Association of two novel proteins, TbMP52 and TbMP48, with the Trypanosoma brucei RNA editing complex

Citation for published version:

Digital Object Identifier (DOI):

Link:
Link to publication record in Edinburgh Research Explorer

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

Published In:
Molecular and Cellular Biology

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.
Association of Two Novel Proteins, TbMP52 and TbMP48, with the *Trypanosoma brucei* RNA Editing Complex

**ASWINI K. PANIGRAHI, 1,2 STEVEN P. GYGI, 3 NANCY L. ERNST, 1,2 ROBERT P. IGO, JR., 1,2 SETAREH S. PALAZZO, 1,2 ACHIM SCHNAUFER, 1,2 DAVID S. WESTON, 1,2 NICOLE CARMEAN, 1 REZA SALAVATI, 1,2 RUEDI AEBERSOLD, 3 AND KENNETH D. STUART, 1,2**

Seattle Biomedical Research Institute, Seattle, Washington 98109, 1 and Departments of Pathobiology 2 and Molecular Biotechnology, 3 University of Washington, Seattle, Washington 98195

Received 3 August 2000/Returned for modification 29 September 2000/Accepted 19 October 2000

RNA editing in kinetoplastid mitochondria inserts and deletes uridylates at multiple sites in pre-mRNAs as directed by guide RNAs. This occurs by a series of steps that are catalyzed by endoribonuclease, 3′-terminal uridylyl transferase, 3′-exoribonuclease, and RNA ligase activities. A multiprotein complex that contains these activities and catalyzes deletion editing in vitro was enriched from *Trypanosoma brucei* mitochondria by sequential ion-exchange and gel filtration chromatography, followed by glycerol gradient sedimentation. The complex size is approximately 1,600 kDa, and the purified fraction contains 20 major polypeptides. A monoclonal antibody that was generated against the enriched complex reacts with an ~49-kDa protein and specifically immunoprecipitates in vitro deletion RNA editing activity. The protein recognized by the antibody was identified by mass spectrometry, and the corresponding gene, designated *TbMP52*, was cloned. Recombinant TbMP52 reacts with the monoclonal antibody. Another novel protein, TbMP48, which is similar to TbMP52, and its gene were also identified in the enriched complex. These results suggest that TbMP52 and TbMP48 are components of the RNA editing complex.

Several mitochondrial RNAs are posttranscriptionally edited in kinetoplastid protozoa by the insertion and deletion of uridylates (U’s) at multiple sites, to produce mature mRNAs. RNA editing creates initiation and termination codons and the likely functional open reading frames (ORFs). Indeed, translation of edited RNA has recently been directly demonstrated (11). The RNA editing appears to regulate mitochondrial respiration in different life cycle stages of *Trypanosoma brucei*. The insertion and deletion of U’s is directed by small RNAs that are called guide RNAs (gRNAs). The editing occurs by a series of enzymatic steps. These steps include gRNA-directed cleavage of the pre-mRNA by endoribonuclease, U addition or removal at the 3′ end of the 5′ cleavage product by 3′-terminal uridylyl transferase (TUTase) or 3′-exoribonuclease, respectively, and ligation of 5′ and 3′ cleavage products by RNA ligase (reviewed in references 6, 13, and 28).

RNA editing occurs in association with a ribonucleoprotein complex which sediments at 20S in glycerol gradients (4, 22). Fractionation and hence partial purification of the complex by glycerol gradient and liquid chromatographic techniques have been reported (4, 18, 22, 24). For the most part, these preparations were insufficient to identify specific proteins that are part of the editing complex. However, Rusché et al. (24) suggested that a complex of eight proteins could catalyze editing. They concluded that three of these proteins were adenyllylatable and suggested that they represented the editing RNA ligase, although the role of these proteins has not yet been demonstrated. Indeed, little progress has been made on the definitive identification of proteins that are components of the editing complex. Three *T. brucei* mitochondrial proteins, gBP21 (15), DEAD box protein mHEL61p (19), and REAP1 (18), were identified as candidate components of the editing complex. In addition, two *T. brucei* mitochondrial poly(U) binding proteins, TBRGG1 (30) and RBP16 (10), were identified and suggested to have a role in RNA editing. Knockout of both gBP21 alleles (i.e., null mutations) had no effect on RNA editing in bloodstream-form *T. brucei* in vivo, indicating that gBP21 is not essential for editing (16). However, knockout of both mHEL61 alleles resulted in slow-growing insect procyclic forms. These cells are capable of in vitro editing but have a >70% reduction in edited mRNAs in vivo, which is restored upon reexpression of mHEL61p (19). These data suggest that mHEL61p may be a component of the editing complex, although not an essential one. Similar assays of the other candidate editing complex proteins have not yet been published.

The difficulty in identifying the protein components of the RNA editing complex reflects the apparent low cellular abundance of the complex, the low sensitivity of the in vitro editing assays, and the uncertainty that assays of endonuclease, exonuclease, TUTase, and RNA ligase are specific for activities associated with the intact complex. These factors, in addition to contamination from protein adsorption during fractionation, made protein identification by conventional microsequencing difficult. However, mass spectrometric analysis has been useful for identifying proteins that are present in small amounts and in mixtures of proteins (17). It was successfully used to identify components of multiprotein complexes, such as the U1 snRNP from the yeast *Saccharomyces cerevisiae* (21). Indeed, in organisms where the complete genome sequence is available, mass spectrometry can be used to identify the gene for virtually any protein that can be visualized by conventional staining methods. Few genomic sequence data were available for *T. brucei* until recently, as sequence data from the genome...
sequencing projects have been accumulating rapidly in the databases.

In this study, we report the biochemical fractionation of the RNA editing complex from \textit{T. brucei} mitochondria. The fractionation was monitored using the in vitro deletion editing assay in an attempt to purify the complex that is capable of all steps of editing. The editing complex was isolated by sequential ion-exchange and gel filtration chromatography followed by sedimentation on a glycerol gradient. Two novel related proteins in the most purified fraction and their genes were identified using capillary liquid chromatography-tandem mass spectrometry (LC-MS/MS) and by comparison to the \textit{T. brucei} genome sequence database. They were designated TbMP52 and TbMP48, based on the predicted mass of the preprocessed protein. One monoclonal antibody (MAb) from a panel that was generated against the isolated complex was specific for TbMP52 in Western analyses of native and recombinant protein. This MAB also immunoprecipitated the in vitro deletion editing activity. These data strongly suggest that TbMP52 and TbMP48 are components of the editing complex.

MATERIALS AND METHODS

Cell growth and isolation and fractionation of mitochondrial proteins. \textit{T. brucei} procyclic forms (strain IStaR 1.7a) were grown to log phase in vitro as described previously (29). The mitochondrial vesicles were isolated (9) and stored at 

-70°C. Mitochondria from 5.1 \times 10^{11} cells were lysed in 55 ml of buffer SP-A (10 mM Tris [pH 7.0], 10 mM MgCl2, 50 mM KCl, 1 mM dithiothreitol [DTT]) containing the protease inhibitors leupeptin (10 μg/mL), pepstatin (5 μg/mL), and Pefabloc (1 μM). The lysis was carried out using 0.5% Triton X-100 for 15 min at 4°C with bidirecional mixing. The lysate was centrifuged at 17,500 \times g for 30 min at 4°C. The cleared lysate was filtered through 0.2-μm-pore-size membranes and loaded onto a 10-ml SP Sepharose HR column (Pharmacia) at a 1-ml/min flow rate. All chromatographic steps were carried out using an automated fast protein liquid chromatography system (LKB-Pharmacia) at 4°C. The unbound proteins were washed away with 5 column volumes of buffer SP-A. The bound proteins were eluted with 80 ml of a linear salt gradient of 50 to 330 mM KCl, followed by an 80-ml linear gradient to 1 M KCl at a buffer SP-A. The bound proteins were eluted with 80 ml of a linear salt gradient (Sequentially fractionated by SP Sepharose, Q Sepharose, and Superose 6 columns) using the enzyme-linked immunosorbent assay (ELISA). All ELISA-positive samples were screened by Western blot analysis using SP and Q Sepharose fractionated编辑 complexes as the antigens. The contents of 15 ELISA- and Western blot-positive hybridomas were subcloned by dilution. The supernatants were screened for MAB production by both ELISA and Western analyses.

Western blot analysis. RNA editing complex that was partially purified by SP and Q Sepharose chromatography as described above was used as the antigen for Western analysis. One microgram of protein sample was separated by sodium dodecyl sulfate–10% polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto a nitrocellulose filter. The filter was blocked overnight with 5% nonfat milk powder in PBST (10 mM phosphate buffer [pH 7.2], 150 mM NaCl, 0.05% Tween 20) at 4°C. It was washed three times with PBST and incubated with tissue culture supernatant diluted 1:100 in PBST. The incubation was carried out at room temperature for 2 h with gentle shaking. The filter was washed three times with PBST and incubated with 1 h with horse-radish peroxidase-conjugated anti-mouse immunoglobulin G (IgG) (Bio-Rad) at a 1:2,000 dilution in PBST. The filter was washed four times with PBST and developed with an ECL kit (Amersham) according to the manufacturer’s instructions. To assay the distribution of TbMP52 in the glycerol gradient, 5 μl of sample from each fraction was separated on an SDS-PAGE gel and examined by Western analysis.

Immunoprecipitation of editing complex. Immunomagnetic beads (Dynabeads M-450; Dynal) coated with goat anti-mouse IgG were coupled to MABs from tissue culture supernatants. A total of 4 \times 10^{10} beads were incubated with 1 ml of tissue culture supernatant and 1% bovine serum albumin at 4°C with bidirectional mixing for 1 h. The beads were washed three times with immunoprecipitation buffer (IP buffer; 10 mM Tris [pH 7.2], 10 mM MgCl2, 200 mM KCl, 0.1% Triton X-100). The mitochondria were lysed in IP buffer with 0.5% Triton X-100, cleared by centrifugation, and fractionated on a 10- to 30%-glycerol gradient as described above. The antibody-bound beads were incubated with the mitochondrial 20S fraction (50 μg of proteins) in 1× IP buffer and 1% bovine serum albumin for 1 h at 4°C with bidirectional mixing. The beads were washed four times (each wash of 5 min duration with bidirectional mixing) with IP buffer and twice with buffer containing 25 mM HEPES (pH 7.9), 10 mM MgCl2, 50 mM KCl, 0.5% Triton X-100, and Q Sepharose chromatography as described above was used as the antigen for Western analysis. One microgram of protein sample was separated by sodium dodecyl sulfate–10% polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto a nitrocellulose filter. The filter was blocked overnight with 5% nonfat milk powder in PBST (10 mM phosphate buffer [pH 7.2], 150 mM NaCl, 0.05% Tween 20) at 4°C. It was washed three times with PBST and incubated with tissue culture supernatant diluted 1:100 in PBST. The incubation was carried out at room temperature for 2 h with gentle shaking. The filter was washed three times with PBST and incubated with 1 h with horse-radish peroxidase-conjugated anti-mouse immunoglobulin G (IgG) (Bio-Rad) at a 1:2,000 dilution in PBST. The filter was washed four times with PBST and developed with an ECL kit (Amersham) according to the manufacturer’s instructions. To assay the distribution of TbMP52 in the glycerol gradient, 5 μl of sample from each fraction was separated on an SDS-PAGE gel and examined by Western analysis.

Immunoprecipitation of editing complex. Immunomagnetic beads (Dynabeads M-450; Dynal) coated with goat anti-mouse IgG were coupled to MABs from tissue culture supernatants. A total of 4 \times 10^{10} beads were incubated with 1 ml of tissue culture supernatant and 1% bovine serum albumin at 4°C with bidirectional mixing for 1 h. The beads were washed three times with immunoprecipitation buffer (IP buffer; 10 mM Tris [pH 7.2], 10 mM MgCl2, 200 mM KCl, 0.1% Triton X-100). The mitochondria were lysed in IP buffer with 0.5% Triton X-100, cleared by centrifugation, and fractionated on a 10- to 30%-glycerol gradient as described above. The antibody-bound beads were incubated with the mitochondrial 20S fraction (50 μg of proteins) in 1× IP buffer and 1% bovine serum albumin for 1 h at 4°C with bidirectional mixing. The beads were washed four times (each wash of 5 min duration with bidirectional mixing) with IP buffer and twice with buffer containing 25 mM HEPES (pH 7.9), 10 mM MgCl2, 50 mM KCl, 0.5% Triton X-100, and Q Sepharose chromatography followed by ion-exchange and gel filtration chromatography followed by sedimentation on a glycerol gradient. Two novel related proteins in the most purified fraction and their genes were identified using capillary liquid chromatography-tandem mass spectrometry (LC-MS/MS) and by comparison to the \textit{T. brucei} genome sequence database. They were designated TbMP52 and TbMP48, based on the predicted mass of the preprocessed protein. One monoclonal antibody (MAB) from a panel that was generated against the isolated complex was specific for TbMP52 in Western analyses of native and recombinant protein. This MAB also immunoprecipitated the in vitro deletion editing activity. These data strongly suggest that TbMP52 and TbMP48 are components of the editing complex.
Cloning and expression of TbMP52 and TbMP48. The entire ORF representing TbMP52 was amplified from genomic DNA with primers Tb10-2261P (ACT GCA GAT GCA ACT CCA AAG G) and Tb10-3690M (A GA ATT CGC AGT AT CAT TCG CC) (the restriction sites are italicized). The amplified DNA was cloned into the pGEM-T easy vector (Promega), and the sequence was confirmed. The insert was released with PstI and EcoRI enzymatic digestion and cloned into the pRSET C vector (Invitrogen) cut with the same enzymes. Similarly, the TbMP48 ORF was PCR amplified with the primer pairs Tb07-1423P (TGG ATC CGT AAG ATG TTG CGT C) and Tb07-2679M (AGA ATT CGC AGT AT CAT TCG CC) (the restriction sites are italicized). The amplified DNA was cloned into the pGEM-T easy vector (Promega), and the sequence was confirmed. The insert was released with PstI and EcoRI enzymatic digestion and cloned into the pRSET C vector (Invitrogen) cut with the same enzymes. Similar procedures were used for cloning the TbMP48 ORF. The plasmids were transformed into BL21 DE3-LysS cells, and recombinant proteins were expressed with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) induction. The Escherichia coli cells expressing recombinant proteins were separated by SDS-PAGE and either stained with Coomassie blue or transferred onto nitrocellulose paper. These were reacted with individual MAb supernatants (1:100 dilution) as described above. E. coli cells transformed with vector only were used as negative controls.

Subcellular localization of TbMP52 by immunofluorescence. Procytoc-form trypanosomes were fixed onto microscope slides, and immunofluorescence reactions were carried out as previously described (1). Briefly, the cells were incubated with MAB P3C1-G2 (tissue culture supernatant diluted 1:10 in 1X phosphate-buffered saline [PBS]) for 1 h, washed, and incubated with fluorescein isothiocyanate-conjugated goat anti-mouse IgG (diluted 1:200 in PBS) for 1 h. The washed cells were treated with 4,6-diamidino-2-phenylindole (DAPI; 0.5 μg/ml in PBS) to stain DNA. Fluorescence was observed with a Nikon fluorescence microscope equipped with the appropriate filters.

Nucleotide sequence accession numbers. The nucleotide and protein sequences have been submitted to GenBank and SWISS-PROT with accession numbers AY009110, AY009111, P82863, and P82864.

RESULTS

Enrichment of RNA editing complex. The RNA editing complex was enriched from T. brucei mitochondrial lysate by sequential fractionation by two ion-exchange columns, a gel filtration column, and glycerol gradient sedimentation (Fig. 1). The fractionation of the functional complex was monitored using the in vitro assay for deletion editing (26). The peak of in vitro editing activity from cleared mitochondrial lysate, which was prepared as described in Materials and Methods, eluted from the SP Sepharose (cation-exchange) column at about 200 mM KCl (Fig. 1A). Similarly, the peak of the editing activity from the pooled SP Sepharose fractions containing the editing activity eluted from the Q Sepharose (anion-exchange) column at about 225 mM KCl (Fig. 1B). There was no detectable editing activity in the fractions that did not bind to these columns. The editing activity from the pooled active fractions from the Q Sepharose column eluted from the Superose 6 (gel filtration) column with a peak which centered at an elution

Cloning and expression of TbMP52 and TbMP48. The entire ORF representing TbMP52 was amplified from genomic DNA with primers Tb10-2261P (ACT GCA GAT GCA ACT CCA AAG G) and Tb10-3690M (A GA ATT CGC AGT AT CAT TCG CC) (the restriction sites are italicized). The amplified DNA was cloned into the pGEM-T easy vector (Promega), and the sequence was confirmed. The insert was released with PstI and EcoRI enzymatic digestion and cloned into the pRSET C vector (Invitrogen) cut with the same enzymes. Similarly, the TbMP48 ORF was PCR amplified with the primer pairs Tb07-1423P (TGG ATC CGT AAG ATG TTG CGT C) and Tb07-2679M (AGA ATT CGC AGT AT CAT TCG CC) (the restriction sites are italicized). The amplified DNA was cloned into the pGEM-T easy vector (Promega), and the sequence was confirmed. The insert was released with PstI and EcoRI enzymatic digestion and cloned into the pRSET C vector (Invitrogen) cut with the same enzymes. Similar procedures were used for cloning the TbMP48 ORF. The plasmids were transformed into BL21 DE3-LysS cells, and recombinant proteins were expressed with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) induction. The Escherichia coli cells expressing recombinant proteins were separated by SDS-PAGE and either stained with Coomassie blue or transferred onto nitrocellulose paper. These were reacted with individual MAB supernatants (1:100 dilution) as described above. E. coli cells transformed with vector only were used as negative controls.

Subcellular localization of TbMP52 by immunofluorescence. Procytoc-form trypanosomes were fixed onto microscope slides, and immunofluorescence reactions were carried out as previously described (1). Briefly, the cells were incubated with MAB P3C1-G2 (tissue culture supernatant diluted 1:10 in 1X phosphate-buffered saline [PBS]) for 1 h, washed, and incubated with fluorescein isothiocyanate-conjugated goat anti-mouse IgG (diluted 1:200 in PBS) for 1 h. The washed cells were treated with 4,6-diamidino-2-phenylindole (DAPI; 0.5 μg/ml in PBS) to stain DNA. Fluorescence was observed with a Nikon fluorescence microscope equipped with the appropriate filters.

Nucleotide sequence accession numbers. The nucleotide and protein sequences have been submitted to GenBank and SWISS-PROT with accession numbers AY009110, AY009111, P82863, and P82864.

RESULTS

Enrichment of RNA editing complex. The RNA editing complex was enriched from T. brucei mitochondrial lysate by sequential fractionation by two ion-exchange columns, a gel filtration column, and glycerol gradient sedimentation (Fig. 1). The fractionation of the functional complex was monitored using the in vitro assay for deletion editing (26). The peak of in vitro editing activity from cleared mitochondrial lysate, which was prepared as described in Materials and Methods, eluted from the SP Sepharose (cation-exchange) column at about 200 mM KCl (Fig. 1A). Similarly, the peak of the editing activity from the pooled SP Sepharose fractions containing the editing activity eluted from the Q Sepharose (anion-exchange) column at about 225 mM KCl (Fig. 1B). There was no detectable editing activity in the fractions that did not bind to these columns. The editing activity from the pooled active fractions from the Q Sepharose column eluted from the Superose 6 (gel filtration) column with a peak which centered at an elution
position corresponding to about 1,600 kDa compared to globular protein size standards (Fig. 1C and data not shown). The buffer used for size fractionation on the Superose 6 column contained 200 mM KCl in order to avoid nonspecific association of other proteins with the editing complex. We observed that the complex eluted from Superose 6 with the same apparent mass when a buffer containing up to 300 mM KCl was used. However, it eluted with a greater apparent mass (in void volume) with a buffer containing 50 mM KCl (data not shown), perhaps indicating its association with other proteins or complexes or self aggregation. The pooled peak of deletion editing activity from the Superose 6 column was very similar to that of the glycerol gradient fractions that showed the greatest editing activity. Examination of glycerol gradient fraction 7, which contains the greatest editing activity, reveals 20 major polypeptides with apparent masses of 20, 22, 29, 34, 42, 43, 44, 45, 47, 49, 50, 53, 55, 57, 69, 72, 90, 99, 106, and 114 kDa. While the 22-, 44-, and 90-kDa protein bands stained more intensely than the other proteins, this may or may not reflect the ratio of the proteins in the complex, since different proteins stain at different intensities with silver nitrate. In addition, editing was not detected in fractions 2 to 4, although many of these 20 proteins were present, and there was more editing in fraction 9 than in fraction 5, although it had less protein (Fig. 1D and 2).

Perhaps there is a smaller proportion of complete (i.e., fully functional) complexes in the upper (lower sedimentation value) fractions than in fraction 9.

**Immunoprecipitation of editing complex using MAbs.** A panel of MAb was generated using the enriched complex as the immunogen, to generate reagents that are specific to components of the editing complex and that can be used to identify and characterize these components. A total of 19 independent MAb were isolated, and they were directed against seven different proteins, based on Western analyses. One MAb, P3C1-G2, reacted with an ~49-kDa protein in Western analyses and immunoprecipitated in vitro deletion editing activity (Fig. 3A and B). The immunoprecipitations and washes were performed with the 20S fraction of glycerol gradient-fractionated mitochondrial lysate using 0.1% nonionic detergent and either 200, 300, or 400 mM KCl (to reduce nonspecific associations) as described in Materials and Methods. The editing activity did not immunoprecipitate with MAb 58 (Fig. 3B, lane 2), which is directed against the *T. brucei* mitochondrial pyruvate dehydrogenase E2 subunit (reference 1 and unpublished results) or with MAb P7D9-A12, which reacts with a protein in an unidentified 40S complex. Even material that was immunoprecipitated with 400 mM KCl was able to catalyze in vitro editing, although the activity was reduced by 30% compared to the activity of the material that was immunoprecipitated with 200 mM KCl (Fig. 3B, lane 4). The immunoprecipitated material also contained component editing activities (precleaved insertion editing, endonuclease, TUTase, 3'-exouridylylase, and RNA ligase activities [results not shown]). This demonstrates that MAb P3C1-G2 can immunoprecipitate the active editing complex.

We also tested the ability of MAb P3C1-G2 to immunoprecipitate the 50- and 57-kDa adenylylatable proteins, which were reported to be RNA ligases with a possible role in RNA editing (25). These proteins were also reported to cofractionate with editing activities (18, 24). The material that was immunoprecipitated by MAb P3C1-G2 contained two adenylylatable proteins (Fig. 3C). In addition, MAb P3C1-G2 was able to immunoprecipitate two proteins that were adenylylated prior to the immunoprecipitation, even in buffer containing 400 mM KCl and 0.1% Triton X-100 (Fig. 3D). Essentially no adenylylatable proteins or preadenylated proteins immunoprecipitated with MAb 58 (1), which was used as the negative control. The protein from each step of purification was separated by SDS-PAGE and stained with silver nitrate. Results for protein size standards (M), cleared mitochondrial lysate (Cr), and pooled editing activity-positive fractions from SP Sepharose (SP), Q Sepharose (Q), and Superose 6 (S6) columns are presented. Glycerol gradient fractions 1 to 10 (fraction 1 is at the top) are shown (fractions 11 to 23 are not shown since essentially no protein was detected in these fractions). The numbers on left indicate the sizes of molecular mass markers, in kilodaltons. The most purified editing activity-positive fraction from the glycerol gradient (7) shows 20 major polypeptide bands.
control. Unlike the data in earlier reports, the major adenylylatable and preadenylylated proteins had apparent sizes of 44 and 50 kDa. Western analysis with MAb P3C1-G2 of fractions from a 10-to-30% glycerol gradient of mitochondria that were lysed in buffer with 200 mM KCl showed the bulk of the 49-kDa protein sediments at 20S, similar to the deletion editing activity (Fig. 1D and 3E).

Identification of TbMP52 and TbMP48 proteins. The 49-kDa protein that reacts with MAb P3C1-G2 and its corresponding gene were identified by a combination of high-performance LC-MS/MS and analyses of \( T. brucei \) genome sequence data. Mass spectrometric analyses were performed on both the fraction containing the editing complex and on individual bands that were cut from SDS-PAGE gels (e.g., Fig. 2, lanes 5 and 6). LC-MS/MS is highly sensitive and can allow identification of proteins in mixtures (17). Protein was digested in gel with trypsin, and the resulting peptides were fractionated by on-line capillary liquid chromatography and eluted directly into the mass spectrometer, where they were analyzed. Peptide masses and amino acid sequences were determined by automated selection and fragmentation of specific ionized peptides. This entailed switching between MS and MS/MS modes and resulted in the CID MS/MS spectrum of each peptide. Individual CID spectra were compared to predicted spectra from sequence databases by using computer algorithms to identify the gene and hence the protein. This analysis identified multiple peptides that corresponded to multiple CID MS/MS spectra from the same protein.

An example of a CID spectrum of a peptide that corresponds to a tryptic peptide with the amino acid sequence VH GTNFGIYLLNQGDHEVVR from the 49-kDa protein is shown in Fig. 4A. The results of a database search that compared this CID spectrum with CID spectra predicted from \( \text{Trypanosoma} \) sequences are shown in Table 1. This search identified a candidate matching peptide sequence in a 1,410-nucleotide ORF in The Institute for Genomic Research \( \text{T. brucei} \) genome sequencing project database. The confidence level for this match (at positions 50,980 to 52,389; accession no.}

FIG. 3. Immunoanalysis of the editing complex using a MAb. (A) MAb P3C1-G2 raised against the purified complex reacts with an ~49-kDa protein in Western analysis of a partially purified complex. (B) MAb P3C1-G2 specifically immunoprecipitates editing activity from the 20S mitochondrial fraction (see Materials and Methods for details) (lanes 3 and 4). Input RNA and edited RNA from which they are derived are indicated. MAb 58 (1), the negative control, did not immunoprecipitate these activities (lane 2). Editing activity immunoprecipitated with 200 and 400 mM KCl (lanes 3 and 4, respectively). The positive deletion editing control using the mitochondrial 20S fraction is also shown (lane 1). (C) MAb P3C1-G2 immunoprecipitates both 50- and 44-kDa adenylylatable proteins. MAb 58, which was used as a negative control, essentially did not immunoprecipitate the adenylylation activity. The control (+ve) for these proteins using the mitochondrial 20S fraction is shown. (D) MAb P3C1-G2 also immunoprecipitated both the 50- and 44-kDa proteins from the mitochondrial 20S fraction in a buffer containing 400 mM KCl following their adenylylation. (E) Western analysis of glycerol gradient-fractionated cleared mitochondrial lysate using MAb P3C1-G2.

FIG. 4. Identification of TbMP52. (A) Sample tandem mass spectrum derived by CID of a peptide precursor ion, \( m/z \) 1135.0 (bottom), and the peptide sequence predicted by SEQUEST (top) showing b- and y-type ions (above and below the sequence, respectively). (B) Amino acid sequence of the complete ORF identified by 11 tryptic peptide matches with CID spectra. The identified peptides are underlined (at two different positions, two and three peptides were contiguous). The dashed underline indicates the probable N-terminal peptide in the mature protein that is nontryptic, and the double underline indicates the peptide with the highest correlation score (Table 1). The predicted mitochondrial targeting signal is italicized.
TABLE 1. SEQUEST output file of the TbMP52 CID spectrum showing the correlation between peptides from the trypanosome database and the observed peptide CID spectrum.

<table>
<thead>
<tr>
<th>(M+H)+</th>
<th>Score</th>
<th>Database reference</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2269.5</td>
<td>4.9026</td>
<td>7330318</td>
<td>(K)VLGTPNGTNIQNGDHEVVR</td>
</tr>
<tr>
<td>2268.7</td>
<td>2.7907</td>
<td>2K0.TV</td>
<td>(S)IRTTRLW KD VDVGCR PVC</td>
</tr>
<tr>
<td>2267.7</td>
<td>2.6511</td>
<td>2312.TF</td>
<td>(Q)SSSYTGETINL VIVKR VV K</td>
</tr>
<tr>
<td>2268.8</td>
<td>2.5782</td>
<td>4827345</td>
<td>(T)KKLRQJNF KGS GLPV V L</td>
</tr>
<tr>
<td>2269.8</td>
<td>2.5587</td>
<td>5198286</td>
<td>(K)SKHM AHR IV LGV IN TPL</td>
</tr>
</tbody>
</table>

* Cross-correlation score.

b GenBank sequence identification number.

c, modified methionine.

7330318 and AC013484.9 [GenBank]; chromosome IX clone RPC193-1L12) was quite high. Similar analyses identified 10 additional tryptic peptides with matches to this ORF (Fig. 4B). Thus, the probability that this gene corresponds to the 49-kDa protein is very high. The protein predicted from the ORF has a mass of 52 kDa, but its N-terminal sequence has an amphiphilic helix that is predicted by Gene Runner (Hastings Software, Inc.) and a mitochondrial targeting signal that is predicted by the PSORT II algorithm (http://psort.ims.u-tokyo.ac.jp/form2.html). In addition, the CID spectrum of a peptide from the 49-kDa protein identified the peptide (T)YMPLPDQSDSPFYIE DLPSESR from this ORF (Fig. 4B).

The N-terminal amino acid Y in the peptide could be generated if in vivo processing removed the signal peptide, and the amino acid T, being nontryptic, may be part of the signal peptide. These data are consistent with localization of the 49-kDa protein within the mitochondrion after cleavage of the mitochondrial signal peptide. The protein is designated TbMP52, for “T. brucei mitochondrial protein,” and has a preprocessed molecular mass of 52 kDa; the gene is designated TbMP52. The first 44 amino acids in the N-terminal region may form a signal peptide, and this predicts a 48-kDa mature protein, which is consistent with its apparent migration size of 49 kDa. Immunofluorescence analysis with MAb P3C1-G2 showed staining of the mitochondria, indicating that TbMP52 is localized in this organelle (Fig. 5).

Another gene that is related to TbMP52 and its corresponding protein, which copurifies with the editing activity, were identified by a combination of mass spectrometry and database analyses (Fig. 6). A search of the Sanger Center T. brucei genome project database identified a 1,251-nucleotide ORF (positions 98742 to 97492 of contig TRYP1.0.7383 of chromosome l) that has significant homology to TbMP52 (Fig. 6A). This ORF encodes a 48-kDa protein, and similar to TbMP52, the N-terminal sequence of the 48-kDa protein predicts a mitochondrial targeting signal sequence. There is a deletion/insertion site near the N terminus of TbMP48 relative to TbMP52 that appears to be located within the signal peptide. LC-MS/MS analysis of an ~44-kDa protein from the SDS-PAGE gel (Fig. 2, lane 7) identified 13 peptides that correspond to peptides predicted from this ORF (Fig. 6B). A peptide with the amino acid sequence (F)VGGDGFIFER was identified from near the N terminus of the predicted protein (Fig. 6B and C and Table 2). The amino acid (F) N-terminal to V, since it is not tryptic, may be part of the mitochondrial targeting signal. Thus, the first 17 amino acids, which also can fold into an amphiphilic helix (results not shown), are predicted to form the mitochondrial targeting signal. Hence, this second protein, designated TbMP48, also appears to be localized in the mitochondrion after removal of the signal peptide, resulting in a 45-kDa mature protein. The mature TbMP48 migrates as an ~44-kDa band in SDS-PAGE gels.

Additional BLAST searching of the National Center for Biotechnology Information database revealed that the Leishmania major gene L5701.8 (accession no. AAC24666 [20]) has significant homology to TbMP52 at both the nucleic acid and predicted amino acid sequence levels, while the relationship of TbMP48 to TbMP52 is evident only at the predicted amino acid sequence level. L5701.8 has 66% identity and 78% similarity to TbMP52 at the amino acid level, while TbMP48 has 41% identity and 60% similarity to TbMP52. Thus, L5701.8 appears to be an ortholog of TbMP52, while TbMP48 is a paralog of TbMP52. Database searches for homology, motif searches, a ProfileScan of the PROSITE database, and a search of the BLOCKS database revealed no homologs of these proteins, nor any functional motifs. Thus, these proteins appear to be novel.

**Association of TbMP52 and TbMP48 with the RNA editing complex.** The association of the 45- and 48-kDa proteins with the editing complex was assessed by immunoprecipitation, since the cofractionation of these proteins with the editing activity suggests that they may both be associated with the editing complex. Recombinant TbMP52 (rTbMP52) and rTbMP48 proteins were expressed in E. coli with an N-terminal six-His tag. Abundant expression of these proteins was evident as bands in SDS-PAGE gels of IPTG-induced E. coli that stained intensely with Coomassie blue (Fig. 7A). None of the MAbs that we isolated after immunization with the purified editing complex reacted with rTbMP52, while rTbMP48 was recognized by all of the MAbs specific for TbMP48 (Fig. 7B). This further confirms that the TbMP52 gene encodes the protein identified by mass spectrometry analysis of the ~49-kDa protein present in the purified editing complex. This observation, along with the finding that MAb P3C1-G2 specifically immunoprecipitates RNA editing activity from the 20S mitochondrial fraction in high

FIG. 5. Immunofluorescence with MAb P3C1-G2, which is specific for TbMP52. Procyclic T. brucei cells were stained with MAb (A) and DAPI (B), showing the nucleus and smaller kinetoplast.
KCl concentrations (Fig. 3B), shows that TbMP52 is tightly associated with the RNA editing complex. Peptides corresponding to both TbMP52 and TbMP48 were identified in the immunoprecipitated editing complex by mass spectrometry analysis (results not shown).

DISCUSSION

This study reports enrichment of the RNA editing complex from *T. brucei* mitochondria and the identification of two genes that encode 45- and 48-kDa (mature) proteins that are tightly associated with the RNA editing complex. The alignment indicates amino acids that are conserved (+), semiconserved (:), and partially conserved (.) among all these proteins. The first N-terminal peptide (dashed underline) is nontryptic, and the 17 amino acids at the N terminus (italicized) are a predicted mitochondrial targeting signal. (C) CID spectrum of the likely N-terminal peptide of TbMP48.

KCl concentrations (Fig. 3B), shows that TbMP52 is tightly associated with the RNA editing complex. Peptides corresponding to both TbMP52 and TbMP48 were identified in the immunoprecipitated editing complex by mass spectrometry analysis (results not shown).

FIG. 6. (A) Alignment of predicted amino acid sequences of TbMP52 and related proteins TbMP48 and L5701.8. Potential N-terminal amino acids are in bold. The TbMP48 gene sequence is from contig TRYP1.0.7383 of chromosome I (Sanger Center *T. brucei* database) and the L5701.8 ORF is from *L. major* chromosome 1 (20). The alignment indicates amino acids that are conserved (+), semiconserved (:), and partially conserved (.) among all these proteins. (B) Predicted amino acid sequence of TbMP48 showing the 12 tryptic peptides (two of which were contiguous) that were identified by mass spectrometric analysis (underlined). The first N-terminal peptide (dashed underline) is nontryptic, and the 17 amino acids at the N terminus (italicized) are a predicted mitochondrial targeting signal. (C) CID spectrum of the likely N-terminal peptide of TbMP48.
TABLE 2. SEQUEST output file for the CID spectrum of TbMP52 N-terminal peptide

<table>
<thead>
<tr>
<th>(M+H)+</th>
<th>Scorea</th>
<th>Database reference Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1037.1</td>
<td>3.9293</td>
<td>TRYP1.0.7383 (F)VGGDGSIIFER</td>
</tr>
<tr>
<td>1037.2</td>
<td>2.9684</td>
<td>6AA9.TR (V)VGGHGTIFFER</td>
</tr>
<tr>
<td>1037.2</td>
<td>2.8651</td>
<td>1395S223 (A)RFFEAGNVP</td>
</tr>
<tr>
<td>1037.1</td>
<td>2.7272</td>
<td>18L2.22.TF (R)VDSSGKMERT</td>
</tr>
<tr>
<td>1037.1</td>
<td>2.7390</td>
<td>trypE64.p1p (S)VDAYM*IGH</td>
</tr>
</tbody>
</table>

a Cross-correlation score.  
b GenBank sequence identification number.

associated with the editing complex. The complex was fractionated by a combination of cation-exchange, anion-exchange, and gel filtration chromatography followed by glycerol gradient sedimentation using the in vitro deletion editing assay to monitor purification. The enriched complex contains all of the catalytic activities that are associated with editing. These include the gRNA-directed endoribonuclease, 3'-TUTase, 3'-exonuromyldase, and RNA ligase (adenylylation) activities that are predicted by models of the general mechanism of editing (2, 14, 26). A MAb that was prepared by using the enriched complex is specific for both native and recombinant 48-kDa protein and also immunoprecipitates the in vitro deletion editing activity as well as the associated catalytic activities.

The fractionation reported here differs from previously reported purification of the editing complex (18, 24). This study used in vitro editing to monitor purification rather than following purification of adenylylatable proteins. It also enriched the complex sequentially with SP Sepharose (cation-exchange), Q Sepharose (anion-exchange), and Superose 6 (gel filtration) columns followed by sedimentation in a glycerol gradient. The purification achieved could not be calculated based on specific activity since the in vitro editing assay is not linear (our unpublished results). However, purification from cleared mitochondrial lysate is estimated to be ∼8,500-fold based on protein recovery and substantially more based on total cellular protein. The in vitro editing assays require that all steps in editing occur. The intent of the purification approach used in this study was to select for the fully functional complex and to avoid activities that may resemble but not be part of the editing complex.

The complex that is fully functional in editing has an apparent mass of 1,600 kDa, based on gel filtration chromatography, and it sediments at ∼20S. Several of the editing-associated activities, such as endonuclease, TUTase, and RNA ligase activities, elute in a second peak with an apparent mass of 500 kDa (Panigrahi et al., unpublished results). Madison-Antenucci et al. (18) reported two complexes with masses of ∼700 and ∼450 kDa. The smaller masses may reflect the differences in isolation protocol used in the study. In addition, the isolated ∼1,600-kDa complex is capable of in vitro editing in the present study, while this was not shown for the complex isolated by Madison-Antenucci et al. (18). The complex isolated here contained gRNAs and preedited mRNAs, similar to that described by Madison-Antenucci et al. (18) but unlike the complex reported by Rusché et al. (24).

The composition of the fully functional editing complex is not known. The most purified fraction reported here contains 20 major proteins (Fig. 2) that together total 1,140 kDa, but at least three of these proteins may be contaminants, and editing complex proteins may not all be present in a 1:1 stoichiometry (Panigrahi et al., unpublished). The most purified fraction reported by Rusché et al. (24) contained eight major proteins, and those authors suggested that three of the proteins may represent RNA ligase based on their ability to be adenylylated. The most purified fraction isolated by Madison-Antenucci et al. (18) contained 13 major proteins. Since abundance of proteins in a purified fraction is not definitive evidence that a protein is a component of the editing complex, determination of which are components of the complex will await independent evidence. It seems unlikely that the fully functional editing complex contains only eight polypeptides, as suggested by Rusché et al. (24). Those authors suggested that three of the proteins may represent RNA ligase and that editing requires proteins that perform the endoribonuclease, exoribonucleolytic, and TUTase functions. It is also likely that some proteins are involved in RNA-RNA positioning, annealing, and unwinding functions. Indeed, the complex isolated here contains gRNAs and preedited mRNAs, indicating the presence of RNA binding proteins. A helicase activity appears to be associated with RNA editing based on cofractionation (4) and experiments showing that mitochondrial helicase null mutants have reduced editing (19). Thus, helicase may be a component of the editing complex, although not absolutely essential for editing. The pre-mRNA and gRNA must be bound by the complex, perhaps by the anchor duplex, and the 5' cleavage fragment must be retained after cleavage by the endoribonuclease. It also seems likely that proteins are needed to maintain the structure of the editing complex in order to position the catalytic sites of the editing enzymes and relocate the pre-mRNA and gRNA as each site is edited.

Several proteins have been suggested to be components of the editing complex based on approaches other than purification.
tion of the complex. Several proteins which specifically cross-link to gRNA upon UV irradiation have been described. Read et al. (23) demonstrated cross-linking of gRNA to 25- and 90-kDa proteins upon incubation with mitochondrial extract. The 90-kDa protein was shown to be specific for oligo(U), as is the case for the 16-kDa RBP16 (10) and the 75-kDa TBRGG1 (30). Allen et al. (1) identified a 55-kDa gRNA binding protein in an enriched RNA editing fraction. Wang et al. (personal communication) identified 50- and 70-kDa proteins in a partially purified (sequential SP Sepharose- and Q Sepharose-fractionated) editing complex fraction that specifically bind gRNA. The roles of these proteins in RNA editing have not been demonstrated. Indeed, while MAbs against gBP21, the 25-kDa protein that specifically binds gRNA, immunoprecipitate editing activity, this immunoprecipitation is prevented by micrococcal nuclease treatment (1), and gBP21 null mutants perform editing normally (16). Thus, while it appears that gBP21 is an RNA binding protein, its role in editing is uncertain. It does not appear to be an essential component of the editing complex and thus may have an accessory role, such as bringing gRNA to the complex. Alternatively, it may have precipitated editing activity simply based on its affinity for RNA. The 110-kDa glutamate dehydrogenase specifically cross-links to RNA in Leishmania tarentolae (3), but the knock-out of the gene for this protein in T. brucei had no effect on editing (5). The roles of the other RNA binding proteins have not been tested. REAP1 was identified using MAbs made to an ~40S fraction from glycerol gradient fractionation of a mitochondrial lysate. This antibody recognizes proteins in the ~20S and ~40S fractions, where in vitro editing is detected, and upon incubation with the mitochondrial fraction, the antibody inhibits in vitro insertion editing, suggesting that this protein may be a functional component of the complex (18). At this time, the role of REAP1 in RNA editing has not been determined.

The two proteins and corresponding genes that we identified here using a combination of mass spectrometry and immunoprecipitation and Western analyses are candidate components of the editing complex. A MAb that was raised against the purified native complex is specific for a 48-kDa protein that is primarily localized in the 20S fraction in a glycerol gradient of total cleared mitochondrial lysate (Fig. 3E), where editing primarily sediments. The MAb also immunoprecipitates in vitro deletion editing activity from the 20S fraction in addition to the precleaved insertion editing, endonuclease, TUTase, 3'-exonucleolyase, and RNA ligase activities that are associated with editing. Furthermore, it immunoprecipitates the adenylation activity and adenylylated proteins (Fig. 3C and D), which have been suggested to be the editing-associated RNA ligases. The immunoprecipitation was specific, since it occurred even in 400 mM KCl with 0.1% nonionic detergent and did not occur with antibodies that do not react with proteins in the fraction which contains the most in vitro editing activity. The 45-kDa protein is related to the 48-kDa protein and also cofractionates with the in vitro deletion editing activity. Both the 45- and 48-kDa proteins were identified in the immunoprecipitated sample by mass spectrometry (Punigrahi et al., unpublished). The genes for each of these proteins predict a mitochondrial targeting signal sequence, and TbMP52 was localized in mitochondria using the MAb. Thus, taken together the evidence strongly suggests that both proteins are associated with the editing complex. TbMP52 is an ortholog of the L. major gene L5701.8 (20), which has no known function. Database searching found no other homologs or functional motifs in the databases, and thus we are not able to assign a function to these proteins. Further biochemical and genetic studies on the roles of these proteins in RNA editing are in progress. Preliminary results suggest that TbMP52 has RNA ligase activity (Schnaufer et al., unpublished results).

Overall no protein has been shown to be essential for RNA editing. Of the candidates identified to date, a functional association with RNA editing has been demonstrated only for RNA helicase, mHEL61p (19). While several other candidates have been identified by a variety of approaches, functional analyses such as knockout and genetic modification in vivo have yet to show a definitive role in RNA editing or have yet to be done. However, the list of candidates has increased and the criteria for a possible role in editing have improved, suggesting that functions in editing will be demonstrated for several proteins in the near future. For example, what are the specific functions of the individual protein components of the complex, what is the structure of the editing complex and does it have subunits, and what is the biogenesis of the editing complex? Such studies are likely to make progress toward answering these and other questions about the editing complex.

ACKNOWLEDGMENTS

We thank Barbara Morach, Brian Panicucci, and RoseMary Reed for technical help, Bingbing Wang for sharing unpublished results, Elizabeth Wayner for MAb production, and Peter Myler for helpful suggestions. We thank Najib M. El-Sayed for providing sequence information prior to publication. Sequencing of the T. brucei genome was accomplished as part of the Trypanosoma Genome Network with support from The Wellcome Trust and NIAID. R.I. was supported by NIH postdoctoral fellowship AI10312. This work was supported in part by NIH grants RR1823 and AI141109 to R.A. and AI14102 and GM42188 to K.S.

REFERENCES