RNA editing in *Trypanosoma brucei* is a mitochondrial RNA processing reaction that results in the insertion and deletion of uridylic acid residues into otherwise untranslatable mRNAs. The process is directed by guide RNAs which function to specify the edited sequence. RNA editing in vitro requires mitochondrial protein extracts and guide RNAs have been identified as part of high molecular weight ribonucleoprotein complexes. Within the complexes, the RNAs are in close contact with several mitochondrial proteins and here we describe the isolation and cloning of a gRNA-interacting polypeptide from *Trypanosoma brucei*. The protein was named gBP21 for guide RNA-binding protein of 21 kDa. gBP21 shows no homology to proteins in other organisms, it is arginine-rich and binds to gRNA molecules with a dissociation constant in the nanomolar range. The protein does not discriminate for differences in the primary structures of gRNAs and thus likely binds to higher order structural features common to all gRNA molecules. gBP21 binding does not perturb the overall structure of gRNAs but the gRNA/gBP21 ribonucleoprotein complex is more stable than naked guide RNAs. Although the protein is arginine-rich, the free amino acid or an arginine-rich peptide were not able to inhibit the association to the RNAs. In contrast, the gRNA-gBP21 complex formation was sensitive to potassium chloride. This U-tail has been suggested to be the reservoir for the U-addition/deletion reaction or to transiently stabilize the interaction with the pre-mRNAs. The mitochondrial machinery that catalyzes RNA editing is not known today, however, there is substantial evidence that suggests the participation of mitochondrial polypeptides in the process. Guide RNAs as well as pre-edited mRNAs are assembled in high molecular weight ribonucleoprotein (RNP) complexes (reviewed in Ref. 3) and RNA editing in vitro requires mitochondrial protein extracts (4–7). Associated with the RNP complexes are various enzymatic activities such as endonuclease (8, 9), RNA ligase (8–10), RNA helicase (11, 12), and terminal uridylyltransferase (8, 12, 13), which have been suggested to be catalytic components of different steps of the RNA editing process. Corell et al. (12) determined an apparent S value of 20 S for a mitochondrial uridylic uracil deletion activity.

Potential candidates of the editing machinery are polypeptides that are in close contact with gRNA molecules. Several gRNA-binding proteins have been detected in kinetoplastid organisms, based on zero distance cross-linking experiments. In *Leishmania tarentolae*, two potential gRNA-binding proteins with molecular masses of 18 and 51 kDa have been isolated and cloned (14) and the p51 was identified to be homologous to aldehyde dehydrogenase. The p18 protein displayed no homology to known polypeptides. In *Crithidia fasciculata* polypeptides with apparent molecular masses of 33, 65, and 88 kDa showed binding specificity to gRNAs, provided that the RNAs were tailed with 3′ oligo(U) extensions (15). This was interpreted as a form of sequence specific binding to the post-transcriptionally added U-tail of gRNAs. In *Trypanosoma brucei*, several proteins with apparent molecular masses ranging from 124 to 9 kDa were found to cross-link to gRNA molecules (15–18). The binding of the various proteins could be blocked by increasing the monovalent cation concentration and only three proteins with apparent molecular masses of 90, 21, and 9 kDa were stable at ≥100 mM potassium chloride. This suggested a dependence on ionic interactions in the interaction of these polypeptides to the gRNAs. Similar to the situation in *C. fasciculata*, the 90-kDa polypeptide was, in addition, reliant on the presence of an oligo(U) extension at the gRNAs 3′ ends, thus displaying features of sequence specificity.

All the aforementioned gRNA-cross-linking studies in *T. brucei* identified a prominent cross-linking polypeptide; PCR, polymerase chain reaction; nt, nucleotide.

---

**Trypanosoma brucei** gBP21

AN ARGinine-RICH MITochondrial PROTEIN THAT BINDS TO GUIDE RNA WITH HIGH AFFINITY*

(Received for publication, August 29, 1996)

Johannes Köller, Ulrich F. Müller, Beate Schmid, Andreas Missel, Volker Kruft‡, Kenneth Stuart§§, and H. Ulrich Göringer**

From the Laboratorium für molekulare Biologie, Genzentrum der LMU München am MPI für Biochemie, 82152 Martinsried, Germany, §PE Applied Bioysystems, Brunnenweg 13, 64331 Weiterstadt, Germany, the §Seattle Biomedical Research Institute, Seattle, Washington 98109, and the ¶Pathobiology Department, SC38, University of Washington, Seattle, Washington 98195

Mitochondrial gene expression in kinetoplastid organisms requires a RNA processing reaction series known as RNA editing. The process is characterized by the site-specific insertion and deletion of exclusively uridylic acid residues into otherwise encrypted pre-mRNAs to create full-length mRNA molecules for polypeptide synthesis (reviewed in Ref. 1). Key molecules in the encrypted pre-mRNAs to create full-length mRNA molecules are guide RNA molecules. The process is characterized by the site-specific insertion of uridylic acid residues into otherwise untranslatable mRNAs. The process is directed by guide RNAs which function to specify the edited sequence. RNA editing in vitro requires mitochondrial protein extracts and guide RNAs have been identified as part of high molecular weight ribonucleoprotein complexes. Within the complexes, the RNAs are in close contact with several mitochondrial proteins and here we describe the isolation and cloning of a gRNA-interacting polypeptide from *Trypanosoma brucei*. The protein was named gBP21 for guide RNA-binding protein of 21 kDa. gBP21 shows no homology to proteins in other organisms, it is arginine-rich and binds to gRNA molecules with a dissociation constant in the nanomolar range. The protein does not discriminate for differences in the primary structures of gRNAs and thus likely binds to higher order structural features common to all gRNA molecules. gBP21 binding does not perturb the overall structure of gRNAs but the gRNA/gBP21 ribonucleoprotein complex is more stable than naked guide RNAs. Although the protein is arginine-rich, the free amino acid or an arginine-rich peptide were not able to inhibit the association to the RNAs. In contrast, the gRNA-gBP21 complex formation was sensitive to potassium chloride. This U-tail has been suggested to be the reservoir for the U-addition/deletion reaction or to transiently stabilize the interaction with the pre-mRNAs. The mitochondrial machinery that catalyzes RNA editing is not known today, however, there is substantial evidence that suggests the participation of mitochondrial polypeptides in the process. Guide RNAs as well as pre-edited mRNAs are assembled in high molecular weight ribonucleoprotein (RNP) complexes (reviewed in Ref. 3) and RNA editing in vitro requires mitochondrial protein extracts (4–7). Associated with the RNP complexes are various enzymatic activities such as endonuclease (8, 9), RNA ligase (8–10), RNA helicase (11, 12), and terminal uridylyltransferase (8, 12, 13), which have been suggested to be catalytic components of different steps of the RNA editing process. Corell et al. (12) determined an apparent S value of 20 S for a mitochondrial uridylic uracil deletion activity.

Potential candidates of the editing machinery are polypeptides that are in close contact with gRNA molecules. Several gRNA-binding proteins have been detected in kinetoplastid organisms, based on zero distance cross-linking experiments. In *Leishmania tarentolae*, two potential gRNA-binding proteins with molecular masses of 18 and 51 kDa have been isolated and cloned (14) and the p51 was identified to be homologous to aldehyde dehydrogenase. The p18 protein displayed no homology to known polypeptides. In *Crithidia fasciculata* polypeptides with apparent molecular masses of 33, 65, and 88 kDa showed binding specificity to gRNAs, provided that the RNAs were tailed with 3′ oligo(U) extensions (15). This was interpreted as a form of sequence specific binding to the post-transcriptionally added U-tail of gRNAs. In *Trypanosoma brucei*, several proteins with apparent molecular masses ranging from 124 to 9 kDa were found to cross-link to gRNA molecules (15–18). The binding of the various proteins could be blocked by increasing the monovalent cation concentration and only three proteins with apparent molecular masses of 90, 21, and 9 kDa were stable at ≥100 mM potassium chloride. This suggested a dependence on ionic interactions in the interaction of these polypeptides to the gRNAs. Similar to the situation in *C. fasciculata*, the 90-kDa polypeptide was, in addition, reliant on the presence of an oligo(U) extension at the gRNAs 3′ ends, thus displaying features of sequence specificity.

All the aforementioned gRNA-cross-linking studies in *T. brucei* identified a prominent cross-linking polypeptide; PCR, polymerase chain reaction; nt, nucleotide.

---

*This work was supported in part by grants from the Bundesministerium für Bildung und Forschung and the Deutsche Forschungsgemeinschaft (to H. U. G.). The costs of publication of this article were thereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U61382.

¶Supported by National Institutes of Health Grant GM42188.

§To whom correspondence should be addressed: Genzentrum der LMU München am MPI für Biochemie, Am Klopferspitz 18, 82152 Martinsried, Germany, Tel.: 89-8578-2475; Fax: 89-8578-3810; E-mail: goeringe@biochem.mpg.de.

1 The abbreviations used are: gRNA, guide RNA; RNP, ribonucleoprotein; PCR, polymerase chain reaction; nt, nucleotide.
range of 21–26 kDa, appearing as a broad radioactive signal in SDS containing polyacrylamide gels (15–18). Despite the differences in the molecular mass estimation and also some differences in the experimental conditions, the identified characteristics of that cross-link were very similar with one obvious difference: Leegwater et al. (15) reported a dependence on a long U-tail on the gRNAs whereas Köller et al. (16) found cross-linking even with a tail-less gRNA substrate. There is no apparent explanation for this difference but all other features support the interpretation that this cross-link is based on the interaction of gRNAs to the same protein or the same set of proteins. A stimulating new aspect concerning this cross-link was only recently provided by Corell et al. (12). They identified the cross-link in the same gradient fractions of partial purified mitochondrial extracts that contained a gRNA-dependent uridylylation activity. Although a co-localization is no direct evidence, the result might hint a potential role of this gRNA-binding protein during RNA editing.

In this paper we describe the isolation, cloning, and first characterization of the 21–26 kDa gRNA cross-linking protein. The cross-link was identified to be due to a single polypeptide with a calculated molecular mass of 21 kDa. It was named gBP21 for guide RNA-binding protein of 21 kDa. gBP21 is arginine-rich and shows no homology to known proteins. It is nuclear-encoded and presumably is imported into the mitochondrion via a cleavable presequence of 19 amino acids. The gBP21 gRNA-binding protein during RNA editing.

Cell Growth, Preparation of Nucleic Acids, and Mitochondrial Vesicle Isolation—The procyclic life cycle stage of T. brucei brucei strain IS3aR 1 (19) was grown in SDM-79 as described (20). Bloodstream trypanosomes (clone MTTat 1.2) were grown in HMI-9 medium according to Ref. 21. Genomic DNA was prepared as described in Ref. 21 and poly(A)·mRNA was purified from whole cell RNA preparations (23) using paramagnetic oligo(dT)25 beads (Dynabeads). Mitochondrial vesicles were isolated essentially as described by Harris et al. (24). Vesicle preparations were stored at −80°C in 20 mM Tris-HCl, pH 5.5, 2 mM Na2EDTA, 250 mM sucrose, and 50% (v/v) glycerol.

Isolation of gBP21 Protein—The polypeptide was isolated from mitochondrial extracts prepared at low ionic strength in the presence of Nonidet P-40 (3.3 mM) as described (25). Extracts were cleared from insoluble material by centrifugation at 16,000 × g for 5 min at 4°C and protein concentrations were determined by dye-binding using bovine plasma α-globulin as a standard (26). The purification of gBP21 was followed by monitoring the binding activity to T. brucei gRNA gA6-14 (16). Cleared extracts were precipitated with ammonium sulfate at 4°C and the protein was enriched in the 50–60% (v/v) saturation fraction. The precipitate was dissolved in 10 mM HEPES pH 7.5, 1 mM (NH4)2SO4 and loaded onto a phenyl-Sepharose column equilibrated in the same buffer. Proteins were eluted with a linear gradient of 0–1 mM (NH4)2SO4. These fractions were concentrated by ultrafiltration and dialyzed against 20 mM Tris-HCl, pH 7.5, 100 mM KCl, 2.1 mM MgCl2, 0.1 mM Na2EDTA, 0.5 mM dithiothreitol, and purified by ion exchange chromatography on Mono Q HR 5/5 and Mono S HR 5/5 columns (Pharmacia). The purity of the recombinant gBP21 was verified by microsequencing.

Determination of Equilibrium Dissociation Constants—The binding of gBP21 to gRNA molecules was measured by retention on nitrocellulose filters as described with Witherell and Uhlenbeck (31). RNA molecules (gA6-14, gA6-48, gA6-14/U, gND7-506, RNA1) were synthesized from the N-terminal amino acid sequence as well as primers directed against the spliced leader (SL) sequence and a universal amplification primer (UAP, Life Technologies Inc.). 5′-rapid amplification of cDNA ends: 1) round PCR, T.b.Sal-Sal: GTCGGTGACACCTAAAGGATTTAT- AAAGAAG; 2) round PCR, T.b.Sal-Sal: GACAGTTTCTGATATAAGTTG; 3′-Rapid amplification of cDNA ends: 1) round PCR, T.b.Sl-Sl: GGCTGCTTGCGGGCCGCTTAC; UAP: GGCCACGGCGTCACTAGTAC. 2) round PCR, T.b.Sl-Sl: GGCCACGGCGTCACTAGTAC. PCR products were ligated into plasmid pCR-Script SK (+) (Stratagene) and transformed into Escherichia coli XLI-Blue MRF′ Kan cells. Plasmid DNA was sequenced and the sequence information was used to PCR amplify the full-length gBP21 gene from T. brucei genomic DNA using primers gBP21–5′: CCACCAGGTCCTACGATAGAAG and gBP21–3′: CGCAGGAAGCTTTTACGGTTCCCGGATTCTAG. The resulting PCR fragment of 551 base pairs was cloned into NdeI and HindIII restricted plasmid pTTT-7 (29) resulting in plasmid pTTT-7-gBP21 which was transformed into BL21(DE3) and BL21(DE3) pLYsS cells (30) for expression. Cells were grown at 37°C in 3×YT medium containing 50 μg/ml ampicillin and 50 μg/ml chloramphenicol for plasmid selection. At an A500 of 0.7 expression was induced by adding 0.5 mM isopropyl-1-thio-β-D-galactopyranoside and incubating for additional 3 h at 24°C. E. coli cells were broken open by freeze/ thaw lysis followed by sonication in 50 mM NaH2PO4, pH 7.8, 500 mM NaCl. The cytosolic extract was dialyzed against 20 mM Tris-HCl, pH 7.5, 100 mM KCl, 2.1 mM MgCl2, 0.1 mM Na2EDTA, 0.5 mM dithiothreitol and purified by ion exchange chromatography on Mono Q HR 5/5 and Mono S HR 5/5 columns (Pharmacia). The term "equilibrium" refers to the point at which the concentrations of bound and unbound molecules are in constant state. The term "dissociation constant" refers to the equilibrium constant for the reaction between the bound and unbound molecules.

Circular Dichroism—Circular dichroism measurements were made at 20°C in 6 mM Hepes pH 7.5, 300 mM KCl, 2.1 mM MgCl2, 0.1 mM Na2EDTA, 0.5 mM dithiothreitol, 1 mM ATP, and 6% (v/v) glycerol using a Jasco J-20 CD spectrometer. CD spectra were recorded at a scan rate of 50 nm/min, a bandwidth of 1 nm, a response time of 1 s, and a temperature of 20°C. The instrument was calibrated with a 100% methanol solution. The CD spectra were corrected for the contribution of the buffer used.

Glycerol Gradient Sedimentation—The sedimentation coefficient (s20,w) of gBP21 protein was determined by ultracentrifugation in linear 5–35% (w/v) glycerol gradients in 50 mM Tris-HCl, pH 7.5, 100 mM KCl, 2.1 mM MgCl2, 0.1 mM Na2EDTA, 0.5 mM dithiothreitol. Centrifugation was carried out in a Beckman SW41 Ti rotor at 35,000 rpm for 36 h at 4°C using RNaase A (2.0 S), chymotrypsinogen (2.5 S), ovalbumin (3.5 S), and bovine serum albumin (4.3 S) as standard proteins.

Circular Dichroism—CD spectra of gBP21, gA6-14, and the gBP21:gA6-14 complex were measured at 20°C in 6 mM Hepes pH 7.5, 50 mM KCl, 2.6 mM MgCl2, 0.1 mM Na2EDTA, and 6% (v/v) glycerol. Spectra were recorded from 300 to 195 nm with a data acquisition every 0.1 nm. Scans were repeated 10 times and averaged. Mean molar residual ellipticities (θm) were calculated per mole of amino acid or nucleotide monomer and secondary structure contents were estimated according to Provencher (33). For the CD spectra of the gBP21:gA6-14 complexes equimolar amounts of protein and gA6-14 RNA were incubated at 4°C for 2 h in 10 mM Hepes, 100 mM KCl, 0.5 mM MgCl2, 0.1 mM Na2EDTA, 0.5 mM dithiothreitol, 0.1% Tween 80, and 20% (v/v) glycerol.
bated for 30 min at 27 °C and the spectrum was corrected for the presence of the protein by subtracting the curve for the free polypeptide.

UV Melting Curves—Absorbance versus temperature profiles of gA6-14 gBP21 RNP complexes were recorded at 260 nm in 50 mM potassium cacodylate, pH 7.2, 50 mM KCl, 2.6 mM MgCl2, 0.1 mM Na2EDTA, and 6% (v/v) glycerol using a thermoelectrically controlled Perkin Elmer Lambda 16 spectrophotometer. The temperature was increased at a heating rate of 2 °C/min at temperatures between 10 and 95 °C. For complex formation equimolar amounts of gA6-14 and gBP21 (70 nM) were incubated for 30 min at 27 °C in the above described buffer. 

Tm values were determined from derivative plots of absorbance versus temperature as 0.5 ΔA260.

RESULTS

Purification and Cloning of T. brucei gBP21—The protein was enriched from low salt detergent extracts of T. brucei mitochondria isolated from the procyclic life cycle stage of strain IsTaR 1 (19). As an assay system for the enrichment we used the demonstrated ability of the polypeptide to specifically bind to gRNA molecules, which was monitored in an UV light induced label transfer experiment as described in Ref. 16. The procedure, as outlined below, routinely yielded 100 pmol of purified gBP21 from 20 mg of total mitochondrial protein.

A cleared mitochondrial lysate (Fig. 1, lanes mt extract) was first fractionated by the stepwise addition of ammonium sulfate. The majority of the gRNA cross-linking activity (>90%) precipitated in the 50–60% (w/v) saturation fraction but this fraction still contained the bulk of mitochondrial proteins (Fig. 1, lanes 50–60% AS). We further separated this sample by hydrophobic interaction chromatography using a phenyl-Sepharose column. Polypeptides were eluted with a linear ammonium sulfate gradient and the gRNA binding activity eluted between 0.5 and 0.3 M (NH4)2SO4. Three major polypeptides remained in this fraction (Fig. 1A, lane HIC) but only the protein band with an apparent molecular mass of 26 kDa overlapped with a broad but strong radioactive signal when cross-linked to added radiolabeled gA6-14 RNA (Fig. 1B, lane HIC). To exclude that more than one protein co-migrated within that protein band, we performed two-dimensional gel electrophoresis with the 26-kDa protein separating at a pI of 9.5 (data not shown). Gels were blotted and the protein spot was microsequenced. Forty-three amino acids were obtained from the N terminus of the protein and this sequence information was used to PCR amplify the corresponding gene from T. brucei genomic DNA following the procedure outlined under "Materials and Methods."

The gene encoded for an open reading frame of 621 nt (Fig. 2A), the putative 19-amino acid presequence is denoted in bold italic. The sequence stretch verified by N-terminal sequencing and translation initiation and termination codons are underlined. The polyadenylation site as well as sites for the addition of the miniexon sequence are denoted by a small diamond. B, helical wheel projection of the putative mitochondrial targeting sequence of gBP21. Hydrophobic amino acids are underlined. The broken line separates the hydrophilic domain of the helix from the hydrophobic part to demonstrate its amphiphilic characteristic. The helix projection was calculated for an angle of 100°.

Fig. 1. Enrichment of gBP21 from T. brucei mitochondrial extracts. Panel A, electrophoretic analysis of proteins from a mitochondrial (mt) extract, a 50–60% (w/v) ammonium sulfate (AS) fraction and an isolate from the hydrophobic interaction chromatography (HIC) on phenyl-Sepharose. Separation was performed in SDS containing polyacrylamide gels (12% w/v) and stained with Coomassie Brilliant Blue. Size markers are annotated on the left in kDa and the position of gBP21 is shown by an arrowhead. B, UV cross-linking activity of the same protein isolates as in A demonstrating the specific enrichment of the polypeptide. Cross-linking was performed with 32P-labeled gA6-14 RNA and analyzed on 12% (w/v) SDS containing polyacrylamide gels followed by autoradiography.

Fig. 2. Nucleotide and predicted amino acid sequence of gBP21. A, the putative 19-amino acid presequence is denoted in bold italic. The sequence stretch verified by N-terminal sequencing and translation initiation and termination codons are underlined. The polyadenylation site as well as sites for the addition of the miniexon sequence are denoted by a small diamond. B, helical wheel projection of the putative mitochondrial targeting sequence of gBP21. Hydrophobic amino acids are underlined. The broken line separates the hydrophilic domain of the helix from the hydrophobic part to demonstrate its amphiphilic characteristic. The helix projection was calculated for an angle of 100°.
long peptide, rich in hydroxylated and positively charged amino acids but not containing negatively charged amino acids. It can be folded into an amphiphilic helix thus displaying features of a mitochondrial import sequence (Fig. 2B).

The processed protein is 187 amino acids long with a calculated molecular mass of 21,125 Da. It was named gBP21 for guide RNA-binding protein of 21 kDa. It contains single cysteine and tryptophan residues and 18 arginines as the most abundant amino acid (13.3 molecular mass%). The polypeptide is hydrophilic (62% of the amino acids are hydrophilic) with a calculated pI of 9.7. We were not able to identify membrane spanning domains (34, 35) and only three very short hydrophobic stretches exist. The extreme C terminus (13 amino acids) is the most hydrophobic part of the polypeptide with seven charged and five hydroxylated amino acids separated by a single valine residue. Sequence comparison did not show significant homologies to other polypeptides in the various data bases and no specific protein motif was identified using the MOTIF program of the GCG software (28). In particular, no amino acid sequence known to confer single- or double-stranded RNA binding specificity was identified (36), with the possible exception of three arginine-rich sequence stretches (residues 82–89, 101–112, and 194–204) showing some similarity to arginine-rich motifs in other RNA-binding proteins (residues 82–89, 101–112, and 194–204) showing some similarity to arginine-rich motifs in other RNA-binding proteins.

A Northern blot analysis revealed that gBP21 was expressed as a 1.1-kilobase mRNA in procyclic and bloodstream forms, the major life cycle stages of the parasite (Fig. 3).

Expression of gBP21 in E. coli Cells—A gBP21 coding sequence, equivalent to the processed protein was cloned into expression vector pTT7-7 (29) and transformed into E. coli strain BL21 (DE3) (30). This construct allowed the expression of recombinant (r) protein after transcription induction with isopropyl-1-thio-β-D-galactopyranoside and as shown in Fig. 4A, large amounts of r-gBP21 accumulated already after 1 h of induction. Bacterial growth was performed at 24 °C to avoid formation of inclusion bodies containing inactive r-gBP21. Small amounts of the protein were expressed even before transcription induction (Fig. 4A, lane 0), which subsequently was avoided by using BL21 (DE3) pLysS cells as the parental strain (30). The protein was isolated from bacterial lysates by ion exchange chromatography taking advantage of the cationic nature of the polypeptide for purification (see “Materials and Methods”). One hundred ml of bacterial culture yielded approximately 1 mg of r-gBP21 with a purity of ≥95% as judged from silver-stained SDS containing polyacrylamide gels. Recombinant gBP21 preparations were checked for their cross-linking activity with synthetic, radiolabeled gRNAs (Fig. 4B). The cross-link properties were identical to what we had observed using mitochon-
mRNA; gA6-48, another A6 specific gRNA; and gND7-506, a NADH subunit 7-specific gRNA molecule. As shown in Fig. 6A we measured typical saturation-type binding isotherms for all three RNAs with a maximal retention on the filters varying between 35 and 50%. Guide RNA, gA6-14/U, a RNA construct lacking the 3'-terminal U-tail but otherwise identical to gA6-14 (16), showed a very similar binding curve when compared to the three uridylated RNAs, indicating that the 3'-terminal bases of the gRNAs had no strong influence on the binding of the protein. We also tested a "non-gRNA" transcript, termed RNA1 (16), which was of similar length to gRNAs (56 nt). It showed only 5% retention at the same assay conditions, demonstrating the binding specificity of gBP21 for gRNA molecules in line with previous data (16). Equilibrium $K_d$ values for the formation of gBP21-gRNA complexes were determined in a Scatchard analysis. In these experiments the gBP21 concentration was held constant at 100 nM and the RNA concentrations were varied in the neighborhood of the protein concentration. Fig. 6B shows representative examples of two Scatchard plots and the derived $K_d$ values for the tested gRNAs are listed in Fig. 6C. For the uridylated gRNAs (all three guide RNAs contained 10 Us at their 3' end) we determined $K_d$ values in the range of 8–10 nM whereas the non-uridylated gA6-14 molecule (gA6-14/U) had a slightly weaker $K_d$ of 16 nM.

The amount of protein in the r-gBP21 preparations that was active in RNA binding was determined in a direct titration experiment. Increasing amounts of radiolabeled gA6-14 were added to a fixed amount of gBP21 (4 μM) and the data were analyzed in a plot of free versus added total gA6-14 (Fig. 6D). Complete binding occurred until the concentration of gA6-14 reached 0.4 μM from where the concentration of free RNA increased linearly. Assuming one binding site per gBP21 molecule, the data demonstrated that only 0.4 μM binding sites could be saturated from a total of 4 μM. This indicated that only 10% of the gBP21 molecules were active in binding. For different r-gBP21 preparations we determined values between 10 and 30%.

**Binding Inhibition Experiments**—Based on the observation that the sequence of gBP21 contained three arginine-rich sequence elements (see above) which potentially could provide the RNA binding specificity (37), we tested whether arginine would act as a competitor of complex formation (41). As shown in Fig. 7A, even at a concentration of 20 mM arginine, equivalent to a 4 × 10^3-fold molar excess over gBP21, we were not able to see a significant decrease in gRNA binding. Lysine as another positively charged amino acid as well as serine (data not shown) gave similar half-maximal inhibition values ($K_i$ = 20 mM). For the arginine-rich heptapeptide KRTLRR we measured a $K_i > 5$ mM. In contrast, binding experiments at increasing potassium chloride (10–250 mM) and NH$_4$Cl (1–500 mM) concentrations demonstrated the salt sensitivity of the gBP21/gRNA association. Half-maximal inhibition ($K_i$) occurred at 100 mM KCl and 90 mM NH$_4$Cl (Fig. 7B).

**CD and Hyperchromicity Measurements**—Circular dichroism measurements revealed that the gRNA molecules did not change their overall structure upon binding to the protein. Fig. 8A shows the typical A-form RNA spectrum (42) of gA6-14 in its naked form (dotted line); a positive ellipticity around 270 nm and negative elipticities at 240 and 210 nm. When compared to the spectrum in its complexed form with gBP21, no significant changes were detected (Fig. 8A, solid line). This indicated that the gRNA structure remained largely unperturbed within the RNP complex. However, gBP21 binding did result in a stabilization of the gRNA structures. As has been shown before (32), UV melting curves of naked gRNAs are characterized by a main melting transition around 38–40 °C (Fig. 8B, broken line). When complexed with gBP21 this melting transition completely disappeared, indicating 100% binding of the gRNA. In addition, two new transitions were detected (Fig. 8B, solid line); a high temperature $T_m$ around 70 °C and a second transition at 52 °C. Assuming that gBP21 binds only to fully folded gRNA molecules, the $T_m$ at 52 °C presumably reflects the temp-
temperature-dependent dissociation of the gRNA-gBP21 complex and thus demonstrated that the RNP complex is more stable than the naked gRNAs (equivalent to a melting point shift of $10^\circ$C). The $T_m$ at 70°C presumably corresponds to the complete unfolding of dissociated gRNA or maybe a consequence of denatured protein aggregation as suggested by Xing and Draper (43). Last, we detected a general decrease in the hyperchromicity (approximately by a factor of 10), indicative of a strong shielding effect by the protein.

DISCUSSION

Here we describe the identification and first characterization of a novel T. brucei protein which we termed gBP21, for guide RNA-binding protein with a molecular mass of 21 kDa. The molecule is a basic polypeptide, 187 amino acids in length and shows no homology to known proteins. In particular, there is no homology to the p18 gRNA-binding protein identified in L. tarentolae cells (14) although the two proteins are of identical length. gBP21 and p18 share only a 16-amino acid overlap of 37.5% identity, located at the N terminus of the mature gBP21 and the extreme C terminus of p18. Deduced from the sequence of the gBP21 gene, we identified a 19-amino acid long potential N-terminal mitochondrial import sequence. This suggests a cytosolic translation of the gBP21 mRNA to yield a precursor protein that undergoes maturation after mitochondrial import.

The length of the presequence is similar to what has been reported for other putative import sequences in kinetoplastid organisms (14, 44, 45). Given the absence of membrane spanning domains or long hydrophobic stretches within the polypeptide, we propose that gBP21 is a soluble mitochondrial matrix protein.

The processed protein shows an anomalous electrophoretic mobility in SDS containing polyacrylamide gels of 26 kDa versus a calculated 21 kDa. This is likely due to the basic nature of the polypeptide (pI 9.5) causing a reduced electrophoretic mobility at pH 8.8. Our recombinant protein preparations consisted of roughly 50% $\beta$-strand secondary structure elements and only around 10% $\alpha$-helical domains. Whether these values reflect the in vivo folding characteristics of gBP21 is doubtful. Although we avoided denaturing conditions during the isolation of the recombinant protein, only 10–30% of the molecules were active in gRNA binding. This might be due to a lack of obligatory post-translational modifications of the recombinant protein isolates, or to an incorrect folding of r-gBP21 in the E. coli cytosol. The formation of inclusion bodies upon expression of gBP21 at 37°C might argue for the latter explanation. Although the values of active protein seem to be low, they are

![Fig. 6. RNA binding properties of gBP21.](image)

**A**. binding isotherms for radioactively labeled guide RNAs gA6-14, gA6-48, and gND7-506 are shown in comparison to gA6-14/U, a guide RNA lacking its 3′-terminal oligo(U) extension and RNA1, an artificial RNA of similar length to gRNAs (56 nt) (16). The data are derived from nitrocellulose filter binding experiments with the percentage of retained RNA plotted against the amount of gBP21 added to the binding reaction. Protein concentrations were corrected for the amount of active protein in the preparation (see D). Binding experiments were performed at least in triplicate with different RNA and protein preparations. B. equilibrium dissociation constants ($K_d$) were calculated based on the Scatchard equation: $r/[A] = n/K_d - r/K_d$, with the slope of the plot representing: $1/K_d (A$, free RNA), $r = b/gPB21$, $b$, bound RNA. Representative plots are shown for the binding of gND7-506 and RNA1 and the derived $K_d$ values for the tested gRNAs are listed in C. D. example of a titration experiment to determine the amount of active protein within a gBP21 preparation. Plotted is the amount of free gRNA (gA6-14) as a function of the total gRNA concentration added to a fixed amount of gBP21 (4 μM). Free gRNA can be detected at a concentration of 0.4 μM. Assuming one gRNA-binding site per gBP21 molecule this indicated that 10% of the 4 μM gBP21 were competent in gRNA binding.
The polypeptide binds to various gRNA molecules with equilibrium dissociation constants around 10 nM, thereby not discriminating for differences in the primary sequences of the gRNAs. It is therefore likely that higher order folding characteristics common to all gRNAs, as suggested (32), provide the binding site for the protein. The hyperchromicity measurements indicated a strong shielding effect of the protein and NH₄Cl from 1 to 500 mM (in the presence of 10 mM KCl). Half-maximal inhibitory concentrations (Kᵢ) were derived by interpolation from the semilogarithmic plots of RNA retained (%) versus the concentration of competitor or inhibitor.

A guide RNA lacking its 3′ oligo(U) tail bound to the protein almost as strongly as the corresponding molecule containing 10 Us at its 3′ end. This excludes the U-tail as a major determinant of the binding site for gBP21 and supports previous results from UV cross-linking experiments using the same 3′ end truncated form of gA6-14 as a substrate (16). However, since we have studied only one pair of gRNA molecules with and without a U-tail, we cannot exclude that the observed difference reflects more subtle elements in the association of gBP21 to gRNAs. In principle, two explanations can account for the observed effect: first, removal of the U-tail might cause a subtle structural alteration in the gBP21-binding site, thereby weakening the interaction with the protein. Alternatively, some bases of the U-tail might be part of the gBP21 interaction domain and provide a small but defined contribution to the overall Kᵢ. Our current knowledge on the secondary structure folding of T. brucei gRNAs (32) suggest a single stranded, helical conformation for the 3′ tail, namely that pool. A discrepancy still remains regarding the work of Blum and Simpson (49) determined an in vivo length heterogeneity of the oligo(U) extensions in L. tarentolae cells between 5 and 24 Us and it is conceivable that small differences in the Kᵢ values might be sufficient to select a subpopulation of gRNAs out of that pool. A discrepancy still remains regarding the work of Leegwater et al. (15). They reported a strict requirement of a T. brucei 26-kDa gRNA-specific cross-link on the presence of an oligo(U) extension. In line with their view, we feel it is unlikely...
that this cross-link represents a different mitochondrial protein. However, their data are not supported by the work presented here, since gA6-14/U preparations still bound to gBP21 with a $K_d$ in the nanomolar range.

Interestingly, from our CD measurements we were not able to resolve any structural alteration of the complexed gRNAs when compared to the free RNA structures. Obviously, no gross structural changes are induced within the gRNA molecules to accommodate the association with gBP21. However, the formed RNP complexes have more stable characteristics than naked gRNA. UV melting experiments indicated a $T_m$ of 52°C for the gA6-14-gPB21 complex in contrast to only 38°C for the naked gA6-14 molecule. Thus, gBP21 binding results in a stabilization of the gRNA molecules. The possibility that a transient interaction of the protein might have induced a structural rearrangement within the RNA molecules is not supported on two counts. First, the CD measurements did not show such a structural alteration and second, a secondary structure with a melting point of 52°C cannot be formed for gA6-14, based on the very high A/U content of the RNA molecule (32).

At the moment we can only speculate as to the molecular nature of the gBP21-gRNA association. None of the known protein motifs known to mediate RNA binding (reviewed in Ref. 36) such as RGG (arginine/glycine/glycine) boxes, RRM domains (RNA recognition motif), or KH modules (50) are present in gBP21. The protein does contain three arginine-rich sequence stretches but free arginine and also a synthetic arginine-rich hexapeptide (KRTLRR) were not capable of inhibiting the association to gRNAs. This is in contrast to what has been reported for other RNA-binding arginine-rich proteins (38, 51, 52) and therefore, it seems unlikely that the gBP21-gRNA interaction is based on specific arginine contacts. However, we cannot exclude that the entire sequence context of one or even all three arginine-rich sequences are required for gRNA binding. Support for a RNA-protein interaction, at least in part, based on ionic contacts comes from the salt sensitivity of the gBP21-gRNA interaction. Both, KCl and NH$_4$Cl at concentrations around 100 mM were able to inhibit the formation of the RNP complex. This can be interpreted as the presentation of a defined geometry of negative charges on the surface of the gRNAs, which is mapped by an exact array of positive charges on the exterior of gBP21. These features are in line with previously published work describing the salt sensitivity of gRNA-containing mitochondrial RNP complexes (25) and gRNA-protein cross-links (16) as well as with the salt sensitivity of the in vitro RNA editing activity in T. brucei mitochondrial extracts (4). Our CD measurements of recombinant gBP21 protein identified a high content of $\beta$-strand secondary structure. Although it seems questionable that this represents the native folding of gBP21 (see discussion above), we have to consider that a platform of $\beta$-strands, as identified in several other protein interactions (53), might be used as the basal binding surface for the gRNA molecules.

None of the identified characteristics of gBP21 allowed us to deduce a biological function for the protein. A participation during the editing process is conceivable since the polypeptide binds specifically and with high affinity to gRNAs, key components in the editing process. Further support for an involvement during RNA editing comes from the observation that a protein-gRNA cross-link of similar apparent size co-localizes with an in vitro RNA editing activity in gradient fraction of T. brucei mitochondrial extracts (12). The ability of gBP21 to cross-link to added synthetic gRNAs within an E. coli cell extract as well as in a purified form suggests, that the binding of gRNAs to gBP21 is not dependent on the interaction of other mitochondrial proteins. Although we cannot exclude a stimulatory effect by other polypeptides, this might indicate that gBP21 is an early assembly component to gRNAs which form several high molecular weight RNP complexes in T. brucei mitochondria (8, 18). In contrast, it is equally possible that gBP21 by binding to gRNAs prevents the base pairing interaction to the pre-mRNAs. In this case the protein would exclude gRNAs from the editing reaction unless it is specifically removed by other components of the editing machinery. Preliminary experiments to determine the on and off rates for the gBP21-gRNA interaction seem to support such a scenario since the on rates vastly exceed the off rates.2 Lastly, gBP21 could simply be a polypeptide that binds to gRNAs to regulate their turnover within the mitochondria. Experiments to corroborate the various possibilities are currently being performed.

Acknowledgments—We thank all members of the Göringer laboratory for helpful comments. A. Souza is thanked for critical reading of the manuscript and for stimulating discussions. G. Norskau is acknowledged for technical assistance and L. Moroder and E. Weyer-Stingl for their help performing the CD measurements. T. Seebeck kindly provided a plasmid encoding T. brucei tubulin genes.

REFERENCES

1. Simpson, L., and Thiemann, O. H. (1995) Cell 81, 837–840
2. Benne, R. (1992) Mol. Biol. Rep. 16, 217–227
3. Göringer, H. U., Köller, J., and Shu, H. H. (1995) Parasitology Today 11, 265–267
4. Seiwert, D. S., and Stuart, K. (1994) Science 266, 114–117
5. Frech, G. C., Bakalar, N., Simpson, L., and Simpson, A. M. (1995) EMBO J. 14, 178–187
6. Seiwert, D. S., Heidmann, S., and Stuart, K. (1996) Cell 84, 831–841
7. Pillier, K. J., Ruschö, L. N., and Solnner-Webb, B. (1996) J. Biol. Chem. 271, 4013–4019
8. Pollard, W. V., Harris, M. E., and Hajduk, S. L. (1992) EMBO J. 11, 4429–4438
9. Pillier, K. J., Decker, C. J., Ruschö, L. N., and Solnner-Webb, B. (1995) Mol. Cell. Biol. 15, 2925–2932
10. Sabatini, R., and Hajduk, S. L. (1995) J. Biol. Chem. 270, 7233–7240
11. Misset, A., and Göringer, H. U. (1994) Nucleic Acids Res. 22, 4050–4056
12. Corell, R. A., Read, L. K., Riley, G. H., Nellissey, J. K., Allen, T. E., Kable, M. L., Wachal, M. D., Seiwert, C. S., Myler, P. J., and Stuart, K. (1996) Mol. Cell. Biol. 16, 1410–1418
13. Bakalar, N., Simpson, A. M., and Simpson, L. (1989) J. Biol. Chem. 264, 18679–18686
14. Bringaud, F., Peris, M., Zen, K. H., and Simpson, L. (1995) Mol. Biochem. Parasitol. 71, 65–79
15. Lee, S. T., Speiger, D., and Benne, R. (1996) Eur. J. Biochem. 227, 780–786
16. Köller, J., Norskau, G., Paul, A. S., and Stuart, K., and Göringer, H. U. (1994) Nucleic Acids Res. 22, 1988–1995
17. Read, L. K., Göringer, H. U., and Stuart, K. (1994) Mol. Cell. Biol. 14, 2629–2639
18. Zhu, H. H., Stuart, K., and Göringer, H. U. (1995) Biochem. Biophys. Acta 1291, 349–359
19. Stuart, K., Gobright, E. J., Nill, M., Thomaslow, M., and Abagian, H. (1984) J. Biol. Chem. 70, 747–754
20. Bruns, R. and Schönenberger, V. (1979) Acta Trop. 36, 289–292
21. Hirumi, H., and Hirumi, K. (1994) Parasitol. Today 10, 80–84
22. Mazzio, R. M., Baxter, M. L., Robertson, B. D., and Selkirk, M. E. (1991) in Parasite Antigens, Parasite genes: A Laboratory Manual for Molecular Parasitology, pp. 118–119, Cambridge University Press, UK
23. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156–159
24. Harris, M. E., Moore, D. R., and Hajduk, S. L. (1990) J. Biol. Chem. 265, 11368–11376
25. Göringer, H. U., Kaslowsky, D. J., Morales, T. H., and Stuart, K. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 1775–1780
26. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
27. McPherson, M. J., Jones, K. M., and Gurr, S. J. (1992) in PCR: A Practical Approach (McPherson, M. J., Quirke, P., and Taylor, G. R., eds) pp. 171–186, IRL Press, Oxford, UK
28. Devereux, J., Haebeli, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 887–895
29. Tañor, S., and Richardson, C. C. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 1074–1078
30. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorf, J. W. (1990) Methods Enzymol. 185, 60–89
31. Witherell, G. W., and Uhlenbeck, O. C. (1989) Biochemistry 28, 71–76
32. Schmidt, B., Riley, G. H., Stuart, K., and Göringer, H. U. (1995) Nucleic Acids Res. 23, 3083–3102
33. Provencher, S. W. (1984) Comput. Phys. Comm. 27, 229–242
34. Fasman, G. D., and Gilbert, W. A. (1990) Trends Biochem. Sci. 15, 89–92
35. Jahnig, F. (1990) Trends Biochem. Sci. 15, 92–95
36. Burd, C. G., and Dreyfuss, G. (1994) Science 265, 615–620

$^2$ U. F. Müller and H. U. Göringer, unpublished data.
Molecular Cloning of a gRNA-binding Protein in T. brucei

37. Lazinski, D., Grzadzielska, E., and Das, A. (1989) Cell 59, 207–218
38. Tan, R., and Frankel, A. D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 5282–5286
39. Vioque, A., Arnez, J., and Altman, S. (1988) J. Mol. Biol. 202, 835–848
40. Green, C. J., Rivera-León, R., and Vold, B. S. (1996) Nucleic Acids Res. 24, 1497–1503
41. Tao, J., and Frankel, A. D. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 2723–2726
42. Gray, D. M., Liu, J.-J., Ratliff, R. L., and Allen, F. S. (1981) Biopolymers 20, 1337–1382
43. Xing, Y., and Draper, D. E. (1995) J. Mol. Biol. 249, 319–331
44. Efstron, P. N., Torri, A. P., Engman, D. M., Donelson, J. E., and Englund, P. T. (1993) Mol. Biochem. Parasitol. 59, 191–200
45. Priest, J. W., and Hajduk, S. L. (1995) Biochim. Biophys. Acta 1269, 201–204
46. Talbot, S. J., and Altman, S. (1994) Biochemistry 33, 1399–1405
47. Spierer, P., Bogdanov, A. A., and Zimmermann, R. A. (1978) Biochemistry 17, 5394–5398
48. Karn, J., Gait, M. J., Curcher, M. J., Mann, D. A., Mikaelian, I., and Pritchard, C. (1994) in RNA-Protein Interactions (Nagai, K., and Mattaj, I. W., eds) pp. 192–220, IRL press, Oxford, UK
49. Blum, B., and Simpson, L. (1996) Cell 62, 391–397
50. Siomi, H., Matunis, M. J., Michael, W. M., and Dreyfuss, G. (1993) Nucleic Acids Res. 21, 1193–1198
51. Greenbaum, N. (1996) Structure 4, 5–9
52. Sundquist, W. I. (1996) Nature Struct. Biol. 3, 8–11
53. Mattaj, I. W., and Nagai, K. (1995) Nature Struct. Biol. 2, 518–522