

The surface coat of procyclic *Trypanosoma brucei*: Programmed expression and proteolytic cleavage of procyclin in the tsetse fly

Alvaro Acosta-Serrano*[†], Erik Vassella[‡], Matthias Liniger[‡], Christina Kunz Renggli[§], Reto Brun[§], Isabel Roditi[‡], and Paul T. Englund*[¶]

*Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205; [†]Institut für Zellbiologie, Universität Bern, CH-3012 Bern, Switzerland; and [§]Swiss Tropical Institute, CH-4002 Basel, Switzerland

Communicated by M. Daniel Lane, Johns Hopkins University School of Medicine, Baltimore, MD, December 22, 2000 (received for review November 29, 2000)

Trypanosoma brucei, the protozoan parasite causing sleeping sickness, is transmitted by a tsetse fly vector. When the tsetse takes a blood meal from an infected human, it ingests bloodstream form trypanosomes that quickly differentiate into procyclic forms within the fly's midgut. During this process, the parasite loses the 10⁷ molecules of variant surface glycoprotein that formed its surface coat, and it develops a new coat composed of several million procyclin molecules. Procyclins, the products of a small multigene family, are glycosyl phosphatidylinositol-anchored proteins containing characteristic amino acid repeats at the C terminus [either EP (EP procyclin, a form of procyclin rich in Glu-Pro repeats) or GPEET (GPEET procyclin, a form of procyclin rich in Glu-Pro-Glu-Glu-Thr repeats)]. We have used a sensitive and accurate mass spectrometry method to analyze the appearance of different procyclins during the establishment of midgut infections in tsetse flies. We found that different procyclin gene products are expressed in an orderly manner. Early in the infection (day 3), GPEET2 is the only procyclin detected. By day 7, however, GPEET2 disappears and is replaced by several isoforms of glycosylated EP, but not the unglycosylated isoform EP2. Unexpectedly, we discovered that the N-terminal domains of all procyclins are quantitatively removed by proteolysis in the fly, but not in culture. These findings suggest that one function of the protease-resistant C-terminal domain, containing the amino acid repeats, is to protect the parasite surface from digestive enzymes in the tsetse fly gut.

Trypanosoma brucei, the protozoan parasite that causes sleeping sickness, has a major impact on health in parts of sub-Saharan Africa. The trypanosome's life cycle alternates between a mammalian host and the tsetse fly *Glossina*. There are two forms in the mammalian bloodstream, a proliferating long slender form and a nonproliferating short stumpy form that is preadapted for survival in the fly. When the fly ingests bloodstream trypanosomes, stumpy forms differentiate into procyclic forms within the lumen of the peritrophic matrix in the midgut (1). Approximately 4–7 days after infection, the trypanosome population expands (2), and they migrate to the ectoperitrophic space between the matrix and the gut epithelium. The parasites progress through several rounds of differentiation and migration, eventually becoming proliferative epimastigote forms in the salivary glands. These give rise to infectious metacyclic forms.

When the stumpy forms differentiate to the procyclic forms, the variant surface glycoprotein coat is totally replaced within a few hours with a new coat composed of procyclins (3, 4). Procyclins are unusual glycosyl phosphatidylinositol (GPI)-anchored proteins. There are two major forms of procyclin, EP and GPEET, which differ mainly in the type of amino acid repeats in their C-terminal domains. EP procyclins have Glu-Pro repeats (5, 6), and GPEET procyclins have Gly-Pro-Glu-Glu-Thr repeats (7). In *T. brucei* AnTat 1.1, the fly-transmissible trypanosome used in this study, the procyclin repertoire contains six different species (E.V. and I.R., unpublished data). Whereas the

EP procyclins (EP1–2, EP2–1 and EP3–2, -3, and -4) have 21 to 27 EP repeats, the GPEET procyclin (GPEET2) has five GPEET repeats instead of the six found in some other *T. brucei* strains (8, 9) (see Figs. 2 and 3 for sequences of the AnTat 1.1 procyclins). Procyclins have several posttranslational modifications. All procyclins except EP2 and GPEET contain a single homogeneous *N*-glycan, Man₅-GlcNAc₂, in their N-terminal domains (10, 11), and GPEET procyclin is phosphorylated on the threonine residues of its GPEET repeats (9, 12, 13). The GPI anchors of all procyclins contain a large heterogeneous side chain of branched poly-*N*-acetylglucosamine units terminated with sialic acid residues (10).

The function of procyclins is unknown, although they contribute to the establishment of strong infections in the fly vector. Parasites in which all of the EP procyclin genes are deleted (14) or that have no surface procyclin because of a defect in GPI synthesis (15) are less efficient at establishing infections in flies. There is now good evidence that procyclic trypanosomes can switch their surface coats, changing from one form of procyclin to another. These changes in expression seem to depend on the trypanosome strain and on experimental culture conditions (10–12, 16). Of relevance to this paper is the recent finding that during infection of a tsetse fly by bloodstream trypanosomes the first procyclic forms detected expressed both GPEET and EP procyclins. In contrast, from day 7 onwards, only EP procyclins were found (17). EP procyclins are also found on epimastigote forms in the salivary glands (18) but are lost when the parasite reacquires a variant surface glycoprotein coat in the metacyclic form. EP and GPEET are also coexpressed when trypanosomes differentiate *in vitro*. Both classes of procyclin continue to be synthesized when trypanosomes are cultured in the presence of glycerol, whereas cells cultured in glycerol-free medium repress GPEET with kinetics similar to those of trypanosomes in the fly (17).

Until recently, there has been little biochemical analysis of the procyclin repertoire. It has been impossible to separate the different procyclin species because of their similar structures and their heterogeneous GPI anchors. Although EP and GPEET procyclins can be distinguished by monoclonal antibodies, and

Abbreviations: EP procyclin, a form of procyclin rich in Glu-Pro repeats; GPEET procyclin, a form of procyclin rich in Glu-Pro-Glu-Glu-Thr repeats; GPI, glycosyl phosphatidylinositol; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; aq.HF, 48% aqueous hydrofluoric acid.

[†]Present address: Division of Biological Chemistry and Molecular Microbiology, Wellcome Trust Biocentre, University of Dundee, Dundee DD1 5EH, Scotland, United Kingdom. E-mail: a.f.acostaserrano@dundee.ac.uk.

[¶]To whom reprint requests should be addressed. E-mail: penglund@jhmi.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.041611698. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.041611698

the glycosylated isoforms EP1 and EP3 bind Con A (16, 19), there has been no straightforward way to distinguish between all of the different procyclins. However, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) has proven to be a powerful method in which all procyclins can be easily identified (9, 11, 16). Amazingly, it is not even necessary to purify the procyclins to homogeneity for this analysis. A simple organic extraction procedure, followed by removal of the GPI anchors by hydrofluoric acid (HF) treatment (11), provides a sample with purity adequate for MS. Furthermore, the method is sensitive, requiring very small amounts of protein sample.

We therefore decided to use MS to identify the procyclins expressed during an infection of a tsetse fly. This approach is technically demanding, particularly early in infection, when each fly provides only $\approx 10,000$ parasites. Nevertheless, we have been able to identify procyclin species expressed at different times during the course of midgut infections. In addition, we unexpectedly discovered that N-terminal domains of procyclins expressed in the tsetse are quantitatively removed by proteolysis. In contrast, the C-terminal domains, containing the amino acid repeats, remain intact. These findings provide important clues to procyclin function.

Materials and Methods

Trypanosomes. We used pleomorphic *T. brucei* AnTat1.1 (20, 21), a stock able to differentiate efficiently and that can be transmitted cyclically by tsetse flies (21–23). The stumpy bloodstream forms were harvested from agarose plates or isolated from infected mouse blood (23). For *in vitro* differentiation to procyclic forms, stumpy bloodstream forms were resuspended (2×10^6 cells/ml) in SDM-79 supplemented with 10% FCS (24) in the presence or absence of 10 mM glycerol. Differentiation was triggered by the addition of 6 mM *cis*-aconitate and a drop in temperature to 27°C (25). Cultures were maintained in the same medium that was used for differentiation (17). *T. brucei* 427 procyclic forms (26) were used as a positive control in immunofluorescence experiments.

Infection of Tsetse Flies and Isolation of Midgut Forms. Pupae of *Glossina morsitans morsitans* were from the International Atomic Energy Agency (Seibersdorf, Austria). Teneral (newly hatched) flies were infected with short stumpy bloodstream forms during the first blood meal after emergence (17). Midguts were isolated from infected flies and disrupted by mechanical force (17). Alternatively, midguts were isolated from infected flies on ice, transferred to SDM-79 containing 10% FCS and a protease inhibitor mixture (Complete, Mini, Roche Molecular Biochemicals, used according to the manufacturer's instructions) at 4°C and disrupted in this medium. Large tissue fragments were removed by filtration through filter paper (Schleicher & Schuell), and the cells in the filtrate were washed with PBS containing protease inhibitors.

Purification of Procyclins. For MS analysis, procyclins were extracted from 2×10^7 parasites cultured *in vitro* or from $1-6 \times 10^6$ parasites isolated from tsetse flies. The parasites were freeze-dried and stored at -80°C . After thawing, they were immediately delipidated twice with 100 μl of chloroform/methanol/water, 10:10:3 (v:v:v). After centrifugation (5 min, 10,000 rpm, room temperature), the insoluble pellet was dried in an N_2 stream and procyclins extracted (three times) with 200 μl of 9% 1-butanol. After another centrifugation (5 min, 10,000 rpm, room temperature), the 1-butanol fractions, containing procyclins, were dried in a Speed-Vac (11).

Preparation of Procyclins for MS Analysis. To remove GPI anchors from procyclins, dried 1-butanol extracts were dephosphorylated with 25 μl of 48% aq.HF (Aldrich) for 12 h at 0°C (9). Samples

Table 1. C-terminal fragments of EP procyclins from trypanosomes isolated from fly 7–28 days after infection (Fig. 3B)

Ion no.	Procyclin protein	No. EP repeats	Mass	Days present	Tentative assignment
1	EP1-2*	25	5,871	7, 14, 21, 28	P ³⁵ -G ⁸⁶ -EtN
2	EP1-2*	25	6,083	7, 14, 21, 28	P ³³ -G ⁸⁶ -EtN
3	EP1-2*	25	6,299	7, 14, 21, 28	T ³¹ -G ⁸⁶ -EtN
4	EP1-2*	25	7,903	14, 21, 28	D ²⁷ -G ⁸⁶ -EtN
5	EP1-2*	25	8,018	7, 21, 28	S ²⁶ -G ⁸⁶ -EtN
6	EP1-2	25	8,178	21, 28	D ²⁴ -G ⁸⁶ -EtN
7	EP3-3	25	8,220	7, 14, 21, 28	S ²⁴ -G ⁸⁶ -EtN
8	EP3-2	21	6,898	7, 14, 21, 28	D ²⁶ -G ⁷⁶ -EtN
9	EP3-2	21	7,016	28	D ²⁵ -G ⁷⁶ -EtN
10	EP3-4	27	6,322	7, 14, 21, 28	P ³⁵ -G ⁹⁰ -EtN
11	EP3-4	27	6,532	14, 28	P ³³ -G ⁹⁰ -EtN

Because of high background and impossibility of detecting N-terminal fragments in positive ion mode after mild acid hydrolysis (11), these assignments are tentative.

*EP3-3, if present, could also contribute to this ion.

were then frozen in dry ice-ethanol, dried in a Speed-Vac, and resuspended in 5 μl of 0.1% trifluoroacetic acid (TFA). 1-butanol extracts from uninfected guts were submitted to the same 48% aqueous HF (aq.HF) dephosphorylation procedure. For some mass spectra, aq.HF-treated procyclins were further treated with mild acid hydrolysis [40 mM TFA (Pierce), 100°C, 15 min] to quantitatively cleave EP proteins at DP bonds (11) or to partially cleave GPEET proteins at a DG bond (16). After cooling to 0°C, samples were dried in a Speed-Vac.

MALDI-TOF-MS Analysis. Spectra were acquired in a PerSeptive Biosystems (Framingham, MA) Voyager-DE mass spectrometer calibrated with insulin, thioredoxin, and apomyoglobin. Aq.HF-treated 1-butanol fractions (1 μl containing $3-7 \times 10^5$ cell equivalents from parasites isolated from flies, or $1-2 \times 10^6$ cell equivalents from parasites cultured *in vitro*) were cocrystallized with 1 μl of α -cyano-4-hydroxycinnamic acid (Sigma) as the matrix. Spectra were collected in the negative ion mode, and masses were calculated by using the program PROTEININFO (<http://prowl.rockefeller.edu>). Figs. 2–3 and Table 1 report theoretical masses that agreed, within about 1 Da, with those measured by MS. Procyclin ions contain a C-terminal ethanolamine residue derived from the GPI anchor.

Immunofluorescence. Immunofluorescence was performed on acetone-fixed cells (17). The monoclonal antibody TRBP1/247 (mAb 247), which binds to the dipeptide repeat of EP procyclin (27), was used at a dilution of 1:500. TBRP1/346 (mAb 346), which binds to the first 20 amino acids of mature EP procyclin (27), was used at 1:1,000.

Results

Previous immunofluorescence studies on procyclin expression in the tsetse fly (17) had not revealed which EP procyclins are present or whether the different EP isoforms undergo a programmed shift in expression. To address these issues, we infected tsetse flies with *T. brucei* AnTat 1.1, and at days 3, 7, 14, 21, and 28, we dissected 30–50 guts and harvested trypanosomes. We then isolated procyclins and removed their GPI anchors.

Because the procyclin pellets were larger than those from comparable numbers of cultured trypanosomes, we had to determine whether they were contaminated with gut proteins that could confuse the MS analysis. Therefore, we analyzed the contents of guts from uninfected flies. We subjected the contents of these guts (equivalent in number to those in the standard

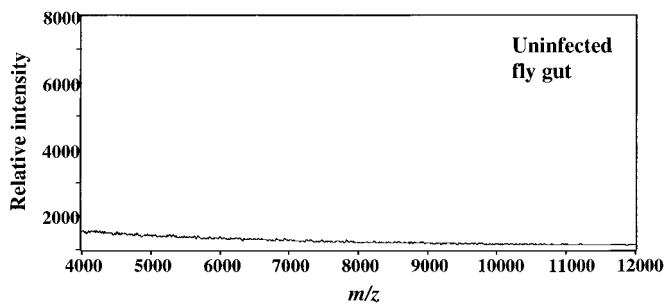


Fig. 1. MS of contents of uninfected fly guts. Guts were isolated 3 days after feeding with noninfected blood, and contents were treated by the procyclin extraction/GPI removal protocol. A 1-butanol extract equivalent to five guts was analyzed.

trypanosome isolation) to the procyclin extraction/GPI removal protocol and analyzed these products by MS. In a sample taken from flies 3 days after they were fed a single noninfected blood meal, we detected no proteins in the m/z range characteristic of procyclins (Fig. 1). We obtained the same result by using several dilutions of the 3-day extract and from a sample taken from 11-day-old flies that had received several noninfected blood meals (not shown). We concluded that contaminants from the gut should not confuse procyclin analyses.

Procyclins Expressed 3 Days After Infection of a Tsetse Fly. Before analyzing material from infected flies, we performed another control experiment, evaluating procyclins from trypanosomes cultured *in vitro* in the presence of 10 mM glycerol [conditions that support expression of GPEET procyclin (17)]. MS analysis of this sample (Fig. 2A) shows a major $[M-H]^-$ pseudomolecular ion of m/z 5,629, corresponding to GPEET2 procyclin. We confirmed the presence of GPEET2 by identifying a fragment lacking the N-terminal sequence VIVK (“-4” fragment; m/z 5,189), which is characteristic of GPEET proteins (9, 11, 12). Furthermore, as expected for GPEET procyclins, both the full length and the “-4” fragment are accompanied by ions separated by an m/z value of 80. These represent the well-characterized family of phosphorylated species (9).

We then analyzed tsetse-derived trypanosomes isolated 3 days after infection. In contrast to the procyclins from cultured parasites, the trypanosomes dissected from the fly contain a major ion of m/z 4,476, much smaller than that of GPEET2, although it is accompanied by phosphorylated species (m/z 4,556, 4,636, and 4,716; Fig. 2B). To test the possibility that the m/z 4,476 ion represented a proteolytic fragment of GPEET2 procyclin, we subjected a sample to mild acid hydrolysis, a treatment that partially cleaves GPEET procyclin at its D¹³-G¹⁴ bond (16). In a control experiment, mild acid treatment of GPEET2 (isolated from parasites cultured *in vitro*) partially generated a cleavage product of m/z 4,232 (together with its

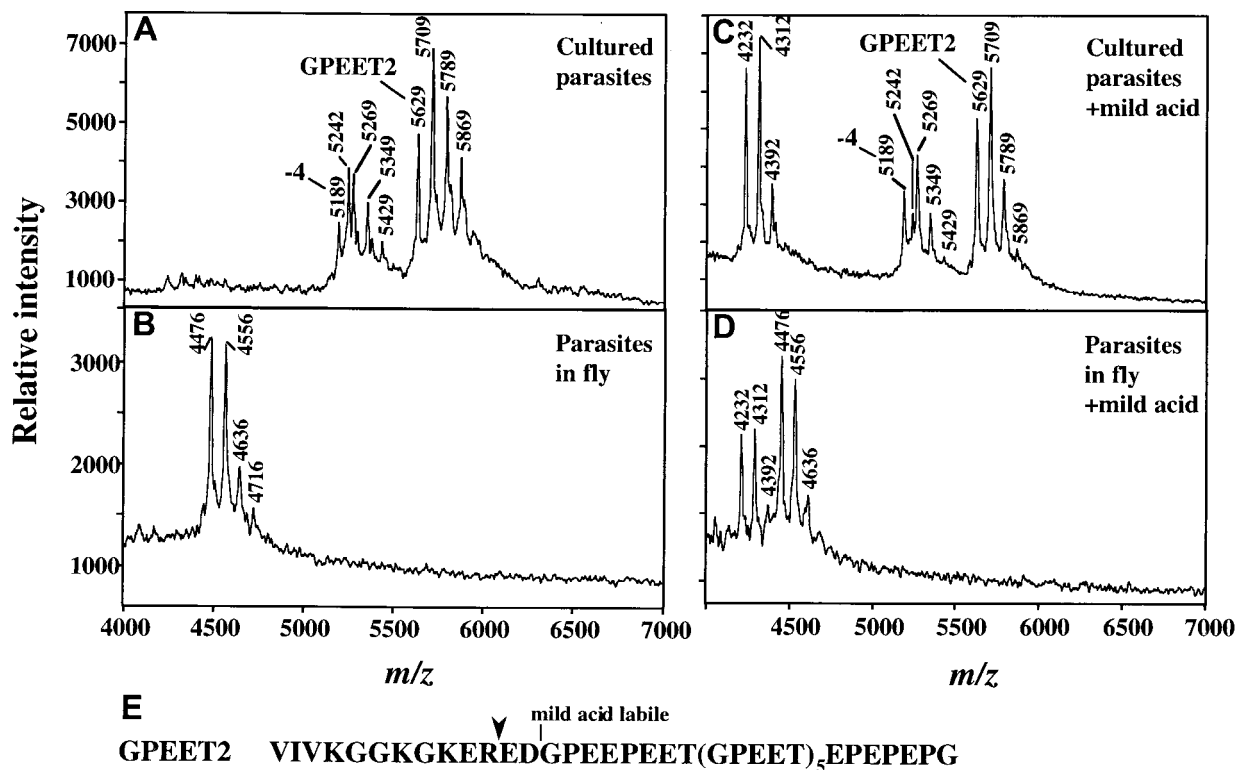


Fig. 2. MS of procyclins from trypanosomes cultured *in vitro* (A and C) or from tsetse flies three days after infection (B and D). (A) Analysis of aq.HF-treated-procycylins from $\approx 10^6$ parasites cultured *in vitro* in the presence of 10 mM glycerol. The ion at m/z 5,629 is full length GPEET2, and it is accompanied by phosphorylated species (m/z 5,709, 5,789, and 5,869). The ion at m/z 5,189 is GPEET2 procyclin truncated by 4 residues, and it is also accompanied by phosphorylated species (m/z 5,269, 5,349, and 5,429). The ion at m/z 5,242 (also present in C) is an unknown polypeptide. Low levels of EP procyclin (EP1–2 and EP3–4) are detected in the 9,000–11,000 m/z range (not shown). (B) Analysis of aq.HF treated-procycylins from $\approx 7 \times 10^5$ parasites extracted from tsetse flies 3 days after infection. The ion at m/z 4,476 is a proteolytic fragment of GPEET2 procyclin (see E for cleavage site), and it is accompanied by ions of phosphorylated species (m/z 4,556, 4,636, and 4,716). (C) Mild acid hydrolysis of aq.HF-treated procycylins from cultured trypanosomes, identical to those in A. (D) Mild acid hydrolysis of the aq.HF treated-procycylins from tsetse-derived trypanosomes, identical to those in B. Note in both C and D, the ion of m/z 4,232, representing the C-terminal fragment generated by partial mild acid cleavage at D¹³ (16). This ion is accompanied by phosphorylated species (m/z 4,312 and 4,392). No fragments of EP procycylins were detected in D. (E) Sequence of GPEET2 procyclin. Downward arrowhead, site of proteolytic cleavage in the fly midgut.

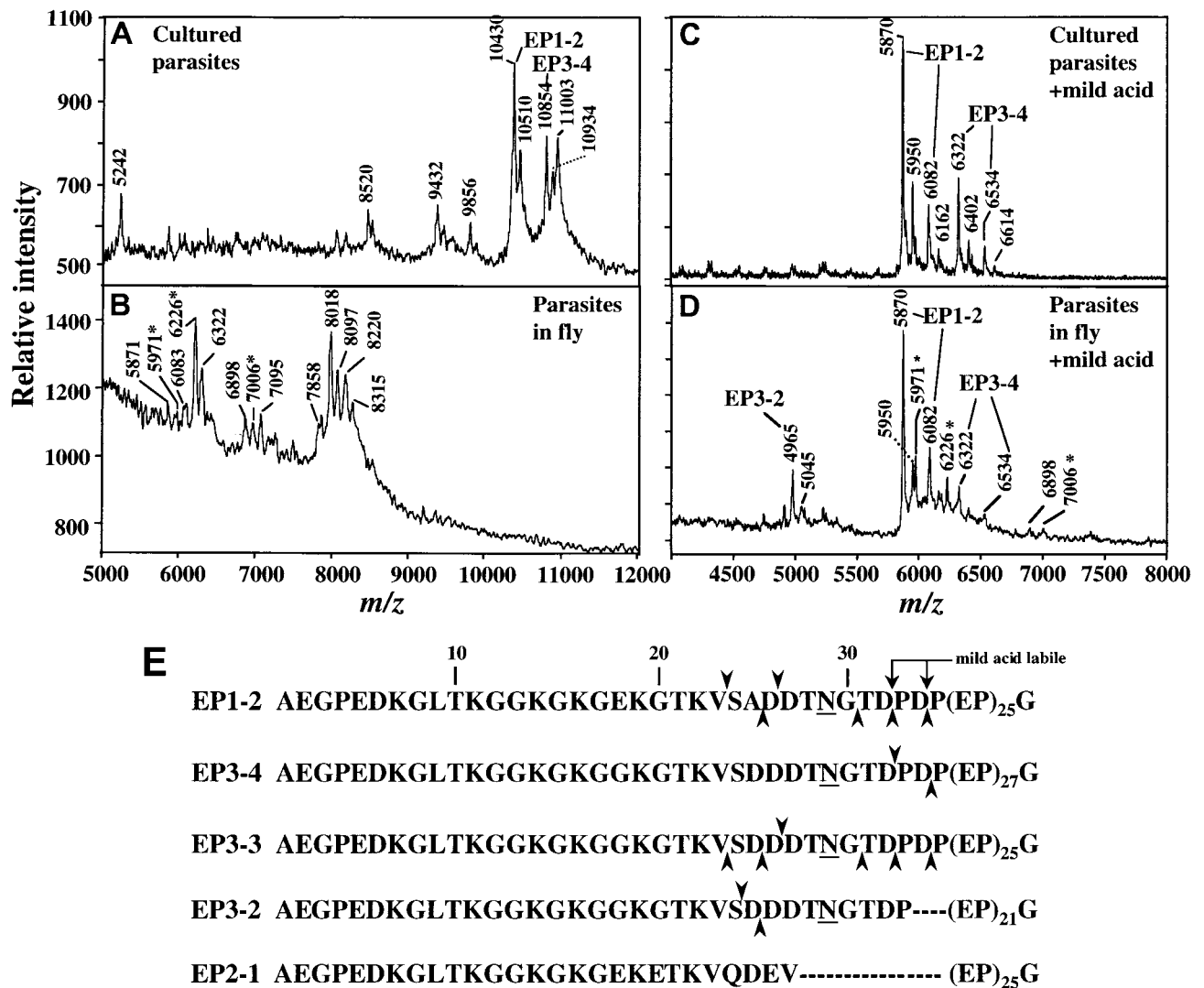


Fig. 3. MS of procyclins from trypanosomes cultured *in vitro* (A and C) or from tsetse flies 7 days after infection (B and D). (A) Analysis of aq.HF-treated-procyclins from $\approx 10^6$ parasites cultured *in vitro* in the absence of glycerol to promote expression of EP procyclin. Ions at m/z 10,430 and 10,854 are from full-length EP1-2 and EP3-4. Ions at m/z 10,510 and 10,934 are from the same procyclin species, except there is a residual phosphate group, derived from the GPI anchor, linked to the C-terminal ethanolamine. These phosphorylated species are not detected after longer HF treatment (11). Ions at m/z 9,432 and 9,856 are from EP1-2 and EP3-4 procyclins, respectively, which are missing 10 amino acids from the N terminus (11). The m/z ions 5,242 and 8,520 are polypeptide contaminants, and that at m/z 11,003 is probably KMP-11, a protein that frequently contaminates procyclin (33). (B) Analysis of aq.HF-treated procyclins from $\approx 3 \times 10^5$ parasites extracted from infected tsetse flies at 7 days after infection. See Table 1 for assignments of ions. No ions derived from GPEET2 were detected. Ions with asterisks in B and D are probably contaminants (see text). (C) Mild acid hydrolysis of aq.HF-treated-procyclins, from cultured trypanosomes, identical to those in A. (D) Mild acid hydrolysis of the aq.HF-treated-procyclins, from tsetse-derived-trypanosomes, identical to those in B. Each of the C-terminal fragments after mild acid hydrolysis (C and D) is paired with a corresponding species phosphorylated on ethanolamine (with m/z increased by 80 Da). (E) Primary sequence of mature EP procyclins, with numbering starting at the N terminus of the mature protein. Upward arrowheads indicate cleavages detected at day 7. Downward arrowheads indicate cleavages detected at days 14–28 (Table 1). Underlines indicate glycosylation sites (N^{29}).

phosphorylated species; Fig. 2C). That we detected the same partial cleavage product after mild acid hydrolysis of the protein isolated from tsetse gut trypanosomes (Fig. 2D) conclusively proves that the m/z 4,476 ion is a fragment of GPEET2 procyclin. Its mass precisely matches that of GPEET2 cleaved at $\text{R}^{11}\text{-E}^{12}$, generating a fragment lacking 11 residues from its N terminus (see sequence in Fig. 2E). This truncated form is observed only in the fly, indicating that midgut proteases quantitatively cleave GPEET2 at a unique site.

Procyclins Expressed 7–28 Days After Infection of a Tsetse Fly. Our next objective was to analyze procyclins from trypanosomes from a later stage, isolated from tsetse guts 7 days after infection. As a control experiment, Fig. 3A shows a MS of procyclins from

trypanosomes cultured *in vitro* in the absence of glycerol, a condition in which procyclin expression is limited to the EP isoforms (17). We detected two major ions at m/z 10,430 and 10,854, corresponding to the EP1-2 and EP3-4 proteins. However, the m/z 10,430 peak could also contain EP3-3 (calculated $[\text{M-H}]^-$, 10,402 Da), as AnTat 1.1 trypanosomes contain the gene for this species. EP3-3 and EP1-2 differ in mass by 28 Da. Fig. 3B shows proteins extracted from trypanosomes 7 days after infection of the tsetse. No intact EP or GPEET procyclins are present. Instead, there are three clusters of ions in the m/z range 6,000 to 8,500.

To determine whether the fragments in Fig. 3B actually derive from procyclin, we used mild acid hydrolysis to promote quantitative cleavage of EP procyclins. This treatment cleaves at

either of the two DP bonds at the junction of the N-terminal domain and the EP-repeat region (see Fig. 3E for location of these bonds). Fig. 3C shows a control spectrum of the mild acid products of EP-procycylins from cultured trypanosomes (same sample used for Fig. 3A). Two of the ions (m/z 5,870 and 6,082) represent the C-terminal fragments P(EP)₂₅G-EtN (EtN, C-terminal ethanolamine) and PDP(EP)₂₅G-EtN, respectively, derived from EP1–2 (or from EP3–3, if present). Likewise, the ions at m/z 6,322 and 6,534 match the predicted values for the EP3–4 fragments P(EP)₂₇G-EtN and PDP(EP)₂₇G-EtN, respectively. When we conducted the same analysis of procycylins isolated from tsetse-derived trypanosomes (like those in Fig. 3A), we found the same C-terminal fragments detected in procycylins from cultured trypanosomes (Fig. 3D), indicating that EP1–2 and EP3–4 (and possibly EP3–3), are expressed in the tsetse. In addition, we found the fragment P(EP)₂₁G-EtN (m/z , 4,965), indicating that the tsetse-derived trypanosomes also express EP3–2, an isoform not detected in cultured parasites.

Characterization of procyclylin species in the mild acid hydrolysis experiment (Fig. 3D) facilitated analysis of the peaks in Fig. 3B, and the identities of most are described in Table 1. They are mainly cleavage products of EP-procyclylin (see Fig. 3E for location of cleavage sites). Some of the ions detected after HF treatment (e.g., m/z 6,226 in Fig. 3B) were resistant to mild acid (these are marked with asterisks in Fig. 3C and D). These fragments could derive from EP2–1 procyclylin, which has no DP bonds and therefore resists mild acid cleavage. However, the fragments marked with asterisks correspond neither to intact EP2–1 protein nor to fragments of it. Therefore, they likely derive from other parasite proteins that are also expressed 7 days after infection or from tsetse proteins induced in response to infection.

In summary, all of these data indicate that at day 7 after infection, tsetse-derived trypanosomes exclusively express glycosylated EP isoforms that are quantitatively cleaved at different positions near the junction of the N-terminal domain and the EP-repeats. We identified a similar distribution of EP procycylins, with only slightly different cleavage patterns, in procycylins isolated from trypanosomes at 14, 21, and 28 days after infection (Table 1).

The tsetse midgut contains proteases (28–30), raising the possibility that the cleavages shown in Figs. 2E and 3E could have occurred during isolation of the trypanosomes from the fly gut. However, we observed the same pattern of cleavage whether or not the trypanosome isolation was done at room temperature or at 4°C and whether or not protease inhibitors and FCS were added during the isolation. The protease inhibitors and serum completely blocked cleavage of procycylins by gut proteases *in vitro* (M.L. and I.R., unpublished data). Therefore, we conclude that procyclylin cleavage occurs naturally within the fly midgut during the course of the infection.

Immunofluorescence Confirms Absence of Procyclylin N-Terminal Domains. We next used immunofluorescence to confirm that EP procycylins from trypanosomes residing in the tsetse midgut are proteolyzed. We used two monoclonal antibodies, one specific for the first 20 amino acids of EP-procyclylin (mAb 346) and another that binds to the EP repeats (mAb 247). Probing trypanosomes from flies 14 days after infection, we found that the mAb 247 bound strongly (Fig. 4 Lower Left), whereas mAb 346 bound poorly (Fig. 4 Upper Left; arrows indicate cells). The weak signal in this panel is most likely because of intracellular EP procyclylin that had not been exposed to midgut proteases. As a control, we found that trypanosomes cultured *in vitro*, where procycylins are not proteolyzed, bind both antibodies equally well (Fig. 4 Right). Therefore, these results support those obtained by MS.

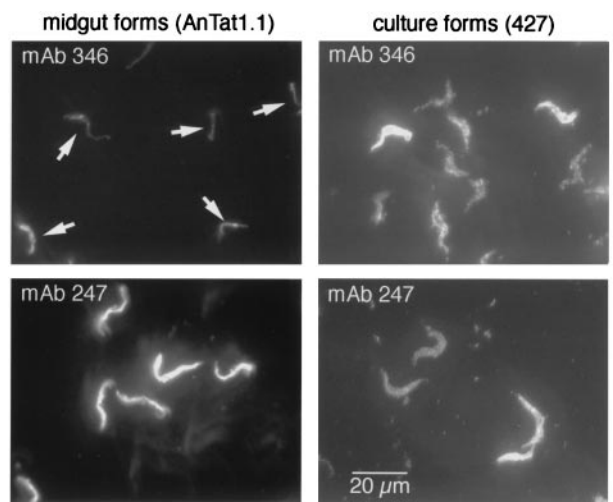


Fig. 4. Immunofluorescence analysis of parasites isolated from tsetse flies. Trypanosomes from tsetse flies 14 days after infection were processed for immunofluorescence and stained with antibodies specific for the EP N terminus (mAb 346) or the EP repeat sequence (mAb 247). Procyclylin culture forms (427) were used as a positive control. The same exposure times and settings were used for all four images. Arrows indicate midgut-derived trypanosomes that were weakly staining for mAb 346 [the location of cells was confirmed by 4',6'-diamidino-2-phenylindole staining (not shown)]. Stock 427, which expresses high levels of GPEET (12), commonly shows heterogeneous staining with anti-EP antibodies (unpublished observations).

Discussion

MALDI-TOF-MS has proven a powerful method for analysis of *T. brucei* procycylins in cultured trypanosomes (9, 11, 16), and in this paper in parasites isolated from the tsetse. Using this method, we have made two major findings. First, we identified procyclylin species expressed in infected tsetse flies from days 3 to 28 after infection. Second, we found that both GPEET and EP were quantitatively cleaved during the infection, removing most of their N-terminal domains. It is surprising that we detected no EP procycylins in tsetse-derived trypanosomes 3 days after infection, even though these species had been detected in infected flies with antibodies specific for the EP repeat (17). Because it is likely that EP and GPEET procycylins can be detected by MS with comparable sensitivities (11), it is possible that low-level expression of multiple EP isoforms, individually below the threshold of detection by MS, can collectively contribute to antibody binding.

GPEET2 procyclylin is strongly expressed 3 days after infection of the tsetse, and it is phosphorylated in the fly as it is in culture (12). However, by day 7, all of the GPEET2 had disappeared and had been replaced by a family of glycosylated EP procycylins (EP1–2, EP3–2, EP3–4, and possibly EP3–3). These are the same EP procycylins expressed *in vitro* (Fig. 3C), except for EP3–2, which was detected in tsetse-derived but not in cultured parasites. We did not see shifts in expression of EP procyclylin during the time period studied. It is striking that neither tsetse-derived trypanosomes (between days 3 and 28 after infection) nor cultured parasites expressed EP2 procyclylin, the only EP procyclylin that lacks *N*-glycosylation. EP2 is rarely expressed in cultured trypanosomes, and we have detected it in abundance only in a *T. brucei* glycosylation mutant that is resistant to killing by Con A (11, 16). We did not have adequate trypanosomes to evaluate procycylins in 1 day infections. At day 27, we dissected salivary glands, hoping to identify the procyclylin species present on epimastigotes. Again, we could not isolate sufficient parasites to detect a procyclylin signal.

It was completely unexpected to learn that most of the N-terminal domains of GPEET and EP procyclins were quantitatively cleaved by proteolysis during a fly infection. No cleavage has been detected in cultured trypanosomes except for the partial removal of VIVK, the N-terminal residues of GPEET procyclin (9, 11, 12). The cleavage in the fly of GPEET2 is at a unique position (Fig. 2E) resembling a trypsin cleavage site. The situation is much more complicated at later times of infection, when EP is expressed. Although all EP procyclins are cleaved at multiple sites (Fig. 3E), the specificity of cleavage is much more complex than that of GPEET procyclin, raising the possibility that the trypanosomes are exposed to a different set of proteases at this time. If this speculation is correct, the difference could be caused by the fact that new proteases are induced at later stages of infection, or that the trypanosomes at day 7 reside in a different compartment, after having traversed the peritrophic membrane. Adult tsetse flies have at least six classes of proteolytic enzymes in the gut, including trypsin-like enzymes, aminopeptidases, and carboxypeptidases (28–30). We do not know whether there is a single proteolytic attack on each procyclin molecule, or whether the released N-terminal products are further fragmented.

As discussed above, procyclins are required for efficient infection of a tsetse fly (14, 15). Our study suggests that one function of procyclins could be to provide protection against the powerful gut proteases. Although most of the N-terminal domains of the procyclins are removed by proteolysis, the amino acid repeat sequences are resistant to attack by proteases. Together with the highly glycosylated GPI anchors (10), these

repeat domains could form a shield that safeguards susceptible proteins on the cell surface from proteolysis. In contrast to procyclic forms, long slender bloodstream forms are killed in the fly midgut within a few hours (31) and are highly susceptible to proteases *in vitro* (32). Interestingly, the same proteases that kill the slender form trigger the differentiation of the stumpy form to the procyclic form (32). Differentiation is synchronous and occurs with rapid kinetics. Thus, proteases also have the potential to regulate development in the fly.

Even though much of the N terminus is removed from EP procyclin, the N-glycan persists at least in some molecules. Therefore it is possible that this glycan could serve as a ligand in some interaction with tsetse proteins. Finally, it is possible that the procyclin N-terminal fragments, removed by proteolysis, exert some signaling function. They could, for example, participate in a density-sensing mechanism, preventing uncontrolled growth of the parasite, or they could deliver a signal for differentiation or migration. Alternatively, they might influence the response of the fly to infection. Further studies are needed to investigate these possibilities.

We thank Viiu Klein for technical support, Terry Pearson (University of Victoria) for providing antibodies, and Dr. Luger (International Atomic Energy Agency) for tsetse pupae. This research was supported by National Institutes of Health Grant AI21334 (to P.T.E.) and by a grant from the Swiss National Science Foundation (31–50932.97) (to I.R.). A.A.S. was supported in part by Consejo Nacional de Investigaciones Científicas y Tecnológicas (Venezuela), and he thanks Judith Campos for stimulation during the course of this work.

- Vickerman, K. (1985) *Br. Med. Bull.* **41**, 105–114.
- Van Den Abbeele, J., Claes, Y., van Bockstaele, D., Le Ray, D. & Coosemans, M. (1999) *Parasitology* **118**, 469–478.
- Ziegelbauer, K., Quinten, M., Schwarz, H., Pearson, T. W. & Overath, P. (1990) *Eur. J. Biochem.* **192**, 373–378.
- Matthews, K. R. & Gull, K. (1994) *J. Cell Biol.* **125**, 1147–1156.
- Roditi, I., Carrington, M. & Turner, M. (1987) *Nature (London)* **325**, 272–274.
- Mowatt, M. R. & Clayton, C. E. (1987) *Mol. Cell Biol.* **7**, 2838–2844.
- Mowatt, M. R., Wisdom, G. S. & Clayton, C. E. (1989) *Mol. Cell Biol.* **9**, 1332–1335.
- Mowatt, M. R. & Clayton, C. E. (1988) *Mol. Cell Biol.* **8**, 4055–4062.
- Mehlert, A., Treumann, A. & Ferguson, M. A. (1999) *Mol. Biochem. Parasitol.* **98**, 291–296.
- Treumann, A., Zitzmann, N., Hulsmeier, A., Prescott, A. R., Almond, A., Sheehan, J. & Ferguson, M. A. (1997) *J. Mol. Biol.* **269**, 529–547.
- Acosta-Serrano, A., Cole, R. N., Mehlert, A., Lee, M. G., Ferguson, M. A. & Englund, P. T. (1999) *J. Biol. Chem.* **274**, 29763–29771.
- Bütikofer, P., Ruepp, S., Boschung, M. & Roditi, I. (1997) *Biochem. J.* **326**, 415–423.
- Bütikofer, P., Vassella, E., Ruepp, S., Boschung, M., Civenni, G., Seebeck, T., Hemphill, A., Mookherjee, N., Pearson, T. W. & Roditi, I. (1999) *J. Cell Sci.* **112**, 1785–1795.
- Ruepp, S., Furger, A., Kurath, U., Renggli, C. K., Hemphill, A., Brun, R. & Roditi, I. (1997) *J. Cell Biol.* **137**, 1369–1379.
- Nagamune, K., Nozaki, T., Maeda, Y., Ohishi, K., Fukuma, T., Hara, T., Schwarz, R. T., Sutterlin, C., Brun, R., Riezman, H. & Kinoshita, T. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 10336–10341. (First Published August 22, 2000; 10.1073/pnas.180230697)
- Acosta-Serrano, A., Cole, R. N. & Englund, P. T. (2000) *J. Mol. Biol.* **304**, 633–644.
- Vassella, E., Van den Abbeele, J., Bütikofer, P., Renggli, C. K., Furger, A., Brun, R. & Roditi, I. (2000) *Genes Dev.* **14**, 615–626.
- Vickerman, K., Tetley, L., Hendry, K. A. & Turner, C. M. (1988) *Biol. Cell* **64**, 109–119.
- Hwa, K. Y., Acosta-Serrano, A., Khoo, K. H., Pearson, T. & Englund, P. T. (1999) *Glycobiology* **9**, 181–190.
- Le Ray, D., Barry, J. D., Easton, C. & Vickerman, K. (1977) *Ann. Soc. Belg. Med. Trop.* **57**, 369–381.
- Delauw, M. F., Pays, E., Steinert, M., Aerts, D., Van Meirvenne, N. & Le Ray, D. (1985) *EMBO J.* **4**, 989–993.
- Vassella, E. & Boshart, M. (1996) *Mol. Biochem. Parasitol.* **82**, 91–105.
- Vassella, E., Reuner, B., Yutzky, B. & Boshart, M. (1997) *J. Cell Sci.* **110**, 2661–2671.
- Brun, R. & Schönenberger, M. (1979) *Acta Trop.* **36**, 289–292.
- Brun, R. & Schönenberger, M. (1981) *Z. Parasitenkd.* **66**, 17–24.
- Cross, G. A. & Manning, J. C. (1973) *Parasitology* **67**, 315–331.
- Richardson, J. P., Beecroft, R. P., Tolson, D. L., Liu, M. K. & Pearson, T. W. (1988) *Mol. Biochem. Parasitol.* **31**, 203–216.
- Cheeseman, M. T. & Gooding, R. H. (1985) *Insect Biochem.* **6**, 677–680.
- Van Den Abbeele, J. & Declair, W. (1992) in *Tsetse Control, Diagnosis, and Chemotherapy Using Nuclear Techniques* (International Atomic Energy Agency, Vienna), IAEA-TECDOC-634, pp. 91–103.
- Osir, E. O., Abubakar, L. & Imbuga, M. O. (1995) *Parasitol. Res.* **81**, 276–281.
- Turner, C. M., Barry, J. D. & Vickerman, K. (1988) *Parasitol. Res.* **74**, 507–511.
- Sbicego, S., Vassella, E., Kurath, U., Blum, B. & Roditi, I. (1999) *Mol. Biochem. Parasitol.* **104**, 311–322.
- Bridge, M. A., Zhou, Q., Koop, B. F. & Pearson, T. W. (1998) *Mol. Biochem. Parasitol.* **91**, 359–363.