Molecular & Biochemical Parasitology 168 (2009) 34–42

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Analysis of expressed sequence tags from the four main developmental stages of *Trypanosoma congolense*☆

Jared R. Helm[a,1], Christiane Hertz-Fowler[b,1], Martin Aslett[b], Matthew Berriman[b], Mandy Sanders[b], Michael A. Quail[b], Marcelo B. Soares[c,2], Maria F. Bonaldo[c,2], Tatsuya Sakurai[d], Noboru Inoue[d], John E. Donelson[a,*]

[a] *Department of Biochemistry, Carver College of Medicine, University of Iowa, Iowa City, IA 52242, USA*
[b] *Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK*
[c] *Department of Pediatrics, Carver College of Medicine, University of Iowa, Iowa City, IA 52242, USA*
[d] *National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080-8555, Japan*

ARTICLE INFO

Article history:
Received 16 April 2009
Received in revised form 15 June 2009
Accepted 16 June 2009
Available online 25 June 2009

Keywords:
Metacyclic
Epimastigote
Procyclic
Bloodstream
Variant surface glycoprotein
Cysteine protease
Hypothetical proteins

ABSTRACT

Trypanosoma congolense is one of the most economically important pathogens of livestock in Africa. Culture-derived parasites of each of the three main insect stages of the *T. congolense* life cycle, *i.e.*, the procyclic, epimastigote and metacyclic stages, and bloodstream stage parasites isolated from infected mice, were used to construct stage-specific cDNA libraries and expressed sequence tags (ESTs or cDNA clones) in each library were sequenced. Thirteen EST clusters encoding different variant surface glycoproteins (VSGs) were detected in the metacyclic library and 26 VSG EST clusters were found in the bloodstream library, 6 of which are shared by the metacyclic library. Rare VSG ESTs are present in the epimastigote library, and none were detected in the procyclic library. ESTs encoding enzymes that catalyze oxidative phosphorylation and amino acid metabolism are about twice as abundant in the procyclic and epimastigote stages as in the metacyclic and bloodstream stages. In contrast, ESTs encoding enzymes involved in glycolysis, the citric acid cycle and nucleotide metabolism are about the same in all four developmental stages. Cysteine proteases, kinases and phosphatases are the most abundant enzyme groups represented by the ESTs. All four libraries contain *T. congolense*-specific expressed sequences not present in the *Trypanosoma brucei* and *Trypanosoma cruzi* genomes. Normalized cDNA libraries were constructed from the metacyclic and bloodstream stages, and found to be further enriched for *T. congolense*-specific ESTs. Given that cultured *T. congolense* offers an experimental advantage over other African trypanosome species, these ESTs provide a basis for further investigation of the molecular properties of these four developmental stages, especially the epimastigote and metacyclic stages for which it is difficult to obtain large quantities of organisms. The *T. congolense* EST databases are available at: http://www.sanger.ac.uk/Projects/T_congolense/EST_index.shtml.

1. Introduction

The *Trypanozoon* subgenus of African trypanosomes are protozoan parasites of major socio-economic impact in sub-Saharan Africa. Included in this group are the three "*brucei*" subspecies, which are transmitted by tsetse flies [1]. Two of the subspecies, *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense*, cause human sleeping sickness in East Africa and in West/Central Africa, respectively, and the third, *T. b. brucei*, is responsible for a small amount of livestock disease in Africa but does not survive in human serum [2,3]. Most laboratory-based research on African trypanosomes during the past four decades has been conducted on *T. b. brucei* because it does not pose as much of a laboratory safety concern as the two human-infective subspecies and two developmental stages of its life cycle, the procyclic form in the insect and

Abbreviations: EST, expressed sequence tag; PCF, culture-derived procyclic form; EMF, culture-derived epimastigote form; MCF, culture-derived metacyclic form; BSF, bloodstream form; VSG, variant surface glycoprotein; mVSG, metacyclic VSG; bsfVSG, bloodstream form VSG; CESP, *congolense* epimastigote-specific protein; ESAG, expression-site associated gene; ORF, open reading frame; UTR, untranslated region; nts, nucleotides; aa, amino acids.

☆ *Note:* The sequence data have been submitted to EMBL under the following accession numbers: FN263376–FN292969.

* Corresponding author. Tel.: +1 319 335 7934; fax: +1 319 335 9570.
E-mail address: john-donelson@uiowa.edu (J.E. Donelson).

[1] These two authors contributed equally to the work.

[2] Current address: Children's Memorial Research Center, 2430 N. Halsted Street, Northwestern University, Chicago, IL 60614, USA.

0166-6851/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2009.06.004

the bloodstream form in the mammal, can be readily maintained in culture. The 35-Mb haploid genomic DNA sequence of a specific *T. b. brucei* clone has been determined [4].

The most important trypanosome species for the African livestock industry, however, is *Trypanosoma congolense* [5], which belongs to the *Nannomonas* subgenus and is also transmitted by tsetse flies. *T. congolense* causes nagana, a chronic wasting disease of cattle that is characterized by anemia, weight loss and immunosuppression and is typically fatal if untreated. It has the largest impact by far of any trypanosome species on domestic livestock in Africa. Economic losses in Africa due to *T. congolense* infections of livestock have been conservatively estimated as US$1.3 billion annually [6]. *T. congolense* can be subdivided further on the basis of isoenzyme analysis, repetitive DNA detection and geographic location into three subgroups: Savannah, Forest, and Kilifi (coastal region). The Savannah group is the most prevalent (90% of stocks) followed by the Forest (9%) and Kilifi (1%) groups [7]. A genome sequencing project on *T. congolense* Savannah clone IL3000 [8] is underway [http://www.sanger.ac.uk/Projects/T_congolense/]. The work described here was conducted with this same *T. congolense* clone IL3000.

T. congolense and the *T. brucei* subspecies have similar, but non-identical, life cycles *in vivo*. Bloodstream *T. brucei* occurs as dividing long slender, intermediate forms as well as a non-dividing stumpy form, whereas bloodstream *T. congolense* has the corresponding dividing forms but the stumpy form is less obvious [9]. In the nutrient-rich, 37 °C bloodstream the main challenge for both species is to avoid the mammalian immune system. Their bloodstream forms possess about 10^7 copies of a variant surface glycoprotein (VSG) (~5% of the total cellular protein) [10], and periodically switch from one VSG to another in a successful effort to avoid the mammalian immune responses [11]. When bloodstream trypanosomes are consumed by a tsetse fly during its blood meal, they enter the insect gut and differentiate to the procyclic form, which must adapt to cooler and variable temperatures, new nutrient sources, hostile digestive enzymes and an insect immune system (see [12] for a review of the interactions between African trypanosomes and tsetse flies). This differentiation from bloodstream to procyclic form is associated with (i) morphological changes, (ii) a switch from glycolysis to oxidative phosphorylation for energy metabolism, and (iii) the replacement of the VSG with procyclic-specific surface proteins [13]. The procyclic form and subsequent epimastigote form *T. congolense* contain on their surface a procyclin protein consisting almost exclusively of 13 heptapeptide repeats (EPGENGT) [14], as well as an unrelated invariant protein, GARP (glutamic acid/alanine-rich protein) [15–17]. A *T. congolense* "*congolense* epimastigote-specific protein" (CESP) has recently been reported [18]. The function(s) of these procyclic- and epimastigote-specific proteins have not been explicitly determined [16,18,19], but they may protect against proteases and other enzymes in the tsetse fly midgut and/or they may play a role in determining the sites of subsequent trypanosome development in the fly [16,19].

Studies tracking the progress of *T. congolense* through flies are much less extensive than for *T. brucei*. The development from procyclic to epimastigote form appears to be similar to *T. brucei*, and the most obvious difference is that for unknown reasons *T. congolense* epimastigotes are directed, not to the fly's salivary glands as are *T. brucei* epimastigotes, but to its proboscis and mouth parts where the epimastigotes differentiate into non-dividing, infective metacyclic forms (reviewed in [12]). Similar to *T. brucei*, immunofluorescence studies using monoclonal antibodies suggest these individual *T. congolense* metacyclic organisms possess on their surface one of about 12 different metacyclic VSGs [20,21], instead of the repertoire of as many as 1000 or more VSGs available to bloodstream *T. brucei* [4]. Metacyclic trypanosomes are inoculated during the fly bite into the host, where parasites multiply at the site of the bite and for 5–9 days continue to express metacyclic VSGs [20–22]. The life cycle is complete when the parasites begin to express bloodstream VSGs on their surface and invade the bloodstream and lymphatic system.

We took advantage of the ability to culture the three insect stages of *T. congolense in vitro*, *i.e.*, the culture-derived procyclic (PCF), epimastigote (EMF) and metacyclic (MCF) forms [23,24], and to obtain the bloodstream form (BSF) from laboratory mice infected with the cultured MCF organisms, so that cDNA libraries could be prepared from each of these four developmental stages. We also prepared normalized cDNA libraries from two forms, MCF and BSF. We report here an analysis of expressed sequence tags (ESTs) in each of these six libraries.

2. Methods and materials

2.1. Trypanosomes and culturing conditions

The bloodstream *T. congolense* (Savannah group) IL3000 clone [8], previously obtained from the International Livestock Research Institute (ILRI; Nairobi, Kenya), was maintained at the National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido, Japan. All animal experiments were conducted in accordance with the standards relating to the Care and Management of Experimental Animals at Obihiro University of Agriculture and Veterinary Medicine (No. 14-69).

A frozen stabilate of the BSF IL3000 clone was thawed and aliquots inoculated intraperitoneally into female, 8-week-old BALB/c mice obtained from CLEA Japan Inc. (Tokyo). Infections were monitored via tail nicks and at the first peak of parasitemia the mice were bled by cardiac puncture and blood collected in heparin. Trypanosomes were purified from whole blood on DE-52 anion exchange column chromatography (Whatman plc., Middlesex, UK) [25]. *In vitro* cultures of PCF, EMF and MCF cells derived from these BSF IL3000 cells were obtained following the methods of Hirumi and Hirumi [23]. Briefly, the concentration of the BSF cells was adjusted to 3×10^6 cells/ml in Eagle's minimum essential medium supplemented with 20% fetal bovine serum, 2 mM L-glutamine and 10 mM L-proline. This parasite suspension was transferred to 25 cm^2 culture flasks (10 ml/flask) and the flasks kept in an incubator at 27 °C for a week, during which the BSF cells differentiated to PCF cells. The PCF cells were maintained and amplified by changing medium and preparing subcultures. After 1–2 months, adherent EMF parasites began to appear as small clusters on the bottom surface of the flasks. The clusters increased in size and number, and finally covered the whole bottom surface. When the culture was primarily adherent EMFs on the flasks' bottom surface, the non-adherent MCF parasites began to appear in the supernatant of the confluent EMF cultures [26].

For isolation of the different life-cycle forms, the PCF cells were collected from the culture supernatant by centrifugation at 1500 × *g* for 10 min at 4 °C. They were washed 3 times with phosphate-buffered saline (PBS). EMF cells were collected from confluent culture flasks by scraping the adherent cells from the bottom of the culture flasks. Briefly, the flasks were washed gently three times with 10 ml PBS to remove any remaining PCF cells and newly appearing MCF cells. The EMF cells were removed by scraping from the bottom surface of the flasks and suspending in PBS. The EMF cells were washed three times with PBS by centrifugation at 1500 × *g* for 10 min at 4 °C. The MCF cells (which possess surface VSGs) were separated from the EMF and PCF cells (which do not have surface VSGs) in a mixed culture of MCF, EMF and PCF cells by passing the culture through a DE-52 anion exchange column [23,25]. MF cells passed through the column, while PCF and EMF cells were retained. These *in vitro*-generated MCF cells were used