU LEU SER ASP LEU ARG LEU
                                    1892                                    1922
CAG GTG CAC CAG GAG TAC CTG GAG GCA TTC CGT CGC CTG TAC AAA ACT CTT GGC CAG CTT
GLN VAL HIS GLN GLU TYR LEU GLU ALA PHE ARG ARG LEU TYR LYS THR LEU GLY GLN LEU
                                    1952                                    1982
GTG TAC AAG AAA GAA AAG CGC CTG GAG GAG ATT GAT CGC AAC ATC CGC ACC ACA CAC ATT
VAL TYR LYS LYS GLU LYS ARG LEU GLU GLU ILE ASP ARG ASN ILE ARG THR THR HIS ILE
                                    2012                                    2042
CAA CTG GAG TTT GCC ATT GAG ACC TTT GAC CCC AAC GCG AAA CTA CAC TCC GAC AAG AAG
GLN LEU GLU PHE ALA ILE GLU THR PHE ASP PRO ASN ALA LYS LEU HIS SER ASP LYS LYS
                                    2072                                    2102
AAA GAC CTA TAC AAA CTT CGT GCG CAG GTG GAG GAA GAG TTG GAG ATG CTG AAG GAC AAG
LYS ASP LEU TYR LYS LEU ARG ALA GLN VAL GLU GLU GLU LEU GLU MET LEU LYS ASP LYS
                                    2132                                    2162
ATG GCG CAG GCG TTG GAG ATG TTT GGA CCT ACT GAG GAT GCG CTG AAC CAG GCT GGT ATC
MET ALA GLN ALA LEU GLU MET PHE GLY PRO THR GLU ASP ALA LEU ASN GLN ALA GLY ILE
                                    2192                                    2222
GAT TTT GTT CAC CCT GCT GAG GAG GTT GAG TCC GGC AAC ATG GAT CGC CGC AGC AAG ATG
ASP PHE VAL HIS PRO ALA GLU GLU VAL GLU SER GLY ASN MET ASP ARG ARG SER LYS MET
                                    2252                                    2282
GTG GAG TAC CGT GCA CAC CTG GCG AAG CAG GAG GAG GTG AAG ATT GCC GCG GAG CGC GAG
VAL GLU TYR ARG ALA HIS LEU ALA LYS GLN GLU GLU VAL LYS ILE ALA ALA GLU ARG GLU
                                    2312                                    2342
GAG CTG AAA CGA TCT AAG ATG CTC CAG AGC CAG CAG TAC CGC GGC CGC ACG ATG CCG CAG
GLU LEU LYS ARG SER LYS MET LEU GLN SER GLN GLN TYR ARG GLY ARG THR MET PRO GLN
                                                                    2402
ATC ACT CAG TAG CGCTGCGC TTAAATGTCT TTCATTATAA TCAATGTATA ACCTTTATGT
ILE THR GLN END
                                                                    2462
AGTATTTCAA TCTATGCCGC TGTGTACGTG CACTGCGGTG CCTATCCTTC GGCATTAGAG
                                                                    2522
AGTCACTGTT TGTGTAGATC GTAGCTGCAT GTCTGACATC TGGTTCATTA GTGCCGTTTT
                                                                    2582
GAGTCTGCAC TGTTGTGCAA CTTGATATGC ATCACTGCAC ACCAACCGTGT GTTTCCTCC
                                                                    2642
TTTGGGTGCA CCCTTTAATC CTTGTCTTCT CCTTTTTTGT CTCTTTCCCCC TCGAAAAGG
                                                                    2702
TGTCAAACTA CTGCCGCATA AACTACGGTT ATCCCAAATT TAAGAGAAAGC AATAAAGCA
                                    2732                                    2762
TCA ATG AGC GGA AAG GAA GTT GAA GGT GTT GTG AGT CCT GCG GAC CAG CAG CAG CCA GCC
    MET SER GLY LYS GLU VAL GLU GLY VAL VAL SER PRO ALA ASP GLN GLN GLN PRO ALA
                                    2792                                    2822
GTC CCG GAG GTA ACA GAT ATC ACG CTG GAG GCC GCC CGC AAG CAG AAA ATT CAC AAC CTG
VAL PRO GLU VAL THR ASP ILE THR LEU GLU ALA ALA ARG LYS GLN LYS ILE HIS ASN LEU
                                    2852                                    2882
```

Fig. 7

```
AAG TTG AAG ACC GCC TGC CTT TCG AAT GAG GAA TAT GTC CAG GAC CTG CAC GTA TCC GAG
LYS LEU LYS THR ALA CYS LEU SER ASN GLU GLU TYR VAL GLN ASP LEU HIS VAL SER GLU
                                    2912                                    2942
TGG AGT GAG ACG CAG AAG CAG AAG CTG CAG GCT GCA CAC GAG AAA GCG CAT GAA TTG CTT
TRP SER GLU THR GLN LYS GLN LYS LEU GLN ALA ALA HIS GLU LYS ALA HIS GLU LEU LEU
                                    2972                                    3002
GCC TCA GTG GAG GGT GGG ACG AAG TGG AGC CTG ACA GAG GCG TAT GAC ATC AAG AAG CTG
ALA SER VAL GLU GLY GLY THR LYS TRP SER LEU THR GLU ALA TYR ASP ILE LYS LYS LEU
                                    3032                                    3062
ATG CGC GTC TGT GGT CTT GAG ATG TCT GTG CGT GAA CTG TAC AAG CCG GAG GAC AAG CCA
MET ARG VAL CYS GLY LEU GLU MET SER VAL ARG GLU LEU TYR LYS PRO GLU ASP LYS PRO
                                    3092                                    3122
CAG TTC ATG GAG ATT GTT GCA CTC AAG AAG ACA ATG AAC GAA CTG AAG CAA CAT CAC AAC
GLN PHE MET GLU ILE VAL ALA LEU LYS LYS THR MET ASN GLU LEU LYS GLN HIS HIS ASN
                                    3152                                    3182
AAG ACT CGC ACG GTG TCT TTC ACC GGC ATG ATC GAC AAT GCC ATC GCC AAA CTG GAG AAA
LYS THR ARG THR VAL SER PHE THR GLY MET ILE ASP ASN ALA ILE ALA LYS LEU GLU LYS
                                    3212                                    3242
ATC GAA GAC GAA CTG CGC CGG TCC CAG CTC GAC GCT TCT GAG ATG GCG CAA GTT CCT GTG
ILE GLU ASP GLU LEU ARG ARG SER GLN LEU ASP ALA SER GLU MET ALA GLN VAL PRO VAL
                                    3272                                    3302
GCT GCA CTG AAG AAT ATT GAG GAC ACG ATG AAC GTG GCT GTT GTG CAG ACG GCT CTT CTT
ALA ALA LEU LYS ASN ILE GLU ASP THR MET ASN VAL ALA VAL VAL GLN THR ALA LEU LEU
                                    3332                                    3362
GGG AAC GAG GAG CAG ATC AAA GCC CAA CTT GCA GCC GTT GAG AAG GCG AAC GAA ATC CGT
GLY ASN GLU GLU GLN ILE LYS ALA GLN LEU ALA ALA VAL GLU LYS ALA ASN GLU ILE ARG
                                    3392                                    3422
AAT GTT GCC ATT GCC GAT GGT GAG ATG GCG ATT GCT GAG GAA CAG TAT TAC ATT AAG GCG
ASN VAL ALA ILE ALA ASP GLY GLU MET ALA ILE ALA GLU GLU GLN TYR TYR ILE LYS ALA
                                    3452                                    3482
CAG CTG TTG GAG CAC CTT GTG GAG CTT GTG GCC GAC AAG TTT CGC ATC ATT GGG CAA ACT
GLN LEU LEU GLU HIS LEU VAL GLU LEU VAL ALA ASP LYS PHE ARG ILE ILE GLY GLN THR
                                    3512                                    3542
GAG GAT GAG AAT AAG AGC TTC AGT AAG ATC CAC GAG GTA CAG AAG AAG TCA TTT CAG GAA
GLU ASP GLU ASN LYS SER PHE SER LYS ILE HIS GLU VAL GLN LYS LYS SER PHE GLN GLU
                                    3572                                    3602
TCT GCC TCA ATC AAG GAC GCG AAG CGC CGC CTT AAG CAA CAC TGC GAG GAC GAC CTA CGT
SER ALA SER ILE LYS ASP ALA LYS ARG ARG LEU LYS GLN HIS CYS GLU ASP ASP LEU ARG
                                    3632                                    3662
AAC CTT CAC GAT GCC ATC CAG AAA GCT GAC TTG GAG GAC GCC GAA GCC ATG AAA CGG TTC
ASN LEU HIS ASP ALA ILE GLN LYS ALA ASP LEU GLU ASP ALA GLU ALA MET LYS ARG PHE
                                    3692                                    3722
GCC ACG CAG AAG GAG AAG TCG GAG CGG TTC ATC CAC GAG AAC CTC GAC AAA CAG GAC GAG
ALA THR GLN LYS GLU LYS SER GLU ARG PHE ILE HIS GLU ASN LEU ASP LYS GLN ASP GLU
                                    3752                                    3782
GCA TGG CGT CGC ATT CAG GAA CTG GAG CGC GTG TTG CAG CGC CTT GGG ACG GAG CGT TTT
ALA TRP ARG ARG ILE GLN GLU LEU GLU ARG VAL LEU GLN ARG LEU GLY THR GLU ARG PHE
                                    3812                                    3842
GAA GAG GTG AAG CGC CGT ATT GAG GAG AAC GAC CGC GAG GAG AAG CGT AAG GTG GAG TAC
GLU GLU VAL LYS ARG ARG ILE GLU GLU ASN ASP ARG GLU GLU LYS ARG LYS VAL GLU TYR
                                    3872                                    3902
CAA CAG TTC CTC GAT GTA TGT GGC CAG CAT AAA AAG CTG CTG GAA CTG TCT GTG TAC AAC
GLN GLN PHE LEU ASP VAL CYS GLY GLN HIS LYS LYS LEU LEU GLU LEU SER VAL TYR ASN
                                    3932                                    3962
TGC GAC CTT GCG CTT CGC TGC ATG GGT ATG CTG GAG GAG ATC GTA GCC GAG GGC TGC AGT
CYS ASP LEU ALA LEU ARG CYS MET GLY MET LEU GLU GLU ILE VAL ALA GLU GLY CYS SER
                                    3992                                    4022
GCC GTC AAG TCA CGC CAT GAC AAG ACG AAC GAT GAG TTG TCT GAC CTT CGG CTG CAG GTG
ALA VAL LYS SER ARG HIS ASP LYS THR ASN ASP GLU LEU SER ASP LEU ARG LEU GLN VAL
                                    4052                                    4082
CAC CAG GAG TAC CTG GAG GCA TTC CGT CGC CTG TAC AAA ACT CTT GGC CAG CTT GTG TAC
HIS GLN GLU TYR LEU GLU ALA PHE ARG ARG LEU TYR LYS THR LEU GLY GLN LEU VAL TYR
                                    4112                                    4142
AAG AAA GAA AAG CGC CTG GAG GAG ATT GAT CGC AAC ATC CGC ACC ACA CAC ATT CAA CTG
LYS LYS GLU LYS ARG LEU GLU GLU ILE ASP ARG ASN ILE ARG THR THR HIS ILE GLN LEU
                                    4172                                    4202
```

Fig. 7

```
GAG TTT GCC ATT GAG ACC TTT GAC CCC AAC GCG AAA CTA CAC TCC GAC AAG AAG AAA GAC
GLU PHE ALA ILE GLU THR PHE ASP PRO ASN ALA LYS LEU HIS SER ASP LYS LYS LYS ASP
                                    4232                                    4262
CTA TAC AAA CTT CGT GCG CAG GTG GAG GAA GAG TTG GAG ATG CTG AAG GAC AAG ATG GCG
LEU TYR LYS LEU ARG ALA GLN VAL GLU GLU GLU LEU GLU MET LEU LYS ASP LYS MET ALA
                                    4292                                    4322
CAG GCG TTG GAG ATG TTT GGA CCT ACT GAG GAT GCG CTG AAC CAG GCT GGT ATC GAT TTT
GLN ALA LEU GLU MET PHE GLY PRO THR GLU ASP ALA LEU ASN GLN ALA GLY ILE ASP PHE
                                    4352                                    4382
GTT CAC CCT GCT GAG GAG GTT GAG TCC GGC AAC ATG GAT CGC CGC AGC AAG ATG GTG GAG
VAL HIS PRO ALA GLU GLU VAL GLU SER GLY ASN MET ASP ARG ARG SER LYS MET VAL GLU
                                    4412                                    4442
TAC CGT GCA CAC CTG GCG AAG CAG GAG GAG GTG AAG ATT GCC GCG GAG CGC GAG GAG CTG
TYR ARG ALA HIS LEU ALA LYS GLN GLU GLU VAL LYS ILE ALA ALA GLU ARG GLU GLU LEU
                                    4472                                    4502
AAA CGA TCT AAG ATG CTC CAG AGC CAG CAG TAC CGC GGC CGC ACG ATG CCG CAG ATC ACT
LYS ARG SER LYS MET LEU GLN SER GLN GLN TYR ARG GLY ARG THR MET PRO GLN ILE THR
                                                                    4562
CAG TAG GATT GTGTACTGTA ATTGTATTTT TTGGTTTTTT TTTGAAAGTG GTAGTAGTAT
GLN END
                                                                    4622
TTTAGAGTTG CTACTTACTT ATGAGTTTGA TATGTTTGTT TTTATGTATA ATTTTTTTTT
                                                                    4682
TGAGAAACTG ATGTTTATAT GTATGTGTAT CTTCTTATTT GCGGTTAGAT TAGACGATTT
                                                                    4742
TTAGTGTATT TTTAGTTGGC TGGGGGTTAT TTGGCTGCTA TTTGTTGTAG GATGCGCCGT
                                                                    4802
GCTGACGCTT CTTGCTCTTT TATGTCGCCG AGTTATGCTG ATTGTTGCTC TCTCTATATT
                                                                    4862
TCTTTTGTGC GGTGGTGTTG TTGAAAATTT TTGCTGCTAA TATTGTTGTC ATGTATTTTG
                                                                    4922
TTGTTGTTAC TTGTCTGTGA GTGGTTTCTT TCTTTCTTTT TTTCACTGTG TTTTGTAATA
                                                                    4982
GTATCTTACT TTAGCTTAGG ATTCTAGGTT TGGAGGGATG TGTTATTGCC TGATCAAATA
                                                                    5042
TTTCTGTTGG AGTTGTTAGG TTACTGTTGC ACGTGACTTC CGTATATATC CGTGTTGGCT
                                                                    5102
TTTTCTTTTT GAGTTAAAGT TCTTTTGTTT TGGTATAATC GTGTGTGTGG GATCTGCTGC
                                        5135
ATGTAATGAC CGGATGGTCG AGCCGACGTG GTA
```

Figure 7. Sequence of the PFR genes A and B. Two homologous sequences of 1800 bp code for 600 amino acids with a molecular mass of 69.6 kD. The coding sequence from λ TbPFR72 is underlined. This sequence is accessible in the EMBL/GenBank/DDBJ nucleotide sequence databases under accession number X14819.

nine codons CGC and CGT and the avoidance of AUA and UUA as codons for isoleucin and leucine, respectively.

Secondary structure prediction indicates that the protein coded for by PFR-A and PFR-B may assume a mostly helical conformation throughout its entire length (>80% helicity predicted by the method of Garnier et al., 1978). However, no evidence was found throughout the sequence for a repeating heptad motif which is characteristic for the helical domains of many filamentous proteins and which enables them to form coiled coil structures due to hydrophobic interactions between opposing amino acids along the helix surface (McLachlan, 1984). The amino acid distribution along most of the helical domains of PFR would allow the formation of ion pairs along the helix at a high frequency, which would significantly contribute to a stabilization of this structure (Sundaralingam et al., 1985). This observation further suggests that the PFR protein is in fact largely helical in its native configuration. Amino acids 335–355 may form a calmodulin-binding site (Ericksson-Viitanen and DeGrado, 1987). Similarity searches of the protein sequence library (PIR release 18) revealed that the COOH terminus of PFR (amino acids 587–600; SQQYRGRTMPQITQ) is closely similar to a highly conserved sequence found in all β-tubulins around tyrosine residue 281. In β-tubulin, this region may be involved in dimer recognition and polymerization (Rudolph et al., 1987; Fridovitch-Keil et al., 1987). Its function on PFR remains to be elucidated.

A nucleotide sequence survey of the GenBank/EMBL DNA sequence libraries revealed identities of 77 and 43 bp for PFR-A and PFR-B, respectively, with a previously published partial cDNA sequence from *T. brucei* (pSLc1; Parsons et al., 1984). The pSLc1 sequence was shown to represent the 5′ terminus of a then unidentified mRNA carrying the mini-exon sequence. The 77 nucleotides after the mini-exon, whose sequence had been determined by Parsons et al., are almost identical to nucleotides 499–577 of the PFR-A sequence (including the first eight codons of PFR-A). The only difference between the two sequences is at position 571

(PFR-A) and 2,725 (PFR-B), where Parson's sequence contains a G, while the sequence presented in Fig. 7 contains an A. The shorter extent of similarity between pSLc1 and PFR-B is due to the degeneration of similarity between PFR-A and PFR-B upstream of the coding sequence. Thus, the similarity between pSLc1 and the PFR locus indicates that pSLc1 represents the mRNA transcribed from PFR-A.

PFR-A and PFR-B Code for Distinct mRNAs

To identify potential mRNAs derived from PFR-A and PFR-B, total trypanosomal RNA was analyzed by primer extension sequencing. The two primers were selected to represent the divergent upstream regions of each gene to allow the identification of transcripts from each gene (Fig. 8). The results summarized in Fig. 8 indicate that trypanosomes contain stable mRNAs derived from both genes, PFR-A and PFR-B. Both PFR mRNA species carry the mini-exon sequence at their 5′ termini. The sequence of PFR-A mRNA is identical to the one reported earlier as the 5′ terminus of an unidentified mRNA (Parsons et al., 1984). The nucleotide sequences of the two mRNAs diverged along most of the 5′ leader, while the sequences of the last 19 nucleotides immediately upstream of the initiator AUG, as well as the entire coding sequence, are identical. While the length of the two 5′ leader sequences differs by only 11 nucleotides, Northern blotting experiments (not shown) indicate that also the entire length of both mRNAs must be very similar.

PFR Protein Synthesized In Vitro Migrates as a Doublet in Reduced Form

Much published evidence, as well as our own experiments (e.g. Fig. 4), suggested that the PFR protein migrates as two distinct bands in SDS-gel electrophoresis. However, the evidence presented in Fig. 2 strongly indicates that the more slowly migrating of the two bands represents a reduced derivative of the single PFR protein. This conclusion was now reinvestigated, and confirmed, with PFR protein synthesized in vitro. PFR mRNA was isolated by preparative hybridization of total trypanosomal RNA to filter-bound pTbPFRA DNA. The hybrid-selected mRNA was released from the filters and translated in a rabbit reticulocyte lysate. Total translation products were then solubilized in reducing sample buffer, mixed with purified, unlabeled PFR proteins as internal markers, and fractionated by two-dimensional gel electrophoresis. After transfer to nitrocellulose filters, the marker PFR proteins were visualized by immunostaining, and the filter was then exposed for autoradiographic detection of the PFR proteins synthesized in vitro. The results given in Fig. 9 demonstrate that in vitro synthesis results in two radioactive spots that comigrate with the two spots formed by the authentic marker PFR protein. Thus, PFR protein synthesized in vitro from hybrid-selected mRNA displays the same migration behavior as does the PFR protein extracted from trypanosome cells. The experiment illustrated in Fig. 9 thus confirms that the two PFR protein species observed under reducing conditions represent two conformers of a single polypeptide chain, which is coded for by the genetic locus whose detailed structure has been reported in this study.

Discussion

The PFR structure is a highly ordered three-dimensional network of fibrous proteins which is a specific structural component of the flagella of trypanosomatids and euglenoids. This common structure may reflect a common pathway of evolution of these two groups of flagellates, which are currently thought to have split from the mainstream of eukary-

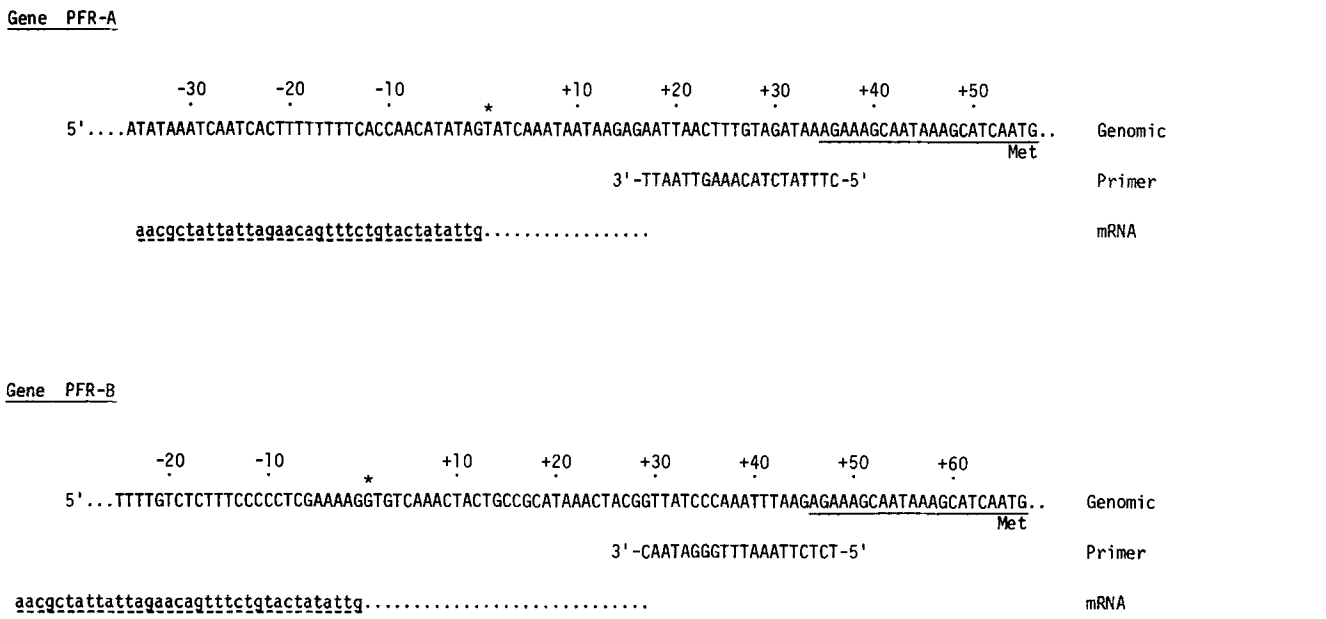

Figure 8. Primer extension analysis of 5′ termini of the PFR mRNAs. The 20-mer primers complementary to the coding strands are indicated. Identical sequences in the noncoding regions of the PFR-A and PFR-B genes are underlined. The asterisk designates the nucleotide to which the mini-exon is joined in the mRNA. The mRNA sequence (identical to the genomic coding strand) is indicated by dots. The sequence of the mini-exon, shown by dashed underline, is from Parsons et al. (1984).

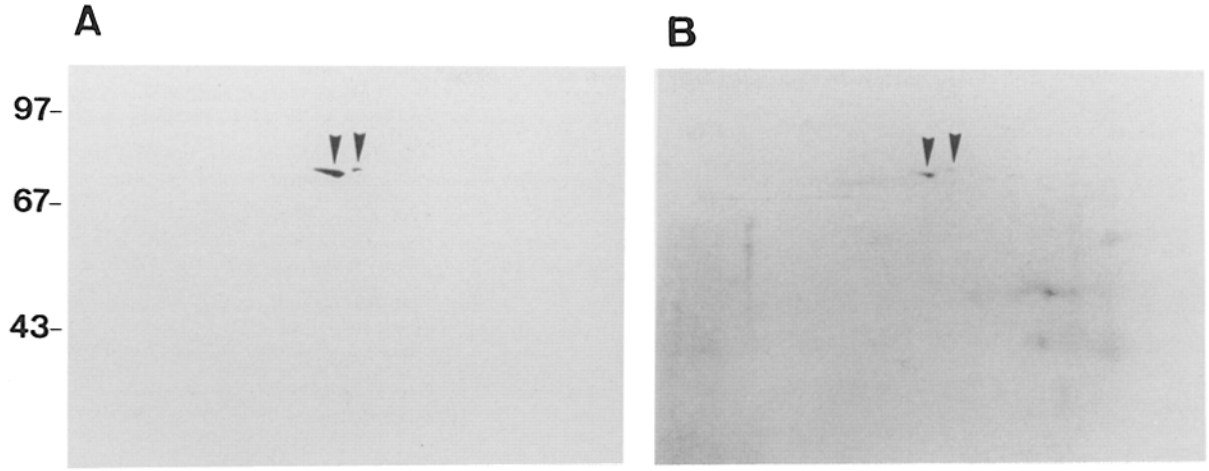

Figure 9. mRNA hybrid selected on pTbPFR-A and translated in vitro gives rise to two proteins that comigrate with authentic PFR in two-dimensional gel electrophoresis. (*A*) Immunoblot of the ^{35}S-labeled in vitro translation products mixed with purified, unlabeled PFR protein (1 μg). (*B*) Autoradiograph of the same filter. Gels are represented with the acidic end to the left.

otic evolution very early on (Sogin et al., 1986). Though the structural details of the PFR differ between the two groups (Hyams, 1982), immunological evidence has been presented that their major proteins may be very similar. The present study represents an analysis of the major structural protein (PFR protein) of the PFR of *T. brucei*.

Antibodies raised against PFR protein were used to isolate the corresponding genes from an expression library. The PFR protein is coded for by a single locus of the trypanosomal genome. DNA sequence analysis of this locus revealed the presence of two closely spaced open reading frames of 1,800 nucleotides length, each, for PFR-A and PFR-B. The nucleotide sequence of both reading frames is identical and, thus, both of them code for an identical protein of 600 amino acids. Outside the coding regions, the similarity between the two genes rapidly degenerates. Upstream of the initial AUG codon, the sequences of PFR-A and PFR-B are identical for an additional 19 bp, but then become dissimilar. No similarity whatsoever is found after the TAG termination codons. A survey of DNA sequence libraries revealed that a previously published 77-bp sequence from the 5′ terminus of a then unidentified mini-exon mRNA (cDNA clone pSLc1; Parsons et al., 1984) is identical to the 5′ terminus of PFR-A, including the first eight codons of this gene.

Primer extension sequencing of the PFR mRNAs revealed that stable mRNAs are generated both from the PFR-A and the PFR-B genes. The mRNAs contain the mini-exon sequence at the 5′ terminus. The sequence of the 5′ terminus of the PFR-A mRNA is identical to the sequence published by Parsons et al. (1984) (see above).

The amino acid composition of PFR calculated from the derived amino acid sequence corresponds very well with that determined biochemically for isolated PFR protein. Also, the calculated molecular mass and isoelectric point are very similar to the values determined experimentally. Secondary structure prediction of the protein indicates a very high content of α-helix. This is in good agreement with the marked resistance of the major part of the protein to protease digestion. The PFR protein bears remarkable similarity to intermediate filament proteins in that (*a*) it is insoluble under a variety of salt and detergent conditions; (*b*) its overall amino acid composition is similar; and (*c*) it has a high content of helical regions. In contrast to intermediate filament proteins, no evidence was found in the sequence of PFR protein for a heptadic arrangement of hydrophobic amino acids, which is characteristic for many helical proteins assuming a coiled coil configuration (McLachlan, 1984). A PFR protein sequence library search indicated that the COOH terminus of PFR is similar to a highly conserved region found in all β-tubulins around tyrosine residue 281. While in β-tubulin this region may be involved in polymerization, its functional significance in the PFR protein remains to be established.

The PFR gene locus codes for a single polypeptide of 600 amino acids, corresponding to a calculated molecular mass of 69 kD. This is at variance with the observation in cell extracts of two distinct, approximately equimolar, bands that migrate on denaturing gels with apparent molecular masses of 73 and 69 kD. This migration is unaltered after exposure to reducing conditions, such as incubation with DTT or tributylphosphine, whereas oxidation leads to a single band migrating at an apparent molecular mass of 69 kD. Translation of hybrid-selected PFR mRNA in vitro similarly generates two polypeptides with different migration properties. Upon two-dimensional gel electrophoresis, the translation products comigrate with authentic PFR protein isolated from cells, forming two distinct spots. This migration pattern of PFR protein synthesized in vitro confirms that the two spots (or bands in one-dimensional electrophoresis) observed by many workers (for review see Cachon et al., 1988) represent in fact two conformations of a single polypeptide chain. In addition, the in vitro translation experiments serve to rule out the possible existence of a second set of PFR genes that might potentially code for the 73-kD protein, but whose nucleotide sequence might have diverged so far as to be no more detectable by hybridization with the PFR gene described in this study. Furthermore, a possible generation of the 73-kD PFR protein species from the 69-kD variety through posttranslational modification is also rendered unlikely by

the results of in vitro translation. In summary, the 73- and 69-kD variants of PFR most likely represent different conformations of the single polypeptide coded for by the PFR locus.

The PFR protein represents a novel type of cytoskeletal protein that is restricted to the trypanosomatids and, most likely, the euglenoids. Considering the fact that the trypanosomatids are parasitic organisms, this parasite-specific structure might not only be of interest in terms of cellular architecture and function, but it also may represent a potential, highly parasite-specific target for trypanocidal drugs.

We are very grateful to U. Kurath for skillful technical assistance and R. Gardi for fine photography work. We would like to thank V. Küng for constructing the gt 11-library of *T. brucei*; A. Schneider, K. Behrens, and T. Wyler for excellent electron microscopy work; L. Rindisbacher for two-dimensional electrophoretic analyses; J. Schaller and U. Kaempfer (Department of Biochemistry, University of Bern, Bern, Switzerland) for determining amino acid compositions; Albert C. W. A. Cornelissen (Max Planck Institute, Tuebingen, FRG) for a generous gift of polyA$^+$ mRNA of bloodstream trypanosomes; and Hans Trachsel (Department of Biochemistry and Molecular Biology, University of Bern, Bern, Switzerland) for many valuable comments on the work and the manuscript.

This study was supported by a grant from the Swiss National Science Foundation and a fellowship to K. Schlaeppi from the Roche Research Foundation.

Received for publication 31 March 1989 and in revised form 13 June 1989.

References

Bartnik, E., M. Osborn, and K. Weber. 1985. Intermediate filaments in non-neuronal cells of invertebrates: isolation and biochemical characterization of intermediate filaments from the esophageal epithelium of the mollusc *Helix pomatia*. *J. Cell Biol.* 101:427–440.

Borst, P., L. H. T. Van der Ploeg, J. M. Van Hoek, J. Tas, and J. James. 1982. On the DNA content and ploidy of trypanosomes. *Mol. Biochem. Parasitol.* 6:13–23.

Brun, R., and M. Schönenberger. 1979. Cultivation and in vitro cloning of procyclic culture forms of *Trypanosoma brucei* in semi-defined medium. *Acta Trop.* 36:289–292.

Cachon, J., M. Cachon, M.-P. Cosson, and J. Cosson. 1988. The paraflagellar rod: a structure in search of a function. *Biol. Cell.* 63:169–181.

Clayton, C. E. 1985. Structure and regulated expression of genes encoding fructose biphosphate aldolase in *Trypanosoma brucei*. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2997–3003.

Cunha, V. L., W. de Souza, and A. Hassen-Voloch. 1984. Isolation of the flagellum and characterization of the paraxial structure of *Herpetomonas megaseliae*. *J. Submicrosc. Cytol.* 16:705–713.

Dente, L., G. Cesareni, and R. Cortese. 1983. pEMBL: a new family of single stranded plasmids. *Nucleic Acids Res.* 11:1645–1655.

DeSouza, W., and T. Souto-Padron. 1980. The paraxial structure of the flagellum of trypanosomatidae. *J. Parasitol.* 66:229–235.

Ericksson-Viitanen, S., and W. F. DeGrado. 1987. Recognition and characterization of calmodulin-binding sequences in peptides and protein. *Methods Enzymol.* 139:455–475.

Evans, D. A., D. S. Ellis, and S. Stamford. 1979. Ultrastructural studies of certain aspects of the development of *Trypanosoma congolense* in *Glossina morsitans*. *J. Protozool.* 26:557–563.

Farina, M., M. Attias, T. Souto-Padron, and W. DeSouza. 1986. Further studies on the organization of the paraxial rod of trypanosomatids. *J. Protozool.* 33:552–557.

Freymuller, E., and E. P. Camargo. 1981. Ultrastructural differences between species of trypanosomatids with and without endosymbionts. *J. Protozool.* 28:175–182.

Fridovitch-Keil, J. L., J. F. Bond, and F. Solomon. 1987. Domains of β-tubulin essential for conserved functions in vivo. *Mol. Cell. Biol.* 7:3792–3798.

Frischauf, A. M., H. Lehrach, A. Poustka, and N. Murray. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170: 827–842.

Fuge, H. 1969. Electron microscopic studies on the intraflagellar structures of trypanosomes. *J. Protozool.* 16:460–466.

Gallo, J.-M., and J. Schrevel. 1985. Homologies between paraflagellar rod proteins from trypanosomes and euglenoids revealed by a monoclonal antibody. *Eur. J. Cell Biol.* 36:163–168.

Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97–120.

Geisler, N., and K. Weber. 1981. Comparison of the proteins of two immunologically distinct intermediate-sized filaments by amino acid sequence analysis: desmin and vimentin. *Proc. Natl. Acad. Sci. USA.* 78:4120–4123.

Gibson, W. C., K. A. Osinga, P. A. M. Michels, and P. Borst. 1985. Trypanosomes of subgenus trypanozoon are diploid for housekeeping genes. *Mol. Biochem. Parasitol.* 16:231–242.

Goldstein, S. F., M. E. Holwill, and N. R. Silvester. 1970. The effects of laser microbeam irradiation on the flagellum of *Crithidia (Stigomonas) oncopelti*. *J. Exp. Biol.* 58:401–409.

Goldstein, L. S. B., R. A. Laymon, and J. R. McIntosh. 1986. A microtubule associated protein in *Drosophila melanogaster*: identification, characterization, and isolation of coding sequences. *J. Cell Biol.* 102:2076–2087.

Hanukoglu, I., and E. Fuchs. 1982. The cDNA sequence of a human epidermal keratin: divergence of sequence but conservation of structure among intermediate filament proteins. *Cell.* 31:243–252.

Huynh, T. V., R. A. Young, and R. W. Davis. 1985. Constructing and screening cDNA libraries in lambda gt 10 and lambda gt 11. *In* DNA Cloning. D. M. Glover, editor. IRL Press Limited, Oxford. 49–77.

Hyams, J. S. 1982. The euglena paraflagellar rod: structure, relationship to the other flagellar components and preliminary biochemical characterization. *J. Cell Sci.* 55:199–210.

Imboden, M., B. Blum, T. DeLange, R. Braun, and T. Seebeck. 1986. Tubulin mRNA of *Trypanosoma brucei*. *J. Mol. Biol.* 188:393–402.

Imboden, M. A., P. W. Laird, M. Affolter, and T. Seebeck. 1987. Transcription of the intergenic regions of the tubulin gene cluster of *Trypanosoma brucei*: evidence for a polycistranic transcription unit in a eukaryote. *Nucleic Acids Res.* 15:7357–7368.

Johnson, D. N., N. R. Silvester, and M. E. J. Holwill. 1979. An analysis of the shape and propagation of waves on the flagellum of *Crithidia oncopelti*. *J. Exp. Med.* 80:299–315.

Kaiser, K., and N. E. Murray. 1985. The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. *In* DNA Cloning. D. M. Glover, editor. IRL Press Limited, Oxford. 1–47.

Kimmel, B. E., S. Samson, J. Wu, R. Hirschberg, and L. R. Yarbrough. 1985. Tubulin genes of the african trypanosome *Trypanosoma brucei rhodesiense*: nucleotide sequence of a 3.7-kb fragment containing genes for alpha- and beta-tubulins. *Gene (Amst.).* 35:237–248.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the heads of bacteriophage T_4. *Nature (Lond.).* 227:680–685.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.

McLachlan, A. D. 1984. Structural implications of the myosin amino acid sequence. *Annu. Rev. Biophys. Biophys. Chem.* 13:167–189.

O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007–4021.

Ozaki, L. S., and Y. M. Traub-Cseko. 1984. Genomic DNA cloning and related techniques. *In* Genes and Antigenes of Parasites. C. M. Morel, editor. Fundacao Oswaldo Cruz, Rio de Janeiro, Brazil. 165–186.

Parnes, J. R., B. Velan, G. Felsenfeld, L. Ramanathan, U. Ferrini, E. Apella, and J. G. Seidman. 1981. Mouse β2-microglobulin cDNA clones: a screening procedure for cDNA clones corresponding to rare mRNAs. *Proc. Natl. Acad. Sci. USA.* 78:2253–2257.

Parsons, M., R. G. Nelson, K. P. Watkins, and N. Agabian. 1984. Trypanosome mRNAs share a common 5'spliced leader sequence. *Cell.* 38:309–316.

Piccinni, E., V. Albergoni, and O. Coppellotti. 1975. ATPase activity in flagella from *Euglena gracilis*: localization of the enzyme and effects of detergents. *J. Protozool.* 22:331–335.

Rudolph, J. E., M. Kimble, H. D. Hoyle, M. A. Subler, and E. C. Raff. 1987. Three *Drosophila* beta tubulin sequences: a developmentally regulated isoform (β3), the testis-specific isoform (β2) and an assembly-defective mutation of the testis-specific isoform ($\beta2t^8$) reveal both an ancient divergence in metazoan isotypes and structural constraints for beta tubulin function. *Mol. Cell. Biol.* 7:2231–2242.

Rüegg, U. Th., and Rudinger, J. 1977. Reductive cleavage of cystine disulfides with tributylphosphine. *Methods Enzymol.* 47:111–116.

Russell, D. G., R. J. Newsam, G. C. N. Palmer, and K. Gull. 1983. Structural and biochemical characterisation of the paraflagellar rod of *Crithidia fasciculata*. *Eur. J. Cell Biol.* 30:137–143.

Saborio, J. L., J. M. Hernandez, S. Narayanswami, R. Wrightsman, E. Palmer, and J. Manning. 1989. Isolation and characterization of paraflagellar rod protein from *Trypanosoma cruzi*. *J. Biol. Chem.* 264:4071–4075.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

Schneider, A., T. Sherwin, R. Sasse, D. G. Russell, K. Gull, and T. Seebeck. 1987. Subpellicular and flagellar microtubules of *T. brucei* contain the same α-tubulin isotypes. *J. Cell Biol.* 104:431–438.

Schneider, A., H. U. Lutz, R. Marugg, P. Gehr, and T. Seebeck. 1988. Spectrin-like proteins in the paraflagellar rod structure of *Trypanosoma brucei*. *J. Cell Sci.* 90:307–315.

Seebeck, T., P. A. Whittaker, M. A. Imboden, N. Hardman, and R. Braun. 1983. Tubulin genes of *Trypanosoma brucei*: a tightly clustered family of alternating genes. *Proc. Natl. Acad. Sci. USA.* 80:4634–4638.

Smith, G. E., and M. D. Summers. 1980. The bidirectional transfer of DNA

and RNA to nitrocellulose or diazobenzyloxymethyl-paper. *Anal. Biochem.* 109:123–129.

Sogin, M. L., H. J. Elwood, and J. H. Gunderson. 1986. Evolutionary diversity of eukaryotic small-subunit rRNA genes. *Proc. Natl. Acad. Sci. USA.* 83:1383–1387.

Souto-Padron, T., W. DeSouza, and J. E. Heuser. 1984. Quick-freeze, deep-etch rotary shadowing of *Trypanosoma cruzi* and *Herpetomonas megaseliae*. *J. Cell Sci.* 69:167–178.

Steinert, P. M., A. C. Sternen, and D. R. Roop. 1985. The molecular biology of intermediate filaments. *Cell.* 42:411–419.

Sundaralingam, M., W. Drendel, and M. Greaser. 1985. Stabilization of the long helix of troponin C by intrahelical salt bridges between charged amino acid side chains. *Proc. Natl. Acad. Sci. USA.* 82:7944–7947.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA.* 76:4350–4354.

Vickerman, K., and T. M. Preston. 1976. Comparative cell biology of the kinetoplastid flagellates. *In* Biology of the Kinetoplastida. W. C. A. Lumsden and D. A. Evans, editors. Academic Press Inc., New York. 35–130.

Wessel, D., and U. I. Flügge. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* 138:141–143.

Young, R. A., and R. W. Davis. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA.* 80:1194–1198.