Identification of novel components of Trypanosoma brucei editosomes

Citation for published version:

Digital Object Identifier (DOI):
10.1261/rna.2194603

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
RNA

Publisher Rights Statement:
Free in PMC.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Identification of novel components of *Trypanosoma brucei* editosomes

ASWINI K. PANIGRAHI,1,2 ACHIM SCHNAUFER,1,2 NANCY L. ERNST,1,2 BINGBING WANG,1,2 NICOLE CARMEAN,1 REZA SALAVATI,1,2 and KENNETH STUART1,2

1Seattle Biomedical Research Institute, Seattle, Washington 98109, USA
2Department of Pathobiology, University of Washington, Seattle, Washington 98195, USA

ABSTRACT

The editosome is a multiprotein complex that catalyzes the insertion and deletion of uridylates that occurs during RNA editing in trypanosomatids. We report the identification of nine novel editosome proteins in *Trypanosoma brucei*. They were identified by mass spectrometric analysis of functional editosomes that were purified by serial ion exchange/gel permeation chromatography, immunoaffinity chromatography specific to the TbMP63 editosome protein, or tandem affinity purification based on a tagged RNA editing ligase. The newly identified proteins have ribonuclease and/or RNA binding motifs suggesting nuclease function for at least some of these. Five of the proteins are interrelated, as are two others, and one is related to four previously identified editosome proteins. The implications of these findings are discussed.

Keywords: RNA editing; protein; mass spectrometry; chromatography; TAP-tag; ribonuclease

INTRODUCTION

Uridylate (U) insertion/deletion editing, which appears to be unique to trypanosomatids, processes most mitochondrial pre-mRNAs to produce mature mRNAs (see Kable et al. 1997; Estevez and Simpson 1999; Stuart et al. 2000; Madison-Antenucci et al. 2002; Stuart and Panigrahi 2002). The editing process is performed by the editosome, a multiprotein complex, which catalyzes the series of coordinated enzymatic steps that result in edited RNA. Small (∼60 nt) mitochondrial guide RNAs (gRNAs) specify the sites of editing and the number of U insertions and deletions. The first enzymatic step in the editing process is gRNA-directed endonucleolytic cleavage of the pre-mRNA at the editing site. Subsequently, U’s are added to the 3’ terminus of the 5’ cleavage product by 3’ Terminal Uridylyl Transferase (TUTase) for insertion or are removed by 3’ exo-uridylase (exo-Uase) for deletion. The U addition or deletion step is followed by ligation of 5’ and 3’ cleavage fragments by RNA ligase. Each gRNA contains sequence information for the editing of multiple sites within a block of 25–35 nt in the mRNA. Most mRNAs require multiple gRNAs to direct the editing of multiple blocks to fully edit the mRNA.

There has been recent progress in characterizing the editosome. The association of editing with a complex was implied by the findings that activities expected for editing sediment at 20–40S as do gRNAs and pre-mRNAs (Pollard et al. 1992; Corell et al. 1996) and two RNA ligases that were identified by adenylation and deadenylation (Sabatini and Hajduk 1995). SDS-PAGE analysis of editosomes purified by various biochemical methods revealed between 7 and more than 20 major protein bands depending on the procedure (Rusché et al. 1997; Madison-Antenucci et al. 1998; Panigrahi et al. 2001a). The genes for the two RNA Editing Ligases (REL) 1 and 2, formerly TbMP52 and TbMP48, respectively, were identified by mass spectrometry of the purified proteins (Panigrahi et al. 2001a) and shown to encode RNA ligases (McManus et al. 2001; Rusché et al. 2001; Schnaufer et al. 2001) and correspond to bands IV and V (Rusché et al. 1997). TbREL1 was shown by gene inactivation studies to be essential for editing and for viability of the parasite (Schnaufer et al. 2001) and similar studies indicated that TbREL2 is not essential for editing or cell survival (Drozdz et al. 2002). Genes for four other editosome proteins, TbMP18, TbMP42, TbMP63, and TbMP81, were also identified by mass spectrometry of purified editosomes (Panigrahi et al. 2001b). These proteins share sequence conservation among themselves to a certain extent and the largest three have zinc fingers. The two largest of these proteins are essential for editing, as inactivation...
of expression of the TbMP63 (band III; Huang et al. 2002) or TbMP81 (Drozdz et al. 2002) gene block editing and result in the loss of the TbREL1 and TbREL2, respectively. A DEAD box helicase, mHEL61p, has a role in editing, as null mutants of *Trypanosoma brucei* have reduced edited mRNAs in vivo (Misset et al. 1997). A mitochondrial 3' TUTase gene was also found to be essential for editing in *T. brucei* by inhibition of its expression (Aphasizhev et al. 2002).

In addition, several RNA-binding proteins, gBP21 (Koller et al. 1997) and the related gBP25 (Blom et al. 2001), as well as REAP1 (Madison-Antenucci et al. 1998), TBRGG1 (Vanhamme et al. 1998), and RBP16 (Hayman and Read 1999), may have roles in RNA editing. At present, it is unclear if or to what extent these proteins are associated with the editosome. The roles of these proteins may be in gRNA processing, annealing of gRNA and mRNA (Muller et al. 2001), and/or transport of RNAs to the editosome (Madison-Antenucci and Hajduk 2001).

This study describes the identification of nine additional editosome proteins, all of which are stably associated with the ∼20S editosome. These proteins are novel but have sequence characteristics suggesting roles in RNA interaction and processing. The editosome proteins identified to date have varying degrees of sequence similarities, which reveal that they occur as pairs and sets of related proteins.

### RESULTS

Editosomes were purified by three different approaches to determine their protein content. They were purified from mitochondrial lysates by sequential column chromatography as previously described (Panigrahi et al. 2001a), from glycerol gradient-fractionated mitochondrial lysates by affinity purification using a MAb specific for editosome protein TbMP63 (Panigrahi et al. 2001b), and from total cell lysates of transgenic trypanosomes by the tandem affinity purification (TAP) procedure (Rigaut et al. 1999). Trypanosomes that express TbREL1 with a C-terminal TAP-tag under the control of a tetracycline-inducible promoter were prepared as described in Materials and Methods. The TAP-tagged protein was efficiently expressed in vivo and glycerol gradient analysis showed incorporation into ∼20S complexes (results not shown).

The TAP-tagged complexes underwent a first affinity chromatography with an IgG column, which binds the protein A segment of the TAP-tag, and were eluted by cleavage with TEV protease. They then underwent a second affinity chromatography with a calmodulin column, which binds the calmodulin-binding protein (CBP) segment of the TAP-tag, and were eluted using EGTA. Editosomes purified by these three methods all contained the four proteins for which we have MAbs, TbMP81, TbMP63, TbREL1 (TAP-tagged in one case), and TbMP42, as shown by Western analysis (Fig. 1A), and a fifth protein TbREL2, along with TbREL1, was identified by adenylation (Fig. 1B). The editosomes purified by these three methods were all functional in full round in vitro deletion editing (Fig. 1C). Hence, all three methods resulted in functional editosomes.

The proteins in the purified editosomes and their corresponding genes were identified by a combination of liquid chromatography tandem mass spectrometry (LC-MS/MS) and DNA sequence database analysis as described.
(Panigrahi et al. 2001b). The editosome proteins were designated TbMPn (Trypanosoma brucei Mitochondrial Protein with n indicating the preprocessed molecular weight of the protein predicted by the open reading frame) and the corresponding genes as TbMPn according to the previous convention (Clayton et al. 1998; Panigrahi et al. 2001a). Analysis of protein bands from MAb affinity-purified editosomes identified 16 proteins (Fig. 1D). A representative collision-induced dissociation (CID) spectrum that was obtained from the mass spectrometer from tryptic peptide VLDLEEVYFR is shown in Figure 2A. This peptide corresponds to position 583–592 on a 90 kD protein predicted by T. brucei ORF CHR1.148 (accession no. CAB95444). Thirteen other tryptic peptides were identified that cover 16.8% of the amino acids in this protein (Fig. 2B). We designated this protein TbMP90. Similarly, TbMP100 (DNA clones 20G11, 32N6-TIGR), TbMP99 (ORF TRYP1.0.000037_106-Sanger), TbMP67 (ORF TRYP1.0.000155_38-Sanger), TbMP61 (ORF TRYP1.0.000155_52-Sanger), TbMP57 (DNA clones 8B08, 105B11, 9C14-TIGR), TbMP46 (DNA clone 28E9-TIGR), TbMP44 (ORF TRYP1.0.000049_1-Sanger), and TbMP24 (ORF TRYP1.0.000155_8-Sanger) were identified by identification of multiple tryptic peptides. Sequence analysis of the TbMP24 and TbMP99 cDNA clones showed that the second and third AUGs are used as the start codons, respectively. As summarized in Table 1, six editosome proteins that we previously identified by LC-MS/MS, TbREL1, TbREL2, TbMP81, TbMP63, TbMP42, TbMP18 (Panigrahi et al. 2001a, 2001b) and the DEAD box protein mHel61p (Missel et al. 1997), were also identified in purified editosomes by multiple peptide matches. Editosomes purified by the three different methods had similar but not identical protein profiles. Of the 16 proteins detected in MAb affinity-purified complexes, two were not detected in editosomes purified by sequential column chromatography and five were not detected in TAP-tag purified editosomes.

FIGURE 2. Identification of editosome protein TbMP90 by LC-MS/MS analysis. (A) CID spectrum of a tryptic peptide generated by the mass spectrometer. The spectrum matches that predicted for the peptide VLDLEEVYFR, both from N to C terminus (b ions) and C to N terminus (y ions), and corresponds to a peptide predicted from the CHR1.148 (TbMP90) gene sequence. (B) Fourteen tryptic peptides (shaded region, peptide 7 and 8 from left overlap) were identified across the protein that covered 16.8% of the sequence.
Other editosome proteins that correspond to genes whose sequences are not yet in the *T. brucei* DNA sequence database may be identified from the LC-MS/MS data. Previous studies showed that the two RNA editing ligase proteins TbREL1 and TbREL2 are related to each other (Panigrahi et al. 2001a; Schnaufer et al. 2001). Four other proteins (TbMP81, TbMP61, TbMP42, and TbMP18) were also shown to have varying degree of sequence similarity and domain conservation among them (Panigrahi et al. 2001b). The editosome protein TbMP24 identified here is related to these four proteins, with the greatest sequence similarity to TbMP42 (29% sequence identity and 42% similarity over 109 amino acids). Similarly, pairwise amino acid sequence comparison and homology searches showed that another set of five proteins, TbMP90, TbMP67, TbMP61, TbMP46, and TbMP44, have some sequence similarity and domain conservation among them (Table 2; Fig. 3A). Of these proteins TbMP90, TbMP67, and TbMP61 are more closely related to each other, as are TbMP46 and TbMP44 (Table 2). Three-way sequence alignment showed the mid region is partially conserved among TbMP90, TbMP67 and TbMP61 (Fig. 3B). Similarly, the mid region of TbMP46 and TbMP44 is partially conserved (Fig. 3C).

Proteins TbMP100 and TbMP99 are similar to each other (overall 28.5/45.8% identity/similarity), and they have a highly conserved C-terminal one-third region (40/61% identity/similarity; Fig. 4). Another novel protein, TbMP57, was identified in editosomes prepared by all three different methods. This protein has sequence similarity to a 108-kD protein 3′ TUTase that was cloned from *T. brucei* (Aphasizhev et al. 2002).

The functions of some editosome proteins have been determined whereas the functions of others are suggested by the motifs that they contain (Table 3). TbREL1 and TbREL2 proteins have significant sequence homology to ligase domains and catalyze RNA ligation (McManus et al. 2001; Rusché et al. 2001; Schnaufer et al. 2001). The TbMP81, TbMP63, and TbMP42 proteins contain C2H2 zinc finger motifs (Panigrahi et al. 2001b), implying molecular interaction. The other two proteins in this related set, TbMP24 and TbMP18, contain RNA-binding domains. Thus, this group of five proteins is likely to be involved in protein–protein and/or protein–RNA interactions. Indeed, it has been shown that TbMP63 interacts with TbREL1 (Panigrahi et al. 2001b) and TbMP81 with TbREL2 in vitro (S.S. Palazzo, unpubl.).

The TbMP61 protein has a ribonuclease III motif (amino acids 187–309) located within the sequence that is most conserved with its related proteins TbMP90 and TbMP67. The catalytic center, key residues E and G that function in substrate binding and Mg enhancement of catalysis, respectively, in *Escherichia coli* (Mian 1997) and flanking sequences are conserved (Fig. 3B). Hence, TbMP61, TbMP90, and TbMP67 may function as nucleases. The TbMP44 protein also contains a ribonuclease III-like motif within this region whereas TbMP46 does not, but contains a potential Pumilio family RNA-binding domain motif (amino acids 263–300; Fig. 3C). TbMP44 has been shown to be essential

### TABLE 1. Proteins identified in editosomes isolated by MAb affinity, column chromatography, and TAP-tag procedures

<table>
<thead>
<tr>
<th>Protein</th>
<th>MAb affinity</th>
<th>Column</th>
<th>TAP-tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>TbMP100a</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>TbMP99a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TbMP90a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TbMP81</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TbMP67a</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>TbMP63</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TbMP61a</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>mHel61p</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbMP57</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbREL1</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbREL2</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbMP46a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TbMP44a</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbMP42</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbMP24a</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TbMP18</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

*Identified in this study

### TABLE 2. Pairwise sequence comparison between editosome proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>TbMP90</th>
<th>TbMP67</th>
<th>TbMP61</th>
<th>TbMP46</th>
<th>TbMP44</th>
</tr>
</thead>
<tbody>
<tr>
<td>TbMP90 817 aa</td>
<td>32/46 (214)</td>
<td>27/42 (332)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TbMP67 596 aa</td>
<td>17.2/28.6</td>
<td>36/52 (287)</td>
<td>22/46 (118)</td>
<td>21/44 (198)</td>
<td></td>
</tr>
<tr>
<td>TbMP61 538 aa</td>
<td>17.9/26.8</td>
<td>23/34</td>
<td>22/37 (275)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TbMP46 414 aa</td>
<td>11.3/17.8</td>
<td>13.5/23.3</td>
<td>17.7/28.6</td>
<td>28/45 (260)</td>
<td></td>
</tr>
<tr>
<td>TbMP44 382 aa</td>
<td>10.7/17.8</td>
<td>12.8/24.1</td>
<td>13.5/22.6</td>
<td>21.9/37.8</td>
<td></td>
</tr>
</tbody>
</table>

The global alignment covers the whole length of both sequences and the local alignment covers the region of greatest similarity. The numbers represent percent identity over percent similarity and the length of the region is indicated in parentheses. aa = amino acids.
for RNA editing and for the structural integrity of the editosome (B. Wang, in prep.).

The TbMP100 and TbMP99 proteins have endonuclease/exonuclease/phosphatase family motifs (pfam03372, probability 2 e-10, and 3 e-10, respectively) suggesting that they may be nucleases. This motif is located in the C-terminal one-third region that is highly conserved between them (Fig. 4). The TbMP57 protein has poly(A) polymerase and nucleotidyl transferase motifs and, as will be described in detail elsewhere, has TUTase activity (N.L. Ernst, unpubl.).

DISCUSSION

This study identified nine additional novel protein components of *T. brucei* editosomes. These proteins along with seven previously identified editosome proteins comprise much, but perhaps not all, of the mass of the stable catalytic core of the editosome. Five of the proteins are interrelated, and another protein is related to four previously identified editosome proteins, making another set of five related proteins. Another two of the newly identified proteins are also related to each other. In addition, another protein is related to a TUTase that is essential for editing. Most of the proteins have motifs suggesting roles in molecular interaction, especially with nucleic acids, and/or nuclease or TUTase activity.

Several criteria indicate that the 16 identified proteins are components of the editosome. The proteins were detected by LC-MS/MS analysis of editosomes isolated with high stringency (salt and nonionic detergent) during MA affinity isolations, indicating that most of these proteins are in the stable core of the editosome. Fifteen proteins were identified in the editosomes that were isolated by at least two independent purification methods, showing stable association.
TbMP100  KILTWVMPDFYSGKPTLQOMGIDCSPZHYPVIAKLIEBEDADVYNEBVFPVPKBFELXDALTPQRYPSOCZHMP4ASPOVYMLIELTRPLVQG...TbMP99  RVTWVQSHSRTPLGQDIDCTSTRYVLAQETYLEQGADTVYMEVPANCKYLSQPVKEKAMTCCBHSAILIPAKQGVOLML-RRSLCIVTS

NYLNPWANTHPVLSMPVULMVMHQT-VHIAANHAAPVTSRNENARTS-DTALREHTMKTGOG-...-DVTVMDFNDTPSNFIMPHEQYQVDCPVP...-

FIGURE 4. Sequence similarities [conserved (|), semiconserved (:), and partially conserved (.)] of the C termini of TbMP100 (amino acids 619–891) and TbMP99 (amino acids 618–906) that belong to the endonuclease/exonuclease/phosphatase family of proteins.

The newly identified nine proteins add to the candidate proteins that perform specific functions in editing, including endonuclease, exo-Uase, and TUTase activities. TbREL1 and TbREL2, which were previously identified (Panigrahi et al. 2001a), catalyze the RNA ligation that occurs during editing (Rusché et al. 2001; Schnaufer et al. 2001). TbMP57 identified here has TUTase activity and sequence similarity to TbREL1 (Panigrahi et al. 2001b). The existence of two sets of five related proteins and two sets of related pairs of proteins, including TbREL1 and TbREL2, also implies that those 14 proteins are editosome components.

The differential composition of the editosomes isolated by different methods may be due to protein dissociation during the chromatography and the effect of the tag on associations of some of the proteins with the complex. TbMP67 was only detected in MAb affinity-isolated editosomes, but its sequence similarity to other editosome proteins suggests that it is an editosome component. Studies in progress indicate interactions among many of these proteins (A. Schnaufer, N.L. Ernst, and S.S. Palazzo, unpubl.) much as TbMP63 has been shown to physically interact with TbREL1 (Panigrahi et al. 2001b). The existence of two sets of five related proteins and two sets of related pairs of proteins, including TbREL1 and TbREL2, also implies that those 14 proteins are editosome components.

The newly identified nine proteins add to the candidate proteins that perform specific functions in editing, including endonuclease, exo-Uase, and TUTase activities. TbREL1 and TbREL2, which were previously identified (Panigrahi et al. 2001a), catalyze the RNA ligation that occurs during editing (Rusché et al. 2001; Schnaufer et al. 2001). TbMP57 identified here has TUTase activity and sequence similarity to TbREL1 (Panigrahi et al. 2001b). The existence of two sets of five related proteins and two sets of related pairs of proteins, including TbREL1 and TbREL2, also implies that those 14 proteins are editosome components.

The newly identified nine proteins add to the candidate proteins that perform specific functions in editing, including endonuclease, exo-Uase, and TUTase activities. TbREL1 and TbREL2, which were previously identified (Panigrahi et al. 2001a), catalyze the RNA ligation that occurs during editing (Rusché et al. 2001; Schnaufer et al. 2001). TbMP57 identified here has TUTase activity and sequence similarity to TbREL1 (Panigrahi et al. 2001b). The existence of two sets of five related proteins and two sets of related pairs of proteins, including TbREL1 and TbREL2, also implies that those 14 proteins are editosome components.

TbMP99 is related and both have an endonuclease/exonuclease/phosphatase motif and hence are candidates for proteins that catalyze the endonuclease and/or exo-Uase activities associated with editing. Because another set of editosome proteins has probable endonuclease function as discussed below, one or both of these proteins may function as an exo-Uase during the editing cycle.

TbMP24 has a S1 domain suggesting that it may be an RNA-binding protein as appears to be the case for the other four members (TbMP81, TbMP63, TbMP42, and TbMP18) of that family (Panigrahi et al. 2001b). Of the five remaining proteins with varying degrees of sequence relatedness, TbMP90, TbMP67, TbMP61, TbMP46, and TbMP44, all except TbMP46 have a sequence related to a ribonuclease III motif. This suggests that they may have nuclease functions associated with editing. This family of proteins cleaves double-stranded RNAs (PROSITE documentation: PDOC00448; http://www.expasy.org). The first enzymatic step in the RNA editing cycle is cleavage of pre-mRNA in the double-stranded pre-mRNA-gRNA duplex; thus endonuclease function(s) would be consistent with the predicted motifs of these proteins. It is also possible that some of these proteins have retained noncatalytic functions such as RNA binding. The presence of six potential nuclease may suggest that nucleases may play roles other than the endonuclease and exoUase steps proposed for the editing mechanism (Kable et al. 1997) such as accommodation of substrate differences that are produced during the editing at multiple sites. In any event, TbMP44 is essential for editing and for editosome integrity (B. Wang, in prep.).

The editosomes that have been characterized here are the more stable complexes, which can catalyze a full round of editing in at least one site. The editosome preparation described by Rusché et al. (1997) contained seven major proteins, of which band II corresponds to TbMP81, III to TbMP63, IV to TbREL1, V to TbREL2, VI to TbMP42, VII to TbMP18 (Huang et al. 2002), and, based on apparent molecular size, I to TbMP44 or TbMP100/TbMP99. Although this preparation catalyzed insertion and deletion editing, some editosome proteins described in this article and elsewhere, including those that appear essential for editing, were not evident to the authors. These proteins include two TUTases (TbMP57 and 108-kD TUTase) and TbMP44, which is essential for editing. It seems likely that such editosome proteins were present but less abundant than the seven major proteins in this preparation, and thus
Some proteins with demonstrated or possible roles in editing were not found associated with the stable editosomes described here. These include the 108-kD TUTase (Aphasizhev et al. 2002), gBP21 (Koller et al. 1997; Allen et al. 1998) and the related protein gBP25 (Blom et al. 2001), RBP16 (Hayman and Read 1999), TbRGG1 (Vanhamme et al. 1998), and REAP1 (Madison-Antenucci et al. 1998). This may reflect a low affinity and/or transient association with the editosome or roles associated with processes other than the catalytic steps of editing. Editing probably entails a dynamic series of events that affect protein association with the editosome and perhaps editosome composition (see Stuart and Panigrahi 2002). Hence, some proteins may function in editing but may not be stable components of the editosome or perhaps not even associated with the editosome. Editosomes and editing activities sediment in glycerol gradients with peaks at \( \sim 20S \) and \( \sim 40S \) (Pollard et al. 1992; Piller et al. 1995; Corell et al. 1996), and the relationship between these complexes is unclear. One possibility is that they represent editosomes in various stages of editing and/or association with proteins with roles in editing. The editing endo- and exonucleases are not yet known although several candidates have been identified in this article. There may be editing endonucleases specific for insertion versus deletion substrates, similar to the TbREL1 and TbREL2 ligases that function in deletion versus insertion editing (Cruz-Reyes et al. 2002), as well as multiple nucleases to accommodate differences in substrate sequences, gRNA interactions, and positions in the editosome that are produced during the editing of multiple sites. The presence of pairs and sets of related proteins, some with clearly related activities, suggests that insertion and deletion editing are physically and functionally separate. In addition, some essential proteins for editing that are not part of the editosome may or may not directly interact with the catalytic complex. Because editing is regulated during the life cycle of \( T. brucei \) (Schnaufer et al. 2002), other complexities are bound to be uncovered. Hence, although the steps in editing and the proteins that catalyze these steps are becoming clarified, additional complexities await elucidation.

### MATERIALS AND METHODS

#### Protein and gene identification

Mitochondrial vesicles were isolated from \( T. brucei \) procyclic cells (strain IsTaR 1.7a) as described (Harris et al. 1990). The mitochondria were lysed with 1% Triton X-100 for 15 min at 4°C with bidirectional mixing, and clarified by centrifugation at 15,000 rpm for 30 min in a microcentrifuge. The soluble sample was fractionated on 10%–30% glycerol gradient, 500 µL fractions were collected from the top, and the peak fractions positive for editosome as determined by Western analysis were pooled (Panigrahi et al. 2001a, 2001b). Editosomes were then immunoprecipitated from

#### TABLE 3. Probable and demonstrated function(s) of \( T. brucei \) mitochondrial editosome proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>Probable motifs/domains/family</th>
<th>Function/probable function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TbMP81 (a)</td>
<td>C2H2 Zn finger (PS0157)</td>
<td>Macromolecule interaction</td>
<td>Panigrahi et al. 2001b; Drozdz et al. 2002</td>
</tr>
<tr>
<td>TbMP63 (a)</td>
<td>C2H2 Zn finger (PS0157)</td>
<td>Macromolecule interaction</td>
<td>Panigrahi et al. 2001b; Huang et al. 2002</td>
</tr>
<tr>
<td>TbMP42 (a)</td>
<td>C2H2 Zn finger (PS0157)</td>
<td>Macromolecule interaction</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP24 (a)</td>
<td>S1 domain (PS0126)</td>
<td>RNA interaction</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP18 (a)</td>
<td>Single-strand binding protein family (PF00436)</td>
<td>Nuclease</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP90 (b)</td>
<td>Ribonuclease III (PS0142)</td>
<td>Nuclease</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP67 (b)</td>
<td>Ribonuclease III (PS0142)</td>
<td>Nuclease</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP61 (b)</td>
<td>Ribonuclease III (PS0142)</td>
<td>Macromolecule interaction</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP44 (b)</td>
<td>Ribonuclease III (PS0142)</td>
<td>Structure/Nuclease</td>
<td>This study; B. Wang, in prep.</td>
</tr>
<tr>
<td>TbREL1(c)</td>
<td>Ligase (CDD-7281)</td>
<td>RNA ligase</td>
<td>Panigrahi et al. 2001a; Schnaufer et al. 2001</td>
</tr>
<tr>
<td>TbREL2(c)</td>
<td>Ligase (CDD-7281)</td>
<td>RNA ligase</td>
<td>Panigrahi et al. 2001a; Schnaufer et al. 2001</td>
</tr>
<tr>
<td>TbMP100 (d)</td>
<td>Endonuclease/exonuclease/phosphatase (PF03372)</td>
<td>Nuclease</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP99 (d)</td>
<td>Endonuclease/exonuclease/phosphatase (PF03372)</td>
<td>Nuclease</td>
<td>This study</td>
</tr>
<tr>
<td>TbMP57</td>
<td>PAP_core, PAP_associated, and Nucleotidyltransferase domains (PS0154, PS0155, PF01909)</td>
<td>Terminal uridylyl transferase</td>
<td>This study; N. Ernst unpubl.</td>
</tr>
<tr>
<td>mHel61p</td>
<td>DEAD box (PF00270), Helicase_C (PF00271)</td>
<td>Helicase</td>
<td>Missel et al. 1997</td>
</tr>
</tbody>
</table>

1Proteins that show sequence similarity between them have been grouped as a, b, c, and d.
2PROSITE (PS), Pfam (PF), and Conserved Domain Database (CDD) accession numbers for the motifs/domains/family are given in parentheses.
3The proteins with macromolecule interaction function may be involved in RNA–protein and/or protein–protein interactions.
Cloning and sequencing of the genes

The TbMP100 ORF was amplified from T. brucei genomic DNA (strain 427) by PCR with primers 4430 (ATAAAGCTTATGGCCAT TGGCTCAGTCAT) and 4331 (AATCTCGACTACCGTTAA CTTCAATGAAA; restriction sites are italicized), and cloned into pGEM-T Easy vector (Promega) and sequenced. Similarly, the TbMP99 ORF was amplified by PCR with primers 3275 (ATAAGCTTATGGCCAT TGGCTCAGTCAT) and 4331 (AATCTCGACTACCGTTAA CTTCAATGAAA; restriction sites are italicized), and cloned into pGEM-T Easy vector (Promega) and sequenced. The TbMP96 ORF was amplified by primers 2425 (GGGAATTCAAGCTTATGGCCAT TGGCTCAGTCAT) and 4331 (AATCTCGACTACCGTTAA CTTCAATGAAA; restriction sites are italicized), and cloned into pGEM-T Easy vector (Promega) and sequenced. The TbMP90 ORF was amplified by primers 4245 (GGAATTCCTAGCTATGGCCAT TGGCTCAGTCAT) and 4246 (TAAACGCTTACCGGTACCGAAATGCGCAGTCG) and 4247 (ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG

TAP-tag purification of the editosome

To create a vector for the inducible expression of C-terminally TAP-tagged proteins in T. brucei, the tag was amplified from plasmid pBS1539 and inserted into plasmid pLew79, generating pLew79TAP. The TbREL1 coding sequence was released from pLew79-TbREL1 (Schnauer et al. 2001) and inserted into pLew79TAP, yielding pLew79-TbREL1TAP. T. brucei cell line 29.13, coexpressing the TET repressor and T7 RNA polymerase was transfected with NotI-linearized pLew79-TbREL1TAP plasmid DNA. Phleomycin-resistant clones were selected and checked for tightly tetracycline-regulated expression. Expression of TbREL1TAP in the recombinant cell lines was induced for 48 h with tetracycline (10 ng/mL). TAP-tagged editing complexes were purified as described (Rigaut et al. 1999) from 2 L of cells harvested at a density of ~20 × 10^6 cells/mL.

In vitro assays

Deletional editing was assayed in vitro using 3′ labeled A6-U5 premRNA substrate with gA6[14]Δ16G gRNA as described (Seiwert et al. 1996). The edited products were detected by polyacrylamide-urea gel electrophoresis and phosphorimaging. TbREL1 and TbREL2 were detected by auto-adenylation in the presence of [α-32P]-ATP as described (Sabatini and Hajduk 1995). The proteins were resolved on 10% SDS-PAGE gels and the radiolabeled proteins were detected by PhosphorImager.

Western blot analysis

The isolated editosome fractions were digested with SDS-PAGE loading buffer and resolved on 10% SDS-PAGE gels. The proteins were transferred onto PVDF membranes, and reacted with MAbs specific for TbMP81, TbMP63, TbREL1, and TbMP42 (Panigrahi et al. 2001a, 2001b). The filter was developed with ECL kit (Amersham) as per the manufacturer’s instructions.

Sequence analysis

The predicted protein sequences were compared with the NCBI nonredundant protein database and EBI database using the BLAST algorithm. The presence of known motifs and domains in the predicted proteins were searched for in the PROSITE, InterPro, BLOCKS, and CDD databases. Homology between the proteins was determined by pairwise sequence comparison using EMBOS and multiple sequences were aligned using the ClustalW algorithm (www.ebi.ac.uk).

Nucleotide sequence accession numbers

The nucleotide sequences have been submitted to GenBank with accession numbers AY228165–AY228173.

ACKNOWLEDGMENTS

We thank S.S. Palazzo for unpublished data and other members of the Stuart laboratory for helpful suggestions. We also thank B. Séraphin for pBS1539 plasmid, E. Wirtz and G. Cross for pLew79 plasmid and T. brucei cell line 29.13, and Najib M. El-Sayed for genomic DNA clones. Sequence data were obtained from The Institute for Genomic Research (http://www.tigr.org) and The Institute for Genomic Research (http://www.tigr.org).
REFERENCES


