Int J Biochem Mol Biol 2013;4(2):73-82
www.ijbmb.org /ISSN:2152-4114/IJBMB1305003

Original Article

Identification of coding sequences from a freshly prepared *Trypanosoma brucei brucei* expression library by polymerase chain reaction

Uthman Okalang[1], Ann Nanteza[2], Enock Matovu[2], George W Lubega[2]

[1]Department of Biology, Faculty of Science and Education, Busitema University, Tororo, Uganda; [2]Department of Biomolecular Resources and Biolaboratory Sciences, Animal Resources and Biosecurity, College of Veterinary Medicine, Makerere University, Kampala, Uganda

Received May 25, 2013; Accepted July 6, 2013; Epub July 29, 2013; Published August 15, 2013

Abstract: Animal African trypanosomiasis (AAT) also known as Nagana is a devastating disease among domestic animals in large parts of Sub-Saharan Africa causing loses in milk and meat production as well as traction power. However, there is currently no commercial vaccine against AAT. The parasites have also developed resistance to some of the drugs in use. Moreover, the use of affordable computer-aided wet bench methods in the search for vaccine and/or new drug targets against this disease have not yet been fully explored in developing countries. This study, therefore, explored the use of PCR to screen a freshly prepared bloodstream form *Trypanosoma brucei brucei* (*T. b. brucei*) expression library for coding sequences followed by bioinformatics analyses specifying the functions and importance of these proteins to parasite survival. Eleven protein coding sequences were identified from twenty nine purified clones. The putative retro transposon hot spot protein 4 (RHSP 4) was the only protein with a fully annotated DNA sequence. All the others were hypothetical or had partial or unqualified annotations. RHSP 4 and pyruvate dehydrogenase E1 component, alpha sub-unit (PDE1α) are involved in aerobic respiration whereas succinyl-Co A-3-ketoacid-coenzyme A transferase mitochondrial precursor (SKTMP) is predicted to be involved in ketone body catabolism. Cystathionine beta-synthase (CBS) and alpha-1,3-mannosyltransferase (αMT) have been predicted in cysteine biosynthesis and vesicular transport respectively. The functions of the hypothetical proteins encountered have neither been experimentally determined nor predicted. We hypothesize that both CBS and PDE1α are good drug targets. Overall, about 300 plates are required to PCR screen the entire Trypanosoma brucei genome in approximately eight months. This method is therefore, applicable and affordable in the search for new drug targets under conditions of limited resources among developing countries.

Keywords: Nagana, drug targets, proteins, expression library, PCR, bioinformatics

Introduction

Trypanosoma brucei brucei (*T. brucei brucei*) is a flagellate protozoan parasite of the genus *Trypanosoma*. It is one of the parasites called trypanosomes in Africa that cause a disease in animals known as African animal trypanosomiasis (AAT) or Nagana. AAT is a devastating disease among domestic animals in large parts of Sub-Saharan Africa [1] threatening approximately forty eight million cattle and responsible for losses in milk, meat and traction power amounting to about $ US 1.0-1.2 billion per annum [2]. This parasite has a life cycle that alternates between the tsetse fly (genus Glossina) and a mammal. In the bloodstream of its mammalian hosts, the parasite evades the immune response by antigenic variation (a continual switching of the variant surface glycoproteins (VSGs) that constitute the surface coat). It has a genome of about 1,000 different VSG genes [3-7].

Antigenic variation has greatly hindered vaccine development efforts [8]. On the other hand, chemotherapy is inadequate due to several reasons, including drug resistance [9]. Identification and characterization of novel *T. b. brucei* drug and vaccine targets are the only foreseeable ways of engaging the problem [10]. *T. b. brucei*

cytoskeleton is believed to be one of the sites that may contain applicable targets. It is predominantly made up of microtubules that form the sub-pellicular corset covering the whole parasite body just beneath the plasma membrane except at the flagella pocket. This microtubule cage is responsible for the maintenance of the cell shape and form. The flagellum also made up of microtubules is used for motility and attachment to the host cell membrane [11, 12]. It is reasonable to expect thousands of these drug and vaccine targets along this microtubule skeleton. Those in the underlying plasma membrane and protoplasm would equally be of interest.

Methods such as Edman sequencing, mass spectrometry, X-ray, Nuclear Magnetic Resonance (NMR) have been used mostly in the developed countries to identify these proteins. However, these methods are expensive and not readily available in developing countries, hence the need for more affordable methods, which can be boosted by the rampant computer-aided technology. Therefore, in this study, we used PCR to screen a freshly prepared bloodstream form *T. b. brucei* expression library for coding sequences that were analyzed by bioinformatics tools.

Materials and methods

Isolation of T. b. brucei from blood by anion exchange chromatography

Ten adult mice were infected with *T. b. brucei* (UTRO 01029B) and blood from them harvested by cardiac puncture at a parasitemia of about 10^7 to 10^8 by matching method [13]. The parasites were then isolated from blood by anion exchange chromatography using DEAE cellulose according to the SIGMA protocol (Sigma Prod. D6418, D0909 and D3764), pelleted, snap frozen in liquid nitrogen and stored at -80°C.

Extraction of mRNA from T. b. brucei

Pure mRNA (5 µg) was extracted from bloodstream form *T. b. brucei* parasites using the magnetic mRNA isolation kit from New England Bio Labs (Cat. No: NEB S1550S) and measured in a NanoDrop ND-1000 UV-Vis spectrophotometer from Thermo Fisher Scientific.

Preparation of the bloodstream form T. b. brucei expression library and titration

Complementary DNA (cDNA) was synthesized using the ZAP Express cDNA Synthesis Kit from Stratagene (Cat. No: 200401). The first cDNA strand was synthesized and used as a template for the synthesis of the second cDNA strand. The cDNA termini were blunted; EcoR I Adapters were ligated using T4 DNA ligase, phosphorylated by T4 polynucleotide kinase and digested using Xho I according to the manufacturer's instructions. To ascertain the presence of cDNA, a 500 bp α-tubulin gene fragment was amplified from 0.5 µl of the cDNA sample in a total reaction volume of 25 µl by PCR using the primer set AAGGATCCCAGAGAAGGCCTACCACGAG and AAAAAGCTTCTCTTCCATACCCTCACCGA as the forward and reverse primers respectively. A control reaction was also set up using genomic DNA extracted from *T. b. brucei*. The products were run on a 2% agarose gel, stained with ethidium bromide and visualized under UV light (gel not shown). The synthesized cDNA was processed to remove contaminating proteins, ligated into Uni-ZAP XR Express® Vector and introduced into the Gigapack® III Gold Packaging Extract to obtain the primary library. This was immediately tittered, amplified and its titer determined using competent XL1-Blue MRF´ cells according to the Stratagene protocol (Cat. No: 200401).

Library plaque generation and PCR amplification of these plaque sequences

Briefly, a warm water-bath was pre-heated to 52°C. Meanwhile, LB agar plates (20 x 150 mm) were placed inside an incubator at 37°C to warm. Eight ml of melted LB/0.7% agarose was aliquoted into each of the two sterile 15 ml Falcon tubes (under sterile conditions) and immediately stored in the warm water-bath (52°C) until needed. Serial dilution (up to 1:1000) of the library (see above) in SM buffer was done and 2 µl of the diluted library (1:1000) added to each of two labeled 15 ml Falcon tubes containing 1.5 ml of competent XL1-Blue MRF´ cells (OD600 of 1.0) at room temperature. The tubes were then incubated at 37°C for 20 minutes in the orbital shaker to allow the phage to bind to the bacterial cells. Workings quickly, a tube containing LB/0.7% agarose was taken from the water-bath, and its contents added to the falcon tube containing the phage

and bacteria. The warm plates were removed from the incubator, labeled and the mixture poured onto them and swirled to distribute the contents evenly. After allowing setting at room temperature (3-5 minutes), the plates were inverted and incubated at 42°C for four hours or until plaques were clearly visible (about 3mm diameter). Thirty six plaques were randomly selected, stubbed out with a wide bore plastic pipette and put into separate Eppendorf tubes containing 200 µl of sterile distilled water. These were then mashed up with the ends of pipette tips (separately for each sample) and stored at 4°C. The T7 and T3 primers (New England Biolabs) were used as the forward and reverse primers respectively. The amplification reactions were prepared in a total reaction volume of 25 µl as follows: 6 µl phage (template), 0.5 µl 100 pg/µl T7, 0.5 µl 100 pg/µl T3, 0.5 µl 10 mM dNTP mix, 0.13 µl 5.0 U/µl Taq DNA polymerase, 2.5 µl x 10 buffer (Mg free), 1.5 µl 25 mM $MgCl_2$ and 13.38 µl double distilled water. PCR was run for 35 cycles with an initial denaturation at 98°C for 2 minutes, cycling conditions at 95°C for 15 seconds, 52°C for 15 seconds, 72°C for 1.5 minutes and final extension at 72°C for 2 minutes. The PCR amplicons were then run on a 2% agarose gel, stained with ethidium bromide and visualized under UV light and samples, which produced single clean bands of at least 100 bp were selected for further study.

Sequencing of PCR products that produced single clean bands

Twenty nine pure PCR amplicons that showed only one band of at least 100 bp were sent for sequencing at the Inqaba biotech laboratory in South Africa. T7 and T3 primers were used for sequencing for comparison of results and reliability. All the amplicons were successfully sequenced.

Bioinformatics analyses of the cDNA sequences

Bioinformatics analyses were done using the public resources at the National Center for Biotechnology Information (NCBI) website. The NCBI VecScreen program was first used to identify contaminating vector sequences, adapters, linkers and primers by comparing the query sequences with those in the UniVec database. Those sequences that belonged to lambda ZAP expression vector, Uni-ZAP XR Express® Vector such as PBluescript SK (-), adapters, linkers and primers used during this study were subsequently deleted to leave behind the unique sequences from the referred cDNAs. Basic Local Alignment Search Tool (BLAST) was set to query the nucleotide database (ntdb) at NCBI with the resulting sequences using blastn program in order to determine the nucleotide (nt) sequences they aligned within this database. Expasy translate tool was used to translate the nt sequences that did not yield any significant hit into amino acid (aa) sequences. These aa sequences were then queried against the peptide database (ptdb) at NCBI using blastx program to determine the corresponding proteins in this database. SignalP from http://www.cbs.dtu.dk/services/SignalP/ was used to predict the presence of signal peptides in the hypothetical proteins encountered to further predict whether the proteins were secretory or non-secretory; SignalP was also used to predict the possible location of all proteins encountered during this study. Additionally, the TriTrypDB from http://tritrypdb.org was queried for orthologs among the kinetoplastids. This is because orthologs are known to retain the same function over the course of evolution. Thus, an efficacious vaccine or drug that would be successfully developed against one kinetoplastid is likely to be effective for the other kinetoplastid (s) that possess (es) ortholog (s) from which the vaccine or drug was designed and developed.

Results

Construction and titration of the amplified T. b. brucei expression library

The purity and quantity of the mRNA samples were determined using a NanoDrop ND-1000 UV-Vis spectrophotometer from Thermo Fisher Scientific. Pooled mRNA (5.0 µg) was extracted and all pooled samples had a 260/280 ratio of between 1.85-2.00 with smooth curves indicating that they were pure and not degraded. The amplification of a 500 bp α-tubulin gene fragment (gel not shown) indicated successful cDNA synthesis. The primary library titer was determined, and the volume needed for amplification determined to be 250.00 µl. The titer of the amplified library was determined as 1.25 x 106 pfu/µl.

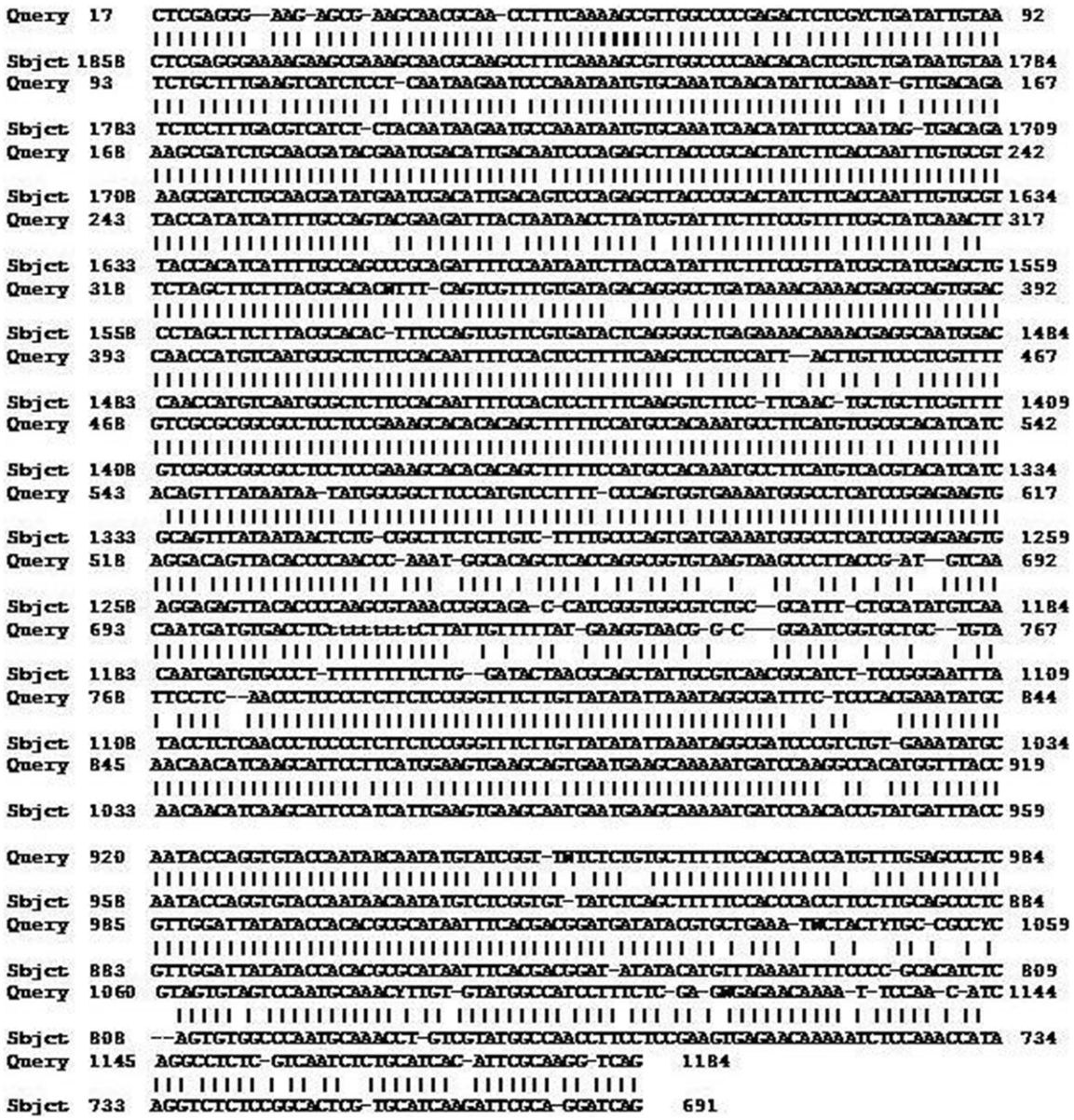

Figure 1. Pwsa for the determination of RHSP 4 (Tb927.1.120). One of the clones matched RHSP4 when BLAST searched in the ntdb at NCBI with Score = 1081 bits (585), Expect = 0.0, Identities = 982/1168 (84%), Gaps = 66/1168 (6%) Strand = Plus/Minus.

Bioinformatics analyses outputs

The only fully annotated DNA sequence hit was RHSP 4. The gene coding for this protein is on chromosome 1 and locus 120. The cDNA isolate corresponded to nt 17-1143 and aligned to segment 1858-716 of the RHSP 4 (**Figure 1**). SignalP analysis predicted RHSP 4 localization in the nucleus of *T. brucei*. Querying the TriTrypDB with the same isolate showed that it's unique to the *T. brucei* complex. The other proteins were hit in the peptide database using the translated cDNA queries as shown in the pwsa (**Figures 2-11**) below. The translated cDNA (aa 1-27) made a hit to PDE1α and aligned on it from aa 241-267 (**Figure 2**). SignalP analysis predicted PDE1α localization in the mitochondria. Querying the TriTrypDB with this cDNA isolate showed orthologs in *L. infantum, L. Mexicana* and *T. cruzi.* The translated cDNA (aa 9-30) made a hit to VSGPx and aligned with it from aa 224-245 (**Figure 3**).

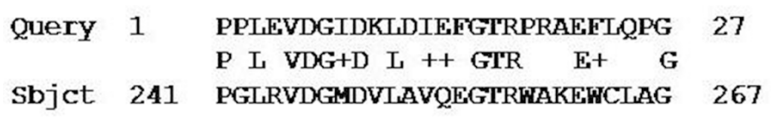

Figure 2. Pwsa for the determination of PDE1α. One of the clones matched *T. brucei* TREU927 PDE1α when BLAST searched in the ptdb at NCBI with Score = 24.6 bits (52), Expect = 0.007, Identities = 12/27 (44%), Positives = 16/27 (59%), Gaps = 0/27 (0%).

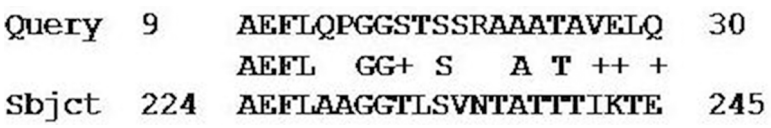

Figure 3. Pwsa for the determination of VSGPx (emb|cad90550.1|). One of the clones matched *T. b. brucei* VSGPx when BLAST searched in the ptdb at NCBI with Score = 22.3 bits (46), Expect = 0.009, Identities = 9/22 (41%), Positives = 13/22 (59%), Gaps = 0/22 (0%).

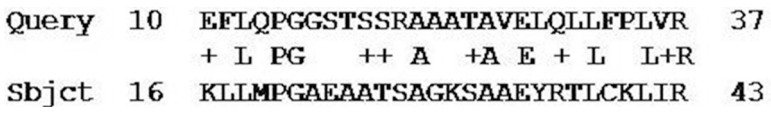

Figure 4. Pwsa for the determination of VSGq (Tb11.v4.0036). One of the clones matched *T. brucei* VSGq when BLAST searched in the ptdb at NCBI with Score = 22.3 bits (46), Expect = 0.009, Identities = 9/28 (32%), Positives = 15/28 (54%), Gaps = 0/28 (0%).

SignalP analysis of this cDNA sequence predicted VSGPx localized in the plasma membrane. Ortho-logs were found in *L. infantum* and *T. cruzi* upon querying this cDNA sequence on the TriTrypDB. The tra-nslated cDNA (aa 10-37) hit VSGq and matched it from aa 16-43 (**Figure 4**). The gene coding for VSGq is on chromosome 11 and locus 0036. SignalP analysis of this cDNA sequence showed that it localizes in the plasma membrane and upon querying it on the TriTrypDB, orthologs were found in *L. mexicana* and *T. cruzi.* The alignment of the translated cDNA query (aa 3-41) to CBS was from aa 16-55 (**Figure 5**). The gene encoding CBS is on chromosome 11 and locus 5400. SignalP analysis of this cDNA sequence showed that CBS has no signal peptide, non-secretory and localizes in the cytoplasm. Upon querying the TriTrypDB, orthologs were found in *L. mexicana* and *T. cruzi.* The alignment of the translated cDNA query (aa 23-51) to αMT was from aa 271-299 (**Figure 6**). The gene coding for αMT is on chromosome 10 and locus 0260. SignalP analysis predicted αMT localization in the endoplasmic reticulum and upon querying the TriTrypDB, orthologs were found in *T. cruzi, L. infantum* and *L. braziliensis.* The alignment of the translated cDNA query (aa 7-28) to SKTMP was from aa 408-429 (**Figure 7**). The gene encoding SKTMP is on chromosome 11 and locus 0290. SignalP analysis of this cDNA predicted that SKTMP localizes in the mitochondria and when the TriTrypDB was queried, orthologs were found in *T. cruzi, L. infantum, L. braziliensis* and *L. major.* The translated cDNA (aa 4-10) hit hypothetical protein I (HP I) and matched its aa 374-410 (**Figure 8**). The gene encoding HP I is on chromosome 7 and locus 4460. SignalP analysis of this cDNA sequence showed that it has a signal peptide and localizes in the mitochondria. Querying this cDNA sequence on the TriTrypDB showed that HP I is unique to *T. brucei.* The alignment of the translated cDNA query (aa 2-16) to hypothetical protein II (HP II) was from aa 2571-2 (**Figure 9**). The gene encoding HP II is on chromosome 6 and locus 620. SignalP analysis of this cDNA sequence showed that it has no signal peptide, non-secretory and localizes in the cytoplasm. Upon querying the TriTrypDB, orthologs were found in *L. Mexicana* and *T. cruzi*. The translated cDNA (aa 14-38) hit and aligned to hypothetical protein III (HP III) from aa 1210-1234 (**Figure 10**). The gene encoding HP III is on chromosome 3 and locus 1420. Signalp analysis of this cDNA sequence showed that it has no signal peptide, non-secretory and localizes in the cytoplasm. Orthologs to HP III were found in *L. infantum*, *L. Mexicana* and *T. cruzi* upon querying the TriTryDB. The alignment of the translated cDNA (aa 4-27) to hypothetical protein IV (HP IV) was from aa 300-326 (**Figure 11**). The gene encoding HP IV is on chromosome 5 and locus 2770. SignalP analysis of this

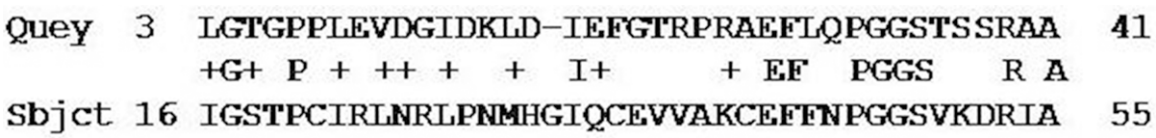

Figure 5. Pwsa for the determination of CBS (Tb11.02.5400). One of the clones matched *T. brucei* TREU927 CBS when BLAST searched in the ptdb at NCBI with Score = 23.9 bits (50), Expect = 0.006, Identities = 11/40 (28%), Positives = 20/40 (50%), Gaps = 1/40 (3%).

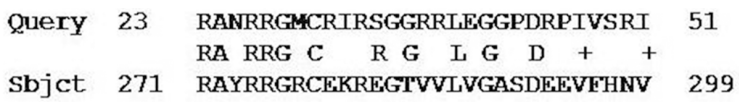

Figure 6. Pwsa for the determination of αMT (Tb10.70.0260). One of the clones matched *T. brucei* TREU927 αMT when BLAST searched in the ptdb at NCBI with Score = 24.3 bits (51), Expect = 0.008, Identities = 11/29 (38%), Positives = 13/29 (45%), Gaps = 0/29 (0%).

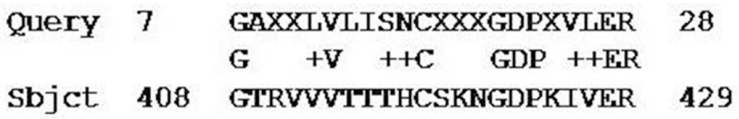

Figure 7. Pwsa for the determination of SKTMP (Tb11.02.0290). One of the clones matched *T. brucei* TREU927 SKTMP when BLAST searched in the ptdb at NCBI with Score = 23.5 bits (49), Expect = 0.005, Identities = 8/22 (36%), Positives = 13/22 (59%), Gaps = 0/22 (0%).

cDNA sequence showed that it has no signal peptide, non-secretory and localizes in the cytoplasm. Orthologs to it were found in *L. infantum* and *T. cruzi* upon querying it against the TriTrypDB.

Discussion

The proteins encountered could be categorized as follows:

Proteins involved in energy production

Both RHSP 4 and PDE1α are involved in energy production through the usual glucose pathway. SKTMP, on the other hand, was inferred from homology search to be a putative protein localized in the mitochondria and involved in ketone body catabolism as an alternative pathway for energy generation [14, 15]. PDE1α is a member of the thiamine diphosphate (ThDP) -requiring 2-oxoacid dehydrogenase complex family and is involved in oxidative decarboxylation of 2-oxoacid substrates to carbon dioxide and acetyl-Co A derivatives [15]. Meanwhile, RHSP 4 is a housekeeping gene with a conserved central region containing an ATP/GTP-binding motif and a ribosomal inserted mobile element insertion site [14]. Immunofluorescence and Western blot analyses with RHSP subfamily-specific immune sera confirmed that RHSP 4 localizes to the nucleus [14] whereas PDE1α was encountered by immunoprecipitation in the identification and characterization of several multi-protein mitochondrial complexes from procyclic form *T. brucei* [16] confirming SignalP prediction that PDE1α localizes in the mitochondria of both procyclic and bloodstream form *T. brucei*. BLAST searching the cDNA isolate of RHSP 4 on the TriTrypDB showed no orthologs in *T. cruzi* but previous Genome Survey Sequence databases indicated that a homologous gene family is present in this parasite [14], confirming bioinformatics analyses as simply predictions, which must be confirmed by experimentation. On the other hand, PDE1α has been predicted to belong to the Thiamine pyrophosphate (TPP) family and has a TPP-binding site for chemical binding, the tetramer and heterodimer interfaces for the binding of polypeptides and the phosphorylation loop region for post-translational modification [17, 18]. It is therefore, plausible to predict that oxidative decarboxylation of 2-oxoacid substrates to carbon dioxide and acetyl-Co A derivatives occurs in *T. brucei* and the organisms with PDE1α orthologs. PDE1α orthologs in the preceding organisms may also have TPP-binding sites for chemical binding, tetramer and heterodimer interfaces for polypeptides' binding and phosphorylation loop regions for post-translational modifications. Thus, the significance of the above functions of RHSP 4 and PDE1α to the

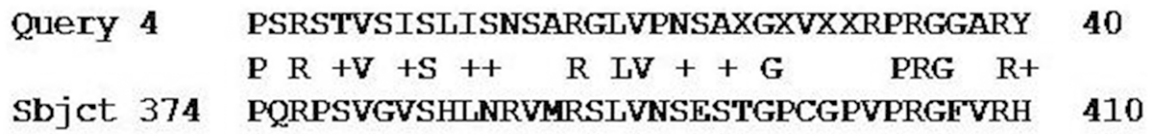

Figure 8. Pwsa for the determination of HP I (Tb927.7.4460). One of the clones matched *T. brucei* TREU927 HP I when BLAST searched in the ptdb at NCBI with Score = 22.3 bits (46), Expect = 0.026, Identities = 12/37 (32%), Positives = 19/37 (51%), Gaps = 0/37 (0%).

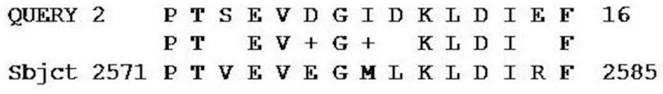

Figure 9. Pwsa for the determination of HP II (Tb927.6.620). One of the clones matched *T. brucei* TREU927 HP II when BLAST searched in the ptdb at NCBI with Score = 24.3 bits (51), Expect = 0.010, Identities = 10/15 (67%), Positives = 12/15 (80%), Gaps = 0/15 (0%).

survival of trypanosomes suggests that these are good targets for drug design and development against trypanosomiasis and diseases caused by organisms possessing the RHSP 4 and PDE1α orthologs.

On the other hand, *T. b. brucei* and organisms with the SKTMP orthologs may survive in the absence of glucose since they have ketone body catabolism as an alternative pathway for energy generation. This makes SKTMP a poor target for drug design and development since this pathway can be evaded by the parasite. However, the role of SKTMP needs to be experimentally determined to ascertain the above predictions.

Proteins involved in anabolism

CBS is the only protein involved in anabolism identified during this study. It has been inferred from electronic annotation that it's involved in the cysteine biosynthesis from serine, cysteine synthase activity, pyridoxal phosphate binding and transferase activity. It was also shown in one of the studies that the growth of *T. brucei* bloodstream forms occurs only in the presence of cysteine either directly added to the medium or reduced from cystine by the action of agents such as Beta-mercaptoethanol and monothioglycerol supplied to the medium as cystine cannot be taken up by bloodstream form trypanosomes [19]. This makes the study of the effects of CBS gene knock down or silencing on trypanosome growth of paramount importance. This could constitute a good candidate for new drug design and development against trypanosomiasis as well as diseases caused by parasites with CBS orthologs.

Proteins involved in antigenic variation

The two proteins identified during this study involved in antigenic variation are VSGq and VSGx. VSGx was inferred from VSGPx whereas VSGq was isolated from plasma membrane sheets of *T. brucei* bloodstream forms by subcellular fractionation in combination with complementary protein separation and identification techniques [20]. There are up to 1,000 different VSG genes within the genome of *T. brucei* [3-7, 17]. They are known for their use by this parasite for the evasion of the host immune system through antigenic variation thus constituting a big challenge to vaccine design and development efforts against trypanosomiasis.

Proteins involved in vesicular transport

Alpha-1,3-mannosyltransferase (αMT) was predicted by groups involved in the sequencing of the genome of the African trypanosome by whole chromosome shotgun (chromosomes 1 and 9 to 11) and bacterial artificial chromosome walking strategies for chromosomes 2 to 8 [17, 18]. Its function has not yet been experimentally determined, however, it has been inferred from electronic annotation that it's involved in hexosyl group's transfer and localizes in the endoplasmic reticulum. Identification and characterization of αMT using a combination of experimental approaches such as monoclonal antibodies, affinity chromatography, immunoprecipitation, mass spectrometry, immunofluorescence, Edman sequencing, X-ray and NMR would provide more significant infor-

```
Query  14     I E F G T R P R A E F L Q P G G S T X S R A A A T  38
              I   F G     P R A     L +   G   S T   S   A +
Sbjct  1210   I L F G A S P R A A V L R G G C S T R S S A V S N  1234
```

Figure 10. Pwsa for the determination of HP III (Tb927.3.1420). One of the clones matched *T. brucei* TREU927 HP III when BLAST searched in the ptdb at NCBI with Score = 23.5 bits (49), Expect= 0.012, Identities = 12/25 (48%), Positives = 14/25 (56%), Gaps = 0/25 (0%).

```
Query  4     GPRAEFLQP---GGSTSSRAAATAVEL   27
               P A  ++P      STSSR+AA +VEL
Sbjct300     DPNAAIVRPRRIEASTSSRSAACSVEL   326
```

Figure 11. Pwsa for the determination of HP IV (Tb927.5.2770). One of the clones matched *T. b. brucei* strain 927/4 GUTAT10.1 HP VI when BLAST searched in the ptdb at NCBI with Score = 22.7 bits (47), Expect = 0.005, Identities = 13/27 (48%), Positives = 17/27 (63%), Gaps = 3/27 (11%).

mation about this enzyme. The function of αMT orthologs could therefore be inferred upon determination of the function of αMT.

Hypothetical proteins

HP I, HP II, HP III and HP IV are conserved with their functions neither experimentally determined nor predicted [17, 18]. The existence of HP III and HP IV was also predicted by scientists involved in the sequencing of the genome of the African trypanosome by whole chromosome shotgun and bacterial artificial chromosome walking strategies [17, 18]. SignalP analysis showed that HP I localizes in the mitochondria meanwhile, HP II, HP III and HP IV occur within the cytoplasm. However, characterization of the plasma membrane sub-proteome of bloodstream form Trypanosoma brucei showed that HP II is found within the plasma membrane (DOI 10.1002/pmic.200700607) and HP IV in the nuclei of procyclic form trypanosomes. The difference in the localization of HP II shows that SignalP solely predicts the possible location of a protein, and only experimentation approaches can confirm, thus this protein most probably localizes in the plasma membrane. The localization of HP IV in the nuclei and cytoplasm of procyclic and bloodstream form trypanosomes respectively suggests that this protein is constitutively expressed. This further suggests that it performs a vital role for the survival of this parasite. The presence of orthologs to HP I, HP II, HP III and HP IV suggest that these proteins have been conserved in these organisms thus the expression, localization and functions of these proteins can be experimentally determined using *T. b. brucei* bloodstream forms and those of their orthologs inferred upon. HP I, HP III and HP IV may constitute candidates for new drug design and development whereas HP II may be for new drug and/or vaccine design and development against AAT and diseases caused by parasites with these protein orthologs.

Identities and positives in this study show how closely similar the queries (nt or aa sequences) are to the genes/proteins they aligned with in terms of structure. The high number of hypothetical proteins (4/11) identified during this study affirms the fact that there are many predicted protein-coding genes (9068) in *T. brucei* [17]. Thus, a large number of genes in trypanosomes are hypothetical and therefore, undetermined. The screening tools used during this study are therefore, reasonably random, picking according to the abundance ratio. The relatively high number of enzymes encountered suggests that there are many vital roles requiring enzymatic activities in trypanosomes. Two VSGs, VSGx and VSGq (2/11) were identified possibly due to the high number (806-1000) of VSGs in *T. brucei* complex used for the evasion of the host immune system. The proteins (eleven) above were determined from twenty nine plaque amplicons randomly selected from two plates each consisting of about one hundred plaques thus can be postulated that 300 plates are required to cover the entire *T. brucei* genome of 9068 predicted genes in about eight months.

In conclusion, this study revealed that this method is feasible as it can pick sequences with novel drug and/or vaccine target value. It is also an affordable method compared to mass spectrometry, X-ray, antibody screening,

NMR and Edman sequencing since the kits, PCR components and computer used are relatively cheap and readily available to the developing world institutions. The construction of the expression library is a well-established method and thus relatively easy. It can readily be repeated many times in an experimental design.

We, therefore, recommend the use of this method coupled with bioinformatics analyses in the search for novel drug and vaccine targets and the study of stage specific gene expression in *T. brucei* and other parasites. The feasibility of CBS and PDE1α as drug candidates should also be investigated by gene knock down or silencing in bloodstream form trypanosomes.

Acknowledgement

This study received support from the Human Frontiers Science Program (HFSP) grant No.RGP0058/2004 to Prof. Keith Gull, Richard McIntosh, LucVanhamme and George W. Lubega. We thank Ms. Monica Namayanja and Mr. Phillip Magambo Kimuda for their technical support.

Address correspondence to: Mr. Uthman Okalang, Department of Biology, Faculty of Science and Education, Busitema University, P.O Box 236, Tororo, Uganda. Phone: +256 772 968176; Fax: +256 454436517; E-mail: uokalang@sci.busitema.ac.ug

References

[1] Aksoy S. Control of tsetse flies and trypanosomes using molecular genetics. Vet Parasitol 2003; 115: 125-45.

[2] Ilemobade AA. Tsetse and trypanosomosis in Africa: the challenges, the opportunities. Onderstepoort J Vet Res 2009; 76: 35-40.

[3] Barry JD, McCulloch R. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. Adv Parasitol 2001; 49: 1-70.

[4] Vanhamme LE, Pays R, McCulloch R, Barry JD. An update on antigenic variation in African trypanosomes. Trends Parasitol 2001; 17: 338-43.

[5] Borst PW, Bitter PA, Blundell I, Chaves M, Cross H, Gerrits F, Van Leeuwen R, McCulloch M, Taylor, Rudenko G. Control of VSG gene expression sites in *Trypanosoma brucei*. Mol Biochem Parasitol 1998; 91: 67-76.

[6] Cross GA, Wirtz LE, Navarro M. Regulation of VSG expression sites transcription and switching in *Trypanosoma brucei*. Mol Biochem Parasitol 1998; 91: 77-91.

[7] Pays E, Nolan DP. Expression and function of surface proteins in *Trypanosoma brucei*. Mol Biochem Parasitol 1998; 91: 3-36.

[8] Cafffrey CR, Scory S, Steverding D. Cysteine Proteinase of Trypanosome Parasites Novel targets for Chemotherapy. Curr Drug Targets 2000; 1: 155-62.

[9] Afework Y, Mäser P, Etschmann B, von Samson-Himmelstjerna G, Zessin KH, Clausen PH. Rapid identification of isometamidium-resistant stocks of *Trypanosoma b. brucei* by PCR-RFLP. Parasitol Res 2006; 99: 253-61.

[10] Zoraghi R, Kunz S, Gong K, Seebeck T. Characterization of TbPDE2A, a novel cyclic nucleotide-specific phosphodiesterase from the protozoan parasite *Trypanosoma brucei.* J Biol Chem 2001 Apr 13; 276: 11559-66.

[11] Hemphill A, Lawson D, Seebeck T. The cytoskeletal architecture of *Trypanosoma brucei.* J Parasitol 1991; 77: 603-612.

[12] Kohl L, Gull K. Molecular architecture of the trypanosome cytoskeleton. Mol Biochem Parasitol 1998; 98: 1-9.

[13] Herbert WJ, Lumsden WHR. Trypanosoma brucei: A rapid "matching" method for estimating the host's parasitemia. Exp Parasitol 1976; 40: 427-431.

[14] Bringaud F, Biteau N, Melville SE, Hez S, El-Sayed NM, Leech V, Berriman M, Hall N, Donelson JE, Baltz T. A new expressed multigene family containing a hot spot for insertion of retroelements is associated with polymorphic subtelomeric regions of *Trypanosoma brucei.* Eukaryot Cell 2002; 1: 137-51.

[15] Reed LJ, Hackert ML. Structure-function relationships in dihydrolipoamide acyltransferases. J Biol Chem 1990; 265: 8971-8974.

[16] Panigrahi AK, Zikova A, Dalley RA, Acestor N, Ogata Y, Anupama A, Myler PJ, Stuart KD. Mitochondrial complexes in *Trypanosoma brucei*. A novel complex and a unique oxidoreductase complex. Mol Cell Proteomics 2008; 7: 534-45.

[17] Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Böhme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC, Arrowsmith C, Atkin RJ, Barron AJ, Bringaud F, Brooks K, Carrington M, Cherevach I, Chillingworth TJ, Churcher C, Clark LN, Corton CH, Cronin A, Davies RM, Doggett J, Djikeng A, Feldblyum T, Field MC, Fraser A, Goodhead I, Hance Z, Harper D, Harris BR, Hauser H, Hostetler J, Ivens A, Jagels K, Johnson D, Johnson J, Jones K, Kerhornou AX, Koo H, Larke N, Landfear S, Larkin C, Leech V, Line A, Lord A, Macleod A,

Mooney PJ, Moule S, Martin DM, Morgan GW, Mungall K, Norbertczak H, Ormond D, Pai G, Peacock CS, Peterson J, Quail MA, Rabbinowitsch E, Rajandream MA, Reitter C, Salzberg SL, Sanders M, Schobel S, Sharp S, Simmonds M, Simpson AJ, Tallon L, Turner CM, Tait A, Tivey AR, Van Aken S, Walker D, Wanless D, Wang S, White B, White O, Whitehead S, Woodward J, Wortman J, Adams MD, Embley TM, Gull K, Ullu E, Barry JD, Fairlamb AH, Opperdoes F, Barrell BG, Donelson JE, Hall N, Fraser CM, Melville SE, El-Sayed NM. The genome of the African trypanosome, *Trypanosoma brucei.* Science 2005; 309: 416-422.

[18] El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Worthey EA, Hertz-Fowler C, Ghedin E, Peacock C, Bartholomeu DC, Haas BJ, Tran AN, Wortman JR, Alsmark UC, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivens AC, Kummerfeld SK, Pereira-Leal JB, Nilsson D, Peterson J, Salzberg SL, Shallom J, Silva JC, Sundaram J, Westenberger S, White O, Melville SE, Donelson JE, Andersson B, Stuart KD, Hall N. Comparative genomics of trypanosomatid parasitic protozoa. Science 2005; 309: 404-9.

[19] Duszenko M, Muhlstadt K, Broder A. Cysteine is an essential growth factor for *Trypanosoma brucei* bloodstream forms. Mol Biochem Parasitol 1992; 50: 269-73.

[20] Bridges DJ, Pitt AR, Hanrahan O, Brennan K, Voorheis HP, Herzyk P, de Koning HP, Burchmore RJ. Characterisation of the plasma membrane subproteome of bloodstream form *Trypanosoma brucei.* Proteomics 2008; 8: 83-99.