Journal of Cell Science 105, 1101-1113 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993



# Molecular cloning and cellular localization of a BiP homologue in *Trypanosoma brucei*

## Divergent ER retention signals in a lower eukaryote

**James D. Bangs[1,*], Lyle Uyetake[1], Marla Jo Brickman[2], Andrew E. Balber[2] and John C. Boothroyd[1]**

[1]Department of Microbiology and Immunology, Stanford School of Medicine, Stanford, CA 94305-5402, USA
[2]Department of Immunology, Duke Universty Medical Center, Durham, NC 27710, USA

*Author for correspondence at present address: Department of Medical Microbiology and Immunology, University of Wisconsin-Madison Medical School, 1300 University Avenue, Madison, WI 53706, USA

## SUMMARY

**Using the polymerase chain reaction with degenerate primers, three new members of the *hsp70* gene family of *Trypanosoma brucei* have been identified. A genomic clone of one of these, *gA*, has been fully sequenced and the corresponding gene product has been characterized using antibody to recombinant gA fusion protein. *gA* is the trypanosomal homologue of *BiP*, an endoplasmic reticulum resident *hsp70* gene family member, based on four lines of evidence: (1) gA protein has 64% deduced amino acid identity with rat BiP; (2) the deduced amino acid sequence has a putative secretory signal peptide; (3) the gA gene product is a soluble luminal resident of a trypanosomal microsome fraction; (4) the gA polypeptide does not cofractionate with mitochondrial markers. Trypanosomes are the most primitive eukaryote yet in which BiP has been identified. The gA polypeptide has been used as a specific marker for the direct visualization of endoplasmic reticulum in trypanosomes by both indirect immunofluorescence and cryoimmuno electron microscopy. The endoplasmic reticulum is seen as a tubular network that extends throughout the cell excluding the flagellum. The C-terminal tetrapeptide of gA is MDDL, which, together with the C-terminal tetrapeptide (KQDL) of a trypanosome protein disulfide isomerase homologue (Hsu et al. (1989) *Biochemistry* 28, 6440-6446), indicates that endoplasmic reticulum retrieval signals in trypanosomes may be as divergent and heterogeneous as any seen in the other eukaryotes yet studied.**

Key words: trypanosomes, BiP/grp78, hsp70, endoplasmic reticulum

## INTRODUCTION

African trypanosomes, parasitic protozoa of the Order Kinetoplastida, are the causative agents of sleeping sickness in humans and nagana in cattle. As judged by 18 S rRNA sequence homology (Sogin et al., 1989), these organisms are among the most primitive of the eukaryotic kingdom, being as distant evolutionarily from yeast as they are from vertebrates. Nevertheless, trypanosomes have provided the first glimpse of eukaryotic processes that have alternatively proven to be either ubiquitous, such as glycosyl phosphatidylinositol membrane protein anchors (Doering et al., 1990), unusual, such as *trans*-splicing (Boothroyd, 1989), or even revolutionary, such as RNA editing (Simpson and Shaw, 1989). The basic cell biology of these organisms, however, has been understudied and, with the exceptions of endocytosis (Balber, 1990) and glycosome biogenesis (Opperdoes, 1990), little is known about protein trafficking in trypanosomes. Although the export of one polypeptide, the variant surface glycoprotein (VSG), has been studied in detail (Bangs et al., 1986) and a few components of the secretory machinery have been identified, including a putative protein disulfide isomerase homologue (Hsu et al., 1989), a critical gap remains in our understanding of a process that can be expected to play an important role in host-parasite interactions.

The first steps in eukaryotic protein export occur during and immediately following the translocation of nascent secretory polypeptides into the endoplasmic reticulum. An important participant in these steps is a protein known as BiP (binding protein) or grp78 (glucose-regulated protein), a soluble luminal resident of the endoplasmic reticulum (ER). BiP is a member of the hsp70 family of heat shock proteins and has been found in all eukaryotes examined to date, including rats (Munro and Pelham, 1986), nematodes (Heschl and Baillie, 1989), yeast (Rose et al., 1989; Normington et al., 1989) and the parasitic protozoan *Plasmodium falciparum*. (Kumar and Zheng, 1992). It functions as a molecular chaperone, assisting in the translocation of secretory polypeptides into the ER (Vogel et al., 1990), in the folding of translocated polypeptides (reviewed by Pelham, 1989) and in the correct assembly of macromole-

cular secretory complexes such as immunoglobulin (Bole et al., 1986). BiP facilitates these related processes by binding reversibly to exposed segments of substrate polypeptide chains (Flynn et al., 1989) and as a consequence of its binding properties it may also serve to prevent the secretion of aberrantly folded or assembled proteins (Bole et al., 1986; Dorner et al.,1987). Its functions have been shown to be essential, at least in yeast where gene knockout yields a non-viable phenotype (Rose et al., 1989; Normington et al., 1989).

BiP, like other soluble ER polypeptides, maintains its subcellular localization by virtue of a tetrapeptide retention signal at the extreme carboxyl terminus of the polypeptide. This signal is typically KDEL in mammals (Munro and Pelham, 1987) and HDEL in *Saccharomyces cerevisiae* (Pelham et al., 1988). Other variations of the XDEL sequence include DDEL in *Kluyveromyces lactis* (Lewis et al., 1990), ADEL in *Schizosaccharomyces pombe* (Pidoux and Armstrong, 1992) and SDEL in *P. falciparum* (Kumar and Zheng, 1992). These retention signals serve as specific ligands for a membrane-bound receptor, the *ERD2* gene product, that has been identified in *S. cerevisiae* (Semenza et al., 1990), *K. lactis* (Lewis et al., 1990) and more recently in *Homo sapiens* (Lewis and Pelham, 1990). Receptor is present in sub-stoichiometric amounts, relative to total luminal ER proteins, and apparently recycles between the ER and a compartment that is transitional between the ER and the Golgi apparatus (Pelham, 1988, 1991) thereby retrieving errant ER residents. Generally speaking, binding is species-specific in that mammalian receptor will not bind the *S. cerevisiae* signal, nor will the yeast receptor function with the mammalian retention signal, in experimental constructs (Pelham et al., 1988; Dean and Pelham, 1990; Lewis and Pelham, 1992). Nevertheless, a limited amount of sequence variation is tolerated, both naturally (including HDEL) and experimentally, in mammalian cells (Andres et al., 1991; Robbi and Beaufay, 1991; Ozawa and Muramatsu, 19923. The *K. lactis* ERD2 protein homologue is of interest, since it will function with HDEL or DDEL sequences, both of which are found on luminal ER proteins in this species (Lewis et al., 1990).

We report here the cloning and characterization of a BiP homologue in African trypanosomes, the most ancient eukaryotic lineage in which this protein has yet been identified. This BiP homologue has been used as a specific marker for the first direct visualization in trypanosomes of the endoplasmic reticulum, which appears as a tubular network that extends throughout the entire cell. The characterization of a bona fide ER resident protein also allows the identification of the putative retention sequences used by trypanosomes. Surprisingly, the deduced carboxy-terminal tetrapeptide of trypanosomal BiP is MDDL, a significant divergence from the ER retention signals seen in other eukaryotes. A gene previously identified in our laboratory that encodes a protein homologous to protein disulfide isomerase, another luminal ER protein, has the deduced carboxy-terminal tetrapeptide KQDL (Hsu et al., 1989). Combined, these results imply the use of multiple ER retention signals in trypanosomes and argue for unusual latitude in the ligand specificity of the trypanosomal receptor.

## MATERIALS AND METHODS

### Trypanosomes and preparation of nucleic acids

Bloodstream form (clone MITat 1.4 expressing VSG 117) and procyclic culture form (corresponding to the insect stage of the parasite life-cycle and derived from MITat 1.4) trypanosomes of the 427 strain of *Trypanosoma brucei brucei* were used for all experiments except for cryoimmuno-electron microscopy, when bloodstream forms of the Wellcome CT strain of *T. b. rhodesiense* (clone DuTat 1.1; Balber and Frommel, 1988) were used. Growth and purification of bloodstream parasites and growth of procyclic culture-form trypanosomes were as described previously (Bangs et al., 1986; Dorn et al., 1991). DNA and poly(A)$^+$ mRNA were purified from bloodstream forms of MITat 1.4 as described previously (Hobbs and Boothroyd, 1990; Bangs et al., 1992).

### Synthethic deoxyoligonucleotides

The following synthetic deoxyoligonucleotides were used for cloning and hybrid arrest experiments (restriction sites are underlined, parentheses indicate degeneracy):

JB8: 5′-GCATCTAGATA(TC)TT(TC)AA(TC)GA(TC)GC-(TCAG)CA(AG)(CA)G-3′ for PCR cloning of *hsp70* homologues, degenerate sense primer with *Xba*I site.

JB9: 5′-CAAAGCTTAC(TCAG)GC(TC)TC(AG)TC(TCAG)-GG(AG)TT(AGT)TT-3′ for PCR cloning of *hsp70* homologues, degenerate antisense primer with *Hin*dIII site.

JB25: 5′-GCCTATCACTTGTCTAGACCCGGGCATCGAC-CTCGGCACA-3′ for PCR amplification of the *gA* coding region for GST fusion, sense primer with *Xma*I site.

JB26: 5′-TCCACTGAGCAACTCTAGACCCGGGTTACA-GAT-CGTCCATGGG-3′ for PCR amplification of the *gA* coding region for GST fusion, antisense primer with *Xma*I site.

LU7: 5′-GCCTATCACTTGTCTAGACCCGGGTATCGAC-CTCGGTACG-3′ for PCR amplification of the *hsp70* coding region for GST fusion, sense primer with *Xma*I site.

LU9: 5′-GGCACTGAGCAACTCTAGACCCGGGCTCGGG-CATACCGCCAGG-3′ for PCR amplification of the *hsp70* coding region for GST fusion, antisense primer with *Xma*I site.

JB43: 5′-GTCAACATTCTTTTTCTTCTTCAGCATGTC-3′ for hybrid arrest, antisense oligonucleotide to *gA* coding region corresponding to amino acids 270-279.

JB44: 5′-GTCCTTACCGGTGCGGGTTCTGAACTCAGT-3′ for hybrid arrest, antisense oligonucleotide to pB1 coding region analogous to JB43.

JB45: 5′-GCTGTCCACCTCCACACGCGCCTCGGGATG-3′ for hybrid arrest, antisense oligonucleotide to *gA* coding region corresponding to amino acids 304-313.

JB46: 5′-TGCATCAATCGCAATGTTGGTGCTTGCTGA-3′ for hybrid arrest, antisense oligonucleotide to pB1 coding region analgous to JB45.

### Cloning of the trypanosomal BiP gene

Genomic trypanosome DNA (MITat 1.4) was amplified by the polymerase chain reaction (PCR) using the degenerate synthetic oligonucleotides primers, JB8 and JB9. Cycle conditions were: denaturation, 30 s at 95°C; annealing, 1 min at 55°C; elongation, 1 min at 72°C. Products of approximately 850 bp and 700 bp were gel-purified and blunt-end cloned into the *Eco*RV site of pBluescript II SK$^-$ (Stratagene Inc., San Diego, CA). Recombinant plasmid DNA from ten transformants of *Escherichia coli* were used as templates for PCR with primers JB8 and JB9. In each case products of identical size to the original inserts, ~700 or ~850 bp, were obtained. Clones fell into three distinct groups based on restriction endonuclease analysis of their respective PCR products: two groups with ~700 bp inserts, pA (pA1 - pA3) and pB (pB1 - pB3); and one group with a ~850 bp insert, pC (pC1 - pC4).

Gel-purified PCR product (primers: JB8/JB9, template: pA1) was used as a probe to screen a genomic *T. brucei* λ phage library (bloodstream stage MITat 1.4 DNA in λEMBL3, generously provided by Dr Tom Beals). Two positive plaques from secondary screening gave 100% positive plaques in tertiary screens and were therefore deemed clonal. One of these (λ5), containing the parental genomic sequence (*gA*) of the pA1 PCR clone, was selected for subcloning and sequencing. Subclones of the *gA* sequence were made directly from λ5 into pBluescript II SK−, including pλA4, a 6 kb *Xma*I fragment containing both 5′ and 3′ non-coding sequences flanking the entire coding region. Further deletion subclones were made as necessary from primary subclones.

## λ Plaque and genomic DNA blotting

λ Plaque lifts were prepared on nitrocellulose filters and screened according to Sambrook et al. (1989) using a high-stringency (0.1× saline sodium citrate, 65°C) final wash (1× = 0.15 M NaCl/0.015 M sodium citrate).

Genomic DNA (3 μg) was digested with various restriction endonucleases, fractionated electrophoretically on 0.8% agarose gels and transferred to Nytran filters (Schleicher and Schuell, Inc., Keene, NH) as described (Hsu et al., 1989). Non-radioactive probes were prepared with an ECL Random Priming Kit (Amersham, Arlington Heights, IL) using gel-purified PCR products as template and hybridization was carried out at 65°C according to the manufacturer's specifications. Stringency washes were as described above and specific hybridization was detected with the ECL Southern Detection System (Amersham) and Kodak XRP film.

## DNA sequencing

DNA sequence was determined by the dideoxy chain termination method (Sequenase kit, United States Biochemical, Corp., Cleveland, OH) using subclones in pBluescript II $SK^-$ as double-stranded templates. All *gA* coding sequence was determined on both strands using primers complementary to either the pBluescript polylinker or known internal *gA* sequence.

## Production of gA and hsp70 fusion proteins

pλA4 and pTb10.A were used as PCR templates with primer combinations JB25/JB26 (*gA*) and LU7/LU9 (*hsp70*), respectively. pTb10.A is a cloned *Eco*RI fragment, containing a single copy of the *T. brucei* 2.5 kb *hsp70* genomic repeat in the pGEM1 vector (Muhich and Boothroyd, 1988). Since *Eco*RI cuts within the *hsp70* coding region, the plasmid was first digested with *Eco*RI and then religated to convert the insert to 2.5 kb closed circles with contiguous coding regions for use as a PCR template.

The amplification products, in each case bearing unique *Xma*I sites at both ends, were digested with *Xma*I, gel purified and ligated into the *Xma*I cloning site of pGEX-2T (Pharmacia, Piscataway, NJ) to generate in-frame fusions of the coding regions with glutathione S-transferase (GST). Expression of the resultant fusion proteins in *E. coli* was dependent on both insert orientation and *lacZ* induction.

Induction of fusion protein synthesis and lysis of induced bacteria was carried out according to the manufacturer's specifications. Each fusion protein, found predominantly in the insoluble fraction, was solublized in urea (Schloss et al., 1988) and renatured by exhaustive dialysis against 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, 5 mM PMSF. Insoluble material was removed by centrifugation (30,000 ***g,*** 30 min, 4°C, JA17 rotor, Beckman Instruments, Palo Alto, CA) and renatured fusion protein was purified from the clarified supernatant by affinity chromatography on glutathione-agarose (Sigma) (Smith and Johnson, 1988).

## Production of anti-gA antibodies

New Zealand White rabbits were immunized with GST:gA fusion protein using the RIBI adjuvant system (RIBI ImmunoChem Research Inc., Halmilton, MT.) according to the manufacturer's instructions. For affinity purification, GST:gA and GST:hsp70 fusion protein were independently coupled (1 mg protein: 1 ml beads) to cyanogen bromide-activated Sepharose (Pharmacia) according to the manufacturer's specifications. Antibodies to GST determinants and cross-reacting hsp70 epitopes were depleted from whole anti-GST:gA serum by adsorption to GST:hsp70-Sepharose. Specific anti-gA antibodies were then affinity purified from the depleted serum on GST:gA-Sepharose. Normal IgG was purified from preimmune serum by Protein A-Sepharose affinity chromatography.

Affinity-purified rabbit anti-VSG antibody was prepared with VSG117 purified by the procedure of Cross (1975). Mouse antiserum raised against GST fusion proteins containing the full length of *T. cruzi* hsp70 and mtp70 were generous gifts from Dr David Engman of Northwestern University, Chicago, IL. Mouse ascites fluid containing monoclonal antibody to purified 38 kDa subunit protein (band 4) of *Crithidia fasciculata* cytochrome *c* reductase was a generous gift from Dr Jeffrey Priest of the University of Alabama at Birmingham.

## Electrophoresis and immunoblotting

Protein samples were fractionated on SDS-PAGE gels and transfered electrophoretically to nitrocellulose membranes (0.2 μm, Schleicher and Schuell). Membranes were blocked with phosphate-buffered saline (PBS: 10 mM $Na_2HPO_4$, 2 mM $KH_2PO_4$, 137 mM NaCl, 8 mM KCl, pH 7.4) containing 0.05% Tween-20 and probed with specific antibodies. After washing, specific binding was detected using an ECL Western Detection System (Amersham) according to the manufacturer's instructions. Prestained molecular mass markers (Bio-Rad Laboratories, Hercules, CA), used on all gels, were: phosphorylase b, 106 kDa; bovine serum albumin, 80 kDa; ovalbumin, 49.5 kDa; carbonic anhydrase, 32.5 kDa; soybean trypsin inhibitor, 27.5 kDa.

## Subcellular fractionation

A modification of the lysis procedure of Harris et al. (1990) was used to disrupt procyclic trypanosomes. All operations were carried out at 4°C and protease inhibitor cocktail (2 μg/ml each, aprotinin, leupeptin, antipain, pepstatin and chymostatin) was included in all solutions except sucrose cushions or when fractions were prepared for subsequent protease protection experiments. Procyclic trypanosomes were washed three times with PBS and resuspended at $0.5\times10^9$ to $1.0\times10^9$/ml in hypotonic buffer (1 mM Hepes-KOH, pH 7.5, 1 mM EDTA). After a 10 min incubation the lysate was passed through a 27 gauge needle three times and then adjusted to isotonicity by the addition of 5× buffer to give final 1× lysis buffer (LB: 25 mM Hepes-KOH, pH 7.5, 100 mM sucrose, 80 mM potassium acetate, 1 mM EDTA). Observation by light microscopy at this point revealed that the cells were totally disrupted. The lysate was separated into low-speed pellet and supernatant fractions (P1 and S1, respectively) by centrifugation (10 min, 1000 ***g***) and the pellet was resuspended in the original volume of LB. The S1 fraction was layered on a cushion of 2 M sucrose in LB and centrifuged (40 min, 100,000 ***g***, Beckman TLS55 rotor). The resultant supernatant fraction (S2), containing cytosol, was removed and the interface (P2), containing crude microsomes, was resuspended in the original volume of LB.

## Protease protection

Crude microsomal fraction (P2, 200 μl, $10^7$ cell equivalents) was treated (30 min, 4°C, final volume 250 μl LB) with increasing concentrations (0 to 200 μg/ml) of TPCK-treated trypsin (Sigma),

in the presence or absence of 0.5% NP-40. Chicken egg white trypsin inhibitor (Sigma type IV-O, 40 μl, 5 mg/ml in LB) was added and, after further incubation (5 min, 4°C), samples (100 μl) were acetone precipitated and analyzed by immunoblotting with anti-gA antibody.

### Alkaline carbonate treatment

Crude microsomal fraction (P2) was prepared and resuspended at $7.5\times10^9$ cell equivalents per ml in LB. A set of two aliquots (50 μl) were diluted 10-fold with LB and another equivalent set was diluted 10-fold with 110 mM $Na_2CO_3$. After incubation (30 min, 4°C), one member of each set received 167 μl of 4 M KCl and the other recieved the same volume of distilled water. After further incubation (30 min, 4°C), the samples were separated into supernatant and pellet fractions by centifugation (40 min, 100,000 ***g***, 4°C, TLS55 rotor). Supernatants were removed, pellets were resuspended in the original volume of LB, and samples (100 μl) of each were TCA precipitated with 20 μg BSA as carrier. The precipitates were rinsed with 70% ethanol and analyzed by immunoblotting with anti-gA antibody.

### Hybrid arrest in vitro translation

[$^{35}$S]methionine-labelled trypanosome polypeptides were synthesized using nuclease-treated rabbit reticulocyte lysates as directed by the manufacturer (Promega, Madison, WI). Bloodstream form (MITat 1.4) poly(A)$^+$ mRNA (2 μg, 10 μl) was denatured (95°C, 2 min) in the presence of combined (JB43/JB45 or JB44/JB46; 25 μM each) antisense oligonucleotides and then snap-chilled on ice. Translation cocktail (40 μl) was added and protein synthesis was initiated by shifting to 30°C. After 60 min, translation was terminated on ice and the reaction mixtures were diluted to 0.5 ml with TEN buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA) containing 1% NP-40, 1% BSA and clarified by centrifugation (10 min, 16,000 ***g***). One-tenth of the supernatant was diluted to 0.5 ml for immunoprecipitation of VSG polypeptides; the remainder was used for immunoprecipitation of gA polypeptides. Lysates were treated (2 h, 4°C) with 2 μg of antibody, either anti-gA or affinity-purified rabbit anti-VSG 117, and Protein A-Sepharose (50 μl) was added (Phamacia, 8% (w/v) in TEN with 1% BSA and 1% NP-40). After incubation (2 h, 4°C), the precipitates were rinsed twice in wash buffer (TEN, 1% NP-40, 0.5 % sodium deoxycholate, 0.1 % sodium dodecyl sulfate) containing 1% BSA, 3 times in wash buffer and once in TEN. Precipitates were analyzed by SDS-PAGE gel electrophoresis followed by fluorography (Entensify, NEN, Research Products, Boston, MA) with Kodak XAR-5 film.

### Immunofluorescence assay

Immunoflourescence on fixed permeabilized cells was performed by a modification of the procedure of Brickman and Balber (1993). All operations were carried out on ice. After washing three times in PBS containing 1 mM glycerol, procyclic trypanosomes were incubated in PBS containing 4% formaldehyde and 0.04% glutaraldehyde for 1 h, and then washed once in PBS. Fixed cells were quenched for 10 min in PBS containing 10 mM ammonium chloride, washed once in PBS, resuspended ($10^7$ cells/ml) in preimmune IgG or affinity-purified anti-gA antibody (10 μg/ml) in PBS, 1% BSA, 0.1% NP-40 and incubated for 1 h. After washing three times in PBS, 5% normal goat serum, cells ($10^7$/ml) were incubated for 1 h in a 1/50 dilution of fluorescein-conjugated goat anti-rabbit IgG (H+L) (Kirkegaard and Perry Laboratories, Inc., Gaithersberg, MD) in PBS, 5% normal goat serum, 1% BSA containing 10 μg/ml Hoechst 33258 dye (bis-benzimide, Sigma, St Louis, MO). Stained cells were washed three times in PBS, air dried on poly-L-lysine-coated slides and mounted in Vectashield (Vecta Laboratories Inc, Burlingame, CA). Fluorescence was visualized under ×100 oil immersion on an Olympus BO71 epi-fluorescence microscope with a PM-10AD photomicrographic adaptor. Photomicrographs were taken with 100 and 400 ASA Kodak T-MAX film.

### Cryoimmuno-electron microscopy

The protocol of Griffiths et al. (1984) was followed with slight modification. Thin sections of cryo-preserved bloodstream trypanosomes, cut on a Reichert Ultracut cryomicrotome, were floated on blocking buffer (1% BSA, 5% normal goat serum in PBS) and then warmed to room temperature for 30 min. Sections were incubated (30 min) in either anti-gA antibody or pre-immune IgG (142 μg/ml each), washed in PBS and then incubated (30 min) in goat anti-rabbit IgG (H+L) antibody conjugated to 10 nm gold particles (Amersham). Grids were washed with PBS and then distilled water, incubated (10 min, 4°C) with 2% methyl-cellulose/0.3% uranyl acetate, and dried before being examined on a Philips transmission electron microscope.

## RESULTS

### Cloning of *T. brucei hsp70* homologues

Degenerate primers complementary to highly conserved portions of the *hsp70* gene family were used to amplify and clone analogous regions of trypanosome *hsp70* gene family members. These primers were selected to flank an internal 4 amino acid sequence, SETL, that is conserved in many known *BiP* genes including those of yeast, rat, *Caenorhabditis* and *Plasmodium* (Rose et al., 1989; Normington et al., 1989; Heschl and Baillie, 1989; Kumar and Zheng 1992). Since all eukaryotes have multiple *hsp70* gene family members (yeast has at least 9; Lindquist and Craig, 1988) multiple PCR clones were expected from this strategy. Overall homology, as well as the presence of the SETL sequence, was expected to identify PCR clones derived from the trypanosome BiP gene.

Ten PCR clones were identified and catagorized into three groups based on restriction endonuclease analysis (see Materials and Methods). Representatives of each group, pA1, pB1 and pC1, were sequenced (Fig. 1) in order to confirm relatedness to the *hsp70* gene family and to identify that clone most likely to be derived from the trypanosome *BiP* gene. All three are clearly of the *hsp70* gene family (Fig. 2). pC1 is apparently derived from the *T. brucei* equivalent of the mitochondrial *hsp70* (*mtp70*) gene previously identified in the related trypanosomatid, *T. cruzi* (Engman et al., 1989), with which it has 92% amino acid identity in the region of amplification. This clone is approximately 200 nucleotides longer than the others due to annealing of the JB9 primer at a related sequence downstream of the intended target. pB1 is derived from a heretofore undescribed trypanosome *hsp70* gene family member and is unlikely to be a *BiP* homologue, on the basis of overall amino acid homology (57% identity with rat *BiP*) and the presence of FSKI in place of SETL. The deduced amino acid sequence of the pA clone has 72% identity with rat *BiP* throughout the region of amplification and, in addition, has the sequence SEKI in place of SETL. Collectively, these data implicate pA1 as most likely to be derived from the trypanosome *BiP* gene. None of the PCR clones are derived from the previously characterized trypanosomal

```
A  (gA)                                                                                                    -261 tctagagagcacataacctctac -239
ttgcaggagccccatggcagcactggaggcgacccctttcttccgccttgccctttccgtgtgccatctgcgttttaccatttccttttccctatttgtccctttcatcgccccaccac -120
acctcctctccacatattttcacgtttcgcgttgttttggaaagataagctaagcatactaaagcacacaaacctacttacctaacactaatattaagttcaagataatccccctcccg   -1
ATG TCG AGG ATG TGG CTG ACC ACT GCA GCG GTG TTC CTG ACT GTG ACG GTT GCA GCC GTC TCA GCA GCA CCC GAA AGT GGC GGC AAG GTG   90
GAA GCA CCA TGC GTG GGC ATC GAC CTC GGC ACA ACA TAC TCC GTT GTG GGT GTG TGG CAG AAG GGT GAT GTG CAT ATC ATC CCG AAC GAG  180
ATG GGT AAC CGC ATC ACA CCT TCC GTC GTC GCC TTT ACC GAC ACA GAG CGG CTG ATC GGT GAC GGT GCG AGG AAT CAG CTT CCA CAG AAT  270
CCG CAT AAT ACA ATC TAC ACC ATC AAG AGG CTG ATT GGC CGC AAG TAC ACG GAT GCG GCA GTG CAG GCT GAC AAG AAA CTG CTG TCG TAT  360
GAG GTC ATT GCG GAC CGC GAC GGG AAG CCG AAG GTA CAA GTG ATG GTG GGT GGG AAA AAG AAA CAG TTC ACA CCA GAA GAG ATC AGT GCC  450
ATG GTG CTG CAG AAG ATG AAG GAA ATT GCA GAA ACG TAC CTT GGC GAG AAA GTG AAA AAC GCT GTT GTA ACG GTA CCT GCG TAC TTC AAT  540
GAC GCA CAA CGG CAG TCG ACA AAG GAT GCT GGG ACC ATC GCC GGT TTG AAC GTA GTG CGC ATC ATC AAT GAG CCC ACC GCA GCC GCC ATC  630
GCA TAT GGG CTA AAC AAA GCC GGT GAG AAG AAT ATC CTG GTG TTC GAT CTT GGT GGT GGT ACC TTT GAT GTG TCA CTG CTG ACA ATC GAT  720
GAG GGC TTC TTC GAA GTT GTG GCG ACA AAC GGT GAT ACG CAC CTT GGT GGT GAG GAC TTT GAT AAC AAC ATG ATG CGA CAC TTT GTG GAC  810
ATG CTG AAG AAG AAA AAG AAT GTT GAC ATC AGT AAG GAC CAA AAG GCA CTG GCA CGT CTT CGC AAG GCA TGT GAG GCT GCG AAG CGA CAG  900
CTG TCG TCT CAT CCC GAG GCG CGT GTG GAG GTG GAC AGC CTT ACG GAG GGC TTC GAT TTC AGC GAG AAA ATC ACA CGT GCG AAG TTT GAG  990
GAG CTG AAC ATG GAC CTC TTC AAG GGG ACA CTC GTG CCC GTG CAA CGT GTG CTG GAG GAT GCG AAG CTG AAG AAG AGT GAC ATC CAC GAG 1080
ATT GTG CTC GTT GGT GGA TCC ACA CGT GTG CCG AAG GTA CAA CAA CTG ATC AGT GAC TTC TTC GGT GGG AAG GAA CTG AAC CGT GGT ATT 1170
AAC CCC GAT GAA GCC GTA GCA TAT GGT GCC GCT GTG CAA GCT GCG GTG CTA ACC GGT GAA AGC GAG GTT GGC GGG CGC GTT GTG CTT GTA 1260
GAT GTG ATC CCA CTG TCG CTT GGT ATT GAA ACT GTA GGT GGT GTC ATG ACA AAA CTC ATC GAG CGT AAC ACA CAG ATC CCC ACC AAG AAG 1350
AGT CAG GTC TTT TCT ACC CAT GCG GAC AAC CAG CCG GGT GTG CTT ATC CAA GTG TAT GAG GGT GAG CGC CAG CTG ACA AAG GAT AAC CGC 1440
CTC CTG GGC AAA TTC GAG TTG TCG GGA ATT CCT CCT GCC GCG CGT GGT GTG CCA CAA ATC GAA GTG ACA TTT GAC GTC GAC GAG AAC AGC 1530
ATC CTG CAG GTG AGC GCT ATG GAC AAG TCG TCT GGA AAG AAG GAA GAG ATC ACC ATC ACC AAC GAC AAG GGT CGC CTG AGC GAG GAG GAG 1620
ATT GAG CGT ATG GTG CGC GAA GCG GCA GAG TTC GAG GAC GAG GAC CGC AAG GTG CGC GAG CGT GTG GAT GCG CGT AAC TCG CTG GAG AGC 1710
GTT GCG TAT TCC CTA CGC AAC CAA GTG AAC GAT AAG GAT AAG CTT GGT GGG AAA CTT GAC CCC AAT GAC AAG GCT GCA GTG GAG ACC GCT 1800
GTT GCG GAG GCG ATC CGC TTC CTT GAC GAA AAT CCG AAC GCC GAA AAG GAG GAG TAC AAA ACT GCC CTG GAA ACT CTG CAG AGC GTG ACC 1890
AAC CCC ATC ATC CAG AAG ACT TAC CAA TCC GCC GGC GGA GGC GAC AAG CCA CAA CCC ATG GAC GAT CTG TAA gtgtactggcaggtgtgtgagtg 1985
gttattattattattattatttataattactgcagcaaatttggttcgtttaaattgtgtgagcagcgagcttactggcagtgtttgtgtgcttgttcaatgatgtagctttgtggct 2104
caatcttttt 2114

B  (pB1)
TAC TTC AAT GAT GCA CAG CGT CAA GCA ACA AAA GAC GCG GGA TCA ATT GCT GGT CTC GAG GTG TTG CGC ATC GTT AAC GAA CCA ACA GCC   90
GCA GCC ATT GCG TAT GGC AGG GAC AGA AGT TCT GAA GGA GCA ATG AAG ACG GTG CTC ATC TTT GAT CTT GGA GGT GGC ACT TTT GAC GTG  180
ACG TTG CTA AAC ATT GAC GGT GGC CTA TTT GAA GTG AGG GCA ACT GCT GGT GAC ACA CAT CTT GGT GGG GAA GAC TTT GAC AGC CGC CTT  270
GTG GAT TAC TTC GCG ACT GAG TTC AGA ACC CGC ACC GGT AAG GAC TTG CGT GGA AAC GCA CGT GCG ATG AGG CGT TTG CGT ACG GCA TGC  360
GAA CGG GTG AAG CGA ACA CTT TCC AGT TCA GCA AGC ACC AAC ATT GCG ATT GAT GCA CTG TAT GAA GGG TTT GAC TTT TTC TCC AAA ATA  450
ACG CGA GCT CGG TTC GAG GAG ATG TGC CGT GAC CAA TTC GAG CGC TGC CTA GAG CCA GTG AGG AAG GTC CTG AAA GAT GCC GAG GTT GAT  540
GCA TCG GCA GTG GAC GAT GTG GTT CTT GTC GGT GGA TCA ACT CGC ATT CCC CGC GTG CAA CAG TTG GTG CAG AAC TTT TTC AAC GGC AAG  630
GAG TTG AAC CGT TCA AAC AAC CCG GAT GAA GCG GT 665

C  (pC1)
TAT TTT AAC GAT GCG CAG CGT CAG GCA ACG AAG GAT GCC GGT ACG ATT GCT GGA CTG AAT GAG ATC CGT GTT GTG AAT GAA CCG ACT GCT   90
GCT GCA CTG GCA TAT GGC TTG GAC AAA ACG AAG GAC AGC TTA ATT GCC GTG TAT GAT CTT GGT GGT GGA ACA TTT GAT ATC TCC GTA CTT  180
GAA ATT GCC GGA GGT GTT TTT GAG GTA AAG GCA ATG AAT GGC GAC ACT CAC CTT GGT GGA GAA GAC TTT GAT CTC TGC CTC TCT GAT CAC  270
ATT CTG GAG GAA TTC CGC AAG ACA TCT GGA ATT GAC TTG AGC AAG GAG CGG ATG GCA CTG CAA CGT ATT CGT GAG GCT GCA GAA AAG GCG  360
AAG TGT GAA CTT TCG ACG ACG ATG GAG ACC CCC GAG GTA ACC CTT CCG TTT ATC ACA GCT AAC CAG GAC GGA GCG CAG GAC GTG CAG ATG  450
ATG GTG AGC CGC AGC AAG TTT GAG TCC CTT GCT GAT AAG CTC GTA CAA CGT TCA CTT GGA CCG TGC AAG CAA TGC ATC AAG GAC GCT GCT  540
GTT GAC CTG AAG GAA ATC TCT GAG GTT GTG CTT GTT GGT GGT ATG ACG CGT ATG CCA AAG GTT GTG GAG GCT GTG AAG CAG TTC TTC GGT  630
CGC GAA CCG TTC CGT GGT GTG AAC TCT GAC GAA GGC GTT GCA CTG GGT GCT GCA ACA CTT GGT GGT GTC CTC CGT GGT GAT GTG AAG GGT  720
CTT GTA TTG CTT GAC GTT ACA CCA CTA TCA CTT GGA ATC GAA ACA CTT GGT GGT GTC TTC ACA CGT ATG ATT CCC AAG AAC ACA AAA ACC  810
CCG ACG AAG CGG T  823
```

**Fig. 1.** Sequences of *T. brucei hsp70* genes. The nucleotide sequences of the *gA* genomic region (A), and the pB1 and pC1 PCR clones (B and C, respectively) are presented. Upstream and downstream noncoding sequences are in lower case and all coding regions are in upper case. Sequences corresponding to PCR primers JB8 and JB9 are underlined. The natural nucleotide sequence in the region corresponding to the primers is known for *gA* only; the genomic sequences in these regions of pB and pC are likely to differ somewhat from those shown. Numbering is relative to the deduced start codon for *gA* and to the end of the upstream primer for pB and pC. Triplet groups represent codons of the deduced open reading frames.

homologues of inducible and constitutively expressed *hsp70* (Lee et al., 1990). The reason for this is unclear but could be due to the limited number of clones analyzed.

Using the pA1 sequence to screen a genomic library, a λ phage clone containing the parental sequence of the PCR clone was identified and subclones encompassing the relevant open reading frame were prepared and sequenced (Fig. 1). The results show that this genomic clone contains the exact nucleotide sequence of the pA1 clone; it is therefore designated as the genomic A (*gA*) coding region. The deduced amino acid sequence (Fig. 2) of the complete open reading frame of the protein indicates a polypetide of 653 amino acids (71.5 kDa) including an amino-terminal hydrophobic sequence (~30 amino acids) that conforms to the general properties of secretory signal sequences (von Heijne, 1990), as would be expected for a resident of the ER. There are no canonical N-linked glycosylation sites. The predicted polypeptide has overall amino acid identities of 64% and 62% with rat and yeast BiP, respectively. The high percentage identity with other BiPs, together with the presence of a signal sequence, strongly suggest that this is the gene for the trypanosomal BiP homologue. If so, this implies that the carboxy-terminal tetrapeptide, MDDL, functions as an ER retrieval signal in trypanosomes.

## Genomic organization of the *gA* gene

Genomic organization in trypanosomes is typified by the absence of *cis* introns (Boothroyd, 1989) and the frequent tandem reiteration of identical protein coding regions (e.g. *hsp70*; Lee et al., 1990). Such reiterations can be confirmed by the generation of a common repeat-sized restriction fragment by any restriction endonuclease that cuts once within the repeat. In order to determine the *gA* genomic organization, the pA1 sequence was used as a probe in Southern analysis (Fig. 3). Five restriction enzymes that cut once within the *gA* region (*Bam*HI, *Bst*EII, *Eco*RI, *Hin*dIII, *Xba*I) generate distinct restriction patterns with no common-sized fragments arguing against a repeat structure. However, a

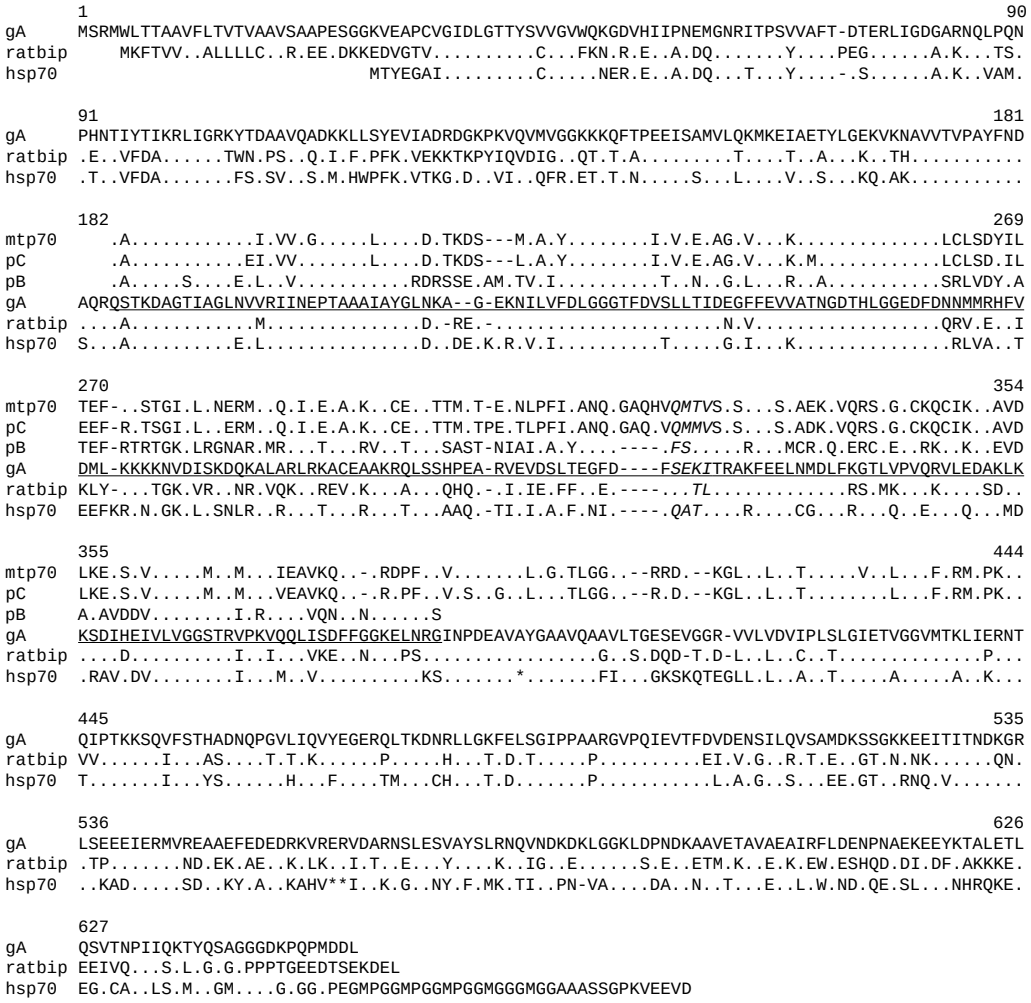


**Fig. 2.** Alignment of deduced amino acid sequences. The deduced amino acid sequences of the gA, pB and pC regions are aligned with the complete published sequences deduced for rat BiP (Munro and Pelham, 1986) and *T. brucei* inducible hsp70 (Lee et al., 1990) and the portion of the *T. cruzi* mtp70 gene product (Engman, 1989) corresponding to the pC sequence. Alignment was performed to maximize identity (indicated by dots) with the gA sequence. Dashes indicate gaps introduced to improve alignment and asterisks indicate ambiguous residues in the published *T. brucei* hsp70 sequence. Amino acids corresponding to the internal gA sequence SEKI are in italics and the portion of the gA sequence amplified in the pA PCR clone (not shown) is underlined. Numbering is relative to the deduced initiation methionine residue of the gA sequence.

pattern (most evident with *Eco*RI, *Hin*dIII and *Xba*I) of two bands of ~1× intensity and a single band of ~2× (or more) intensity is apparent, indicating that there are multiple copies of the *gA* gene per diploid genome. The simplest interpretation is that the *gA* gene is present in at least two loci, each with two alleles, and that one locus is heterozygous.

The non-tandem nature of *gA* genomic organization raises the question of whether the *gA* gene that we have identified is in fact representative of all the copies of the gene family. Digestion with enzymes (*Nco*I, *Pst*I, *Sal*I, *Xba*I/*Hin*dIII) that cut at sites outside the pA1 probe sequence but within or near the coding region all generate single fragments of sizes predicted from the *gA* sequence (Fig.3). These results strongly indicate that all copies are essentially identical throughout the coding region and that the *gA* clone is representative of the gene family. In this regard, it is worth noting that four of six nucleotides comprising the 3′ *Nco*I site fall within the first two codons of the carboxy-terminal tetrapeptide, MDDL, arguing that the proteins coded by all copies of the *gA* gene have at least the first two residues of this putative tetrapeptide retention sequence in common.

## Production of specific anti-gA antibodies

Antibodies were prepared for use in characterizing the polypeptide encoded by the *gA* gene to confirm that it is expressed in trypanosomes and that it has the properties expected of a BiP homologue. Using the glutathione-S-transferase (GST) system, separate recombinant fusion proteins encompassing the entire *gA* coding region except the putative signal sequence (GST:gA), and an equivalent por-

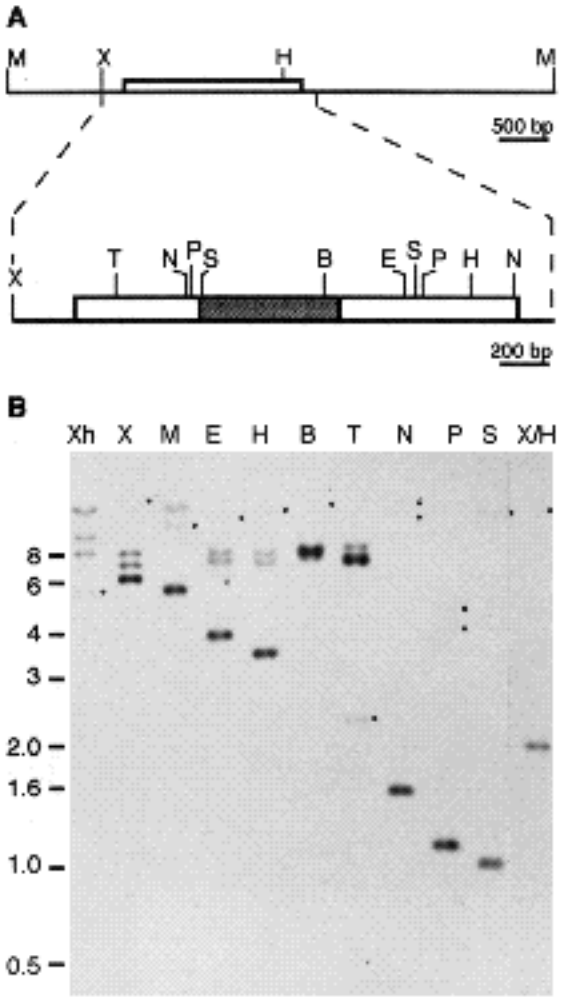


**Fig. 3.** Genomic organization of the *gA* region. (A) Restriction map of the gA region. The coding region is boxed and the region corresponding to the pA PCR clone is shaded. (B) Genomic Southern blot of digested *T. brucei* DNA using the pA region as probe. The scale refers to size in kilobase pairs and dots indicate the positions of faint bands detected by cross-hybridization of the pA probe with genomic pB sequences (see text). Restriction enzymes are: *Bam*HI, B; *Bst*EII, T; *Eco*RI, E; *Hin*dIII, H; *Nco*I, N; *Pst*I, P; *Sal*I, S; *Sma*I, M; *Xba*I, X; *Xho*I, Xh.

tion of the trypanosome *hsp70* gene (GST:hsp70) were prepared. Rabbit anti-GST:gA serum was generated and antibodies specific for GST and common *hsp70* epitopes were depleted from this serum by adsorption to GST:hsp70-Sepharose. Specific anti-gA antibodies were affinity-purified from the depleted serum on GST:gA-Sepharose. This procedure was designed to minimize the possibility of detecting other trypanosomal *hsp70* gene family members due to conserved immunological epitopes.

When tested in immunoblot analysis, anti-gA antibodies reacted with GST:gA, but not GST:hsp70, fusion protein (data not shown) and specifically detected a single polypeptide of about 80 kDa in total cell extracts (Fig. 4) of both cultured procyclic (lanes 1-4) and bloodstream trypanosomes (lanes 5-8). No signal was detected when blots were probed with pre-immune IgG (data not shown). Serial dilutions of the cell extracts indicate the level of expression of the ~80 kDa polypeptide in bloodstream parasites to be ~3-fold higher than that in procyclics (compare lane 1 with lanes 5 and 6). No difference in the levels of gA polypeptide were detected when two strains of bloodstream trypanosomes expressing different VSGs (VSG117 and VSG118) were compared (data not shown).

To confirm that the ~80 kDa polypeptide is the product of the *gA* gene, we tested the ability of two antisense oligonucleotides, complementary to portions of the pA1 sequence, to arrest the synthesis of gA polypeptide in in vitro translations programmed with trypanosome poly(A)$^+$ RNA. As controls, antisense oligonucleotides complementary to analogous regions of the pB1 sequence were also tested. Analogous sets of oligonucleotides were selected for the uniqueness of their sequences, each being identical in no more than 9 out of 30 nucleotides. pB1 was chosen as the negative control, since this was the only trypanosome sequence known to cross-hybridize with the gA sequence under high stringency (Fig. 3 and data not shown). It seemed, therefore, that a putative gB gene product would be the most likely trypanosomal *hsp70* homologue to cross-react immunologically with anti-gA antibodies.

In vitro translation products generated in the presence or absence of oligonucleotides were immunoprecipitated with anti-gA antibody or, as a control for effect on overall translation, anti-VSG antibody and analyzed on SDS-PAGE gels. A fluorograph of a typical experiment is presented in Fig. 5. The oligonucleotide combination complementary to gA sequences (JB43/JB45) dramatically reduced the level

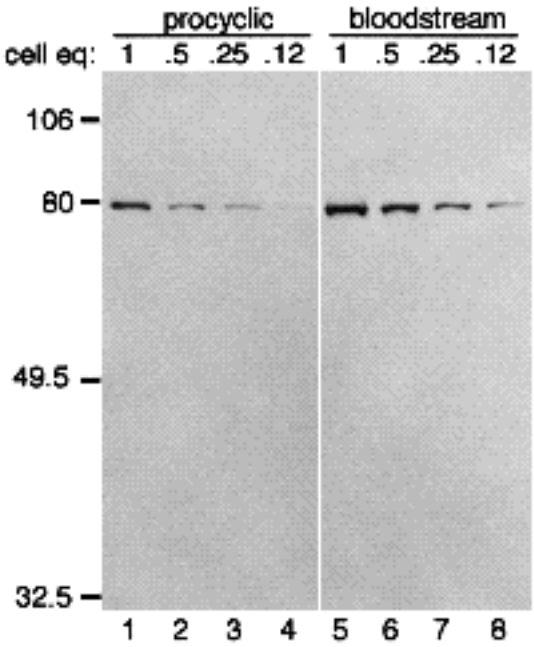


**Fig. 4.** Immunoblot analysis of *T. brucei* extracts with anti-gA antibody. Serial dilutions of whole cell extracts of bloodstream and procyclic trypanosomes were fractionated on SDS-PAGE gels, transferred to nitrocellulose and probed with anti-gA antibody. The number of cell equivalents ($\times 10^{-6}$) loaded per lane are indicated. The scale refers to relative molecular mass in kDa.

of radio-labelled polypeptide detected by anti-gA antibody (lane 2) as compared with the control translation (lane 1). The oligonucleotides complementary to pB sequences (JB44/JB46) had no discernable effect (lane 3) on the synthesis of this translation product. Neither combination of the oligonucleotides significantly impacted the level of translation of an irrelevant polypeptide, VSG 117 (lanes 4-6), nor was the profile of total translation products affected (data not shown).

These data clearly demonstrate that the polypeptide detected by anti-gA antibody is not derived from the pB genomic sequence. Rather, it appears that most, if not all of it, is derived from the *gA* coding region, validating the use of anti-gA antibody as a specific probe of the *gA* gene product.

### The gA polypeptide is a luminal microsomal resident

We wished to determine the subcellular distribution and topology of the gA polypeptide. Procyclic trypanosomes were homogenized by a hypotonic shock procedure and the resultant lysate was fractionated by differential centrifugation into low-speed pellet (P1) and supernatant fractions (S1). The S1 fraction was further separated into pellet (P2, crude microsomes) and supernatant (S2, cytosol) fractions by high-speed centrifugation. Equivalent samples of all fractions were analyzed by immunoblotting with a collection of specific antibodies, including anti-gA and anti-*T. cruzi* hsp70 antibodies.

Following low-speed centrifugation, greater than 50% of the gA polypeptide was detected in the supernatant fraction (Fig. 6a, compare S1 and P1) and when this was subjected to high-speed centrifugation most of the recovered signal was detected in the resulting pellet (Fig. 6a, compare P2 and S2), suggesting that gA polypeptide is associated with microsomal vesicles. The total recovery of gA polypeptide

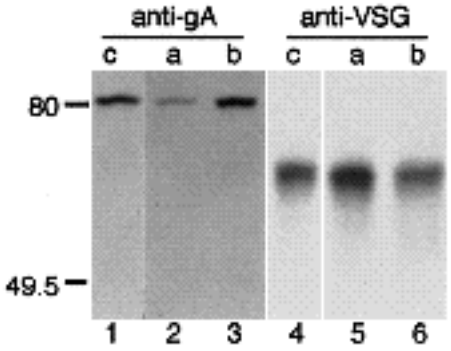


**Fig. 5.** Hybrid arrest in vitro translation of the gA polypeptide. Poly(A)$^+$ RNA from MITat 1.4 bloodstream trypanosomes was translated in vitro in the presence of antisense oligonucleotides complementary to gA sequences (a), complementary to pB sequences (b) or in the absense of any oligonucleotide (c). Radiolabelled polypeptides were immunoprecipitated with either anti-gA antibody (lanes 1-3) or anti-VSG117 antibody (lanes 4-6). Precipitates were analyzed by SDS-PAGE and fluorography. The scale refers to relative molecular mass in kDa. Lanes 1-3 and lanes 4-6, respectively, are from the same exposures and all lanes are from the same gel.

in the S2 and P2 fractions (as a percentage of the S1 fraction) varied somewhat, experimentally, and, while the recovery of gA in Fig. 6a is low, the ratio of gA in P2 relative to S2 is typical. The low recovery of gA is discussed further below. Under identical conditions *bona fide* hsp70 fractionated almost exclusively to the high-speed supernatant as would be expected for a cytosolic protein (Fig. 6b, S2). Based on the results of protease protection experiments (see below), it is likely that the small amount of gA polypeptide in the cytosolic fraction was released from microsomes that were disrupted by the lysis procedure.

Fractions were also analyzed with antibody directed against two mitochondrial markers, mtp70 of *T. cruzi* and cytochrome *c* reductase of *C. fasciculata*, a related kinetoplastid protozoan. mtp70, a luminal protein, was distributed almost equally between the low-speed pellet and supernatant fractions (Fig. 6c, P1 and S1). However, the mtp70 detected in the low-speed supernatant (Fig. 6c, S1) was recovered exclusively in the cytosolic fraction (Fig. 6c, S2) and probably results from disruption of a portion of the mitochondrial population. Cytochrome *c* reductase, an integral inner mitochondrial membrane complex, was found almost exclusively in the low-speed pellet (Fig. 6d, P1) and the small amount in the low-speed supernatant was recovered in the microsomal fraction (P2) as expected. The fractionation patterns of these two markers are both consistent with each other and qualitatively different from that of the gA polypeptide (Fig. 6a) eliminating the mitochondrion from consideration as the site of gA localization. In addition, the differences in fractionation pattern and electrophoretic mobility of the three hsp70 polypeptides detected by these antibodies indicate that the anti-gA antibody does not cross-react to any significant extent with either the cytosolic hsp70 or mtp70 homologues of *T. brucei*.

The poor recovery of gA in the high-speed fractions in Fig. 6b is troubling. This loss is probably not due to preferential proteolysis of gA in the S2 fraction because ample protease inhibitors were included in each step and the recoveries of two other markers in the S2 fraction, hsp70 (Fig. 6b, compare S1 and S2) and mtp70 (Fig. 6c, compare S1 and S2), were not significantly affected. The most likely explanation is poor recovery of microsomal fraction from the sucrose cushion used in this experiment. In either case, the immunolocalization experiments presented below fully support the non-cytoplasmic localization of the gA polypeptide.

The resistance of membrane-associated gA polypeptide to exogenously added protease was investigated. Crude microsomes were treated with increasing amounts of trypsin (0 - 200 μg/ml) in the presence or absense of nonionic detergent and then analyzed by immunoblotting (Fig. 7a). Treatment with trypsin alone, at levels as high as 200 μg/ml, had no effect on the amount of gA polypeptide detected. Addition of NP-40, however, resulted in the trypsin-dependent conversion of the gA polypeptide to fragments of about 45 and 29 kDa. The standard interpretation of these data is that the gA polypeptide is physically protected from exogenous protease by an intact membrane bilayer. It is likely that the gA tryptic fragments correspond to the 44 kDa and 30 kDa chymotryptic fragments of mammalian hsc70

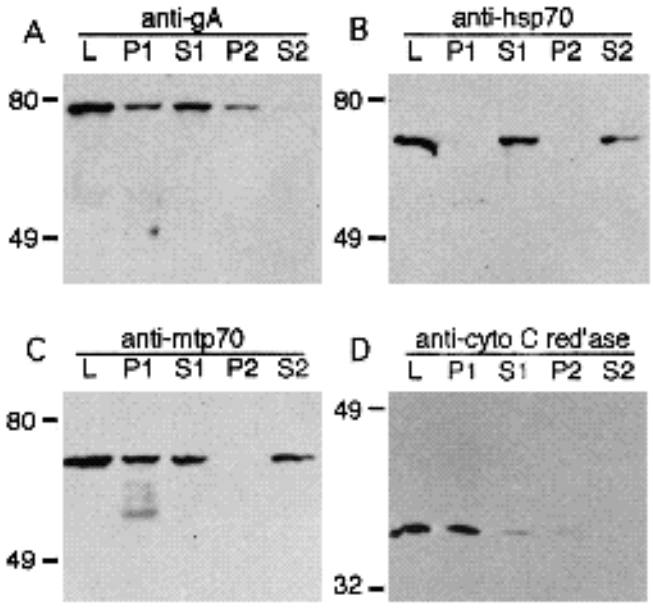


**Fig. 6.** Subcellular fractionation of trypanosomes. Hypotonic lysates (L) of procyclic trypanosomes were fractionated by differential centrifugation into low-speed pellet (P1), postmitochondrial supernatant (S1), crude microsomes (P2) and cytosol (S2). Equivalent samples of each fraction were electrophoresed on SDS-PAGE gels, transferred to nitrocellulose and probed with anti-gA antibody (A), anti-*T. cruzi* hsp70 antibody B), anti-*T. cruzi* mtp70 antibody (C) or anti-*C. fasciculata* cytochrome *c* reductase antibody (D). Cell equivalants per lane are: (A) $1\times10^{-6}$; (B) $1\times10^{-6}$; (C) $2\times10^{-6}$; and (D) $5\times10^{-6}$. The scale refers to relative molecular mass in kDa.

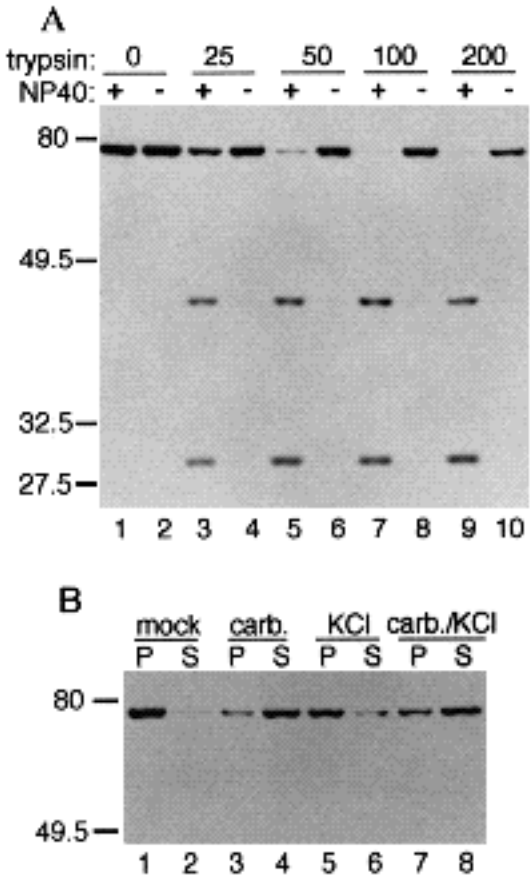


**Fig. 7.** Topology of the gA polypeptide. (A) Protease protection. Crude microsomal fraction (as in P2 of Fig. 6), prepared from procyclic trypanosomes, was subjected to treatment with either protease (A) or alkaline carbonate (B). In each case, equivalent samples of treated microsomes were fractionated by SDS-PAGE and analyzed by immunoblotting with anti-gA antibody. Scales refers to relative molecular mass in kDa. (A) Microsomes were subjected to treatment with the indicated amounts of exogenously added trypsin in the presence (+; lanes 1, 3, 5, 7 and 9) or absence (−; lanes 2, 4, 6, 8 and 10) of 0.5% NP-40. (B) Microsomes were treated with either lysis buffer (mock), 0.1 M sodium carbonate (carb.), 1 M KCl (KCl) or with 0.1 M sodium carbonate/1 M KCl (carb./KCl) and then separated into soluble (S) or membrane fractions (P) by high-speed centrifugation.

(Chappell et al., 1987) and yeast BiP (Tokunaga et al., 1992) that have been characterized as the ATPase and peptide binding domains, respectively (Chappell et al., 1987). The small amount of gA polypeptide found in the cytosolic fraction is sensitive to trypsin (200 μg/ml) in the absence of detergent (data not shown), consistent with the interpretation that this signal represents gA polypeptide that has been released from disrupted microsomes.

The topology of the gA polypeptide was further investigated by alkaline carbonate treatment of crude microsomes (Fig. 7b), a procedure that converts closed membrane vesicles to open sheets (Fujiki et al., 1982). Luminal contents that are either soluble or peripherally associated with membranes are released by this procedure and are readily separated from membranes by centrifugation. When compared with mock treatment (lanes 1 and 2), alkaline carbonate released substantial amounts of gA polypeptide into the supernatant fraction (lanes 3 vs 4), as would be expected for a luminal protein. High-salt washing (lanes 5 vs 6) of crude microsomes had no effect on the distribution of the gA polypeptide, indicating that it is not associated via ionic interactions with the cytoplasmic face of microsomal vesicles. Sequential alkaline carbonate and high-salt treatment (lanes 7 vs 8) released no more gA polypeptide than carbonate alone. In a similar assay, NP-40 releases the gA polypeptide to the same extent as carbonate treatment (data not shown). These data are consistent with the protease resistance of the gA polypeptide and argue that the gA polypeptide is a soluble luminal resident of intact microsomal vesicles.

## Immunolocalization of the gA polypeptide

The fractionation experiments demonstrate that the gA polypeptide resides in a microsomal organelle but give no indication as to its overall subcellular distribution. To address this issue we used anti-gA antibody as a probe for gA cytolocalization by both indirect immunofluorescence assay (Fig. 8) and cryoimmunoelectron microscopy (Fig. 9).

For immunofluorescence, fixed permeabilized procyclic trypanosomes were double-stained with Hoechst dye (Fig.

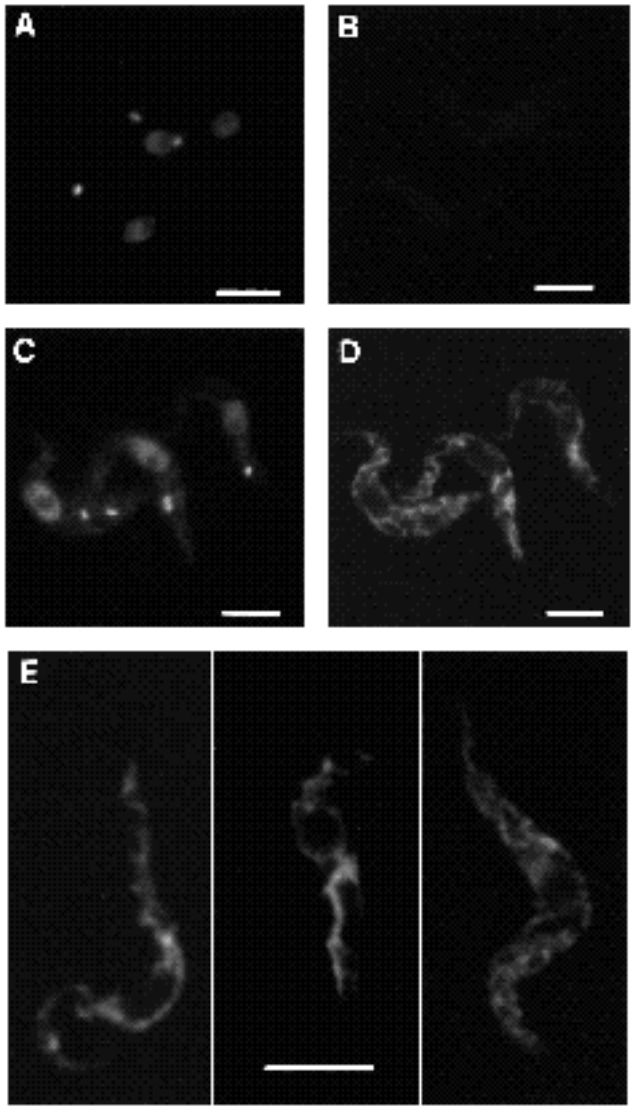


**Fig. 8.** Immunofluorescence localization of the gA polypeptide. Fixed permeablized procyclic trypanosomes were stained with either anti-gA antibody or pre-immune IgG followed by FITC-conjugated goat anti-rabbit IgG and Hoechst 33258. (A and B) Same field of cells visualized for Hoechst staining (A) or staining with pre-immune IgG (B). (C and D) Same field of cells visualized for Hoechst staining (C) or staining with anti-gA antibody (D). (E) Higher magnification of selected cells visualized for staining with anti-gA antibody. Bars, 5 μm.

8A and C) to identify the nucleus and the kinetoplast, the prominent mitochondrial genome that is characteristic of all kinetoplastid protozoa (Englund et al., 1982). This facilitates orientation within the trypanosome: the small well-defined kinetoplast is toward the posterior end of the cell (defined by direction of motility) near the flagellar pocket, and the large, diffusely staining nucleus is midway between the anterior and posterior ends. Anti-gA antibodies labelled all cells with equal intensity (Fig. 8C and D) and the staining was specific as no labelling was detected in cells treated with pre-immune IgG (Fig. 8B vs D). Labelling was apparent throughout the cells, excluding the nucleus and the flagellum, and had a network-like or reticular pattern that could be seen more clearly at higher magnification (Fig. 8E, right). Frequently, tubules could be detected as extensions toward either end of cells (Fig. 8E, left and center) and in some cells distinct perinuclear staining was apparent (Fig. 8E, center). Immunofluorescence was also performed, with similar results, using bloodstream trypanosomes (not shown).

At the higher resolution of cryoimmunomicroscopy, anti-gA antibody was observed to bind throughout the cells to eptiopes associated with internal membranes (Fig. 9a and b). These membranes typically appear as long extended cisternae that have the morphology expected of endoplasmic reticulum. Gold particles were frequently seen in the area subtending the region of flagellar adherence (Fig. 9c) and were occasionally seen at the nuclear membrane (Fig. 9b, arrowhead). Labelling was highly specific; no staining was detected with pre-immune IgG (Fig. 9d). In addition, gold was rarely seen on any other readily identifiable morphological structures (e.g. glycosomes, kinetoplast, flagella or flagellar pocket). These results are completely consistent with the immunofluorescence data and confirm the organellar localization of the gA polypeptide.

## DISCUSSION

We have identified genes for three new members of the *hsp70* gene family in *T. brucei*, including the trypanosomal homologue of *BiP*. This raises to five the total number of known trypanosomal *hsp70* homologues. Two of these newly identified genes are represented by partial PCR clones (pB and pC) and one is represented by a full-length genomic clone (*gA*). The pB PCR clone is derived from a novel *hsp70* homologue that is different from either of the two *hsp70* genes previously identified in *T. brucei*, the closely related heat-inducible and constitutive *hsp70* genes (Lee et al., 1990), and nothing is known concerning its function. The pC PCR clone is apparently derived from the *T. brucei* homologue of *MTP70*, a mitochondrial *hsp70* found in *T. cruzi*. (Engman et al., 1989). mtp70 is physically associated with the kinetoplast, the mitochondrial genome of trypanosomes, and may be involved in the replication of this complex DNA structure. It is also homologous to Ssc1p, a mitochodrial hsp70 in yeast that has been implicated in protein import (Kang et al., 1990). Whether mtp70 also functions in mitochondrial protein import in trypanosomes is unknown.

The third gene, designated *gA*, has been characterized in detail here and, on the basis of several independent lines of evidence, we conclude that it is the trypanosomal homologue of *BiP*. First, the deduced amino acid sequence shows strong similarity (~64% identity) with BiP from other evolutionarily distant eukaryotes. Second, the sequence predicts the presence of an amino-terminal sequence that has the general properties of a secretory signal peptide as would be expected for an ER resident protein (von Heijne, 1990). Third, cell fractionation and topology experiments (Figs 6 and 7) clearly indicate that the encoded polypeptide is a soluble resident of intact microsomal vesicles and are consistent with localization in the endoplasmic reticulum. In

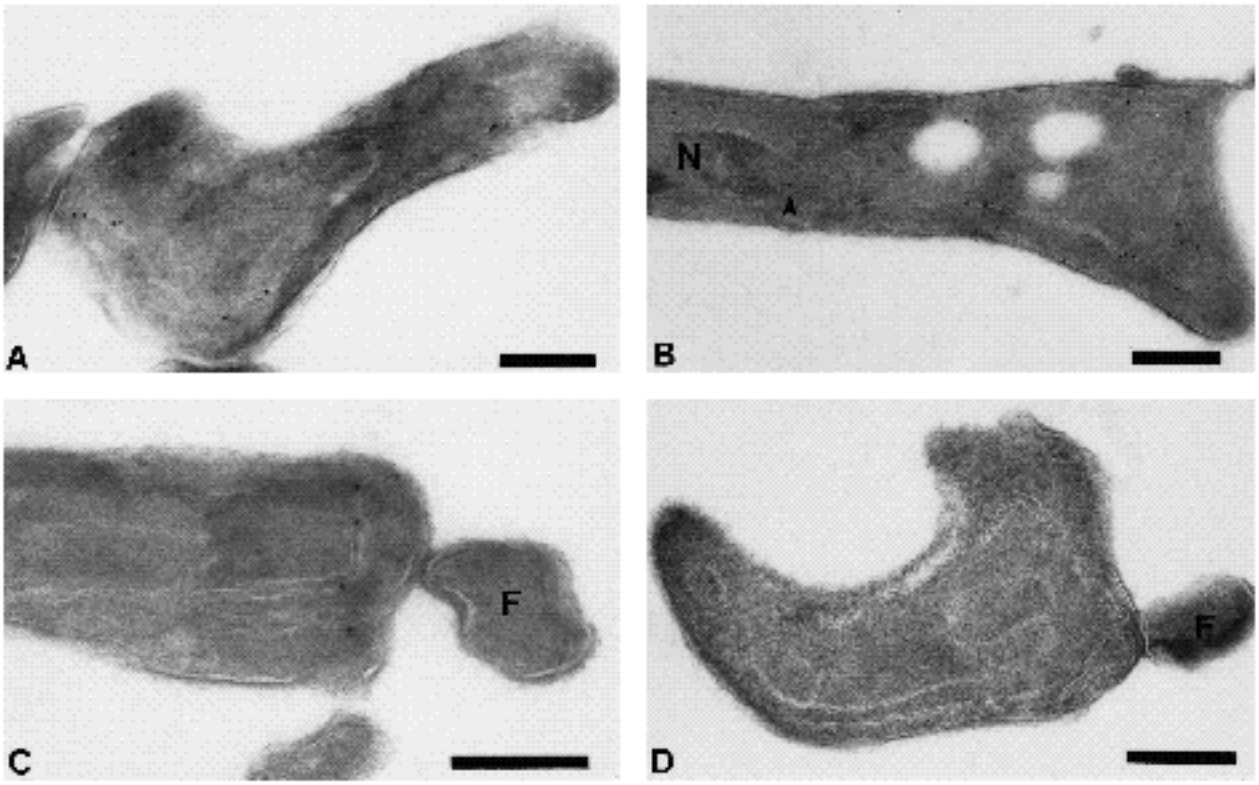


**Fig. 9.** Cryoimmunoelectron microscopy. Cryosections of bloodstream trypanosomes were immunostained with either anti-gA antibody (A-C) or pre-immune IgG (D). Flagellum, F; nucleus, N. Arrowhead indicates gold localized to the nuclear envelope. Bars, 0.2 μm.

addition, the fractionation pattern of the gA polypeptide is qualitiatively and quantitatively distinct from that of both mtp70 and cytochrome *c* reductase, ruling out the possibility that the gA polypeptide is a mitochondrial resident. The gA polypeptide is referred to hereafter as BiP.

BiP in other eukaryotes, and presumably in trypanosomes, is an essential part of the secretory machinery. In this regard, it is interesting that the level of BiP detected in bloodstream parasites is several-fold higher than in procyclic (insect stage) trypanosomes. One possible explanation of this is the enormous metabolic commitment of bloodstream cells to the biosynthesis of VSG. A homodimer comprising 10-20% of total cell protein, this surface antigen forms a uniform coat enveloping the entire cell that is essential for the survival of the parasite in an immunocompetent host (Cross, 1975). Procyclic trypanosomes also have a major surface antigen, procyclin, but this is a monomer and is only ~1% of total cell protein (Clayton and Mowatt, 1989). In addition, the doubling time of bloodstream trypanosomes (~6 hours), and hence the rate at which they must synthesize a new surface coat, is two to three times faster than that of procyclics. The need to fold and assemble so much protein correctly in so little time may account for the elevated levels of BiP in bloodstream trypanosomes.

The characterization of a resident of the endoplasmic reticulum provides us with a specific marker for direct visualization of this organelle in trypanosomes. Immunofluorescence assays (Fig. 8) show that, as in other eukaryotes, the ER is a tubular membranous network that extends throughout the cell, excluding the nucleus and flagellum, and this is confirmed by the general pattern of staining seen in cryoimmunomicroscopy (Fig.9). In addition, we detect a significant amount of anti-BiP staining in subpellicular membranes in the region of flagellar attachment (Fig. 9c). A unique quartet of subpellicular microtubules and an associated single membrane cisternum, the flagellum-associated reticulum, arise near the flagellar pocket and run the length of the cell together, just beneath the pellicle where the flagellum adheres to the trypanosome body (Taylor and Godfrey, 1969; Vickerman, 1969). Our results are consistent with the suggestion of these authors that the flagellum-associated reticulum is a specialized extension of the endoplasmic reticulum. The functional significance of the attachment of this cisternum to the cytoskeleton is not known.

The predicted carboxy-terminal tetrapeptide of trypanosomal BiP is MDDL, indicating that this sequence serves as an ER retrieval signal in trypanosomes, assuming that a typical mechanism of ER retention is operative in these lower eukaryotes. This is one of the most divergent retrieval signal yet observed in nature, probably because this is the most ancient eukaryotic lineage in which BiP has been characterized. In addition, a deduced trypanosomal amino acid sequence that is homologous to another ER resident, protein disulfide isomerase, has the carboxy-terminal tetrapeptide KQDL (Hsu et al., 1989). If this sequence is truly that of an ER resident, it implies that at least two retrieval signals must function in trypanosomes, MDDL and KQDL. The natural use of multiple tetrapeptide retention signals in other species has been observed (Pelham, 1989;

Robbi and Beaufay, 1991; Ozawa and Muramatsu, 1993) and trypanosomes may provide yet another example. Such variation could represent acceptable tolerance in the ligand:receptor interaction that is central to the process of retrieval or it may be that trypanosomes simply utilize a minimal signal, --DL, that is completely independent of the residues in positions 1 and 2. Alternatively, trypanosomes may have multiple receptors, each with a different tetrapeptide specificity. The presence of multiple retention receptors in a single organism is not without precedent; two *ERD2* homologues, albeit with apparently identical ligand specificity, have been identified in humans (Lewis and Pelham, 1992; Hsu et al.,1992).

Our results allow for the first time the direct visualization of the endoplasmic reticulum in this organism and provide a useful marker for the analysis of processes in which this organelle plays a central role, such as protein secretion and the biosynthesis of glycosyl phosphatidylinositol lipid structures. The further characterization of markers for other compartments of the secretory pathway, coupled with the recent advent of technology for the expression of foreign genes in trypanosomes will provide the tools necessary to address issues of protein targeting that were previously unassailable in this lower eukaryote.

We thank Dr Tom Beals for the genomic library, and Drs David Engman and Jeffery Priest for providing antibodies. We also thank Drs Kasturi Haldar and Heidi Elmendorf for thoughtful discussions. This work was supported in part by the MacArthur Foundation, the Burroughs Wellcome Fund, and National Institutes of Health grants AI21025 to J.C.B. and AI28427-01A4 to A.E.B.; J.D.B. was supported in part by United States Public Health Service grant 2T32CA09302-13 awarded by the National Cancer Institute, Department of Health and Human Services. Nucleotide sequences have been deposited in the GenBank data base under the following accession numbers: gA, L14477; pB1, L14478; pC1, L14479.

## REFERENCES

**Andres, D. A., Rhodes, J. D., Meisel, R. L. and Dixon, J. E.** (1991). Characterization of the carboxyl-terminal sequences responsible for protein retention in the endoplasmic reticulum. *J. Biol. Chem.* **266**, 14277-14282.

**Balber, A. E.** (1990). The pellicle and the membrane of the flagellum, flagellar adhesion zone, and flagellar pocket: functionally discrete surface domains of the bloodstream form of African trypanosomes. *Crit. Rev. Immunol.* **10**, 177-201.

**Balber, A. E. and Frommel, T. O.** (1988). *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense:* concanavillin A binding to membrane and flagellar pocket of bloodstream and procyclic forms. *J. Protozool.* **35**, 36-41.

**Bangs, J. D., Andrews, N., Hart, G. W. and Englund, P. T.** (1986). Posttranslational modification and intracellular transport of a trypanosome variant surface glycoprotein. *J. Cell Biol.* **103**, 255-263.

**Bangs, J. D., Crain, P. F., Hashizume, T., McCloskey, J. A. and Boothroyd, J. C.** (1992). Mass spectrometry of mRNA cap 4 from trypanososmatids reveals two novel nucleosides. *J. Biol. Chem.* **267**, 9805-9815.

**Bole, D. G., Hendershot, L. M. and Kearney, J. F.** (1986). Posttranslational association ofimmunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas. *J. Cell Biol.* **102**, 1558-1566.

**Boothroyd, J. C.** (1989). *Trans*-splicing of RNA. In *Nucleic Acids and Molecular Biology*, vol. 3 (ed. F. Eckstein and D. M. J. Lilley.) pp. 216-230. Springer-Verlag, Berlin, Heidelberg.

**Brickman, M. J. and Balber, A. E.** (1993). *Trypanosoma brucei rhodesiense*: membrane glycoproteins localized primarily in endosomes and lysosomes of bloodstream forms. *Exp. Parasitol.* **76**, 324-344.

**Chappell, T. G., Konforti, B. B., Schmid, S. L. and Rothman, J. E.** (1987). The ATPase core of a clathrin uncoating protein. *J. Biol. Chem.* **262**, 746-751.

**Clayton, C. E. and Mowatt, M. R.** (1989). The procyclic acidic repetitive proteins of *Trypanosoma brucei*: purification and post-translational modification. *J. Biol. Chem.* **264**, 15088-15093.

**Cross, G. A. M.** (1975). Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei. Parasitology* **71**, 393 417.

**Dean, N. and Pelham, H. R. B.** (1990). Recycling of proteins from the Golgi compartment to the ER in yeast. *J. Cell Biol.* **111**, 369-377.

**Doering, T. L., Masterson, W. J., Hart, G. W. and Englund, P. T.** (1990). Biosynthesis of glycosyl phosphatidylinositol membrane anchors. *J. Biol. Chem.* **265**, 611-614.

**Dorn, P. L., Amin, R. A. and Boothroyd, J. C.** (1991). Inhibition of protein synthesis results in super-induction of procyclin (PARP) RNA levels. *Mol. Biochem. Parasitol.* **44**, 133-140.

**Dorner, A. J., Bole, D. G. and Kaufman, R. J.** (1987). The relationship of N-linked glycosylation and heavy chain-binding protein with the secretion of glycoproteins. *J. Cell Biol.* **105**, 2665-2674.

**Englund, P. T., Hajduk, S. L. and Marini, J. C.** (1982). The molecular biology of trypanosomes. *Annu. Rev. Biochem.* **51**, 695-726.

**Engman, D. M., Kirchoff, L. V. and Donelson, J. E.** (1989). Molecular cloning of mtp70, a mitochondrial member of the hsp70 family. *Mol. Cell. Biol.* **9**, 5163-5168.

**Flynn, G. C., Chappel, T. G. and Rothman, J. E.** (1989). Peptide binding and release by proteins implicated as catalysts of protein assembly. *Science* **245**, 385-390.

**Fujiki, Y. S., Fowler, S., Shio, H., Hubbard, A. and P. Lazarow.** (1982). Polypeptide and phospholipid composition of the membrane of rat liver peroxisomes: comparison with the endoplasmic reticulum and mitochondrial membranes. *J. Cell Biol.* **93**, 103-110.

**Griffiths, G.,McDowall, A., Back, R. and Dubochet, J.** (1984). On the preparation of cryosections for immunocytochemistry. *J. Ultrastruct. Res.* **89**, 65-78.

**Harris, M. E., Moore, D. R. and Hajduk, S. L.** (1990). Addition of uridines to edited RNAs in trypanosome mitochodria occurs independently of transcription. *J. Biol. Chem.* **265**, 11368-11376.

**Heschl, M. F. P. and Baillie, D. L.** (1989). Characterization of the hsp70 multigene family of *Caenorhabditis elegans. DNA* **8**, 233-243.

**Hobbs, M. R. and Boothroyd, J. C.** (1990). An expression-site-associated gene family of trypanosomes is expressed in vivo and shows homology to a variant surface glycoprotein gene. *Mol. Biochem. Parasitol.* **43**, 1-16.

**Hsu, M. P., Muhich, M. and Boothroyd, J. C.** (1989). A developmentally regulated gene of trypanosomes encodes a homologue of rat protein disulfide isomerase and phosphoinositol-phospholipase C. *Biochemistry* **28**, 6440-6446.

**Hsu, V. W., Shah, N. and Klausner, R. D.** (1992). A brefeldin A-like phenotype is induced by overexpression of a human ERD-2-like protein, ELP-1. *Cell* **69**, 625-635.

**Kang, P., Ostermann, J., Shilling, J., Neupert, W., Craig, E. A. and Pfanner, N.** (1990). Requirement for hsp70 in the mitochondrial matrix for the translocation and folding of precursor proteins. *Nature* **348**, 137-142.

**Kumar, N. and Zheng, H.** (1992). Nucleotide sequence of a *Plasmodium falciparum* stress protein with similarity to mammalian 78 kDa glucose-regulated protein. *Mol. Biochem. Parasitol.* **56**, 353-356.

**Lee, M. G., Polevere, R. I. and Van der Ploeg, L. H. T.** (1990). Evidence for segmental gene conversion between a cognate *hsp70* gene and the temperature-sensitively transcribed *hsp70* genes of *Trypanosoma brucei. Mol. Biochem. Parasitol.* **41**, 213-220.

**Lewis, M. J. and Pelham, H. R. B.** (1990). A human homologue of the yeast HDEL receptor. *Nature* **348**, 162-163.

**Lewis, M. J. and Pelham, H. R. B.** (1992). Ligand-induced redistribution of a human KDEL receptor from the golgi complex to the endolasmic reticulum. *Cell* **68**, 353-364.

**Lewis, M. J. and Pelham, H. R. B.** (1992). Sequence of a second human KDEL receptor. *J. Mol. Biol.* **226**, 913-916.

**Lewis, M. J., Sweet, D. J. and Pelham, H. R. B.** (1990). The *ERD2* gene determines the specificity of the luminal ER protein retention system. *Cell* **61**, 1359-1363.

**Lindquist, S. and Craig, E. A.** (1988). The heat-shock proteins. *Annu. Rev. Genet.* **22**, 631-677.

**Muhich, M. L. and Boothroyd, J. C.** (1988). Polycistronic transcripts in typanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. *Mol. Cell. Biol.* **8**, 3837-3846.

**Munro, S. and Pelham, H. R. B.** (1986). An hsp70-like protein in the ER: identity with the 78kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* **46**, 291-300.

**Munro, S. and Pelham, H. R. B.** (1987). A C-terminal signal prevents secretion of luminal ER proteins. *Cell* **48**, 899-907.

**Normington, K., Kohno, K., Kozutsumi, Y., Gething, M. and Sambrook, J.** (1989). *S. cerevisiae* encodes an essential protein homologous in sequence and function to mammalian BiP. *Cell* **57**, 1223-1236.

**Opperdoes, F.** (1990). The glycosome of trypanosomes and leishmania. *Biochem. Soc. Trans.* **18**, 729-731.

**Ozawa, M. and Muramatsu, T.** (1993). Reticulocalbin, a novel endoplasmic reticulum resident $Ca^{++}$-binding protein with multiple EF-hand motifs and a carboxyl-terminal HDEL sequence. *J. Biol.Chem.* **268**, 699-705.

**Pelham, H. R. B.** (1988). Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *EMBO J.* **7**, 913-918.

**Pelham, H. R. B.** (1989). Control of protein exit from the endoplasmic reticulum. *Annu. Rev. Cell. Biol.* **5**, 1-23.

**Pelham, H. R. B.** (1991). Recycling of proteins between the endoplasmic reticulum and the Golgi complex. *Cur. Opin. Cell Biol.* **3**, 585-591.

**Pelham, H. R. B., Hardwick, K. G. and Lewis, M. J.** (1988). Sorting of soluble ER proteins in yeast. *EMBO J.* **7**, 1757-1762.

**Pidoux, A. and Armstrong, J.** (1992). Analysis of the BiP gene and identification of an ER retention signal in *Schizosaccharomyces pombe*. *EMBO J.* **11**, 1583-1591.

**Robbi, M. and Beaufay, H.** (1991). The COOH terminus of several liver carboxylesterases targets these enzymes to the lumen of the endoplasmic reticulum. *J. Biol. Chem.* **266**, 20498-20503.

**Rose, M. D., Misra, L. M. and Vogel, J. P.** (1989). *KAR2*, a karyogamy gene, is the yeast homologue of the mammalian BiP/GRP78 gene. *Cell* **57**, 1211-1221.

**Sambrook, J., Fritsch, E. F. and Maniatis, T.** (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

**Schloss, P., Hermans-Borgmans, I., Betz, H. and Gundelfinger, E. D.** (1988). Neuronal acetylcholine receptors in *Drosophila*: the ARD protein is a component of a high-affinity alpha-bungarotoxin binding complex. *EMBO J.* **7**, 2889-2894.

**Semenza, J., Hardwick, K. G., Dean, N. and Pelham, H. R. B.** (1990). *ERD2*, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell* **61**, 1349-1357

**Simpson, L. and Shaw, J.** (1989). RNA editing and the mitochondrial cryptogenes of kinetoplastida protozoa. *Cell* **57**, 355-366.

**Smith, D. B. and Johnson, K. S.** (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31-44.

**Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonoso, R. A. and Peattie, D. A.** (1989). Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. *Science* **243**, 75-77.

**Taylor, A. E. F. and. Godfrey, D. G.** (1969). A new organelle of bloodstream salivarian trypanosomes. *J. Protozool.* **16**, 466.

**Tokunaga, M., Kawamura, A. and Kohno, K.** (1992). Purification and characterization of BiP/Kar2 protein from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 17553-17559.

**Vickerman, K.** (1969). On the surface coat and flagellar adhesion in trypanosomes. *J. Cell Sci.* **5**, 163.

**Vogel, J. P., Misra, L. M. and Rose, M. D.** (1990). Loss of BiP/GRP78 function blocks translocation of secretory proteins in yeast. *J. Cell Biol.* **110**, 1885-1895.

**von Heijne, G.** (1990). The signal peptide. *J. Memb. Biol.* **115**, 195-201.

(Received 11 March 1993 - Accepted 26 April 1993)