

Mass Spectrometric Analysis of the Editosome and Other Multiprotein Complexes in *Trypanosoma brucei*

Aswini K. Panigrahi, Thomas E. Allen,* and Kenneth Stuart

Department of Pathobiology, University of Washington, and Seattle Biomedical Research Institute, Seattle, Washington, USA

Paul A. Haynest† and Steven P. Gygi‡

Department of Molecular Biotechnology, University of Washington, Seattle, Washington, USA

The composition of the editosome, a multi-protein complex that catalyzes uridine insertion and deletion RNA editing to produce mature mitochondrial mRNAs in trypanosomes, was analyzed by mass spectrometry. The editosomes were isolated by column chromatography, glycerol gradient sedimentation, and monoclonal antibody affinity purifications. At least 16 proteins form the catalytic core of the editosome, and additional associated proteins were identified. Analyses of mitochondrial fractions identified several non-editosome proteins and multi-protein complexes. These studies contribute to the functional annotation of *T. brucei* genome. (J Am Soc Mass Spectrom 2003, 14, 728–735) © 2003 American Society for Mass Spectrometry

The identification of the editosome, a multi-protein complex that catalyzes the type of RNA editing that is characteristic of trypanosomatids (for reviews see [1–3]), and its components benefited greatly from the use of mass spectrometry [4, 5]. This type of RNA editing produces mature functional mitochondrial (mt) mRNAs by the post-transcriptional insertion and deletion of numerous uridylates (Us). The editing occurs by a series of coordinated enzymatic steps including endonucleolytic cleavage of the pre-mRNA followed by the addition or removal of one or more Us at the 3' end of the resultant 5' cleavage fragment for insertion or deletion editing, respectively, and the ligation of this processed 5' RNA fragment with the 3' RNA fragment. The sites of U insertion and deletion as well as the number of inserted and deleted Us are specified by small guide RNAs (gRNAs). Both the pre-mRNAs and gRNAs are encoded in the unusual mtDNA of trypanosomatids, which consists of two distinct classes of molecules called maxicircles and minicircles that are intercatenated into a complex DNA network. The max-

icircles are homologous to mtDNAs of other organisms and encode the pre-mRNAs while the minicircles encode the gRNAs. Each (unicellular) trypanosomatid cell contains a single mitochondrion with about 50 identical maxicircles that encode 17 proteins, and thousands of minicircles that only encode gRNAs. The minicircles are of the same size but are very heterogeneous in sequence, which provides for the extensive editing. Editing can be very extensive, entailing the insertion of hundreds and deletion of tens of Us from a single mRNA and hence the use of multiple gRNAs.

The definitive identification of the editosome entailed identification of some of its protein components by mass spectrometry, the demonstration that they were part of the editosome [4, 5], and the demonstration that the inactivation of expression of editosome protein gene(s) resulted in loss of editing and in some cases editosomes, while gene reactivation resulted in reacquisition [6–8]. Sequence similarity and homology searches showed that most editosome proteins occur as related pairs or sets [4, 5, 9]. An early indication of this possibility was the detection of two editing-associated RNA ligases [4, 6] and possible functional differences between ligase activities [10]. The two RNA Editing Ligases, TbREL1 and TbREL2 [6], a 3' terminal uridylyl transferase (TUTase) (Ernst et al., unpublished), a likely U specific exonuclease (Schnauffer et al., unpublished), and a RNA helicase [11] were shown to be components of the stable complex that catalyzes editing [9]. However, functional compensation appears to have compli-

Published online May 21, 2003

Address reprint requests to Dr. K. Stuart, Seattle Biomedical Research Institute, 4 Nickerson St., Seattle, WA 98109, USA. E-mail: kstuart@u.washington.edu

*Current address: Department of Pathology, Colorado State University, Fort Collins, CO.

†Current address: Torrey Mesa Research Institute of Syngenta, San Diego, CA.

‡Current address: Department of Cell Biology, Harvard Medical School, Boston, MA.

cated the determination of the roles and essentiality of some editosome proteins. For example, loss of TbREL1 but not TbREL2 blocks editing suggesting that TbREL1 can compensate for the loss of TbREL2 but the reverse is not the case [6, 8]. Some other proteins may play roles in editing while not being a stable part of the editosome. These may include a 108 kDa protein with TUTase activity that is essential for editing [12], and RNA binding proteins gBP21 [13] and related protein gBP25 ([14] and J. Lukes personal communication), and other RNA binding proteins RBP16 [15], and REAP1 [16].

We report here analyses including the use of mass spectrometry to identify additional candidate editosome proteins and other mitochondrial proteins, some of which were previously unknown. It describes the use of incomplete genome sequence databases and database comparisons for gene and protein identification. These studies contribute to a comprehensive determination of editosome composition and annotation of the *Trypanosoma brucei* genome sequence database.

Experimental

Cell Growth and Mitochondria Isolation

Trypanosoma brucei procyclic cells (strain IsTaR 1.7a) were grown in vitro and mitochondrial vesicles were isolated by hypotonic lysis of the cell, followed by differential and density gradient centrifugations [17].

Isolation of the Protein Complexes

The mitochondria were lysed with 0.5 or 1% Triton-X 100, and clarified by centrifugation at 17,500 g for 30 min at 4 °C. The supernatant was collected and the proteins and complexes were isolated by different fractionation techniques as shown in Figure 1. (1) The editosomes were isolated from cleared mitochondrial lysate by sequential ion-exchange (SP- and Q-Sepharose, Pharmacia Biotech, Uppsala, Sweden) and gel filtration (Superose 6, Pharmacia Biotech, Uppsala, Sweden) column chromatography [4]. (2) The mitochondrial lysate was also fractionated on 10–30% glycerol gradients at 38,000 rpm for 5h (SW40 rotor, Beckman, Palo Alto, CA) and fractions were collected from the top [4]. The editosomes were immuno-affinity purified from glycerol gradient fractionated mitochondrial lysate using anti-TbMP63 MAb [5]. (3) Another protocol to enrich editing complex by differential ammonium sulfate precipitation followed by gel exclusion chromatography resulted in enrichment of the pyruvate dehydrogenase (PDH) complex. The mitochondrial lysate was treated with 15% ammonium sulfate at 4 °C for 30 min, and clarified by centrifugation at 10,000 g for 30 min. The soluble fraction was further precipitated by 30% ammonium sulfate. The precipitate was dissolved in 1X HHE buffer (25 mM HEPES pH 7.9, 10 mM Mg(OAc)₂, 50 mM KCl, 1 mM EDTA) and further

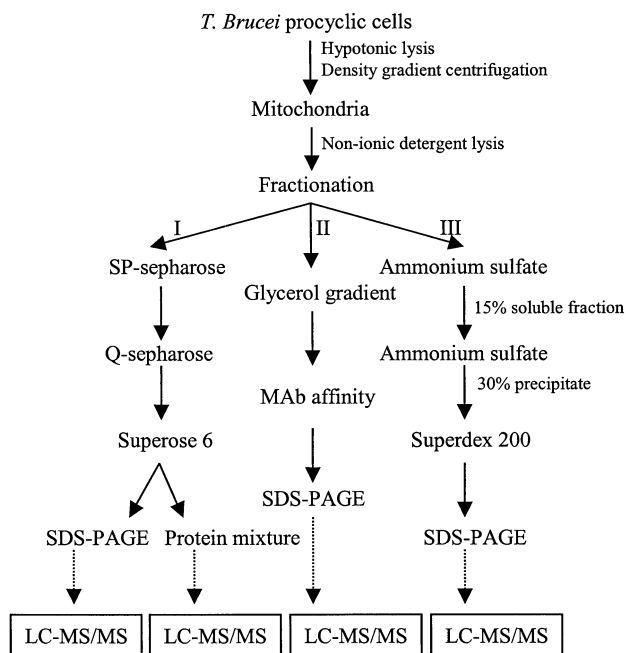


Figure 1. A schematic of the methods used to fractionate the editosome for LC-MS/MS analysis. The solid arrows indicate transfer of pooled active fractions to the next step and dotted arrows indicate preparation for LC-MS/MS.

fractionated by gel filtration chromatography using Superdex-200 column (Pharmacia Biotech, Uppsala, Sweden)[18].

Identification of Proteins

The proteins were analyzed by mass spectrometry either from SDS-PAGE separated gel bands or from the whole mixture. The proteins (10–20 µg) were dissolved in SDS-PAGE buffer, separated on 10% SDS-PAGE gels, and visualized by Coomassie blue or silver staining. The protein bands were excised from the gel and digested with trypsin. The resulting peptides were extracted and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS), by electrospray ionization using LCQ classic or LCQ DECA XP mass spectrometer (Thermo Finnigan, San Jose, CA) ([4] and references therein). Similarly, the complex mixtures of proteins (10 µg) were digested with trypsin in liquid phase and the resulting peptides were separated by C18 reverse phase chromatography (10 cm long, 75 µm i.d. column) and analyzed on line by mass spectrometry. A 30 min linear gradient of 10–60% acetonitrile was used for peptides eluted from gel pieces, and 60 or 90 min gradients were used for peptides derived from digestions of complex protein mixtures. The acquired MS/MS data was analyzed with *T. brucei* nucleotide/protein databases using SEQUEST database search tool for peptide identification. Identification of multiple peptides corresponding to an open reading frame (ORF) identified the protein.

Table 1. Proteins identified in immunoaffinity purified editosomes and their roles

Proteins ^a	Motifs	Functions
TbMP121	None	Putative editosome component
TbMP103	None	Putative editosome component
TbMP100	Endonuclease/exonuclease/phosphatase	Putative exo-Uase
TbMP99	Endonuclease/exonuclease/phosphatase	Putative exo-Uase
TbMP90	Ribonuclease III like	Putative endonuclease
TbMP67	Ribonuclease III like	Putative endonuclease
TbMP61	Ribonuclease III	Putative endonuclease
TbMP46	Pumilio-family RNA binding like	Putative RNA binding
TbMP44	Ribonuclease III like	Required for editing, editosome structure, and cell growth (Wang et al., submitted)
TbMP81	C2H2 zinc finger	Interacts with TbREL2, required for editing, and cell growth (8)
TbMP63	C2H2 zinc finger	Interacts with TbREL1, required for editing, editosome structure, and cell growth (5, 7)
TbMP42	C2H2 zinc finger	RNA/protein interaction
TbMP24	S1 domain	RNA interaction
TbMP18	Single-strand binding protein family	RNA interaction
TbREL1	Ligase	RNA ligase, required for editing, and cell growth (6)
TbREL2	Ligase	RNA ligase, plays role in editing, can be compensated (8)
MHe161p	DEAD box, Helicase C	RNA helicase, plays role in editing, may be compensated (10)
TbMP57	PAP/25A, Nucleotidyl transferase	TUTase
TbRGGm	Glycine rich, RRM	Putative editosome component, RNA interaction

^aThe pairs and sets of proteins that have sequence similarities and homologies among them over blocks of regions are shaded.

Sequence Analysis

The general characteristics of the proteins were predicted by database homology searches. BLAST searches against NCBI and EBI protein databases identified any homologous or related proteins. If no related protein was identified, the protein was designated as novel. The probable functions of the proteins were assigned based on homology searches for motifs and domains in the PROSITE, InterPro, and CDD databases.

Results and Discussion

Identification of Editosome Proteins

Mass spectrometric analyses of the biochemically-purified multi-protein complexes enabled us to identify most protein components of the *T. brucei* editosome despite the incomplete genome sequence database [4, 5, 9]. Criteria used to identify genes for editosome proteins included having predicted mitochondrial-targeting sequences and predicted processed masses corresponding to the mobility of the protein in SDS-PAGE gels. In addition, monoclonal antibodies that were prepared using purified editosomes as antigens and were specific for four proteins were also specific for recombinant proteins from the identified genes. The identified proteins co-sediment at 20S in glycerol gradients, as does *in vitro* editing, and the MAbs immunoprecipitate functional editing complexes [4, 5]. Mass spectrometry analysis of protein bands from editosome preparations identified several proteins, which include two sets of five related proteins and two sets of related pairs of

proteins [9], that have various degrees of sequence similarities and homologies among them (Table 1). These, in addition to two other proteins, contain the core catalysts for editing, the functions of some of which have been identified. However, the ~1600 kDa mass of the editosome that sediments at 20S indicates that additional proteins are associated with it. Additional mass spectrometric analysis of complexes that were purified by immunoaffinity chromatography using a MAbs specific for an editosome protein allowed us to identify three additional candidate editosome proteins. The proteins and corresponding genes were identified by analyses of DNA sequences from the TIGR (<http://www.tigr.org>) and Sanger Institute (<http://www.sanger.ac.uk>) *T. brucei* genome sequencing projects, which are not yet complete. The *T. brucei* databases are periodically searched with unassigned mass spectrometry data in order to identify other candidate editosome genes.

Two of the identified proteins are novel and are designated TbMP121 and TbMP103, based on their pre-processed molecular weight according to a previously described nomenclature [4]. Multiple tryptic peptides that match those predicted from the gene sequence TbMP121 (*T. brucei* chromosome IV clone RPC193-30K5, nucleic acids 87909 to 85108) and TbMP103 (*T. brucei* chromosome XI contig TRYP11.0.004431, nucleic acids 6817 to 3557) were identified and covered 6.1 and 5.4% of the protein sequences, respectively. Both of these proteins have predicted mitochondrial import signals but no known motifs or domains were identified in these proteins.

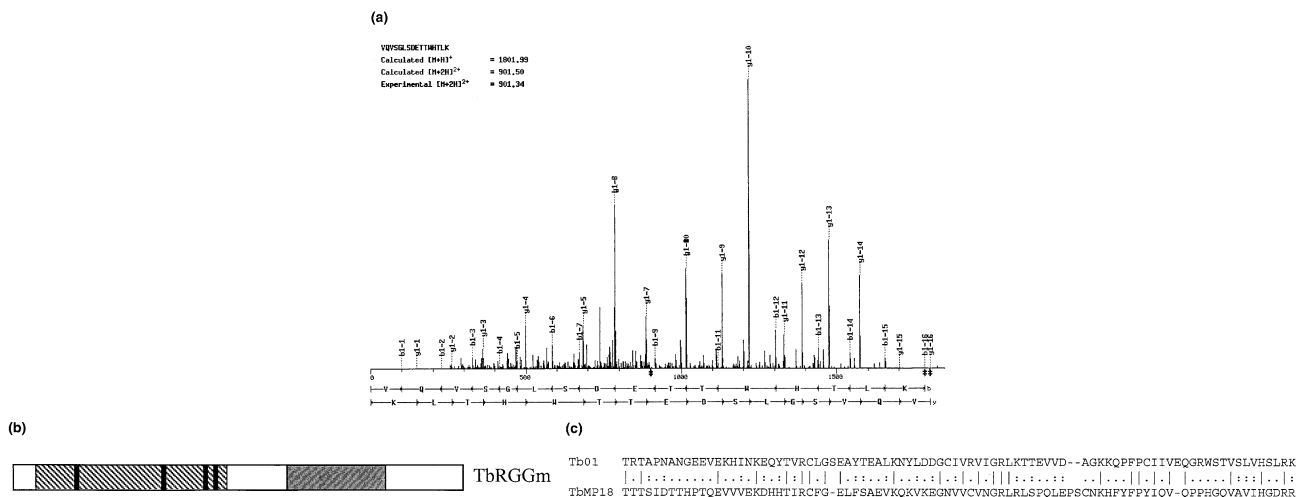


Figure 2. Editosome protein identification and characterization. (a) A typical CID mass spectrum that was acquired by LC-MS/MS analysis. This tryptic peptide VQVSGLSDETTWHILK maps to the TbRGGm protein and matches the predicted spectrum both from N- to C-terminus (b-ions) and C- to N-terminus (y-ions) with a cross-correlation score of 5.14. (b) TbRGGm has a glycine rich N-terminal region (hatched) with 4 RGG repeats (black), and a C-terminal region with a RNA recognition motif (gray). (c) Amino acids that are conserved (perpendicular dash), semi-conserved (semicolon), and partially conserved (dot) between the mid-region of TbMP18 (amino acids 50–135) and a portion of Tb01.

Another protein was identified from peptides that span 15.2% of the protein (*T. brucei* chromosome X gene TRYP10.0.000313_83) and has preprocessed mass of 33 kDa. A representative collision induced dissociation (CID) mass spectrum of a peptide from this protein is shown in Figure 2a. It is an ortholog (overall 75.2% identity, 82.6% similarity) of the 35 kDa *T. cruzi* TcRGGm protein, and thus is designated TbRGGm. TcRGGm, which is an RNA binding protein, has been reported as localized in the nucleus and cytoplasm [19]. Like TcRGGm, TbRGGm has a glycine rich N-terminal region and a RNA recognition motif in the C-terminal half (Figure 2b). The glycine rich region contains four RGG repeats, which is also suggestive of RNA binding, and there are six WGSG_n ($n = 1-7$) repeats between first and second RGG repeats. The glycine rich N-terminal region and presence of RGG repeats resembles mitochondrial protein TbRGG1 [20], a RNA binding protein that co-localizes with RNA editing activity and may have a role associated with the editing process. Whether TbRGGm is associated with the editing needs further investigation.

Roles of Editosome Proteins

One set of five related editosome proteins has three members with zinc fingers and all have oligonucleotide binding motifs or sequences related to these motifs [5]. Two of the proteins, TbMP63 and TbMP81, specifically interact with the TbREL1 [5] and the related TbREL2 ligase proteins, respectively, and enhance their catalytic activities, perhaps by substrate binding (Schnauffer et al., unpublished). Genetic analyses have demonstrated that TbMP81 and TbMP63 are essential for editing [7, 8].

Four members of the other family of five related proteins (TbMP90, TbMP67, TbMP61, and TbMP44) have an RNase III motif or sequence related to this motif [9], Table 1). One or more of these proteins may provide the endonucleolytic function associated with editing but this has not yet been directly shown. These proteins may also play other roles such as RNA binding, RNA translocation, and protein interaction. Inactivation of TbMP44 gene expression results in loss of editosome structural integrity and its disappearance indicating that it is essential to editosome assembly or structural stability (Wang et al., unpublished). The role of TbMP100 that is related to TbMP99 is unknown, but they each have an endonuclease/exonuclease/phosphatase motif and thus may have exonuclease activity. TbMP57 has TUTase activity (Ernst et al., unpublished). Genetic and biochemical studies have shown some of the above proteins specifically interact with each other (Schnauffer et al., unpublished).

A 108 kDa TUTase that is related to the TbMP57 TUTase has not been found associated with the editosome. However, inactivation of 108 kDa TUTase gene expression blocks editing [12] and thus it may add the 3' oligo(U) tails to gRNA or add Us to pre-mRNA, perhaps by transient interaction with the editosome. A DEAD box protein mHel61p that has RNA helicase activity [11] may represent a parallel situation to the TbMP57 and 108 kDa TUTases. mHel61p is associated with the editosome and the inactivation of its gene expression partially inhibits editing [11]. Several related putative RNA helicase proteins were identified in the *T. brucei* database and one has a predicted mitochondrial targeting signal. This protein, designated putHel67, contains a DEAD box motif and C-terminal helicase

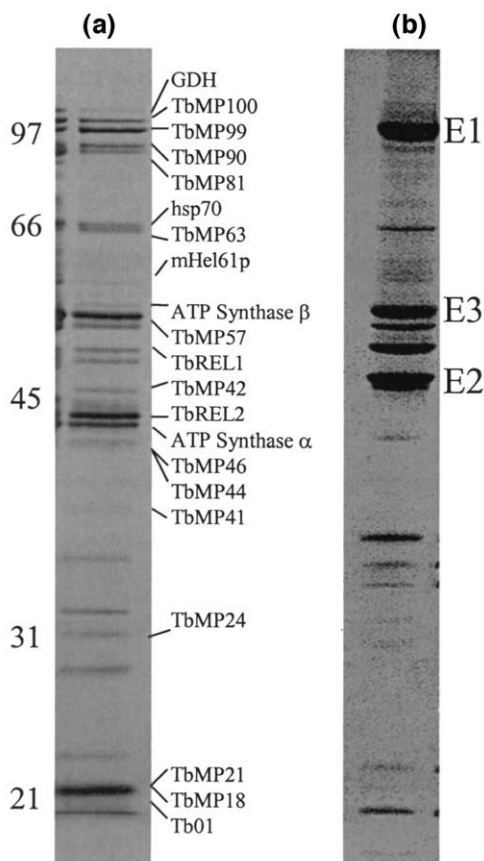


Figure 3. SDS-PAGE protein profile of editosome and PDH enriched fractions. The editosome and non-editosome proteins as well as the positions of the size standards are indicated next to the silver stained gels. (a) editosome fraction and (b) PDH enriched fraction.

domain similar to mHel61p, but peptides corresponding to this protein were not detected in the purified editosomes. It is unclear if this helicase interacts weakly and/or transiently with the editosome or if it has a role in editing.

Additional Putative Editosome Proteins

Based on the mass of the editosome we anticipate that there are editosome proteins in addition to the two families of five proteins, the related TbREL1 and TbREL2 (ligases), related TbMP99 and TbMP100 (possible ExoUase), and TbRGGm, TbMP57 (TUTase), mHel61p (helicase), TbMP103, and TbMP121. The identity of such proteins may come from analyses of mass spectrometry data that has already been collected by comparison with *T. brucei* DNA sequence yet to be placed in the database. Several proteins in addition to the editosome proteins described above were identified in the 1600 kDa editosome fraction that was prepared by sequential ion-exchange and gel filtration chromatography (see Figure 3a). Some of these proteins are probably associated with the editosome and others are

probably present in this fraction because of their high abundance in the cell and/or affinity for RNA. A novel protein designated Tb01, since the complete ORF sequence is not available in the database, has 28.4% identity and 51.1% sequence similarity to editosome protein TbMP18 over an 88 amino acid region (Figure 2c). TbMP41, a novel 41 kDa protein, putative RNA binding protein TbMP21, an ortholog of *Leishmania* p18 (accession CAA83502), and three other novel proteins, whose complete ORF sequence is not yet available, were also identified in this fraction. Tb01 may be associated with the editosome, based on its sequence similarity to TbMP18 that is an integral part of the editosome. Associations of other proteins with the editosome need further investigation. Differences in the protein content of the editosomes that were prepared by the different methods may reflect the effects on protein interactions of the different pH, salt, and non-ionic detergent conditions during fractionation and/or the differential specificity of the MAb affinity chromatography versus the combined ion exchange/gel filtration chromatography.

Other Mitochondrial Proteins and Complexes

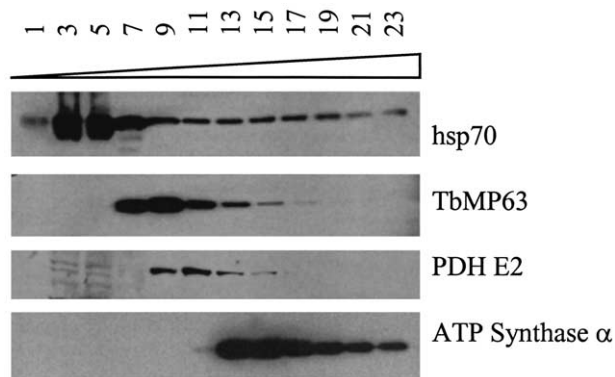
In addition, several other mitochondrial proteins and complexes were identified in this study (Table 2a). These proteins were identified by mass spectrometric analyses of bands from SDS-PAGE gels (Figure 3) and/or of protein mixtures that were prepared as described in the Methods section. LC-MS/MS analysis of the protein bands identified glutamate dehydrogenase, hsp70, ATP synthase α and β subunits in the 1600 kDa editosome fraction, which was prepared by sequential SP-sepharose, Q-sepharose, and Superose 6 column chromatography (see Figure 1) in addition to most of the editosome proteins described above (Figure 3a). Similar analysis of the protein mixture in this 1600 kDa fraction after digestion with trypsin identified additional proteins (Table 2a), most of which appear to be non-editosomal. Western analyses of the glycerol gradient fractionated mitochondrial lysate using MAbs, show that hsp70, the E2 subunit of pyruvate dehydrogenase (PDH), the α subunit of ATP synthase sediment differently than the editosome protein TbMP63 (Figure 4). TbMP63 sedimented at \sim 20S, as previously shown [5], but most hsp70 sedimented near the top of the gradient. However, some hsp70 was detected throughout the gradient, which may reflect its high abundance and/or interactions with other proteins. The E2 subunit of PDH complex sedimented lower in the gradient than TbMP63, indicating that the PDH complex has a somewhat higher S value than the editosome. The α subunit of ATP synthase sedimented even lower in the gradient than did the PDH complex, indicating the S value of this complex is greater than that of the PDH complex. These observations indicate that the proteins are not components of the editosome and that ATP synthase α and PDH E2 are components of other multi-protein com-

Table 2. Mitochondrial complexes and proteins identified by LC-MS/MS

Protein/Complex/Family	Identified sub-unit(s)
(a) Proteins identified in 1600 kDa fraction	
Acyl carrier protein	
ATP synthase complex	α , β
Dihydrolipoamide dehydrogenase	
Dihydrolopanamide succinyl transferase	
Dnak-type molecular chaperon	mtp70, mtp68
Glutamate dehydrogenase	
Heat shock family	hsp70, hsp60
Hydratase/decarboxylase	
Inosine-5'-monophosphate dehydrogenase	
Lipoamide acyltransferase	
Probable 2-oxoglutarate dehydrogenase	
Protenase	
Putative cytochrome c oxidase	VI
Pyruvate dehydrogenase complex	E1, E2, E3
Retrotransposon hot spot	RHS1, ORF1, RHS1a
Tb01	Putative editosome protein
TbMP21	Putative editosome protein
TbMP41	Putative editosome protein
Tryparedoxin peroxidase	
(b) Additional proteins identified in mitochondrial lysate	
Aconitase	
ADP/ATP carrier	
Apocytochrome c1	
gBP21	
gBP25	
Heat shock family	hsp78
Oxidase sub-unit	IV
Phosphoenolpyruvate carboxykinase	
Prohibitin	
RBP16	
Reiske iron-sulfur protein	
Succinate dehydrogenase	
Ubiquitin carboxy-terminal hydrolase	

plexes. These abundant proteins are non-editosomal since they were not detected in immunoaffinity purified editosomes by either Western or LC-MS/MS analyses.

Analysis of a mitochondrial fraction prepared by differential ammonium sulfate precipitation followed by gel filtration chromatography (see Figure 1) identified additional proteins. This fraction contained functional editosomes [18] but its protein profile differed from that of editosomes purified by other means (Figure 3). The mass spectrometric analysis identified the E1, E2, and E3 subunits of the PDH complex, which correspond to three of the most intensely stained pro-

**Figure 4.** Western analysis of the glycerol gradient fractionated cleared mitochondrial lysate. Hsp70, editosome protein TbMP63, the E2 subunit of PDH, and the α subunit of ATP synthase were detected using monoclonal antibodies as described in the Methods. Fraction 1 is at the top of the gradient.

tein bands in the gel (Figure 3b) suggesting that the procedure enriched for the PDH complex.

MS/MS analysis identified several proteins present in the 20S fraction of total mitochondrial lysate. After digesting the proteins with trypsin and separating the peptides on a C18 column, only the TbMP63 and TbMP18 editosome proteins were identified, suggesting that they are in low abundance in the mitochondrion. However, the heat shock proteins and proteins of the ATP synthase complex and other mitochondrial complexes were readily detectable. Proteins of the ATP synthase complex were especially detectable and with higher peptide coverage compared to proteins of other complexes, suggesting their higher relative abundance.

The mitochondrion contains numerous proteins, probably hundreds, most of which have yet to be identified. This is especially true in the three related trypanosomatids (*T. brucei*, *T. cruzi*, and *Leishmania major*) whose genomes are in the process of being sequenced. LC-MS/MS analysis of total mitochondrial lysate led to the identification of many proteins, in addition to those above, some with known or probable functions (Table 2b) and some of which are novel. Many other proteins were identified but since 75% of the probable protein coding genes that have been identified by the trypanosomatid genome projects do not have a known function, only a non-informative numerical designation could be assigned to these proteins (data not shown). The presence of these proteins in the mitochondrial preparations and the presence of a mitochondrial targeting signal on some indicates that most are localized in the mitochondrion.

Non-Mitochondrial Proteins

Non-mitochondrial proteins were also identified in the mitochondrial fraction including α - and β -tubulin, which are primarily present in the cytoskeleton, flagellum, and subpellicular microtubules, as well as the A, B, and C paraflagellar rod proteins and kinetoplastid

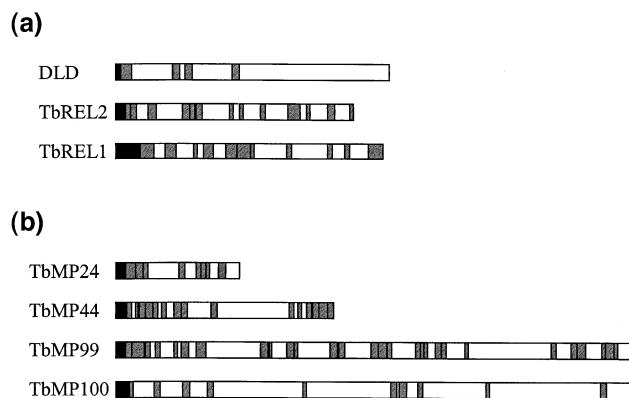


Figure 5. Mitochondrial import signals. The proteins were identified by identification of multiple peptides (gray boxes) across the sequence. (a) The peptide closest to the N-terminus in dihydroliipoamide dehydrogenase (DLD), TbREL2, and TbREL1 is partially tryptic (N-terminal site is non-tryptic) thus identifying the location of the mitochondrial import signal (black box) and cleavage site. (b) The mitochondrial import signal (black box) of editosome proteins TbMP24, TbMP44, TbMP99, and TbMP100 is N-terminal to the most N-terminal fully tryptic peptide (gray box).

membrane protein 4. It also identified organellar proteins such as malate dehydrogenase, hexokinase, fructose-1, 6 biphosphatase, triosephosphate isomerase, fructose-biphosphate aldolase, 6-phospho-1-fructokinase, glycerol dehydrogenase, glyceraldehyde 3-phosphate dehydrogenase, GIM5A, GIM5B, glycerol kinase, and p60 from the glycosome. These are common contaminants of trypanosome mitochondrial preparations and reflect the less stringent purification procedure for preparing mitochondria compared to submitochondrial complexes. The relative abundance of the detected peptides indicates the highest proportion of mitochondrial proteins followed by glycosomal and flagellar proteins in the fraction.

Mitochondrial Targeting Signal

Except for a small number of proteins, mitochondrial proteins are encoded in nuclear DNA, synthesized in the cytoplasm, and imported into the mitochondrion. The mitochondrial proteins contain an import signal at the N-terminus that is cleaved during translocation into the organelle [21]. A variety of algorithms (PSORT, MITOPROT) predict the N-terminal signal, typically from the presence of hydrophobic and basic amino acids that form an amphiphilic helix. The genes for all of the editosome proteins that we have identified to date have such a predicted signal, however, the length of the predicted import signals for the related genes TbREL1 and TbREL2 vary substantially in size (Figure 5a), being 17 and 44 amino acids, respectively [4]. MS/MS data have identified the authentic N-terminus of mitochondrial proteins by identifying a peptide that is very close to the N-terminus of the protein and the predicted mitochondrial targeting signal and has an N-terminal amino acid that results from cleavage during transloca-

tion rather than trypsin digestion. The mitochondrial protein dihydroliipoamide dehydrogenase (DLD) (accession CAA49991) was identified by identification of four peptides, and the N-terminal peptide indicated that the first nine amino acids function as the mitochondrial import signal (Figure 5a). Fully tryptic peptides from editosome proteins TbMP24, TbMP42, TbMP99, and TbMP100 begin at positions 18, 20, 19, and 24, respectively, indicate that the signal sequence is in the small peptides N-terminal to these positions (Figure 5b). Hence, the sizes of the N-terminal mitochondrial targeting signals vary among the proteins, and the mass spectrometry data have identified the regions of the targeting signals and in some cases the cleavage site.

Summary and Perspective

Mass spectrometry has been a very valuable tool for the identification of proteins that are localized in the mitochondrion of trypanosomatids, especially those in multi-protein complexes. It led to the accurate identification of these proteins and components of multi-protein complexes as well as many proteins that are not in complexes and others that are not mitochondrial despite the incomplete trypanosomatid databases. The recent developments in high throughput mass spectrometric analyses of complex mixtures of proteins (see review [22]) provide powerful tools for the identification of large numbers of trypanosomatid proteins when the DNA sequence databases are completed. Determination of the stage of expression, subcellular location, and protein associations in complexes of numerous proteins of the cells' proteome will be a valuable aid for identifying the functions of proteins of these important pathogens. Such information can not only validate computational gene predictions but also help identify the functions of proteins, most of which are unknown, as well as their functional networks.

Acknowledgments

The authors thank Achim Schnauffer, Nancy L. Ernst, Bingbing Wang, and Setty S. Palazzo for unpublished data, Reza Salavati for helpful suggestions, Nicole Carmean for technical assistance, and acknowledge the use of Noreen William's lab for confirming the reactivity of anti-ATP synthase α MAb with the recombinant protein. Sequence data were obtained from The Institute for Genomic Research (<http://www.tigr.org>) and The Sanger Centre (<http://www.sanger.ac.uk>) websites. Mass spectrometric analyses were done in collaboration with Ruedi Aebersold's laboratory. The SBRI Mass spectrometer was purchased with funds from NIH grant RR15916, and this work received support from NIH grant AI14102 to KS.

References

1. Stuart, K.; Panigrahi, A. K.; Salavati, R. RNA Editing in Kinetoplastid Mitochondria. In *RNA Editing: Frontiers in Molecular Biology*; Bass, B. L., Ed.; Oxford University Press: Oxford, 2001; pp. 1–19.

2. Madison-Antenucci, S.; Grams, J.; Hajduk, S. L. Editing Machines: The Complexities of Trypanosome RNA Editing. *Cell* **2002**, *108*, 435–438.
3. Stuart, K.; Panigrahi, A. K. RNA Editing: Complexity and Complications. *Mol. Microbiol.* **2002**, *45*, 591–596.
4. Panigrahi, A. K.; Gygi, S.; Ernst, N.; Igo, R. P., Jr.; Palazzo, S. S.; Schnauffer, A.; Weston, D.; Carmean, N.; Salavati, R.; Aebersold, R.; Stuart, K. D. Association of Two Novel Proteins, TbMP52 and TbMP48, with the *Trypanosoma brucei* RNA Editing Complex. *Mol. Cell Biol.* **2001**, *21*, 380–389.
5. Panigrahi, A. K.; Schnauffer, A.; Carmean, N.; Igo, R. P., Jr.; Gygi, S.; Ernst, N.; Palazzo, S. S.; Weston, D.; Aebersold, R.; Salavati, R.; Stuart, K. D. Four Related Proteins of the *T. brucei* RNA Editing Complex. *Mol. Cell Biol.* **2001**, *21*, 6833–6840.
6. Schnauffer, A.; Panigrahi, A. K.; Panicucci, B.; Igo, R. P., Jr.; Salavati, R.; Stuart, K. D. An RNA Ligase Essential for RNA Editing and Survival of the Bloodstream Form of *Trypanosoma brucei*. *Science* **2001**, *291*, 2159–2162.
7. Huang, C. E.; O'Hearn, S. F.; Sollner-Webb, B. Assembly and Function of the RNA Editing Complex in *Trypanosoma brucei* Requires Band III Protein. *Mol. Cell Biol.* **2002**, *22*, 3194–3203.
8. Drozd, M.; Palazzo, S. S.; Salavati, R.; Igo, R. P., Jr.; Clayton, C.; Stuart, K. D. TbMP81 is Required for RNA Editing in *Trypanosoma brucei*. *EMBO J.* **2002**, *21*, 1791–1799.
9. Panigrahi, A. K.; Schnauffer, A.; Ernst, N. L.; Wang, B.; Carmean, N.; Salavati, R.; Stuart, K. Identification of Novel Components of *Trypanosoma brucei* Editosomes. *RNA*, in press.
10. Cruz-Reyes, J.; Zhelonkina, A. G.; Huang, C. E.; Sollner-Webb, B. Distinct Functions of Two RNA Ligases in Active *Trypanosoma brucei* RNA Editing Complexes. *Mol. Cell Biol.* **2002**, *22*, 4652–4660.
11. Missel, A.; Souza, A. E.; Norskau, G.; Göringer, H. U. Disruption of a Gene Encoding a Novel Mitochondrial DEAD-Box Protein in *Trypanosoma brucei* Affects Edited mRNAs. *Mol. Cell Biol.* **1997**, *17*, 4895–4903.
12. Aphasizhev, R.; Sbicego, S.; Peris, M.; Jang, S. H.; Aphasizheva, I.; Simpson, A. M.; Rivlin, A.; Simpson, L. Trypanosome Mitochondrial 3' Terminal Uridyl Transferase (TUTase): The Key Enzyme in U-Insertion RNA Editing. *Cell* **2002**, *108*, 637–648.
13. Lambert, L.; Muller, U. F.; Souza, A. E.; Göringer, H. U. The Involvement of gRNA-Binding Protein gBP21 in RNA Editing—An In Vitro and In Vivo Analysis. *Nucleic Acids Res.* **1999**, *27*, 1429–1436.
14. Blom, D.; van den Berg, M.; Breek, C. K. D.; Speijer, D.; Muijsers, A. O.; Benne, R. Cloning and Characterization of Two Guide RNA-Binding Proteins from Mitochondria of *Crithidia fasciculata*: gBP27, a Novel Protein, and gBP29, the Orthologue of *Trypanosoma brucei* gBP21. *Nucleic Acids Res.* **2001**, *29*, 2950–2962.
15. Hayman, M. L.; Read, L. K. *Trypanosoma brucei* RBP16 is a Mitochondrial Y-Box Family Protein with Guide RNA Binding Activity. *J. Biol. Chem.* **1999**, *274*, 12067–12074.
16. Madison-Antenucci, S.; Hajduk, S. RNA Editing-Associated Protein 1 is an RNA Binding Protein with Specificity for Preedited mRNA. *Mol. Cell* **2001**, *7*, 879–886.
17. Harris, M. E.; Moore, D. R.; Hajduk, S. L. Addition of Uridines to Edited RNAs in Trypanosome Mitochondria Occurs Independently of Transcription. *J. Biol. Chem.* **1990**, *265*, 11368–11376.
18. Stuart, K.; Kable, M. L.; Allen, T. E.; Lawson, S. Investigating the Mechanism and Machinery of RNA Editing. *Methods* **1998**, *15*, 3–14.
19. Ouaisi, A.; Vergnes, B.; Borges, M.; Guilvard, E. Identification and Molecular Characterization of Two Novel *Trypanosoma cruzi* Genes Encoding Polypeptides Sharing Sequence Motifs Found in Proteins Involved in RNA Editing Reactions. *Gene* **2000**, *253*, 271–280.
20. Vanhamme, L.; Perez-Morga, D.; Marchal, C.; Speijer, D.; Lambert, L.; Geuskens, M.; Alexandre, S.; Ismaïli, N.; Göringer, U.; Benne, R.; Pays, E. *Trypanosoma brucei* TBRGG1, a Mitochondrial Oligo(u)-Binding Protein that Co-Localizes with an in Vitro RNA Editing Activity. *J. Biol. Chem.* **1998**, *273*, 21825–21833.
21. Hauser, R.; Pypaert, M.; Häusler, T.; Horn, E. K.; Schneider, A. In Vitro Import of Proteins into Mitochondria of *Trypanosoma brucei* and *Leishmania tarentolae*. *J. Cell Sci.* **1996**, *109*, 517–523.
22. Griffin, T. J.; Goodlett, D. R.; Aebersold, R. Advances in Proteome Analysis by Mass Spectrometry. *Curr. Opin. Biotechnol.* **2001**, *12*, 607–612.