The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/22246

Please be advised that this information was generated on 2019-10-03 and may be subject to change.
Identification by UV cross-linking of oligo(U)-binding proteins in mitochondria of the insect trypanosomatid *Crithidia fasciculata*

Peter LEEG WATER, Dave SPEIJER and Rob BENNE

E. C. Slater Institute, University of Amsterdam, Academic Medical Centre, Amsterdam, The Netherlands

(Received 12 October/29 November 1994) - EJB 94 1554/1

RNA editing in trypanosomates such as *Leishmania tarentolae*, *Trypanosoma brucei* and *Crithidia fasciculata* is the process of insertion and deletion of U residues at specific sites of mitochondrial transcripts mediated by short guide RNAs (gRNAs) that have a 3′ oligo(U) extension. Here we describe the identification by UV cross-linking of proteins present in mitochondrial extracts from *Crithidia fasciculata* with a high affinity for gRNAs, and the characterization of the binding specificity. A 65-kDa protein binds to gRNAs provided they are equipped with a U tail, to post-transcriptionally labelled mitoribosomal 9S and 12S RNAs that also possess a 3′ terminal stretch of U residues, and to free oligo(U) sequences with a minimal length of 23–29 nucleotides. It does not bind to a number of control RNAs, one of which has an internal U stretch of 13 residues. Poly(U), but not poly(C) or total yeast RNA, efficiently competes for binding to gRNA. Proteins of 88 kDa and 30 kDa also bind to gRNAs with a U tail, to mitochandrial ribosomal RNAs and to oligo(U). These proteins, however, require longer oligo(U) for binding (>39 nucleotides) and they also have an affinity for other U-rich RNAs and poly(C). For comparison, part of the analysis was also carried out with a mitochondrial extract from *Trypanosoma brucei*. In this organism, gRNA-binding proteins of 83 kDa and 64 kDa were found with the same preference for 3′-terminal oligomeric U stretches as the C. fasciculata 65-kDa protein, whereas the binding specificity of a 26-kDa protein resembled that of the C. fasciculata 88-kDa and 30-kDa proteins. The possible involvement of the proteins in the editing process is discussed.

**Keywords.** RNA editing; guide RNA-binding proteins; kinetoplast; UV cross-linking.
gRNA-protein complexes have also been characterized by native-gel-retardation assays in extracts of L. tarentolae (Peris et al., 1994) and T. brucei (Göringer et al., 1994; Read et al., 1994). Moreover, cross-linking experiments have identified several proteins which associate with gRNA and pre-mRNA in T. brucei (Köllner et al., 1994; Read et al., 1994), which has led to the realisation that a 90-kDa protein from T. brucei binds specifically to the gRNA oligo(U) tail. To date, however, no clear role in the editing process could be ascribed to this or any other protein, and their relevance remains obscure.

We are in the process of identifying and characterizing proteins that may play a role in RNA editing in the insect trypanosomatid C. fasciculata and have initiated the analysis of proteins that display high affinity for gRNAs in UV cross-linking assays. In this report, we describe the identification of proteins of 88, 65 and 30 kDa that can be cross-linked to gRNA. We show that this binding solely depends on the presence of the gRNA U tail but that the binding is not specific for gRNA since the three proteins also bind to U-tailed mitoribosomal rRNAs and to free oligo(U). The 65-kDa protein is particularly interesting in that it has no affinity for U-rich control RNAs, not even if these possess an internal homopolymeric U stretch. In this paper, we also present a detailed comparison between the binding specificity of C. fasciculata proteins and that of their counterparts from T. brucei.

MATERIALS AND METHODS

Cell growth and production of mitochondrion fractions. C. fasciculata was grown in batches of 10 l as described by Kleisen et al. (1975), with shaking and aeration to a density of 2×10^7 cells/ml. Mt vesicles were isolated according to the method of Birkenmeyer and Ray (1986). The T. brucei procyclic form was grown as described (Hoellmijer et al., 1981); mt vesicle isolation was essentially as described by Feagin et al. (1987).

In vitro RNA synthesis. cdnas of gRNAs and pre-mRNA were cloned behind the T7 promoter (Arts et al., 1993 and unpublished results). The DNA clones were linearized with appropriate restriction enzymes (DraI for ND7 [FS] and [S]gRNA; Apol to obtain a version of ND7 [FS]gRNA that has no U tail; BamHI for chimeric RNA; HindIII for vector-derived RNA and SnaBI for the pre-mRNA fragment). T7 RNA polymerase reactions (15 μl) contained DNA template (1 μl respective plasmid), 40 mM Tris/HCl, pH 7.9, 6 mM MgCl₂, 2 mM spermidine, 1 mM spermine, 10 mM dithiothreitol, 0.5 mM each ATP, GTP and CTP, 0.1 mM EDTA, 10 mM dithiothreitol, 0.5 mM each ATP, GTP and CTP, 0.1 mM EDTA.

Oligo(U) was eluted from gel slices in double-distilled water at 60°C for 2 h and recovered by phenol extraction and ethanol precipitation. The size of eluted oligo(U) was determined by electrophoresis on a 8% polyacrylamide gel in TBE with 5 μg/ml ethidium-bromide. The RNA was extracted and purified by standard tech-
RESULTS

UV cross-linking of mt proteins to synthetic RNA substrates.

Recently, UV cross-linking has evolved as a powerful technique to identify RNA-binding proteins (Pifhol-Roma et al., 1989). Since our objective is to identify gRNA-binding proteins from mitochondria of C. fasciculata, we have applied this method to mt lysates to which synthetic, uniformly radiolabelled gRNA derivatives and control RNAs were added, obtained from in vitro transcription. UV cross-linking was carried out in the presence of 150 mM KCl to minimize low-affinity protein-RNA interactions. The cross-linked complexes were treated with RNase A to minimize the retarding effect of the RNA during gel electrophoresis. For comparison, the experiments were also performed with mt extracts from T. brucei.

The results of representative experiments are shown in Fig. 1. In Fig. 1, lanes 1 and 7, cross-linking patterns are shown obtained with a synthetic version of the gRNA involved in the editing of the frameshift region (FS) of C. fasciculata ND7 [FS]gRNA that has no 3' U tail (lanes 3), a chimeric molecule of ND7 [FS]gRNA and part of the ND7 pre-mRNA (lanes 4), a fragment of ND7 pre-mRNA (lanes 5) and a vector transcript of 65 nucleotides (lanes 6). (A) Lane 8 ND7 [FS]gRNA was incubated with an aliquot of total cellular extract from C. fasciculata. Details of the substrate RNA preparation and sequence and of the UV cross-linking procedure are given in Materials and Methods.

Fig. 1. UV cross-linking of mt proteins to synthetic (g)RNAs. Aliquots of mitochondrial lysate of C. fasciculata (A) and T. brucei (B) were incubated with radiolabelled synthetic C. fasciculata ND7 [FS]gRNA (lanes 1 and 7), ND7 [FS]gRNA (lanes 2), a truncated ND7 [FS]gRNA that has no 3' U tail (lanes 3), a chimeric molecule of ND7 [FS]gRNA and part of the ND7 pre-mRNA (lanes 4), a fragment of ND7 pre-mRNA (lanes 5) and a vector transcript of 65 nucleotides (lanes 6). (A) Lane 8 ND7 [FS]gRNA was incubated with an aliquot of total cellular extract from C. fasciculata. Details of the substrate RNA preparation and sequence and of the UV cross-linking procedure are given in Materials and Methods.
ly(U), poly(C), unlabelled ND7 [FS]gRNA or total yeast RNA (Fig. 2) and of poly(A) and poly(G). The addition of poly(U) in a 10-fold excess over labelled gRNA indeed significantly reduced gRNA binding to the 65-kDa protein of *C. fasciculata* (Fig. 2A), comparable to the effect of similar amounts of unlabelled ND7 [FS]gRNA, as determined from quantification with a PhosphorImager apparatus. In contrast, the addition of a 1000-fold excess of other competitor RNAs only had a marginal effect (yeast RNA) or none at all [poly(C)] (Fig. 2A). The effects of the addition of poly(U), unlabelled ND7 [FS]gRNA and yeast RNA on the binding of radiolabelled gRNA to the 30-kDa and the 88-kDa polypeptides of *C. fasciculata* were essentially similar to those observed for the 65-kDa protein, although the results varied slightly between experiments. The effect of poly(C) addition was markedly different, though, since for the 30-kDa and the 88-kDa protein, only 3-10-fold more poly(C) than poly(U) was required to obtain approximately the same reduction of binding. Finally, although poly(A) and poly(G) also behaved as relatively inefficient competitors, substantially smaller amounts of these homopolymers than of total yeast RNA were required for complete inhibition of gRNA binding to any of these proteins (10–20-fold less depending on the experiment, data not shown). This result does not necessarily contrast with the hypothesis that the proteins in question preferentially bind to U stretches, given the fact that both poly(A) and poly(G) can basepair with gRNA U tails, which could interfere with their capacity to interact with proteins.

The effect of addition of the different competitor RNAs in the *T. brucei* system is shown in Fig. 2B. The 83-kDa polypeptide from *T. brucei* appeared to be the counterpart of the 88-kDa protein of *C. fasciculata* 65-kDa protein, since its binding to labelled gRNA was also efficiently competed away by the addition of poly(U) and ND7 [FS]gRNA, albeit that higher amounts of competitor were required, and not by poly(C) or yeast RNA (Fig. 2B). The minor bands of 64-kDa and 55-kDa were affected by the addition of the competing RNAs in a similar fashion. The binding specificity of the 26-kDa and 42-kDa proteins, in contrast, resembled that of the *C. fasciculata* 30-kDa and 88-kDa proteins, the most prominent characteristic being the fact that poly(C) is only a marginally less efficient competitor for gRNA binding than poly(U).

**UV cross-linking to oligo(U).** From the results obtained so far, we conclude that trypanosome gRNA-binding proteins interact with oligo(U) sequences, albeit with a different degree of specificity. In order to investigate whether oligo(U) by itself is sufficient for binding, the UV cross-linking assay was performed with gel-purified, 5’-end-labelled poly(U) RNA of varying size. The results of this experiment (Fig. 3) demonstrate that the 65-kDa protein from *C. fasciculata* and the 83-kDa and 64-kDa proteins from *T. brucei* efficiently bind to an oligo(U) size class of 23–29 residues or more (Fig. 3, lanes 2–5). The decrease of the signal observed with larger RNA species is most likely due to the fact that the labelled 5’ end of these longer RNAs is not efficiently protected by bound protein against treatment with RNase A. The 30-kDa and the 83-kDa proteins from *C. fasciculata* bound strongly only to oligo(U)26–26 (Fig. 3, lane 4), oligo(U)55–55 giving a weak signal (Fig. 3, lane 3), while the signal for these proteins also decreases in intensity with longer RNAs. Finally, the *T. brucei* 55–42- and 26-kDa proteins displayed weak binding only to oligo(U)26–26.

These results show that all candidate oligo(U)-binding proteins can indeed be cross-linked to radiolabelled oligo(U) in the presence of an excess amount of yeast RNA.
UV cross-linking to endogenous, metabolically labelled mtRNAs. The experiments described thus far have been performed with in-vitro-produced gRNAs and other synthetic substrates. In order to establish whether the polypeptides that bind to these gRNAs also bind to endogenous gRNAs and other mtRNAs that possess U tails, we applied the UV cross-linking technique to intact mt vesicles from *C. fasciculata*. RNAs present in these vesicles become labelled by the action of endogenous TUTase when [α-32P]UTP is added after the depletion of the endogenous nucleotide pool by an incubation step, essentially as described by Bakalara et al. (1989) for *L. tarentolae* and Harris et al. (1990) for *T. brucei*. The labelling patterns obtained are shown in Fig. 4A. No labelling of RNA is found with [α-32P]GTP under these conditions, which confirms that transcription is arrested (data not shown). A large part of the label is incorporated into RNA species of 60–120 nucleotides, migrating in a pattern very similar to that of gRNAs from *L. tarentolae, T. brucei* and *T. equiperdum* (Bakalara et al., 1989; Harris et al., 1990; Blum and Simpson, 1990; Pollard and Hadiuk, 1991). Interestingly, the size of these RNAs (and therefore the length of the U tail added) is proportional to the UTP concentration used (Fig. 4A), suggesting that, in vivo, the UTP concentration may be one of the determinants of the length of the U tail of gRNAs. As in the other trypanosomatids, under these conditions two large RNAs which comigrate with mt 9S and 12S RNA when electrophoresed on a native agarose gel (data not shown), also become labelled. It has indeed been reported that mt 9S and 12S rRNA of *T. brucei* possess a U tail of 11 and 2–17 residues, respectively (Adler et al., 1991).

To investigate whether mt proteins are associated with the labelled RNAs, the vesicles were irradiated with UV light, lysed with Triton and treated with RNase A followed by gel electrophoresis and autoradiography (Fig. 4B, lane 1). For comparison, proteins UV cross-linked to synthetic ND7[67][g]RNA in our standard assay system were electrophoresed in Fig. 4B lane 4, whereas in lane 3, samples as applied to lanes 2 and 4 were mixed. The results demonstrate that the 88-, 65- and 30-kDa proteins become labelled by this procedure. Slight shifts in the electrophoretic mobility observed when comparing lanes 2 and 4 (Fig. 4B) are due to differences in salt concentrations of the samples (Fig. 4B, lane 3). A third prominent band of 60-kDa was also labelled when [α-32P]GTP was used (Fig. 4B, lane 1) and already present when the RNA was not treated with RNase A (not shown). This band probably represents a nucleotide-binding protein. A fourth band of approximately 54-kDa was not formed reproducibly. Finally, gel-purified metabolically labelled RNAs were added to standard assay systems. As shown in Fig. 4B, lane 7, *C. fasciculata* proteins of 88 (visible upon long exposure of the autoradiogram), 65 and 30-kDa can indeed be UV cross-linked to gRNA isolated from the 60–80-nucleotide region of the gel obtained from metabolic labelling at the lowest UTP concentration (see Fig. 4B, lane 1). However, the same bands are radioactively labelled when RNAs isolated from the 12S and 9S region were added (Fig. 4B, lanes 5 and 6), indicating that these proteins also have affinity for the mt RNAs (see discussion).

**DISCUSSION**

Proteins in mt extracts from *C. fasciculata* and *T. brucei* that bind to oligo(U). We have studied the binding specificity of proteins of 88, 65 and 30 kDa from mitochondria of *C. fasciculata* and of 83, 64, 55, 42 and 26 kDa from *T. brucei* that can be UV cross-linked to synthetic gRNAs from *C. fasciculata*. The results clearly show that the 65-kDa protein of *C. fasciculata* and the 83-kDa and 64-kDa proteins of *T. brucei* are oligo(U)-binding proteins and that short stretches of U are sufficient for binding (Figs. 1 – 3). Moreover, the failure of these proteins to bind to chimeric gRNA:pre-mRNA, in spite of the presence of an internal stretch of 13 U residues (Fig. 1), suggests that the U residues must have a 3'-terminal location. In addition, we showed that the *C. fasciculata* 65-kDa protein efficiently cross-links to 12S and 9S mitoribosomal RNAs that were post-transcriptionally labelled by incubating mt vesicles with [α-32P]UTP (Fig. 4). It has been reported that mt 9S and 12S rRNA from *T. brucei* possess 3' terminal U tails of 11 nucleotides and 2–17 residues, respectively (Adler et al., 1991) and, in *C. fasciculata*, tails of a similar length can be expected. The length of the tails of the RNAs that we used in the experiment was not measured, but it is most likely not dissimilar from that in vivo, considering the fact that low UTP concentrations were used in the labelling procedure that did result in gRNAs of normal size (see Fig. 4A, lane 1). Together, these observations strongly suggest that the presence of a short 3' terminal U tail is by itself sufficient for binding of this class of proteins to an RNA and that no other sequences are involved.

The RNA-binding specificities and efficiencies of the 65-kDa protein of *C. fasciculata* and the 83-kDa protein of *T. brucei* are...
are very similar, making it attractive to assume that they are homologues. If so, their different size is somewhat surprising given the overall similarity of the editing processes in these closely related trypanosomatids (for evolutionary distance between trypanosomatids, see Fernandes et al., 1993; Landweber et al., 1994; Maslov et al., 1994; Lukes et al., 1994). Interestingly, the \textit{T. brucei} protein of 64-kDa has binding characteristics identical to those of the 83-kDa protein, although the amount of cross-linked protein is considerably lower. These observations could reflect species-specific differences in the number and size of these oligo(U)-binding proteins, but it could also be hypothesized that in both trypanosomatids a shorter version is generated by proteolysis of a longer form with a considerable species-specific difference in the extent of the cleavage reaction. If so, it is not likely that the 88-kDa protein of \textit{C. fasciculata} is the longer form of the 65-kDa protein, given its rather different binding specificity. Instead, one would have to assume that the long form is virtually completely cleaved in this organism. Detailed characterisation and (gene)sequence analysis of all the relevant proteins would be required, however, to settle this issue.

Also, the \textit{C. fasciculata} 88-kDa and 30-kDa proteins and the \textit{T. brucei} 26-kDa protein bind to gRNAs with a U tail and to oligo(U), although longer U stretches are required for efficient binding (Figs. 1 and 3). In addition, they bind to other RNAs provided that they are U rich. The ensuing suggestion that U residues, albeit not necessarily 3'-terminally located, are essential for binding of RNA to these proteins also, is supported by the competition experiments shown in Fig. 2. Clearly, however, they are a bit more promiscuous in their binding specificity since poly(C) is an only slightly less efficient competitor than poly(U). Finally, most of the binding characteristics of the \textit{T. brucei} 55-kDa and 42-kDa proteins resemble that of the 26-kDa protein. It is difficult, however, to assess their binding specificity with more precision, due to the fact that the amounts of these proteins that can be cross-linked are rather low and the intensity of the signals varies between experiments (compare Figs 1 and 2).

\textbf{Proteins that bind to homologous \textit{T. brucei} gRNAs.} The \textit{T. brucei} proteins described in this paper that have a high affinity for (heterologous) \textit{C. fasciculata} gRNAs, almost certainly correspond to proteins of similar size that have been recently identified by other groups as high-affinity gRNA-binding proteins in assays utilizing homologous \textit{T. brucei} gRNAs. For example, there can be little doubt that our 83-kDa protein is identical to the 90-kDa protein described by Read et al. (1994) and Köller et al. (1994), the difference in the estimated size resulting from differences in the gel systems used. Like the 83-kDa protein, the 90-kDa protein strongly binds to gRNA at high KCl concentrations, its binding depends on the presence of a U tail and poly(U) is an efficient competitor. The fact that identical results can be obtained with heterologous \textit{C. fasciculata} gRNAs lend further support to the assumption that the U tail by itself is both essential and sufficient for binding of a gRNA to this protein. The suggestion made by Köller et al. (1994) that a tail of one U residue would suffice for binding to the 90-kDa and other proteins is not supported by our data, however, because a truncated \textit{C. fasciculata} ND7[FS]gRNA that has two U residues at its 3' end does not bind. It could be possible that the truncated gRNA of 43 nucleotides that we used is too small to be recognized, but we consider this to be unlikely given the fact that oligo(U) of 23–29 nucleotides, is already efficiently bound. The apparent discrepancy may result from the fact that Köller et al. (1994) used an ill-defined gRNA preparation, obtained by reducing the UTP concentration in the \textit{in vitro} synthesis reaction mixtures, in which the presence of gRNA species with longer U tails was not rigorously excluded.

In the lower molecular-mass range, Read et al. (1994) and Köller et al. (1994) find high-affinity gRNA-binding proteins with sizes and binding characteristics similar to those of our 64-, 55- and 42-kDa proteins. It is not immediately obvious, however, that our 26-kDa protein corresponds to the 21-kDa protein of Köller et al. (1994) and the 25-kDa protein of Read et al. (1994). Although in all cases the protein is the major high-affinity gRNA-binding protein in that size range, differences in binding specificity appear to exist. For example, the other groups report that the binding of gRNA to the 21-kDa and 25-kDa proteins is only marginally affected by the removal of the gRNA U tail or not at all, which is in contrast with our finding that the protein does not bind to a tailless ND7[FS]gRNA. In the hands of Köller et al. (1994), nevertheless, poly(U), but not poly(A), is still an efficient competitor, which is essentially what we found for the 26-kDa protein. Although we cannot exclude the possibility that we are dealing with different proteins, we prefer the view that the use of gRNAs with a different sequence and a considerable difference in (tail) length may account for most of the apparent discrepancies: whereas Köller et al. (1994) used a 62-nucleotide \textit{T. brucei} ATPase subunit 6 gRNA with 3–9 U residues at the 3' end, we used a 50-nucleotide \textit{C. fasciculata} ND7[FS]gRNA with 23 U residues. Most of the data could be explained by assuming that the short U tail of the ATP6 gRNA is not essential for the formation of the binding motif recognized by this protein and that other, as yet unidentified, pyrimidine-rich regions can take its place. In line with this, our 26-kDa protein also binds to RNAs that do not possess long homopolymeric U stretches (Fig. 1B); it binds to oligo(U) rather weakly (Fig. 3) and it also has affinity for poly(C) (Fig. 2). Clearly, more work is required to identify with more precision the RNA sequence elements that determine the binding of this protein.

\textbf{Concluding remarks.} Evidently, affinity for gRNA U stretches is, by itself, no guarantee that the proteins described in this work have a role in the editing process and more experimental evidence is required. Interestingly, glycerol gradient analysis of mt extracts from \textit{C. fasciculata} indicates that the 30-kDa and 65-kDa proteins are present in high molecular-mass complexes, which also contain gRNA but no RNA (data not shown). It has further been reported that the \textit{T. brucei} 25-kDa and 90-kDa proteins associate with gRNP complexes \textit{in vitro} (Read et al., 1994). Oligo(U)-binding proteins therefore appear to be part of large, gRNA-containing (editosomal?) complexes. In addition, it should be emphasized that if the editing machinery contains proteins that recognize all gRNAs, the U tail is likely to be at least a part of the binding site since it is the only sequence element that gRNAs have in common. All these observations make the trypanosomatid oligo(U)-binding mt proteins prime candidates for further investigation. One of our current research projects is, therefore, to purify and further characterize these proteins and the mt components (proteins/RNAs) with which they interact.

We thank Gert Jan Arts, Daniëlle Blom, Janne van den Burg, Paulette Groenewelt, Annet de Haan, Paul Stoof and Rob Woltuis for their contributions at various stages of the work and for stimulating discussions. The research is supported by the Netherlands Foundation for Chemical Research (SON) and the Foundation for Medical and Health Research (MW), which are subsidized by the Netherlands Foundation for Scientific Research (NWO).

\textbf{REFERENCES}

Adler, B. K., Harris, M. E., Bertrand, K. I. & Hajduk, S. L. (1991) Modification of Trypanosoma brucei mitochondrial rRNA by post-


Simpson, L., Maslov, D. & Blum, B. (1993) In RNA editing, the alteration of codon sequences of RNA (Benne, R., ed.) pp. 53–85, Ellis Horwood, Chichester, UK.


